Premier Reference Source

AI Innovation in Medical Imaging Diagnostics

Kalaivani Anbarasan

IGI Global
DISSEMINATOR OF KNOWLEDGE

AI Innovation in Medical Imaging Diagnostics

Kalaivani Anbarasan
Department of Computer Science and Engineering, Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

A volume in the Advances in Medical Technologies and Clinical Practice (AMTCP) Book Series

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

Published in the United States of America by
IGI Global
Medical Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Names: Anbarasan, Kalaivani, 1975- editor.
Title: AI innovation in medical imaging diagnostics / Kalaivani Anbarasan, editor.
Description: Hershey, PA : Medical Information Science Reference, [2020] | Includes bibliographical references and index. | Summary: "This book examines the application of artificial intelligence in medical imaging diagnostics"-- Provided by publisher.
Identifiers: LCCN 2019049195 (print) | LCCN 2019049196 (ebook) | ISBN 9781799830924 (hardcover) | ISBN 9781799830931 (ebook)
Subjects: MESH: Artificial Intelligence | Diagnostic Imaging--methods | Deep Learning | Machine Learning
Classification: LCC RC78.7.D53 (print) | LCC RC78.7.D53 (ebook) | NLM W 26.55.A7 | DDC 616.07/54--dc23
LC record available at https://lccn.loc.gov/2019049195
LC ebook record available at https://lccn.loc.gov/2019049196

This book is published in the IGI Global book series Advances in Medical Technologies and Clinical Practice (AMTCP) (ISSN: 2327-9354; eISSN: 2327-9370)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

All work contributed to this book is new, previously-unpublished material.
The views expressed in this book are those of the authors, but not necessarily of the publisher.

For electronic access to this publication, please contact: eresources@igi-global.com.

Advances in Medical Technologies and Clinical Practice (AMTCP) Book Series

ISSN:2327-9354
EISSN:2327-9370

Editor-in-Chief: Srikanta Patnaik, SOA University, India; Priti Das, S.C.B. Medical College, India

MISSION

Medical technological innovation continues to provide avenues of research for faster and safer diagnosis and treatments for patients. Practitioners must stay up to date with these latest advancements to provide the best care for nursing and clinical practices.

The **Advances in Medical Technologies and Clinical Practice (AMTCP) Book Series** brings together the most recent research on the latest technology used in areas of nursing informatics, clinical technology, biomedicine, diagnostic technologies, and more. Researchers, students, and practitioners in this field will benefit from this fundamental coverage on the use of technology in clinical practices.

COVERAGE

- Clinical Data Mining
- Biomedical Applications
- Nutrition
- Clinical Studies
- Nursing Informatics
- Biomechanics
- Clinical Nutrition
- Clinical High-Performance Computing
- Biometrics
- Telemedicine

IGI Global is currently accepting manuscripts for publication within this series. To submit a proposal for a volume in this series, please contact our Acquisition Editors at Acquisitions@igi-global.com or visit: http://www.igi-global.com/publish/.

The Advances in Medical Technologies and Clinical Practice (AMTCP) Book Series (ISSN 2327-9354) is published by IGI Global, 701 E. Chocolate Avenue, Hershey, PA 17033-1240, USA, www.igi-global.com. This series is composed of titles available for purchase individually; each title is edited to be contextually exclusive from any other title within the series. For pricing and ordering information please visit http://www.igi-global.com/book-series/advances-medical-technologies-clinical-practice/73682. Postmaster: Send all address changes to above address.

Titles in this Series

For a list of additional titles in this series, please visit: http://www.igi-global.com/book-series

Optimizing Health Monitoring Systems With Wireless Technology
Nilmini Wickramasinghe (Swinburne University of Technology, Australia & Epworth HealthCare, Australia)
Medical Information Science Reference • © 2021 • 338pp • H/C (ISBN: 9781522560678) • US $265.00

Deep Learning Applications in Medical Imaging
Sanjay Saxena (International Institute of Information Technology, India) and Sudip Paul (North-Eastern Hill University, India)
Medical Information Science Reference • © 2021 • 274pp • H/C (ISBN: 9781799850717) • US $245.00

Design and Quality Considerations for Developing Mobile Apps for Medication Management Emerging Research and Opportunities
Kevin Yap (La Trobe University, Australia) Eskinder Eshetu Ali (Addis Ababa University, Ethiopia) and Lita Chew (National University of Singapore, Singapore)
Medical Information Science Reference • © 2021 • 256pp • H/C (ISBN: 9781799838326) • US $225.00

Expert Approaches to Health IT Tools in Clinical Practice
Ramgopal Kashyap (Amity University, Raipur, India)
Medical Information Science Reference • © 2021 • 300pp • H/C (ISBN: 9781799840510) • US $245.00

Opportunities and Challenges in Digital Healthcare Innovation
Kamaljeet Sandhu (University of New England, Australia)
Medical Information Science Reference • © 2020 • 261pp • H/C (ISBN: 9781799832744) • US $285.00

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

701 East Chocolate Avenue, Hershey, PA 17033, USA
Tel: 717-533-8845 x100 • Fax: 717-533-8661
E-Mail: cust@igi-global.com • www.igi-global.com

Table of Contents

Detailed Table of Contents

healthcare helps to analyze thousands of different data points and suggest outcomes, provide timely risk factors, optimize resource allocation. Machine learning is playing a critical role in patient care, billing processing to set the target to marketing and sales team, and medical records for patient monitoring and readmission, etc. Machine learning is allowing healthcare specialists to develop alternate staffing models, intellectual property management, and using the most effective way to capitalize on developed intellectual property assets. Machine learning approaches provide smart healthcare and reduce administrative and supply costs. Today healthcare industry is committed to deliver quality, value, and satisfactory outcomes.

Brain tumor discovery and its segmentation from the magnetic resonance images (MRI) is a difficult task that has convoluted structures that make it hard to section the tumor with MR cerebrum images, different tissues, white issue, gray issue, and cerebrospinal liquid. A mechanized grouping for brain tumor location and division helps the patients for legitimate treatment. Additionally, the method improves the analysis and decreases the indicative time. In the separation of cerebrum tumor, MRI images would focus on the size, shape, area, and surface of MRI images. In this chapter, the authors have focused various supervised and unsupervised clustering techniques for identifying brain tumor and separating it using convolutional neural network (CNN), k-means clustering, fuzzy c-means grouping, and so on.

In this chapter, the authors used autoencoder in data preprocessing step in an attempt to improve image representation, consequently increasing classification performance. The authors applied autoencoder to the task of breast lesion classification in mammographic images. Image Retrieval in Medical Applications (IRMA) database

mammography. The chapter focuses on automatic feature selection algorithm for diagnosis of women breast cancer from digital mammographic images achieved through multi-layer perceptron techniques.

Medical images stored in distributed and centralized servers are referred to for knowledge, teaching, information, and diagnosis. Content-based image retrieval (CBIR) is used to locate images in vast databases. Images are indexed and retrieved with a set of features. The CBIR model on receipt of query extracts same set of features of query, matches with indexed features index, and retrieves similar images from database. Thus, the system performance mainly depends on the features adopted for indexing. Features selected must require lesser storage, retrieval time, cost of retrieval model, and must support different classifier algorithms. Feature set adopted should support to improve the performance of the system. The chapter briefs on the strength of local binary patterns (LBP) and its variants for indexing medical images. Efficacy of the LBP is verified using medical images from OASIS. The results presented in the chapter are obtained by direct method without the aid of any classification techniques like SVM, neural networks, etc. The results prove good prospects of LBP and its variants.

This chapter contributes to the study of uncertainty of signal dimensions within a microscopic image of blood sample. Appropriate colocalization indicator classifies the leukocytes in the region of interest having ragged boundaries. Signal transduction has been interpreted using correlation function determined fluorescence intensity in proposed work using just another colocalization plugin (JaCoP). Dependence between the channels in the colocalization region is being analysed in a linear fashion using

Pearson correlation coefficient. Manders split, which gives intensity, is represented in a channel by co-localizing pixels. Overlap coefficients are also being analysed to analyse coefficient of each channel. Li's intensity correlation coefficient is being used in specific cases to interpret the impact of staining.

G. Durgadevi, New Prince Shri Bhavani College of Engineering and Technology, India
K. Sujatha, Dr. M. G. R. Educational and Research Institute, India
K.S. Thivya, Dr. M.G.R. Educational and Research Institute, India
S. Elakkiya, Dr. M.G.R. Educational and Research Institute, India
M. Anand, Dr. M.G.R. Educational and Research Institute, India
S. Shobana, New Prince Shri Bhavani College of Engineering and Technology, India

Magnetic resonance imaging is a standard modality used in medicine for bone diagnosis and treatment. It offers the advantage to be a non-invasive technique that enables the analysis of bone tissues. The early detection of tumor in the bone leads on saving the patients' life through proper care. The accurate detection of tumor in the MRI scans are very easy to perform. Furthermore, the tumor detection in an image is useful not only for medical experts, but also for other purposes like segmentation and 3D reconstruction. The manual delineation and visual inspection will be limited to avoid time consumption by medical doctors. The bone tumor tissue detection allows localizing a mass of abnormal cells in a slice of magnetic resonance (MR).

M. P. Chitra, Panimalar Institute of Technology, India
R. S. Ponmagal, SRM Institute of Science and Technology, India
N. P. G. Bhavani, Meenakshi College of Engineering, India
V. Srividhya, Meenakshi College of Engineering, India

Cloud computing has become popular among users in organizations and companies. Security and efficiency are the two major problems facing cloud service providers and their customers. Cloud data allocation facilities that allow groups of users to work together to access the shared data are the most standard and effective working styles in the enterprises. So, in spite of having advantages of scalability and flexibility, cloud storage service comes with confidential and security concerns. A direct method to defend the user data is to encrypt the data stored at the cloud. In this research work, a secure cloud model (SCM) that contains user authentication and data scheduling approach is scheduled. An innovative digital signature with chaotic secure hashing (DS-CS) is used for user authentication, followed by an enhanced work scheduling

much earlier stage than traditional methods or self-examination.

Preface

Recent advancements in the technology of medical imaging, such as CT and MRI scanners, are making it possible to create more detailed 3D and 4D images. These powerful images require vast amounts of digital data to help with the diagnosis of the patient. Artificial intelligence (AI) must play a vital role in supporting with the analysis of this medical imaging data, but it will only be viable as long as healthcare professionals and AI interact to embrace deep thinking platforms such as automation in the identification of diseases in patients.

AI Innovation in Medical Imaging Diagnostics is an essential reference source that examines AI applications in medical imaging that can transform hospitals to become more efficient in the management of patient treatment plans through the production of faster imaging and the reduction of radiation dosages through the PET and SPECT imaging modalities. The book also explores how data clusters from these images can be translated into small data packages that can be accessed by healthcare departments to give a real-time insight into patient care and required interventions. Featuring research on topics such as assistive healthcare, cancer detection, and machine learning, this book is ideally designed for healthcare administrators, radiologists, data analysts, computer science professionals, medical imaging specialists, diagnosticians, medical professionals, researchers, and students.

Chapter 1 detects Ocular Pathologies from Iris Images using two algorithms and their results are also given in a graphical representation. The two algorithms used for the research study are Blind De-convolution and Fuzzy-C means Clustering.

Chapter 2 deals with the basics of machine learning techniques and different algorithm and their application in Healthcare system. The top applications of machine learning in healthcare are specified and ways to assess the growth of the healthcare industry in 2019 and beyond.

Chapter 3 reports about the various supervised and unsupervised techniques for brain tumor detection and segmentation such as K-nearest neighbor (K-NN), K-means clustering, and also morphological operator and also specified the experimental results.

Chapter 4 performs the Breast Cancer Diagnosis in Mammograms Using Wavelet Analysis, Haralick Descriptors and Auto encoder. The proposed work produced improved accuracy with comparison to the state-of-art techniques. Chapter 5 covers the Feature Selection for classification of Lung CT Images using Random Forest Algorithm. The proposed work produced better results than the existing techniques.

Chapter 6 focused on automatic feature selection algorithm for diagnosis of women breast cancer from digital mammographic images and breast cancer classification are achieved through multi-layer perceptron techniques. The outcome of the paper reduced false positive rate and improved diagnosis accuracy at a greater extent.

Chapter 7 deals with retrieval of medical images based on Content Based Approach to Medical Image Retrieval. In Chapter 8, correlation and analysis of overlapping leukocytes in blood cell images using intracellular markers and colocalization operation are discussed with their experimental results.

Chapter 9 covers the enchodroma Tumor Detection from MRI Images using SVM Classifier.

Medical Images captured through different image modality are to be stored in a cloud storage for the researchers to utilize the images for their research work. So, the approach to Cloud Computing for Medical Image Analysis well discussed in Chapter 10.

In Chapter 11 discuss the segmentation Of Spine Tumour using K-Means and Active Contour and also explains how the feature extraction can be done using Gray Level Co-occurrence Matrix.

The emphasis of Chapter 12 is to give a solution of low cost, accurate, automated, portable cancer screening tool that can be operated by a simple clinician. Unlike mammography, our imaging model can be radiation free, non-touch, not painful and works for women of all ages.

We are very much grateful to the authors and reviewers for their excellent contributions for making this book possible. Our special thanks to IGI Global, Lindsay Wertman and support teams especially for their excellent collaborations. I pledge my gratitude to my college honorable administrators Director, Mrs. Ramya Deepak and Director Academics Dr. V. Deepak for their fullest support and cooperation for the completion of this proposal. I also take this opportunity to thank my Principal, Dr. Ramesh and Dr. S P. Chokkalingam, Program Director for their kind support and motivation throughout the proposal. I also extend my heartfelt thanks to Dr. K. Sujatha, Professor from Dr. MGR University who helped me throughout this journey.

This edited book covers the automatic computer aided diagnosis of spine tumor, women breast cancer, brain tumor, Ocular Pathologies from Iris Images. .Being an interdisciplinary book, we hope it will be useful to a wide variety of readers and will provide useful information to professors, researchers and students.

Chapter 1
Detection of Ocular Pathologies From Iris Images Using Blind De-Convolution and Fuzzy C-Means Clustering:
Detection of Ocular Pathologies

Sujatha Kesavan
Dr. M. G. R. Educational and Research Institute, India

Rajeswary Hari
Dr. M. G. R. Educational and Research Institute, India

Kanya N.
Dr. M. G. R. Educational and Research Institute, India

Karthikeyan V.
Dr. M. G. R. Educational and Research Institute, India

Shobarani R.
Dr. M. G. R. Educational and Research Institute, India

ABSTRACT

The images of disease-affected and normal eyes collected from high-resolution fundus (HRF) image database are analyzed, and the influence of ocular diseases on iris using a reliable fuzzy recognition scheme is proposed. Nearly 45 samples of iris images are acquired using Canon CR-1 fundus camera with a field of view of 45° when subjected to routine ophthalmology visits, and the samples of eye images include healthy eyes, eyes affected by glaucoma, cataract, and diabetic retinopathy. These

DOI: 10.4018/978-1-7998-3092-4.ch001

images are then subjected to various image processing techniques like pre-processing for de-noising using blind de-convolution, wavelet-based feature extraction, principal component analysis (PCA) for dimension reductionality, followed by fuzzy c-means clustering inference scheme to categorize the normal and diseased eyes. It is inferred that the proposed method takes only two minutes with an accuracy, specificity, and sensitivity varying in the range of 94% to 98%, respectively.

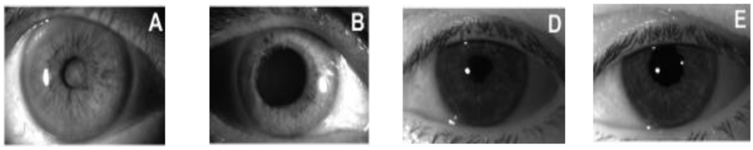
INTRODUCTION

The most accurate method for biometric authentication is iris recognition and is most impressive worldwide, which results in creation of the distinctive identification numbers for the people in India using ADHAAR (Dhooge & de Laey, 1989), or Canadian border control system CANPASS (Roizenblatt et al., 2004). Like any other organ in the human body, the eyes and iris may suffer from various diseases like cataract, acute glaucoma, posterior and anterior synechiae, retinal detachment, rubeosis iridis, corneal vascularization, corneal grafting, iris damage and atrophy and corneal ulcers, haze or opacities. The eye pathologies are separated into five groups based on the impact on iris recognition: 1) healthy without impact), 2) illness detected but still clear and unaffected iris unaffected 3) geometric distortion 4) distortion in iris tissue and 5) obstruction in iris tissue (Aslam et al., 2009; Borgen et al., 2009; Dhir et al., 2010; ISO/IEC 19794-6:2011, 2011; Monro et al., 2009; Rajendra Acharya, 2011; Yuan et al., 2007).

MIRLIN, VeriEye and OSIRIS are the three methods used for iris recognition which is used to find the difference in the average value of the comparison scores inferred between the healthy and disease affected eyes. The comparison scores generated for the disease infected eyes as compared with healthy eyes is not within the tolerable limit when these conventional schemes are used. Variation in the comparison score may mislead in false non-match rate (Budai et al., 2013; McConnon et al., 2012; Neurotechnology, 2012; Odstrcilik et al., 2013; Seyeddain et al., 2014; Smart Sensors Ltd, 2013).

The various ocular diseases were detected using the database. The symptoms and the effects of various ophthalmic disorders are discussed here. Cataract is the common ophthalmic disorder indentified worldwide. The effect of this disease includes blurring of the eye lens causing reduced vision, Figure 1A. This eye disease occurs due to thickening of cornea which prevents the light from entering the lens thereby inhibiting the vision (Aggarwal & Khare, 2015; Canadian Border Services Agency, 2015; Haleem et al., 2015; Sutra et al., 2013; Trokielewicz et al., 2014; Unique Identification Authority of India, n.d.).

Figure 1. Eye Disorders
(Courtesy Department of Ophthalmology of the Medical University of Warsaw).



A). Cataract B). acute glaucoma C). Posterior Synechiae D). Anterior Synechiae

The second kind of eye disorder is acute glaucoma which causes reduction in the space between iris and cornea closing the boundary of the iris on the outer side, hindering the flow of aqueous humor through the trabecular mesh work leading to drastic increase in ocular pressure resulting in loss of vision as in Figure 1B. The third kind is called as Posterior and anterior synechiae which occurs when the iris is partly attached to the lens or to the cornea. This changes the shape of the pupil, with deviation in circular shape, as in Figure 1C and 1D.

Diabetic retinopathy in Figure 2(a), results due to insulin disorders causing diabetes. The blood vessels in the light sensitive region retina are affected. It is because of insufficient supply of oxygen leading to blindness. If this eye disorder is diagnosed at early stage proper treatment can be given preventing blindness. The two major types of retinopathy are non-proliferative and proliferative retinopathy. The less severe type is non-proliferative retinopathy which causes hemorrhage in the retina. This produces a leak in blood serum making the retina wet which leads to diminished vision. The severe type is Proliferative retinopathy which produces new fragile blood vessels on the retina. These vessels frequently bleed into the vitreous, the clear jelly in the center of the eyes causing visual problems. It is treated by laser surgery which will reduce the progression of diabetic retinopathy and at times will reverse visual loss causing permanent damage. If Diabetic retinopathy is identified at early stages a better control of blood sugar can be maintained by ensuring lifestyle modification, including abrupt weight loss, dietary changes and simple exercises (Fuadah, Setiawan, Mengko et al, 2015; Panse et al., 2015; Sachdeva & Singh, 2015; Veras, 2015).

The painless clouding of internal lens of the eye is called as cataract which is shown in Figure 2(b). They block the light from entering the lens, causing blindness over time. Cataracts worsen with time leading to increase in thickness of cornea. Light rays enter the eye through pupil and lens. The function of the lens is to focus the light onto the retina, transmitting the visual signals to the brain through the optic nerve. Clouding of the lens reduces the vision causing blurring of the images at any distance. The patients describe their vision to be foggy, cloudy, or filmy. Intensity of cataracts increases with time and only less light reaches the retina. People with cataracts have difficulty in night driving. They are characterized by double vision and second sight. In this situation, the cataract acts as a stronger lens, temporarily improving the ability to see things at a close distance. Formerly the people who needed reading glasses may no longer need them and also require frequent changes in spectacles as the vision blurring increases with time. Surgery is the only remedy to remove cataracts which is performed for only one eye at a time and may be required if the related vision loss cannot be corrected with glasses or contact lenses. This involves natural altering the cloudy lens with artificial lens. The operation is usually safe and effective (Salam et al., 2015).

Figure 2a. Diabetic Retinopathy

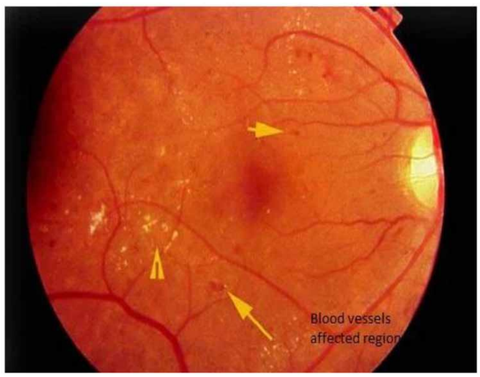
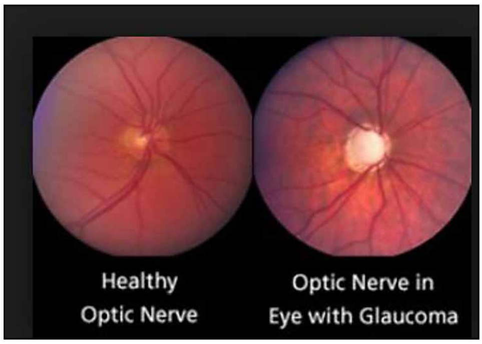


Figure 2b. Cataract

Figure 2c. Glaucoma



Glaucoma (shown in Figure 2c) is a condition where the vision is lost because the optic nerve gets affected due to increase in intraocular pressure (IOP). The two kinds of glaucoma are open-angle and closed angle glaucoma, for long term. African-American elderly people and those who have blood relatives suffer from this condition. Glaucoma does not produce symptoms in the early stages and when affected, the patients notice changes in vision. Timely treatment may inhibit further vision loss but it cannot revert existing vision loss. Glaucoma is treated with prescription eye drops. Occasionally, laser and surgical procedures may be employed. Early diagnosis and treatment can help preserve vision in people affected by glaucoma (Fuadah, Setiawan, Mengko et al, 2015).

LITERATURE SURVEY

Glaucoma is a condition where degeneration of optic nerve fibre takes place leading to decrease in FoV. Due to pressure created in blood vessels, blood and other fluids will be observed in eye, giving the retina an abnormal appearance. This eye disorder is called as Diabetic retinopathy, which may result in damaged blood vessels. Cataract is a clouding of the lens of the eye and occurs frequently in older age groups. An ophthalmologist needs a slit camera lens in the diagnosis of cataract, which may not be possible in rural areas. Hence the health of the sensory vision is provided by the processing the retinal images. To detect the presence of eye diseases many attempts are taken to extract useful information. They are summarized as follows

Diagnosis of Glaucoma Using Automated Texture and Spectra Features Extraction

To detect the presence of glaucoma at early stages, (Yuan et al., 2007) the features like higher order spectra (HOS) and texture descriptors are extracted. These extracted features are given to Support Vector Machine (SVM), Sequential Minimal Optimization (SMO), random forest, and Naive Bayesian (NB) for classification. The advantages include 91% accuracy by five-fold cross validation for images captured using Fundus imaging equipment. The demerit is that only glaucoma is identified.

Glaucoma Classification Using Regional Wavelet Features (RWF)

In this paper (Borgen et al., 2009) eye images are classified as normal and the one affected by glaucoma using Regional Wavelet Features of the Optic Nerve Head (ONH). Instead of global features RWF is more accurate within an accuracy of 93%.

Angle-Closure Glaucoma Feature Detection for Cross Examination

Redundant features are eliminated here (Aslam et al., 2009). For supervised feature selection Minimum Redundancy Maximum Relevance (MRMR) and unsupervised methods using Laplacian Score (Lscore) are used for cross- examination. For classifying Adaboost machine learning classifier is used. The main drawback is that only small data set is used.

Automatic Localization and Contour Detection of Optical Disc

Retinal diseases like Diabetic Retinopathy require identification of optic disc proposed a system using KL divergence matching in order to localize optic disc (Monro et al., 2009), followed by segmentation of main blood vessel. The advantages include location of OD with 92% accuracy with less computation time for histogram analysis. The drawback is that it not efficient for poorly contrasted images.

Retinal Fundus Diseases Diagnosis Using Image Mining

The glaucoma and diabetic retinopathy (Dhir et al., 2010) causes loss of vision is detected using this technique. This method uses Discrete Cosine Transform with k-Nearest Neighbor (k-NN) to classify the normal eyes and eyes affected by glaucoma and diabetic retinopathy with better classification accuracy.

Automatic Identification of Pathological Retinas Using SURF Descriptor and Pattern Recognition

A speedy and robust feature extraction based algorithm was developed to detect the pixels of interest to form visual dictionaries (ISO/IEC 19794-6:2011, 2011). Thereafter k-means clustering algorithm is used to predict whether an eye image is normal or disease affected. This method has two advantages; they include detection of characteristic pixels in the image and are also robust with an accuracy of 95.25%. The major disadvantages include presence of artifacts and loss of information in creation of visual dictionary.

AUTOMATIC SEGMENTATION OF OPHTHALMIC IMAGES FOR AREA CALCULATION OF OPTIC DISC

The appearance of the optic disc changes depending on the severity of Glaucoma condition. The blood vessels present in the eyeball makes the detection difficult. Hence the optic disc region need to be segmented to calculate its area to extract appropriate features (Rajendra Acharya, 2011) so that, early detection of Glaucoma is possible. The is done using an adaptive mask which has multiple sizes and resolutions. The results can be improved using fuzzy logic for segmentation. Further a hardware implementation will help real time application.

Cataract Classification and Grading for Fundus Images Using Computer-Aided Healthcare System

This method analyzes the retinal fundus images. The feature extraction is done to detect the cataract present. Based on severity, cataract is categorized as mild, moderate and severe (Neurotechnology, 2012). Wavelet transform, sketch based methods along with direct cosine transform are used for feature extraction. The main advantage of this method is identification of cataract and non-cataract using spatial features based on the severity of the cataract condition. The limitation of this method is only cataract and its severity is identified and no other eye pathologies are detected by this method.

Automated Diagnosis of Glaucoma for Detection of Optic Disc and Cup from Color Retinal Images

The proposed a system (McConnon et al., 2012) uses Color Fundus Image (CFI) to analyze retinal nerve damage and to detect glaucoma. Segmentation is done to extract the digital CFI optic disc, cup and neuro-retinal rim. Active contour model is used for the detecting cup. CMY color space is used to extract the color information of the pallor region in M channel. Features like vertical Cup to Disc Ratio (CDR), Horizontal to Vertical CDR (H-V CDR), Cup to Disc Area Ratio (CDAR), and Rim to Disc Area Ratio (RDAR) are used for classification by Support Vector Machine (SVM), Naive Bayes (NB) and k-NN. This method is cost effective when compared to Optical Coherence Tomography and Heidelberg Retina Tomography. Even the low quality images are segmented effectively by this method. The k-NN clustering algorithm gives an accuracy of 96.22%. The limitation of this method includes dependence on contour initialization for Geodesic active contour model.

Glaucoma Detection Using Image Processing Techniques

This provides knowledge on different techniques to detect glaucoma using retinal images (Budai et al., 2013). Many methods are compared in tabular form based on pre-processing techniques, classifiers and success rate of the algorithm. Optimal algorithm is proposed to detect glaucoma at early stages.

Calculation of Red Area for Determination for Glaucoma Disease

This method identifies the Red Area Percentage (RAP) for extracting the portions of sclera (Odstrcilik et al., 2013). For this, iris segmentation is done using Circular Hough Transform (CHT). This method is advantageous because it used real time face detection to detect the vessels and redness of the sclera for patients suffering from glaucoma. The extraction of sclera is difficult because the texture of sclera and that of the skin is same, which makes this method difficult.

SVM Based Local Vessel Features for Optic disc localization

This paper proposes a method to detect glaucoma and diabetic retinopathy by analyzing optic disc (Smart Sensors Ltd, 2013). It is dependent on features like blood vessel density and its orientation extracted using SVM. This method produces 98% accurate results in the presence of noise. The limitation of this method is that bright lesions might be detected as optic disc.

Diagnosis of Diabetic Retinopathy by Feature Extraction From the Fundus Images

This paper uses a novel parameter for optic disk detection to assist early stage Diabetic Retinopathy and lesions in fundus image using MAHM algorithm (Seyeddain et al., 2014). This novel parameter is based on the detection of the major vessels and its intersection to approximate the optic disk region. Color properties are used for further analysis which serves as an efficient framework for identification of Diabetic Retinopathy and eye hemorrhages. Using this method only Diabetic Retinopathy is detected.

Automatic Detection of Blood Vessel in Retinal Images

Retinal disease like hypertension of blood vessels is detected using Hessian matrix with Gaussian kernel based convolution (Seyeddain et al., 2014). The is identified by using eigen values of Hessian matrix after the convolving image and helps in identification of both healthy and abnormal retinal images. The demerit of this method is vessel segmentation done without the elimination of Optic disc.

Optimal Combination of Statistical Texture Analysis for Mobile Cataract Detection

An early detection method for cataract is proposed here and this method uses an ophthalmologist with a slit lamp camera on hand using android smart phone with k-NN classifier for statistical texture analysis (Trokielewicz et al., 2014). This method detects cataract with 97% accuracy and the patients need to have a smart phone with them for identification at the initial stage itself. The disadvantage of this method is that only cataract is identified and no information regarding other ophthalmic disorders.

From the elaborate survey done on ophthalmic disorders, it is inferred that many work focuses on detecting either the stages of cataract or diabetic Retinopathy using spatial feature extraction, Circular Hough Transform, Discrete Cosine Transform, Wavelet transform and Hessian matrix with Gaussian kernel based convolution and classifiers like SVM, NB, k-NN, SMO and random forest methods. In applying all these methods it is found that the maximum efficiency achieved is only 97% approximately.

RESEARCH GAP

The detailed literature survey has paved way to improvise the de-noising, feature extraction technique and identification algorithms for diagnosing ophthalmic disorders like cataract, diabetic retinopathy and glaucoma at an early stages so that the proposed algorithm becomes robust in nature. Also the evaluation of the proposed scheme is done with the help of certain performance measures like sensitivity and specificity.

RESEARCH HIGHLIGHTS

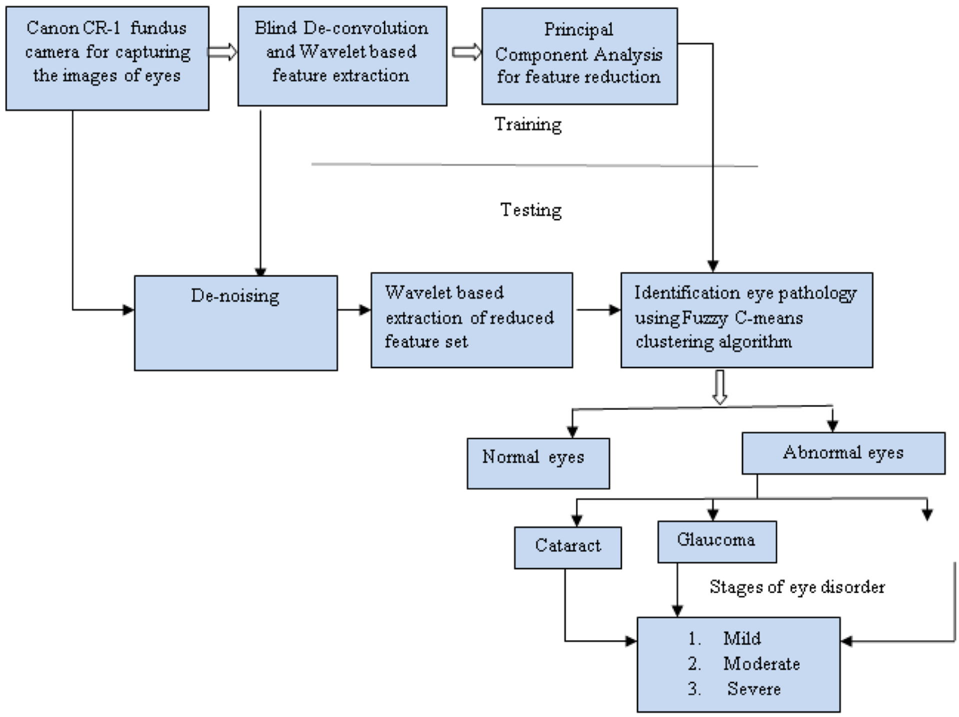
The research highlight includes

- Identification of early, moderate and severe stages in patients suffering from diabetic retinopathy, cataract, glaucoma and to find out the healthy eyes
- Blind de-convolution for blur elimination
- Optimal feature extraction using Wavelet transform
- Principal Component Analysis for feature reduction
- Design of a single robust Fuzzy C-means clustering algorithm for identification of these eye disorders with their corresponding stages

METHODOLOGY

The research gap discussed in section 3 highlighted the need for an automatic system for Identification of the three common eye disorders prevalent among the people. This method helps to detect the various stages of cataract, Glaucoma and diabetic retinopathy at early stages with mild impact on the patient without causing loss of vision, resulting in vision loss if not detected initially. The methodology for the proposed system focuses on detecting three major eye disorders at three different stages of abnormality (mild, moderate and severe) apart from the normal eyes. The block diagram for eye disease identification is depicted in Figure 3. The canon CR-1 fundus camera (Dhooge & de Laey, 1989; Roizenblatt et al., 2004) is used for capturing the images of the eyes with three types of eye diseases. Along with this, some images of the eyes pertaining to normal condition is also captured. These images are preprocessed for noise elimination using wavelet transform followed by feature extraction (wavelet co-efficients). Then the Principal Component Analysis (PCA) is used for decreasing the multi-dimensional feature set to 2D feature set, so that the computational complexity reduces. This feature set is used as inputs for Fuzzy C-means clustering algorithm for diagnosis.

Figure 3. System Architecture for detection of Eye Abnormality



MATERIALS AND METHODS

The various image processing based algorithms used for diagnosing the eye disorders are discussed in this section. The algorithms used include Blind devolution for noise removal, wavelet transform for feature extraction, PCA for feature reduction and finally the fuzzy C- means clustering for identification of the cataract, diabetes retinopathy and glaucoma.

Blind De-convolution Algorithm

The information about distortion is unknown; hence Blind De-convolution Algorithm is used. If the images are corrupted with any blur or noise, then this algorithm restores the image and computes the Point-Spread Function (PSF). Accelerated, damped Richardson-Lucy algorithm is used for noise removal. The values of PSF are taken as one of the inputs for Fuzzy C-means clustering algorithm (Salam et al., 2015).

Step 1: Read the input eye Images corresponding to normal eyes, eyes affected with cataract, diabetic retinopathy and glaucoma
Step 2: Create a blur and make the images of the eyes to be a corrupted one.
Step 3: Use under and oversized PSFs to restore the blurred eye images of various categories
Step 4: The Restored PSF of the normal and abnormal eye images are analyzed
Step 5: Improve the restoration by using true PSF
Step 6: Restore the true PSF for the undersized, oversized and exact PSF for normal and abnormal eyes

Discrete Wavelet Transform

The two dimensional wavelet transform analysis is a remarkable part of image processing methods which has developed tremendous interest for the researchers in analyzing medical images with an idea to expand the assessment in the areas like eye disease detection. In this aspect, the wavelet transform is used for sectional analysis of the images using Haar wavelet for representation of images (Fuadah, Setiawan, & Mengko, 2015).

The images of the eyes pertaining to normal and abnormal (affected by cataract, diabetic retinopathy and glaucoma) categories are well examined, de-noised and divided into various levels by using 2D Discrete Wavelet Transform which uses convolution principle with Haar wavelet. This image decomposition process shown in Figure 4 and is given by $Z = X * I * Y$, where Z is the final matrix of wavelet coefficients, I represents an original image, X represents a matrix of row filters and Y is tagged for matrix of column filters (Lotankar et al., 2015).

The original image when it is decomposed it contains the 'Approximations' which corresponds to the low frequency components. The second part is the 'detailed image' which has three parts namely the horizontal, vertical and diagonal elements. They correspond to high frequency components of the image. Hence the wavelet transform serves an inspiration for utilizing high and low frequency components present in the eye image. The wavelet transform divides the eye image in the time-frequency domain with non uniform tiles (Kumar et al., 2015).

Due to its low computing requirements, the Haar transform has been mainly used for image processing and pattern recognition (Niwas et al., 2016). From this reason two dimensional signal processing is an area of efficient applications of Haar transforms due to their wavelet-like structure. As the original image was processed to wavelet transform analysis, shown in the Figure 5.

Figure 4. Decomposition Process at level 2 for eye image with cataract

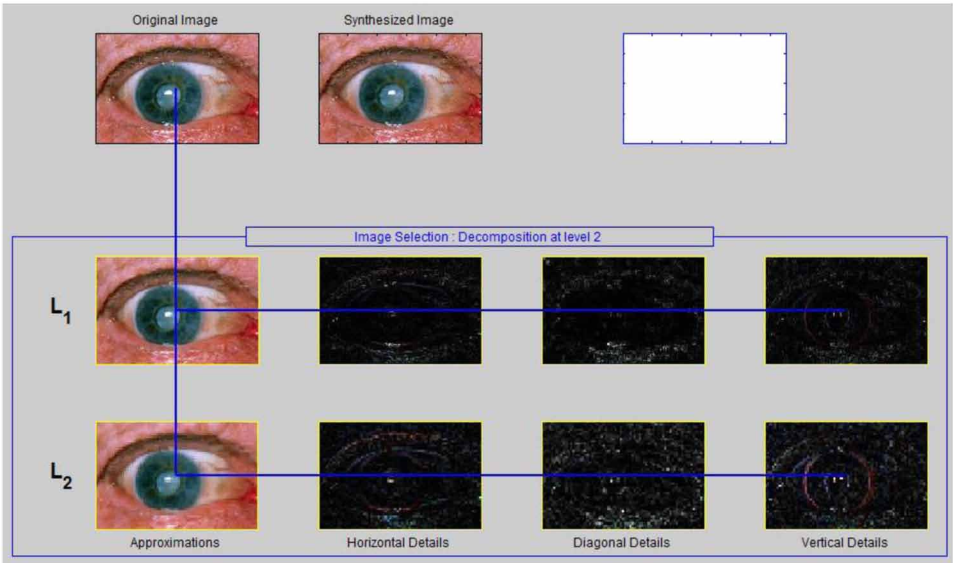
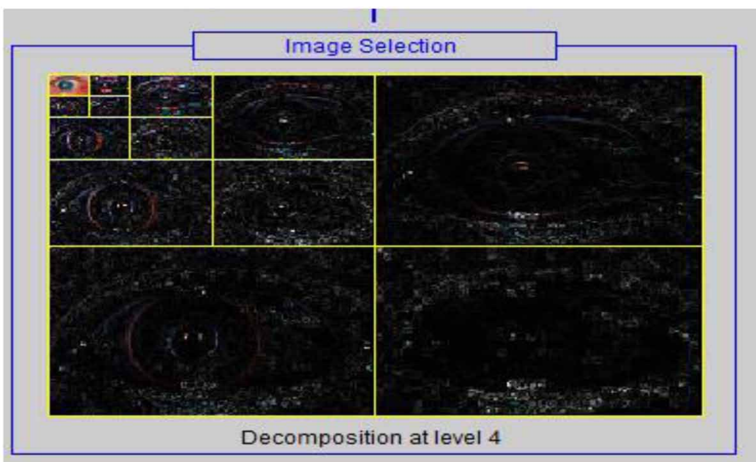


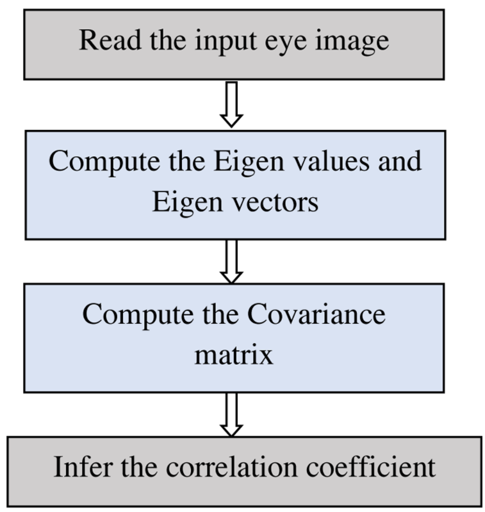
Figure 5. Decomposition of the image affected with cataract - four level



Principal Component Analysis (PCA)

The n-dimensional data extracted from the eye images (both normal and abnormal category) is reduced to two dimensional data using Principal Component Analysis (PCA) which consists of a data set with a correlation value. The PCA computes Eigen values, Eigen vectors based on which the covariance matrix is generated (Naveen Kumar et al., 2016). This is an orthogonal matrix. The dimension reductionality takes place which corresponds to number of the columns in the resultant data. If two values are correlated then its value is +1 or -1. The positive sign denotes that increase in one value will also increase the other value and whereas the negative sign denotes that increase in one value decreases the other value. If both the values are uncorrelated, then the value is ‘0’. The correlation is computed by calculating the covariance matrix which is in turn computed using Eigen values and Eigen vectors as Figure 6.

Figure 6. Flowchart for PCA



Fuzzy C- Means Clustering Algorithm

1. Choose a large value of membership function and classify each feature value into the cluster
2. Obtain the characteristic plot for clustered feature set values and cluster centers extracted from normal and abnormal eye images.

3. The overlap of the membership functions using Fuzzy is inferred between Clusters
4. If there is no overlap between the clusters, then the identification of eye disorder is efficiently done

RESULTS AND DISCUSSION

The data base containing the eye images corresponding to both normal and abnormal categories are acquired using Canon CR-1 fundus camera with a field of view of 45° (Budai et al., 2013). The abnormal categories of images include eye images affected by cataract, diabetic retinopathy and glaucoma. The database consists of totally 121 images of eyes pertaining to normal and abnormal conditions as shown in Table 1. Nearly, 60 images of the eyes corresponding to healthy condition, eyes affected by cataract, diabetic retinopathy and glaucoma are used for training and the remaining 61 images are used for testing the proposed algorithms.

Preprocessing

These images are preprocessed for noise removal using De-convolution algorithm for which the results are shown in Figure 7 (a) to (j) for healthy and abnormal eyes respectively. If noise is present in the captured images, it causes blurring which is reflected in its Power Spectral Function (PSF). The initial, undersized and oversized PSF values for all the disease affected eye images are calculated and tabulated in Table 2 which serves as a feature for identifying the normal and abnormal eye conditions.

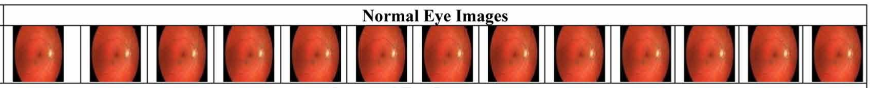
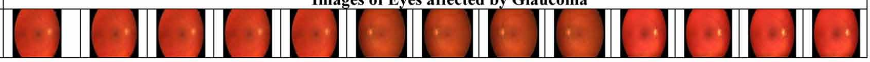
Feature Extraction

The noise removal is followed by feature extraction which includes features like mean, median, maximum intensity, minimum intensity, range, standard deviation, Median absolute deviation, mean absolute deviation, L1 norm, L2 norm and maximum norm respectively. This is done using wavelet tool box in MATLAB.

Feature Reduction Using PCA

The extracted features are in n-dimensional space. Hence it has to be reduced to a 2D feature set using PCA. These Eigen values which help to compute the co-variance matrix and thereby the correlation co-efficients are used as the distinct feature set which is clustered into groups by Fuzzy C-means algorithm.

Table 1. Fundus Data base for Images of Eyes

S. No	Normal Eye Images		
1.			
	Abnormal Eye Images		
	Mild	Moderate	Severe
	Images of Eyes affected by Cataract		
2.			
	Images of Eyes affected by Diabetic Retinopathy		
3.			
	Images of Eyes affected by Glaucoma		
4.			

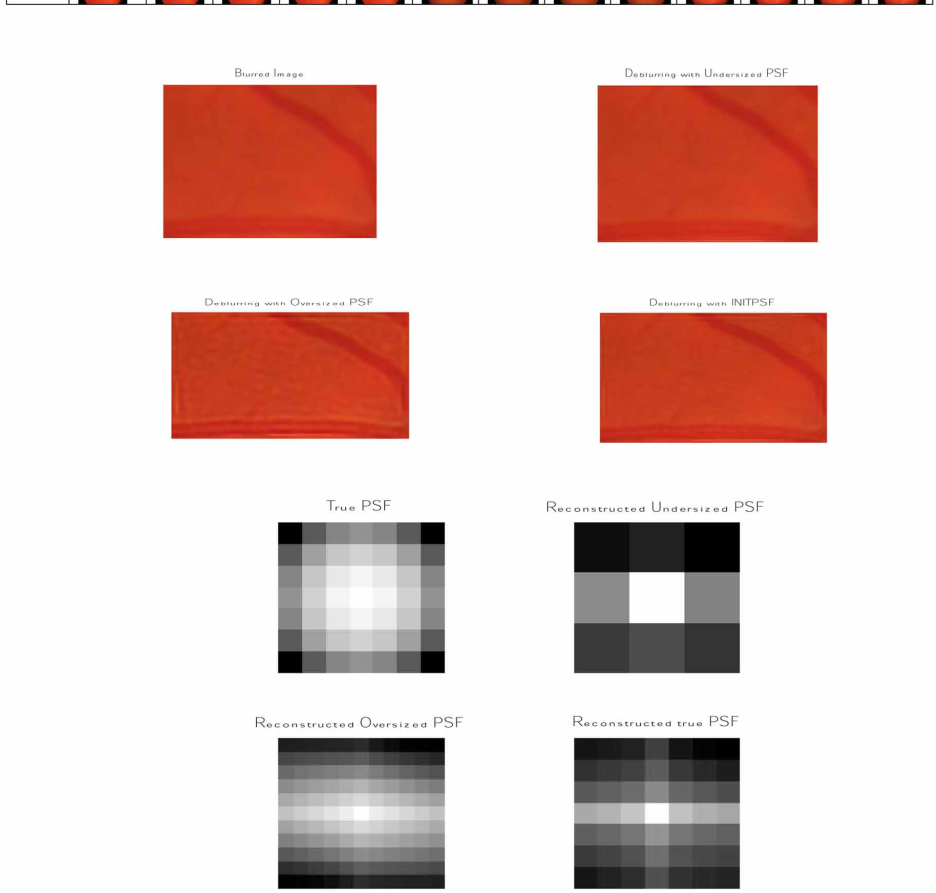


Figure 7a. Results for De-convolution for normal Eyes

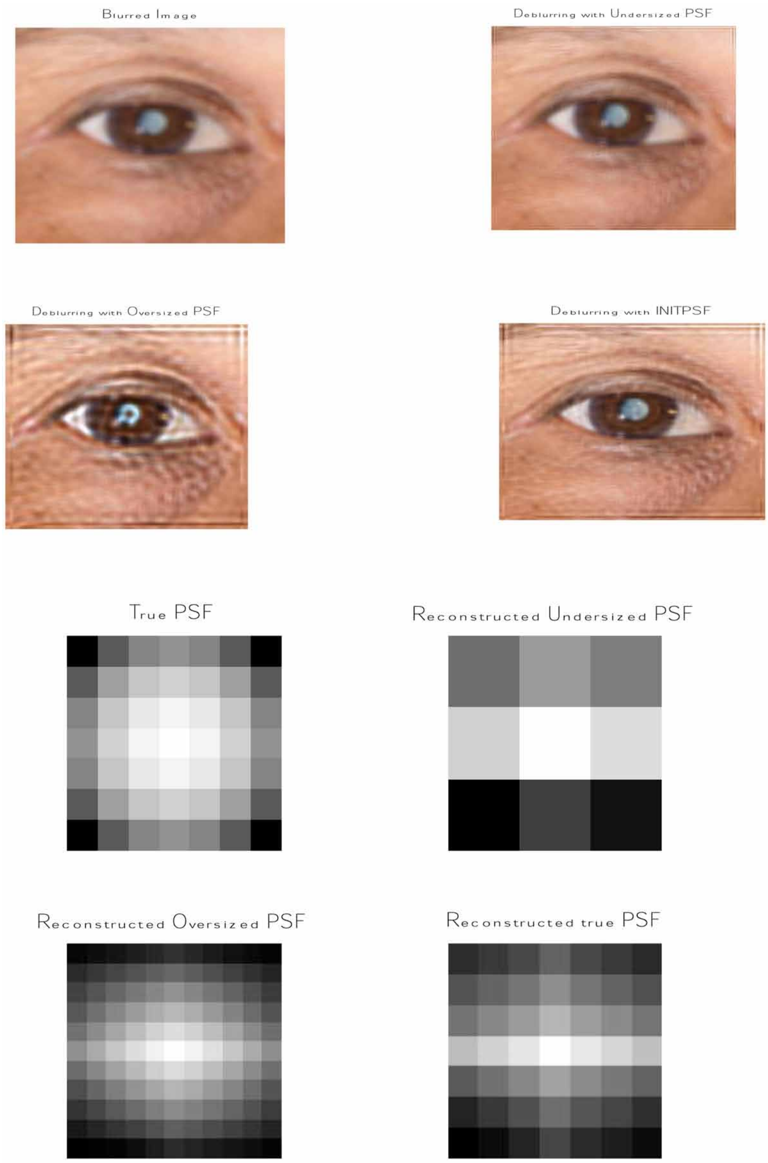


Figure 7b. Results for De-convolution for Eyes affected by Mild Cataract

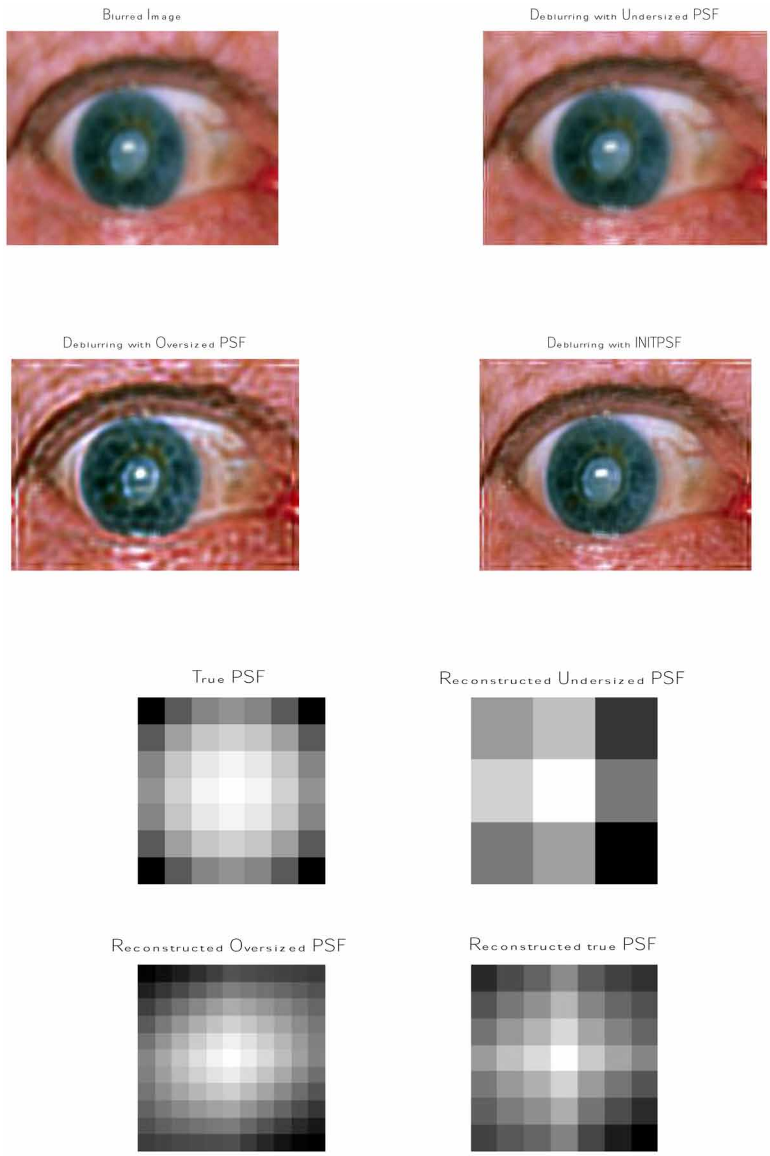


Figure 7c. Results for De-convolution for Eyes affected by moderate Cataract

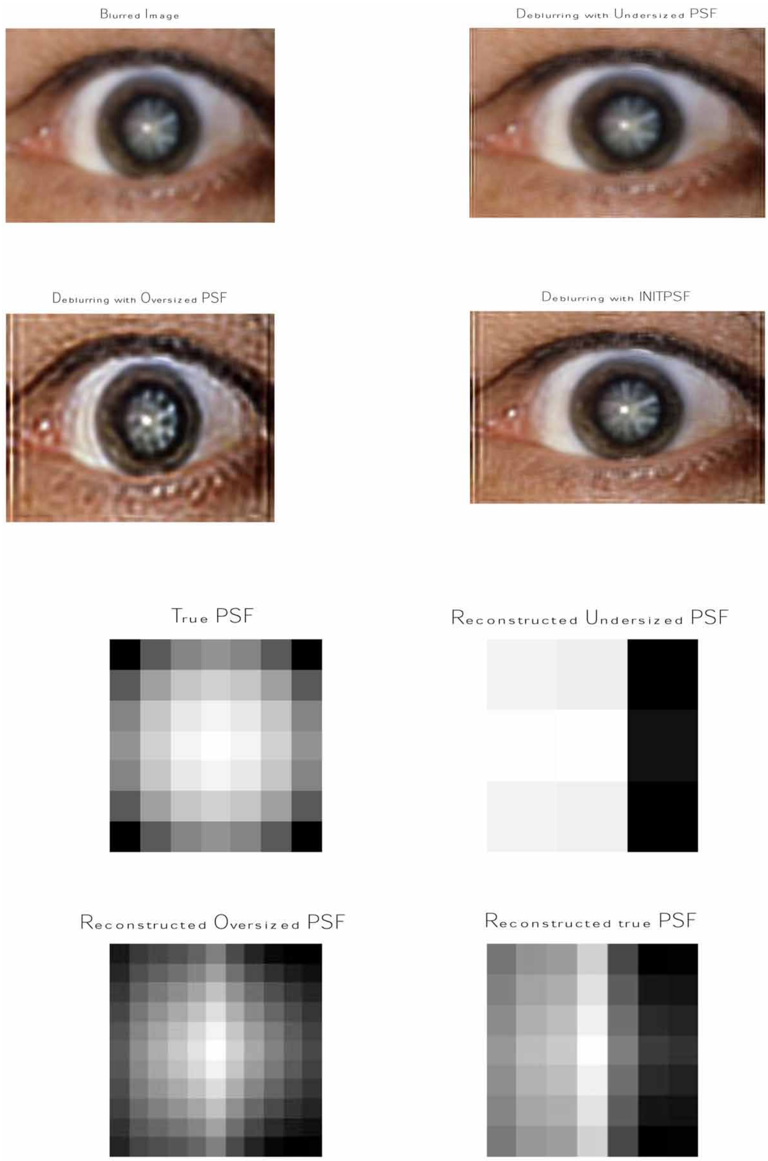


Figure 7d. Results for De-convolution for Eyes affected by Severe Cataract

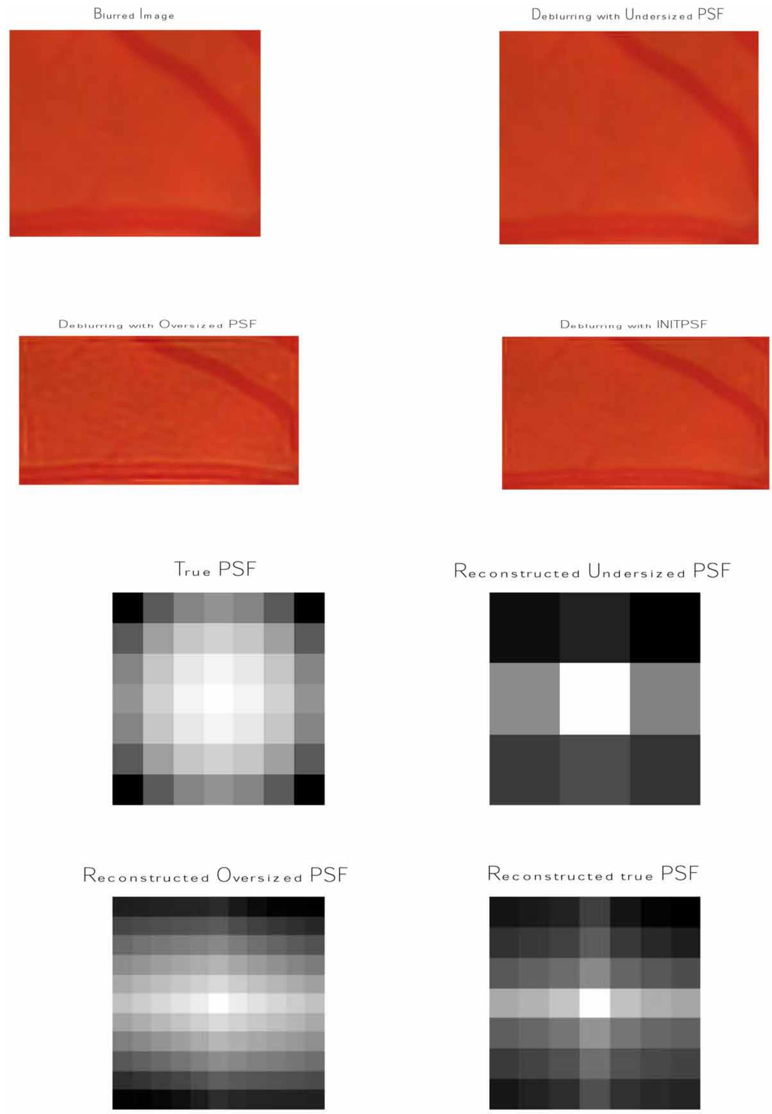


Figure 7e. Results for De-convolution for Eyes affected by mild Glaucoma

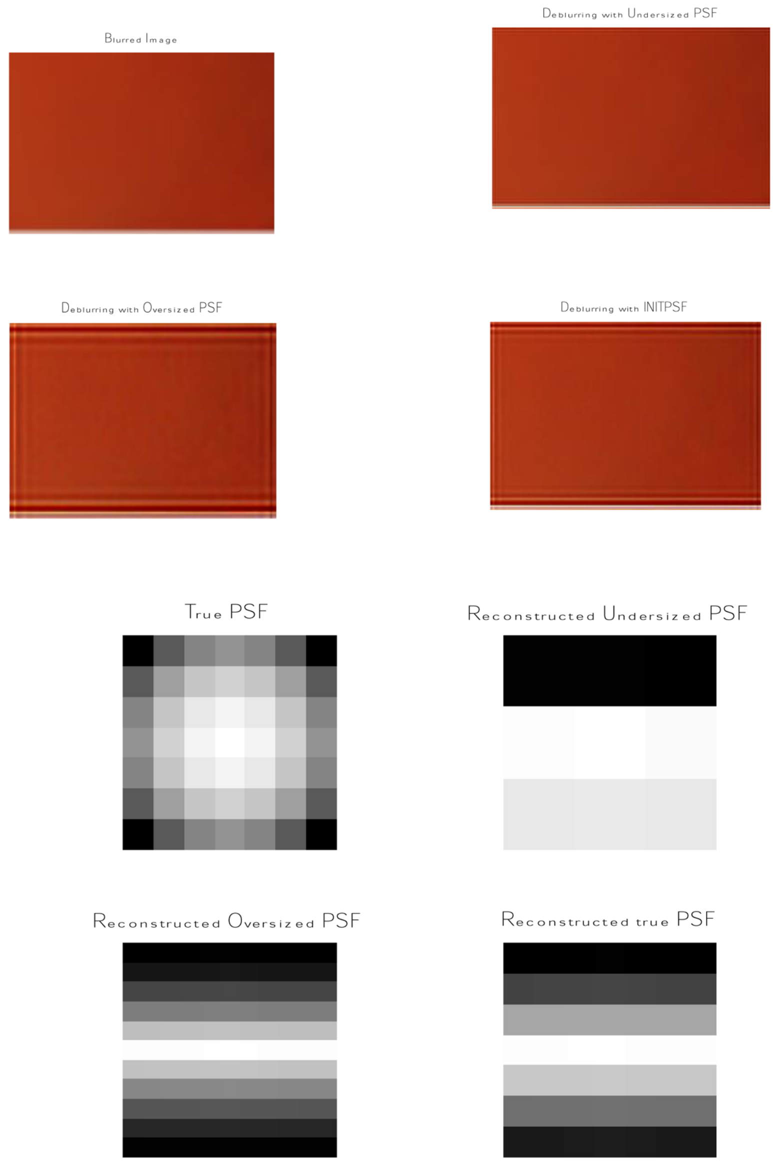


Figure 7f. Results for De-convolution for Eyes affected by moderate Glaucoma

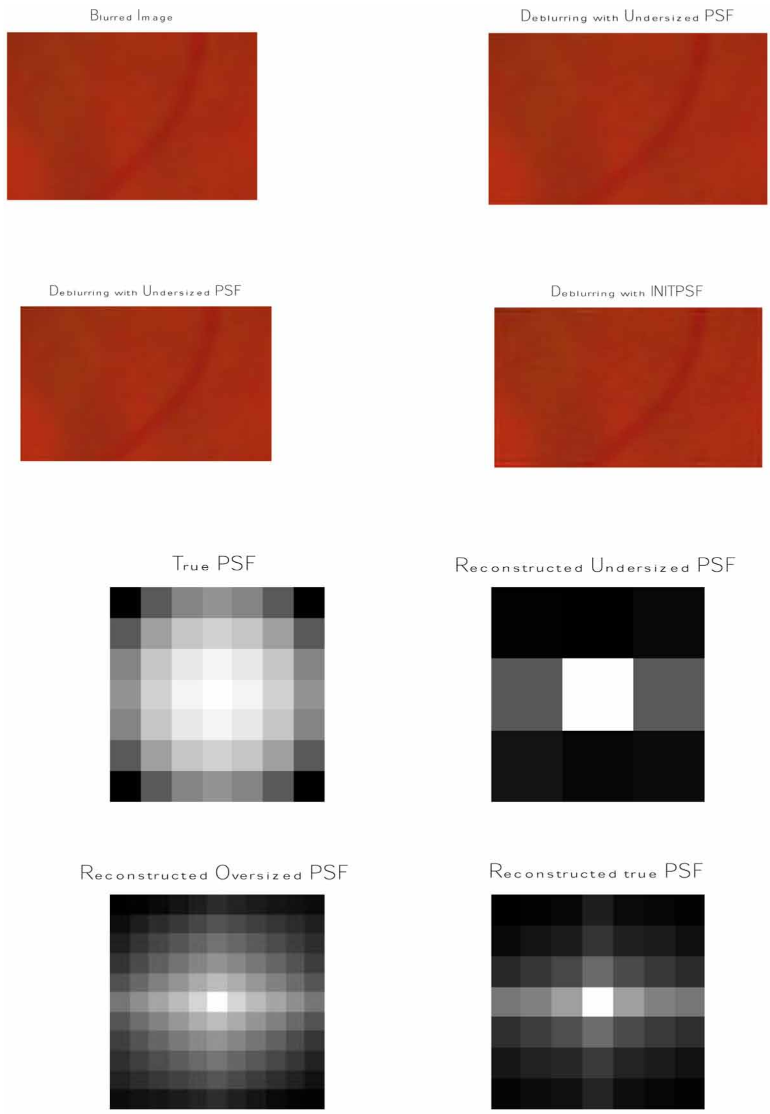


Figure 7g. Results for De-convolution for Eyes affected by severe Glaucoma

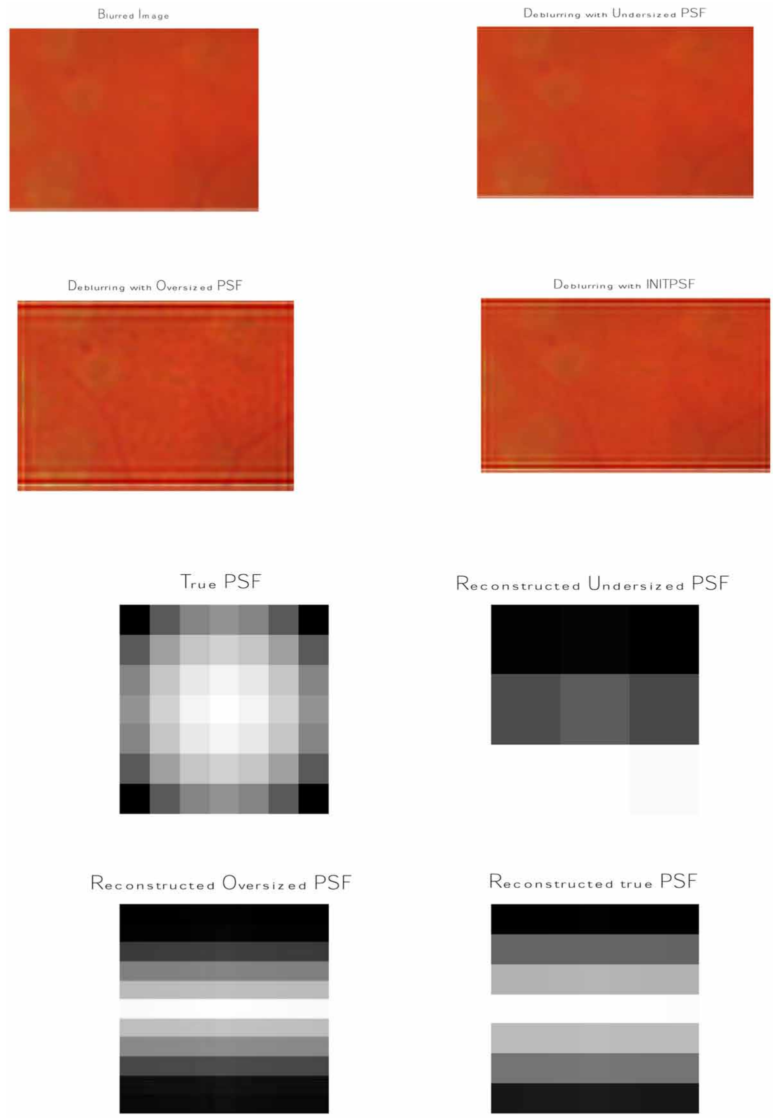


Figure 7h. Results for De-convolution for Eyes affected by mild Diabetic Retinopathy

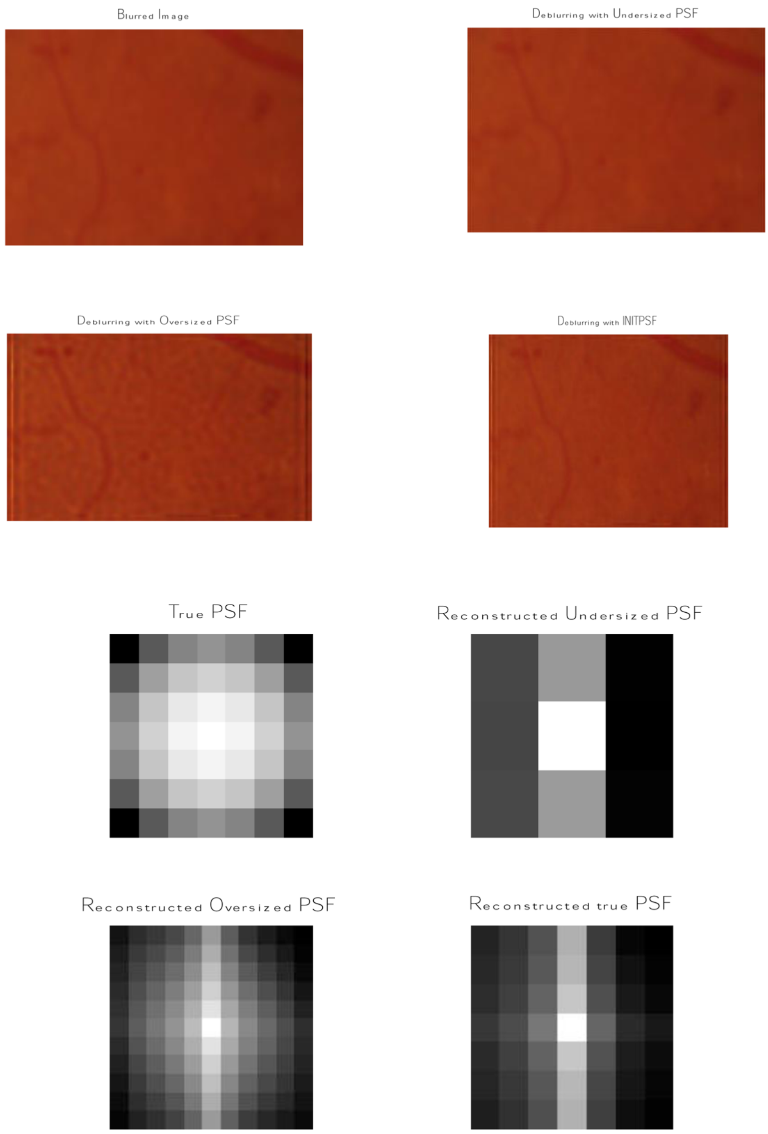


Figure 7i. Results for De-convolution for Eyes affected by moderate Diabetic Retinopathy

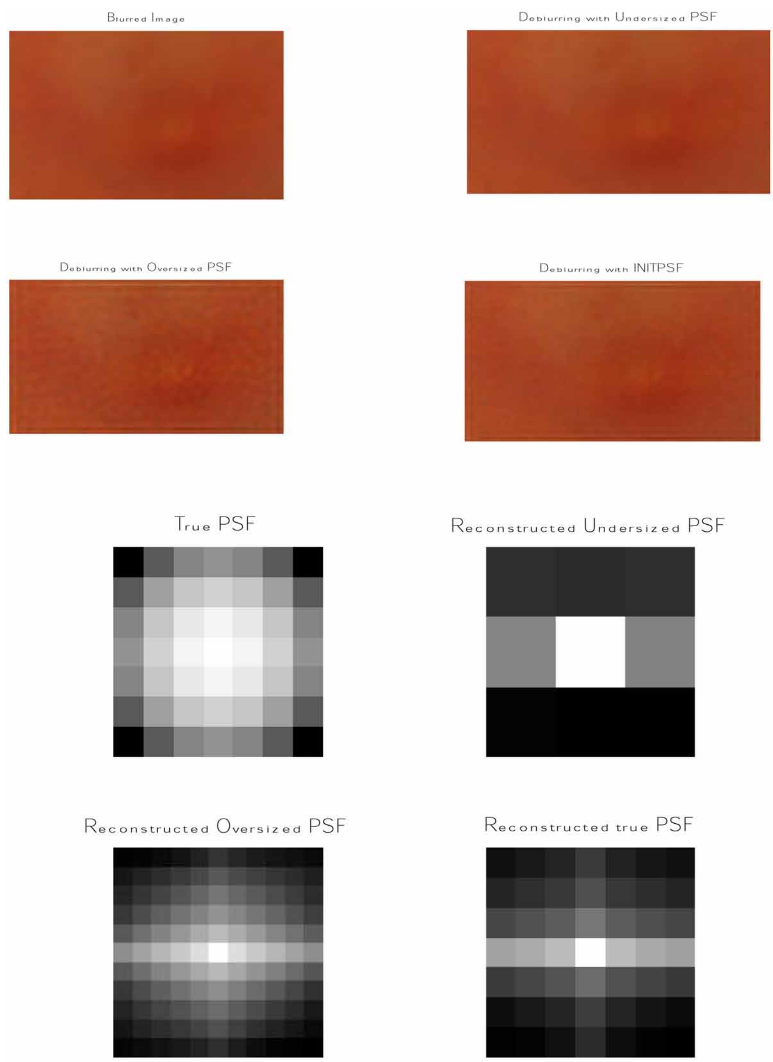


Figure 7j. Results for De-convolution for Eyes affected by severe Diabetic Retinopathy

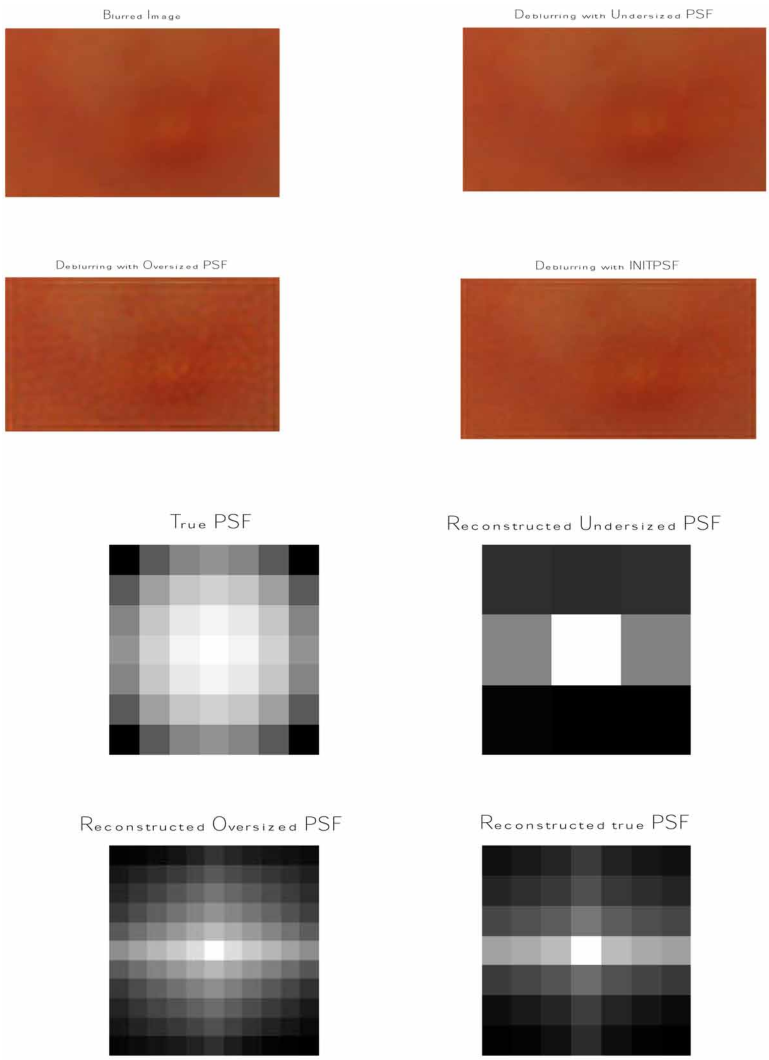


Table 2. PSF values from De-convolution

S. No	PSF Values for Healthy Eyes	Condition of Abnormality	PSF Values for Eyes Affected by Cataract	PSF Values for Eyes Affected by Diabetic Retinopathy	PSF Values for Eyes Affected by Glaucoma
1.	54.7257	Mild	42.9867	39.0857	27.1806
2.	54.7257		42.8554	39.6814	27.7359
3.	54.7257		42.1857	39.1085	27.8245
4.	54.7257		42.9678	39.2681	27.2679
5.	54.7257		42.4855	39.0852	27.3838
6.	54.7257	Moderate	41.6887	38.2866	26.7359
7.	54.7257		41.45854	38.5142	25.2602
8.	54.7257		41.7187	38.8215	25.8067
9.	54.7257		41.8967	38.2681	25.0611
10.	54.7257		41.8554	38.0852	25.1806
11.	54.7257	Severe	40.2478	37.1237	25.7359
12.	54.7257		40.3176	37.2278	24.1806
13.	54.7257		40.4687	37.6791	24.7359
14.	54.7257		40.5687	37.3478	24.1806
15.	54.7257		40.1672	37.5689	24.7359

Figure 8. Feature extraction using wavelet transform for normal eyes: (a). Eyes affected by cataract at mild stage; (b). Eyes affected by cataract at moderate stage; (c). Eyes affected by cataract at severe stage

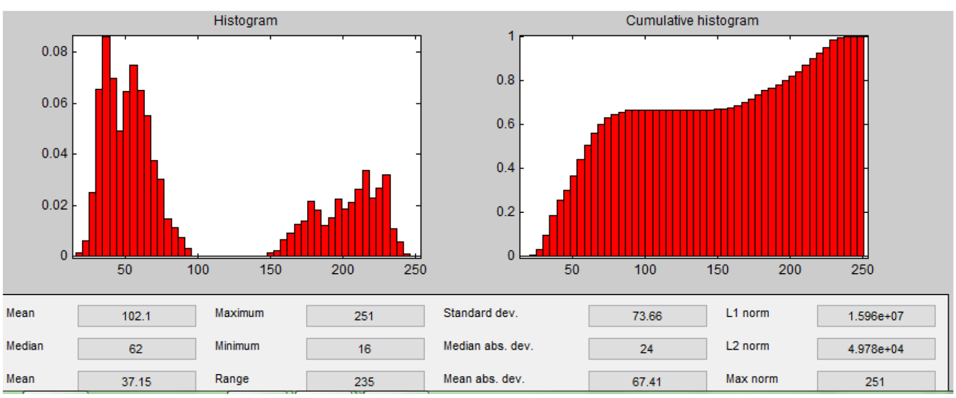
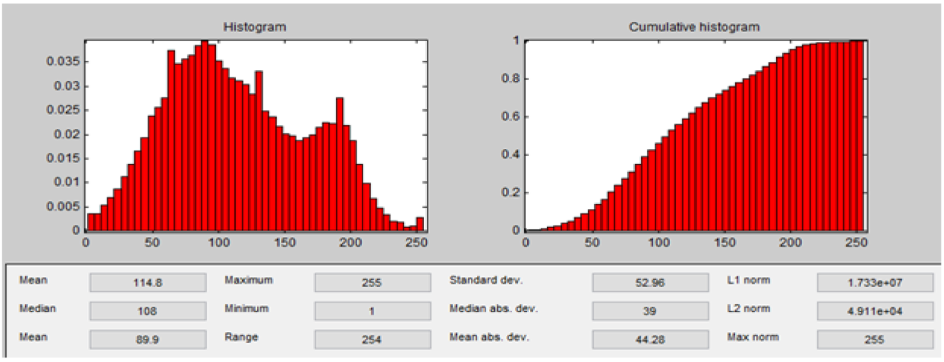
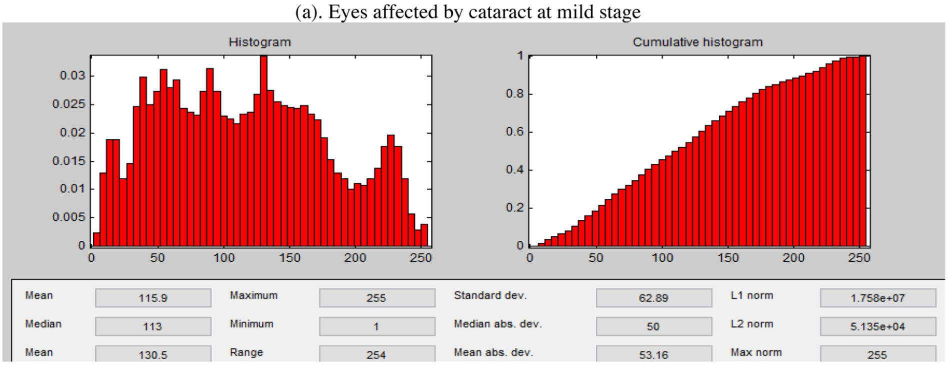


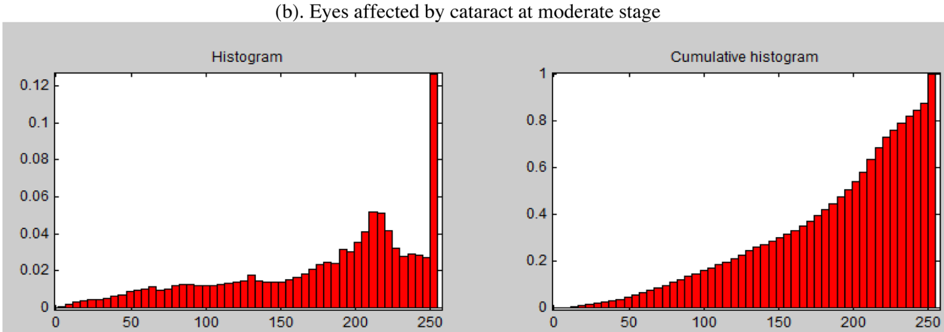
Figure 9. Feature extraction using wavelet transform for eyes affected by cataract: (a). Eyes affected by Diabetic Retinopathy at mild stage; (b). Eyes affected by Diabetic Retinopathy at moderate stage; (c). Eyes affected by Diabetic Retinopathy at severe stage



(a). Eyes affected by cataract at mild stage

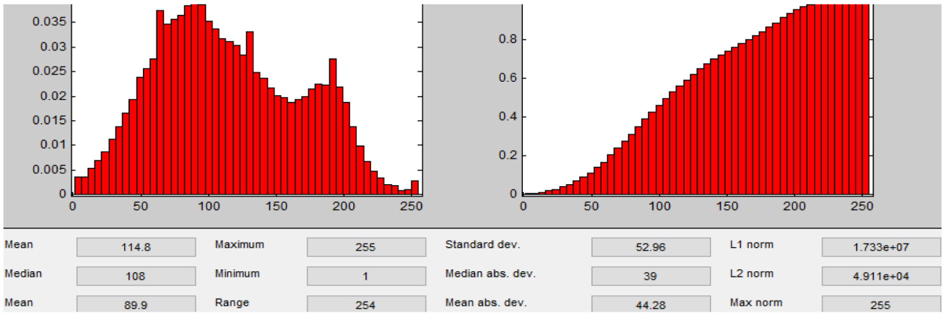


(b). Eyes affected by cataract at moderate stage

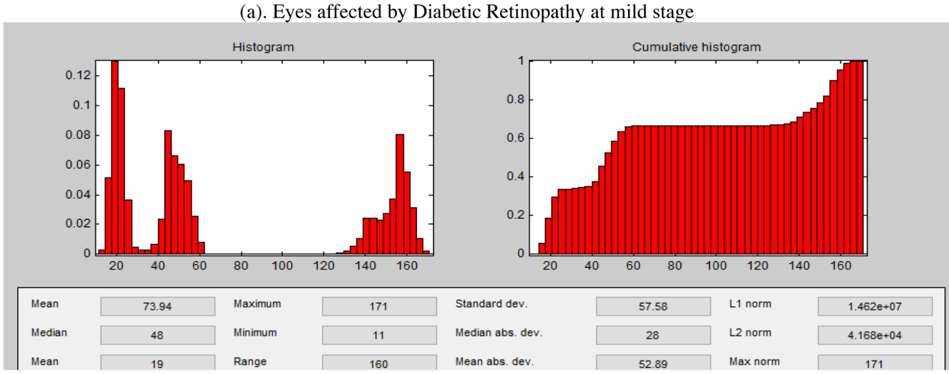


(c). Eyes affected by cataract at severe stage

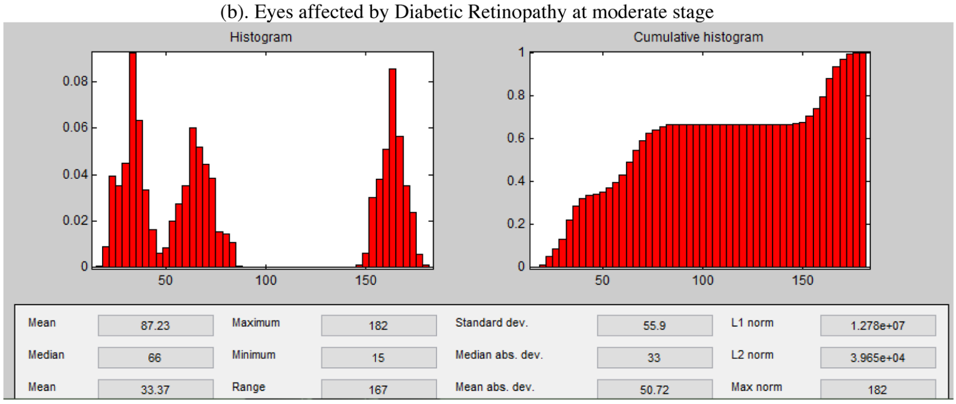
Figure 10. Feature extraction using wavelet transform for eyes affected by Diabetic Retinopathy: (a). Eyes affected by Glaucoma at mild stage; (b). Eyes affected by Glaucoma at moderate stage; (c). Eyes affected by Glaucoma at severe stage



(a). Eyes affected by Diabetic Retinopathy at mild stage

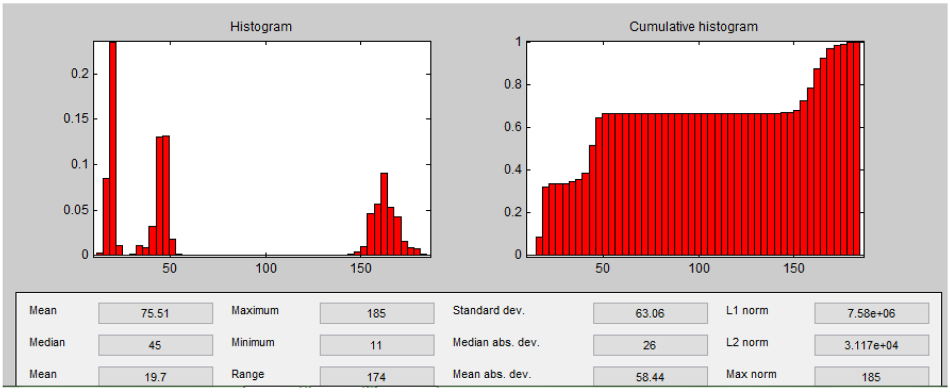


(b). Eyes affected by Diabetic Retinopathy at moderate stage

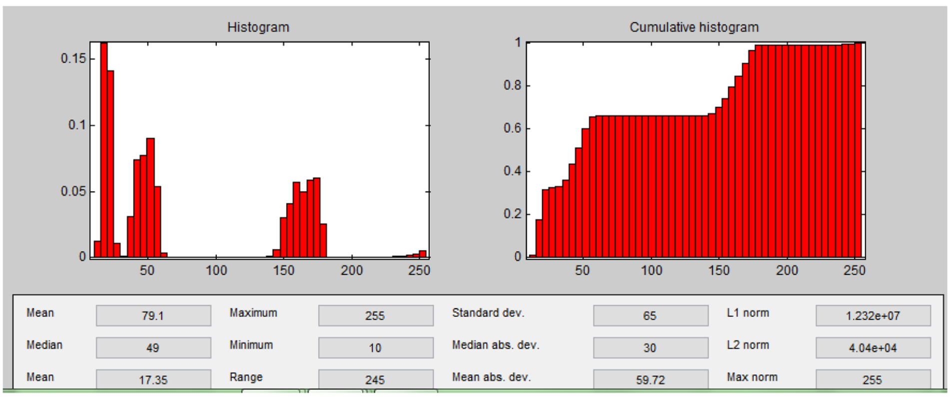


(c). Eyes affected by Diabetic Retinopathy at severe stage

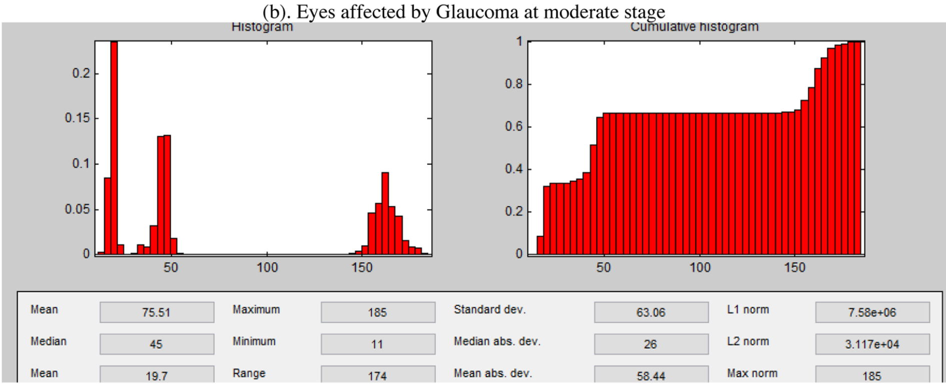
Figure 11. Feature extraction using wavelet transform for eyes affected by Glaucoma



(a). Eyes affected by Glaucoma at mild stage



(b). Eyes affected by Glaucoma at moderate stage



(c). Eyes affected by Glaucoma at severe stage

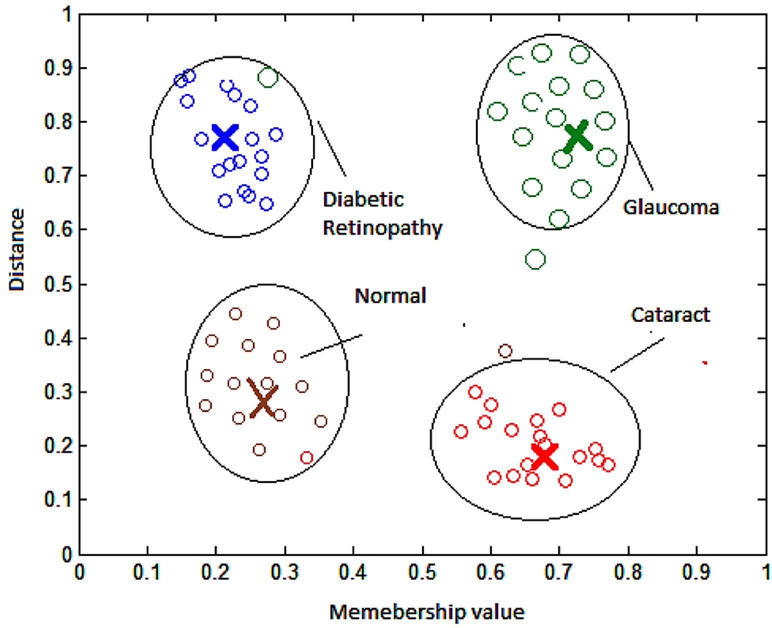
Table 3. Reduced feature set values using PCA

S. No	Correlation Co-efficient for Healthy Eyes	Condition of Abnormality	Correlation co-Efficient for Eyes Affected by Cataract	Correlation Co-efficient for Eyes Affected by Diabetic Retinopathy	Correlation Co-efficient for Eyes Affected by Glaucoma
1.	0. 3435	Mild	0. 2132	0.7085	0. 5425
2.	0. 3524		0. 2546	0.7681	0. 5425
3.	0.4218		0. 2347	0.7915	0. 6542
4.	0.3085		0. 1542	0.7268	0. 6412
5.	0.3681		0. 1542	0.7852	0. 6252
6.	0.3915	Moderate	0. 2422	0.7866	0.5734
7.	0.3268		0. 2253	0.8851	0.6684
8.	0.4171		0.2659	0.8882	0.5745
9.	0.4189		0.2526	0.9826	0.5311
10.	0.4185		0. 1257	0.9808	0.6128
11.	0.4024	Severe	0. 1128	0.9712	0.6769
12.	0.4017		0.2511	0.9722	0.6435
13.	0.4046		0.2586	0.9679	0.5473
14.	0.4056		0.2573	0.9734	0.6486
15.	0.4016		0.2418	0.9689	0.6359

FUZZY C-MEANS CLUSTERING FOR EYE DISEASE IDENTIFICATION

Clustering using Fuzzy C-means algorithm is defined on the basis of extracted features from four categories of eye images grouped as healthy eyes, eyes affected by Glaucoma, Diabetic Retinopathy and Cataract. The output for the fuzzy C means algorithm is shown in Figure 12. During the clustering process using the Eigen values from PCA, the covariance matrix is computed from which the correlation coefficients are computed. The clustering process is dependent on the correlation efficients in Table 3 which indicates that for healthy eyes it is in the range of 0.3-0.4, whereas for the disease affected eyes it approximately in the range of 0.1-0.2 for cataract, 0.7-0.9 for diabetic retinopathy and 0.5-0.6 for eyes affected by glaucoma. This is illustrated in Figure 12.

Figure 12. Output for fuzzy C-means clustering algorithm



Evaluation of the Proposed Technique With Existing Methods

The evaluation results in Table 4, states that the sensitivity, specificity and accuracy is in the range of 94% to 98% for the proposed method in comparison with existing methods.

Table 4a. Evaluation Results for Eye disease Identification during training for the proposed method

Performance Measure	Healthy Eyes	Diabetic Retinopathy	Glaucoma	Cataract
Sensitivity	94.61% ± 0.54%	94.64% ± 0.59%	98.02% ± 0.23%	95.82% ± 0.42%
Specificity	97.50% ± 0.25%	96.19% ± 0.37%	96.38% ± 0.39%	95.44% ± 0.42%
Accuracy	95.39% ± 0.51%	94.45% ± 0.64%	94.97% ± 0.61%	96.65% ± 0.33%

Table 4b. Evaluation Results for Eye disease Identification during testing

Author	Dataset	Sensitivity	Specificity	Accuracy
Budai et. al. [budai09]	Healthy eyes	70.99% ± 4.01%	92.45% ± 0.44%	94.81% ± 0.56%
Budai et. al. [budai09]	Diabetic retinopathy	72.29% ± 3.4%	93.74% ± 0.44%	93.81% ± 0.67%
Budai et. al. [budai09]	Glaucoma	73.34% ± 2.3%	91.45% ± 0.44%	92.81% ± 0.71%
Budai et. al. [budai09]	Cataract	74.99% ± 1.42%	90.65% ± 0.44%	90.81% ± 0.86%
Odstrcilik et. al. [odstrcilik09]	Healthy eyes	78.61% ± 3.92%	97.50% ± 0.65%	95.39% ± 0.61%
Odstrcilik et. al. [odstrcilik09]	Diabetic retinopathy	74.63% ± 5.66%	96.19% ± 0.77%	94.45% ± 0.84%
Odstrcilik et. al. [odstrcilik09]	Glaucoma	79.00% ± 3.18%	96.38% ± 0.69%	94.97% ± 0.61%
Odstrcilik et. al. [odstrcilik09	Cataract	75.82% ± 4.32%	95.44% ± 0.82%	94.65% ± 0.54%
Proposed method	Healthy eyes	93.52% ± 0.54%	96.39% ± 0.25%	94.26% ± 0.51%
Proposed method	Diabetic retinopathy	93.44% ± 0.59%	95.21% ± 0.37%	93.11% ± 0.64%
Proposed method	Glaucoma	97.13% ± 0.23%	95.49% ± 0.39%	93.97% ± 0.61%
Proposed method	Cataract	94.73% ± 0.42%	94.32% ± 0.42%	95.76% ± 0.33%

CONCLUSION

In eye disease identification, an image processing model for detection of different diseases such as glaucoma, cataract and diabetic retinopathy is proposed. A combination of image processing and clustering algorithms are used where the reduced feature set from PCA are extracted from the training dataset and then applied to fuzzy C-means clustering algorithm to identify the exact eye disease or the healthy condition. This method integrates a single approach with the help of correlation coefficient for all three eye diseases and also to detect the normal eye conditions. Different algorithms like De-convolution, wavelet transform, PCA and fuzzy C-means clustering are used for preprocessing, feature extraction, reduction and classification from the fundus database images. Glaucoma and Diabetic Retinopathy are among the leading cause for blindness as compared to Cataract. But detection of these diseases at earliest stage and treatment can aid patient in avoiding vision loss. An automatic eye disease detection system can help by providing accurate and early diagnosis.

REFERENCES

Aggarwal, M. K., & Khare, V. (2015). Automatic localization and contour detection of Optic disc. *2015 International Conference on Signal Processing and Communication (ICSC)*. 10.1109/ICSPCom.2015.7150686

Aloudat & Faezipour. (2016). Determination for Glaucoma Disease Based on Red Area Percentage. *2016 IEEE Long Island Systems*, *Applications and Technology Conference (LISAT)*.

Aslam, T. M., Tan, S. Z., & Dhillon, B. (2009). Iris recognition in the presence of ocular disease. *Journal of the Royal Society, Interface*, *6*(34), 2009. doi:10.1098/rsif.2008.0530 PMID:19324690

Borgen, H., Bours, P., & Wolthusen, S. D. (2009). Simulating the Influences of Aging and Ocular Disease on Biometric Recognition Performance. *International Conference on Biometrics 2009, 8*(8), 857–867. 10.1007/978-3-642-01793-3_87

Budai, A., Bock, R., Maier, A., Hornegger, J., & Michelson, G. (2013). Robust Vessel Segmentation in Fundus Images. *International Journal of Biomedical Imaging*.

Canadian Border Services Agency. (2015). *CANPASS Air*. Available: http: //www.cbsa-asfc.gc.ca/prog/canpass/canpassair-eng.html

Dhir, L., Habib, N. E., Monro, D. M., & Rakshit, S. (2010). Effect of cataract surgery and pupil dilation on iris pattern recognition for personal authentication. *Eye (London, England)*, *24*(6), 1006–1010. doi:10.1038/eye.2009.275 PMID:19911017

Dhooge, M., & de Laey, J. J. (1989). The ocular ischemic syndrome. *Bulletin de la Société Belge d'Ophtalmologie*, *231*, 1–13. PMID:2488440

Elbalaoui, Fakir, Taifi, & Merbohua. (2016). Automatic Detection of Blood Vessel in Retinal Images. *13th International Conference Computer Graphics, Imaging and Visualization*.

Fuadah, Setiawan, & Mengko. (2015). Mobile Cataract Detection using Optimal Combination of Statistical Texture Analysis. *4th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME)*.

Fuadah, Setiawan, Mengko, & Budiman. (2015). *A computer aided healthcare system for cataract classification and grading based on fundus image analysis*. Elsevier Science Publishers B. V.

Haleem, M. S., Han, L., van Hemert, J., & Fleming, A. (2015). Glaucoma Classification using Regional Wavelet Features of the ONH and its Surroundinga. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).*

ISO/IEC 19794-6:2011. (2011). Information technology – Biometric data interchange formats – Part 6: Iris image data.

Kumar, Manjunathand, & Sheshadri. (2015). Feature extraction from the fundus images for the diagnosis of diabetic retinopathy. *International Conference on Emerging Research in Electronics. Computer Science and Technology*.

Lotankar, M., Noronha, K., & Koti, J. (2015). Detection of Optic Disc and Cup from Color Retinal Images for Automated Diagnosis of Glaucoma. *IEEE UP Section Conference on Electrical Computer and Electronics (UPCON).*

McConnon, G., Deravi, F., Hoque, S., Sirlantzis, K., & Howells, G. (2012). Impact of Common Ophthalmic Disorders on Iris Recognition. *2012 5th IAPR International Conference on Biometrics Compendium*, 277–282.

Monro, D. M., Rakshit, S., & Zhang, D. (2009). DCT-Based Iris Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(4), 586–595. doi:10.1109/TPAMI.2007.1002 PMID:17299216

Naveen Kumar, B., Chauhan, R. P., & Dahiya, N. (2016). Detection of Glaucoma using Image processing techniques: A Review. *2016 International Conference on Microelectronics, Computing and Communications (MicroCom).*

Neurotechnology. (2012). *VeriEye SDK, v. 4.3*. Available: https://www.neurotechnology.com/verieye.html

Niwas, Lin, Kwoh, Kuo, Sng, Aquino, & Chew. (2016). Cross-examination for Angle-Closure Glaucoma Feature Detection. *IEEE Journal of Biomedical and Health Informatics*.

Odstrcilik, J., Budai, A., Kolar, R., & Hornegger, J. (2013, June). Retinal vessel segmentation by improved matched filtering: Evaluation on a new high-resolution fundus image database. *IET Image Processing*, *7*(4), 373–383. doi:10.1049/iet-ipr.2012.0455

Panse, N. D., Ghorpade, T., & Jethani, V. (2015). Retinal Fundus Diseases Diagnosis using Image Mining. *IEEE International Conference on Computer, Communication and Control (IC4-2015).* 10.1109/IC4.2015.7375721

Rajendra Acharya, U. (2011, May). Automated Diagnosis of Glaucoma Using Texture and Higher Order Spectra Features. *IEEE Transactions on Information Technology in Biomedicine*, *15*(3).

Roizenblatt, R., Schor, P., Dante, F., Roizenblatt, J., & Jr, R. B. (2004). Iris recognition as a biometric method after cataract surgery. *BioMedical Engineering Online, 3*(2). www.biomedical-engineering-online/com/content/3/1/2

Sachdeva, & Singh. (2015). Automatic Segmentation and Area Calculation of Optic Disc in Ophthalmic Images. *2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS).*

Salam, Akram, Abbas, & Anwar. (2015). Optic Disc Localization using Local Vessel Based Features and Support Vector Machine. *IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE).*

Seyeddain, O., Kraker, H., Redlberger, A., Dexl, A. K., Grabner, G., & Emesz, M. (2014). Reliability of automatic biometric iris recognition after phacoemulsification or drug-induced pupil dilation. *European Journal of Ophthalmology*, *24*(1), 58–62. doi:10.5301/ejo.5000343 PMID:23873488

Smart Sensors Ltd. (2013). *MIRLIN SDK*, *2*, 23.

Sutra, G., Dorizzi, B., Garcia-Salitcetti, S., & Othman, N. (2013, April 23). *A biometric reference system for iris. OSIRIS version 4.1*. Available: http://svnext.it-sudparis.eu/svnview2-eph/ref syst/Iris Osiris v4.1/

Trokielewicz, M., Czajka, A., & Maciejewicz, P. (2014). Cataract influence on iris recognition performance. Proc. SPIE 9290, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments. doi:10.1117/12.2076040

Unique Identification Authority of India. (n.d.). *AADHAAR*. Available: https://uidai.gov.in/what-is-aadhaar.html

Veras, R. (2015). SURF descriptor and pattern recognition techniques in automatic identification of pathological retinas. *2015 Brazilian Conference on Intelligent Systems*.

Yuan, X., Zhou, H., & Shi, P. (2007). Iris recognition: A biometric method after refractive surgery. *Journal of Zhejiang University. Science A*, *8*(8), 1227–1231. doi:10.1631/jzus.2007.A1227

Chapter 2
Machine Learning in Healthcare

Debasree Mitra
https://orcid.org/0000-0003-3723-9499
JIS College of Engineering, India

Apurba Paul
JIS College of Engineering, India

Sumanta Chatterjee
JIS College of Engineering, India

ABSTRACT

Machine learning is a popular approach in the field of healthcare. Healthcare is an important industry that provides service to millions of people and as well as at the same time becoming top revenue earners in many countries. Machine learning in healthcare helps to analyze thousands of different data points and suggest outcomes, provide timely risk factors, optimize resource allocation. Machine learning is playing a critical role in patient care, billing processing to set the target to marketing and sales team, and medical records for patient monitoring and readmission, etc. Machine learning is allowing healthcare specialists to develop alternate staffing models, intellectual property management, and using the most effective way to capitalize on developed intellectual property assets. Machine learning approaches provide smart healthcare and reduce administrative and supply costs. Today healthcare industry is committed to deliver quality, value, and satisfactory outcomes.

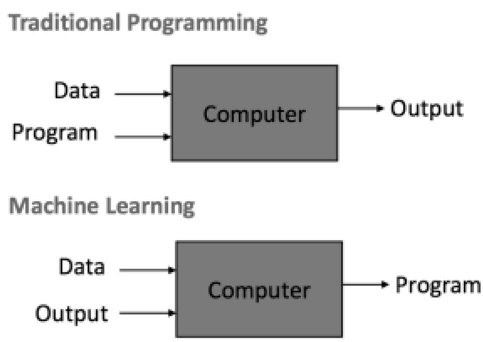
DOI: 10.4018/978-1-7998-3092-4.ch002

INTRODUCTION

What Is Machine Learning?

Machine learning (ML) explores algorithms that learn from data, builds models data and that model used for prediction, decision making or solving task. A computer program is to learn from experience E with respect to some class of task T and performance P. There are two components in ML i.e. learning module and reasoning module. Learner module takes input as experienced data and background knowledge and builds model. Models are used by reasoning module and reasoning module comes up with solution to the task and performance measure. Machine Learning algorithms can generate a mathematical model based on experience data known as training data to predict or decisions.

Figure 1. Traditional Programming vs Machine Learning



Machine learning algorithms are used in diagnose disease, banking system, healthcare, email filtering, and computer vision, data mining, robot control, Natural Language Processing, Speech Recognition, Machine Translation, Business Intelligence, Fraud Detection, Consumer sentiment etc where it is very helpful to develop an algorithm of specific instructions for performing the task. Machine learning is related to statistics and probability, which focuses on making predictions using computers.

What Is Healthcare?

Healthcare is the upgradation of health via technology for people. Health care is delivered by health professionals in allied health fields. Physicians and physician associates are a part of these health professionals. Dentistry, pharmacy, midwifery, nursing, medicine, optometry, audiology, psychology, occupational therapy, physical therapy and other health professions are all part of health care. It includes work done in providing primary care, secondary care, and tertiary care, as well as in public health.

Access to health care may vary across countries, communities, and individuals, largely influenced by social and economic conditions as well as health policies. Providing health care services means "the timely use of personal health services to achieve the best possible health outcomes" (Anthony & Bartlet, 1999). Factors to consider in terms of healthcare access include financial limitations (such as insurance coverage), geographic barriers (such as additional transportation costs, possibility to take paid time off of work to use such services), and personal limitations (lack of ability to communicate with healthcare providers, poor health literacy, low income) (Langley, 1996). Limitations to health care services affects negatively the use of medical services, efficacy of treatments, and overall outcome (well-being, mortality rates).

Health care systems are organizations established to meet the health needs of targeted populations. According to the World Health Organization (WHO), a well-functioning health care system requires a financing mechanism, a well-trained and adequately paid workforce, reliable information on which to base decisions and policies, and well maintained health facilities to deliver quality medicines and technologies (Muller & Guido, n.d.).

An efficient health care system can contribute to a significant part of a country's economy, development and industrialization. Health care is conventionally regarded as an important determinant in promoting the general physical and mental health and well-being of people around the world. An example of this was the worldwide eradication of smallpox in 1980, declared by the WHO as the first disease in human history to be completely eliminated by deliberate health care interventions.

Purpose of Machine Learning in Healthcare

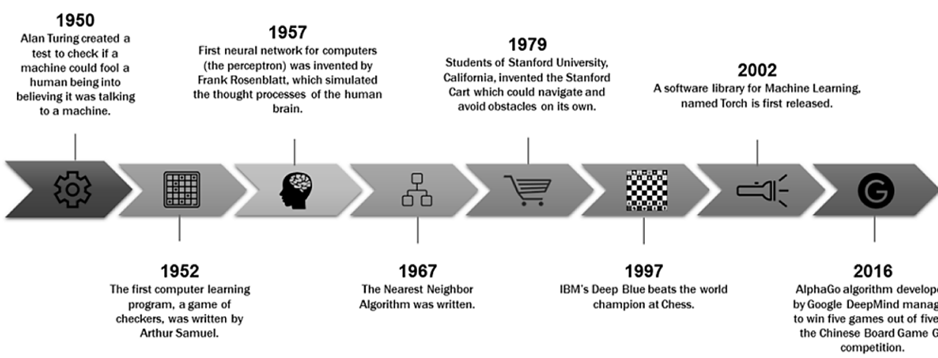
Machine learning has virtually endless applications in the healthcare industry. Today, machine learning is helping to streamline administrative processes in hospitals, map and treat infectious diseases and personalize medical treatments.

The healthcare sector has long been an early adopter of and benefited greatly from technological advances. These days, machine learning (a subset of artificial intelligence) plays a key role in many health-related realms, including the development of new medical procedures, the handling of patient data and records and the treatment of chronic diseases. As computer scientist Sebastian Thrum told the New Yorker in a recent article titled "A.I. Versus M.D., "Just as machines made human muscles a thousand times stronger, machines will make the human brain a thousand times more powerful."

Machine Learning Timeline in Healthcare

Artificial Intelligence is a broad scientific discipline combination with mathematics and computer science that goals to understand and develop an automated systems that provide the properties of human intelligence. Machine learning is a sub discipline of Artificial Intelligence where machine algorithms learn by training phase and predicts with test data. Machine learning uses a broader set of statistical techniques. Newer techniques such as Deep Learning are based on models artificial neural network with more accurate results for more complex data. Following diagram shows different evolution phase of machine learning in healthcare.

Figure 2. Machine learning timeline

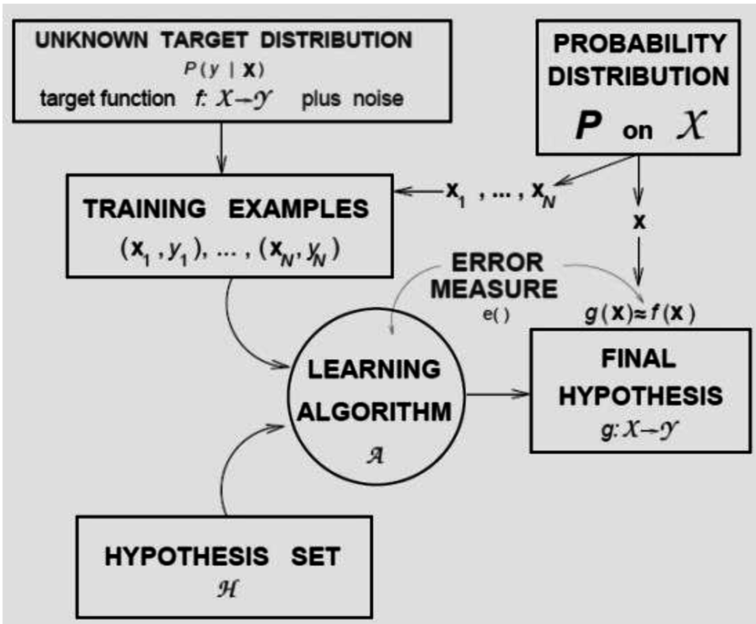


EXPLORATION OF HEALTHCARE DATA

The World Health Organization (WHO) collects and shares data on global health for its 194-member countries under the Global Health Observatory (GHO) initiative. Source users have options to browse for data by theme, category, indicator and by country. The metadata section allows for learning how data is organized. These healthcare datasets are available online or can be downloaded in CSV, HTML, Excel, JSON, and XML formats.Apart from that there are many resources are available like UCI,Kraggle,NCBI, Center for Disease Control (CDC) etc. Some popular datasets are available UCI Heart Decease Data,Diabetes, Breast Cancer Data, Lymphography Data Set,Lung Cancer,SPECT Heart Data Set,SPECTF Heart Data Set,Thyroid Disease,Mammographic Mass, EEG Database and many more.

MACHINE LEARNING TECHNIQUES IN HEALTHCARE

Figure 3. Structure of Machine Learning



Hypothesis

A hypothesis an explanation for some incidents or events. It is a provisional idea, an educated guess that requires some evaluation. A good hypothesis is testable; it can be either true or false. In science, a hypothesis must be falsifiable, meaning that there exists a test whose outcome could mean that the hypothesis is not true. The hypothesis must also be framed before the outcome of the test is known. A good hypothesis fits the evidence and can be used to make predictions about new observations or new situations. The hypothesis that best fits the evidence and can be used to make predictions is called a theory, or is part of a theory. But in Hypothesis in Science can be described as provisional explanation that fits the evidence and can be confirmed or disproved. Statistical hypothesis tests are techniques used to calculate a critical value called an "effect." The critical value can then be interpreted in order to determine how likely it is to observe the effect if a relationship does not exist. If the likelihood is very small, then it suggests that the effect is probably real. If the likelihood is large, then we may have observed a statistical fluctuation, and the effect is probably not real. For example, we may be interested in evaluating the relationship between the means of two samples, e.g. whether the samples were drawn from the same distribution or not, whether there is a difference between them. One hypothesis is that there is no difference between the population means, based on the data samples. This is a hypothesis of no effect and is called the null hypothesis and we can use the statistical hypothesis test to either reject this hypothesis, or fail to reject (retain) it. We don't say "accept" because the outcome is probabilistic and could still be wrong, just with a very low probability.

LEARNING ALGORITHM

Supervised Learning

Supervised learning indicates presence of a supervisor or teacher. There will be a teacher algorithm. Supervised learning is a learning technique in which we teach or train the machine using data which is labelled. That means, there should be an answer regarding each question. The machine is provided with a new set of data so that supervised learning algorithm analyses the training data and produces a correct results from labeled data.

Unsupervised Learning

Unsupervised learning is the training of machine using information that is not labeled and allows the algorithm to produce results on that information without guidance. Here the functions of machine are to classify unsorted information according to similarities, patterns and differences without any prior training of data. There is no teacher algorithm like supervised learning algorithm.

Semi Supervised Learning

Semi-supervised learning is a class of machine learning techniques that also make use of unlabeled data for training Here a small amount of labeled data is mixed with a large amount of unlabeled data. Semi-supervised learning falls between unsupervised learning (i. e. without any labeled data) and supervised learning (with training data).

Reinforcement Learning Algorithm

Reinforcement learning algorithm is training from particular situation on particular action upon it. It finds the best possible behaviour or path of various software and machines considering a specific situation.

FINAL HYPOTHESIS

Hypothesis sets and machine learning algorithms generate a final hypothesis or sometimes we called it is a model.

Goal

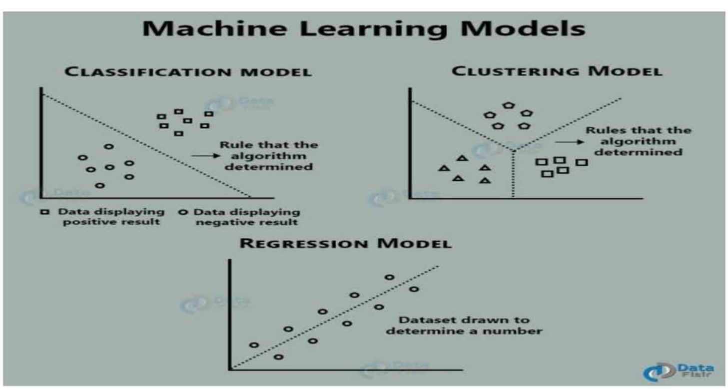
Model will be compared with predicted and actual data in test by accuracy,error rate, precision etc. Those model will give maximum accuracy and precision that will be selected in prediction to reach the goal.

DIFFERENT MODELS OF MACHINE LEARNING

There are three types of models in machine learning.

1. Classification model: This model trained with categorized data
2. Clustering model: This model identify patterns of some group of data
3. Regression Model: This model is to find the predictor between dependent and independent data.

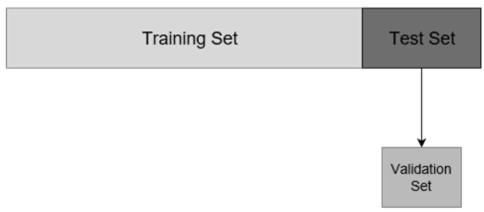
Figure 4. Machine Learning Models



Validation and Evaluation

Most machine learning engineers divide their data into three portions: training data, cross-validation data and testing data. The training data is used to make sure the machine recognizes patterns in the data, the cross-validation data is used to ensure better accuracy and efficiency of the algorithm used to train the machine and the test data is used to see how well the machine can predict new answers based on its training.

Figure 5. Training set, Test Set and validation Set



SUPERVISED LEARNING TECHNIQUES

Decision Trees

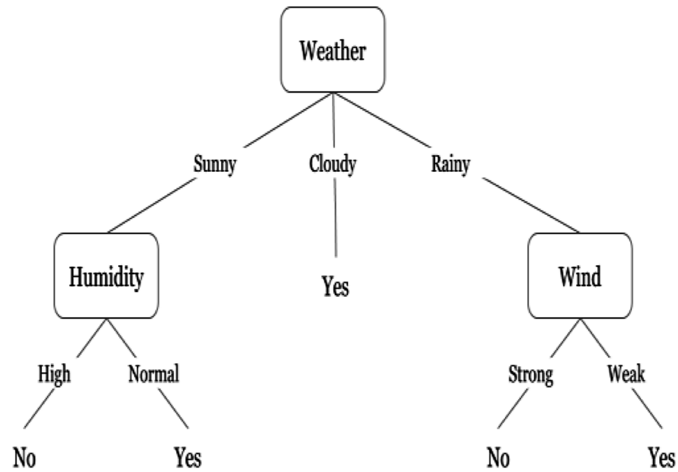
Decision tree makes regression or classification models in the form of a structure similar to that of a tree. Decision tree breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is developed in an incremental manner. The final result is a tree with only two types of nodes, decision nodes and leaf nodes.

Naïve Bayesian Algorithm

In machine learning we are often interested in selecting the best hypothesis (H) given data (D).In a classification problem, our hypothesis (H) may be the class to assign for a new data instance (D).One of the easiest ways of selecting the most probable hypothesis given the data that we have that we can use as our prior knowledge about the problem. Bayes' Theorem provides a way that we can calculate the probability of a hypothesis given our prior knowledge.

Bayes' Theorem is stated as: $P(H|D) = (P(D|H) * P(H)) / P(D)$

Figure 6. Decision tree



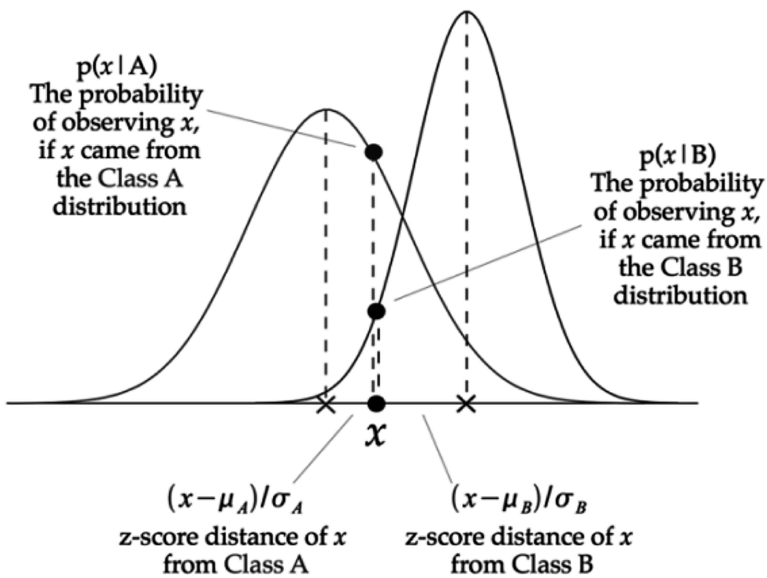
Where P(H|D) is the probability of hypothesis H given the data D. This is called the posterior probability. P(D|H) is the probability of data D given that the hypothesis H was true.P(H) is the probability of hypothesis H being true (regardless of the data). This is called the prior probability of H.P(D) is the probability of the data (regardless of the hypothesis). We are interested in calculating the posterior probability of P(H|D) from the prior probability p(H) with P(D) and P(D|H).After calculating the posterior probability for a number of different hypotheses, you can select the hypothesis with the highest probability. Naive Bayes can be extended to real-valued attributes, most commonly by assuming a Gaussian distribution.This extension of naive Bayes is called Gaussian Naive Bayes. Other functions can be used to estimate the distribution of the data, but the Gaussian (or Normal distribution) is the easiest to work with because you only need to estimate the mean and the standard deviation from your training data.

In the below figure two event A and B is represented by Gauss Naïve bayes Classification techniques by probability distribution function.

Neural Networks

Artificial Neural Networks are the most popular machine learning algorithms nowadays. The invention of these Neural Networks took place in the 1970s.But they have become popular due to the recent increase in computation tool like Python, R, MATLAB etc.

Figure 7. Naïve bayes Classifier



The neurons in human nervous system are able to learn from the past data and similarly the ANN is also capable to learn from the past data or trained data and provide responses in the form of predictions or classifications. ANNs are nonlinear statistical models which establishes a complex relationship between the inputs and outputs to discover a new pattern. A variety of tasks such as optical character recognition, face recognition, speech recognition, machine translation as well as medical diagnosis makes use of these artificial neural networks.

The basic concept is based upon three layer: Input Layers(IL),Hidden layers(HL),Output layers(OL).The input layer receives the input information in the form of various texts, numbers, audio files, image pixels, etc. Hidden Layers is the middle layer where some mathematical computations are done . These hidden layers can be single or multiple. Output Layer provides the result that we obtain through rigorous computations performed by the middle layer.

Support Vector Machine

Support Vector Machines (SVMs) are a classification strategy. SMVs work by transforming the training dataset into a hyperplane. There are two support vectors which easily define classes, and their margins, which are the lines parallel to the hyperplane defined by the shortest distance between a hyperplane and its support vectors. SVMs are able to classify both linear and nonlinear data.The distance between the hyperplane and the nearest data point from either set is known as the margin. The goal is to choose a hyperplane with the greatest possible margin between the hyperplane and any point within the training set, giving a greater chance of new data being classified correctly.

Figure 8. ANN

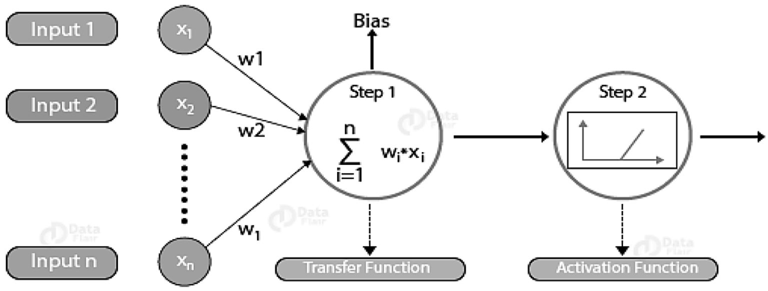
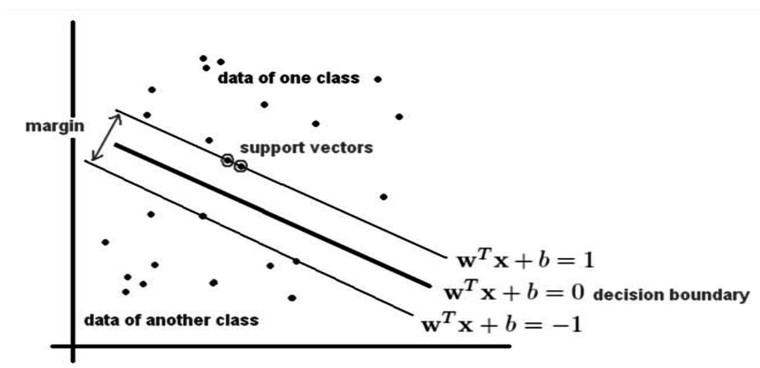


Figure 9. Support Vectors Machine



UNSUPERVISED LEARNING

Gaussian mixture Model

In this model we use the concept of normal distribution on subpopulation and Gaussian distribution on overall population. Gaussian Mixture Model (GMM) does not require the data to which the subpopulation belongs. This allows the model to learn the subpopulations automatically. Since we do not know the assignment of the subpopulation, it comes under unsupervised learning. With the help of GMMs, one can extract the features from speech data, healthcare data etc track the multiple factors in cases where there are a number of mixture components and the means that predict location of objects in a dataset. Many datasets can be easily modeled with the help of Gaussian Distribution. Therefore, one can assume that the clusters from different Gaussian Distributions. The core idea of model is that the data is modeled with several mixtures of Gaussian Distributions.

The single dimension probability density function of a Gaussian Distribution is as follows –

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\left(x-\mu\right)^2}{2\sigma^2}}$$

μ = Mean
σ = Standard Deviation
π = 3.14159…
e = 2.71828…

Hidden Markov Model

Hidden Markov models (HMMs) was developed by the Russian mathematician Andrey Andreyevich Markov in the early 1970s. HMMS are based on the theory of Bayes. HMMs are statistical models to capture hidden information from observable sequential symbols. They have many applications in sequence analysis in genomic. In a HMM, the system being modelled is assumed to be a Markov process with unknown parameters, and the challenge is to determine the hidden parameters from the observable parameters. A good HMM accurately models the real world source of the observed real data and has the ability to simulate the source. A lot of Machine Learning techniques are based on HMMs have been successfully applied to problems including healthcare data, speech recognition, optical character recognition, computational biology etc. HMMs are popular for their robust statistical foundation, conceptual simplicity and malleability; they are adapted fit diverse classification problems. In Computational Biology, a hidden Markov model (HMM) is a statistical approach that is frequently used for modelling biological sequences.

Principal Components Analysis

Principal component analysis (PCA) is a statistical procedure . PCA uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. This transformation is defined in such a way that the first principal component has the largest possible variance and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors are an uncorrelated orthogonal basis set. PCA is sensitive to the relative scaling of the original variables.

Figure 10. Basic Structure of Hidden Markov Model

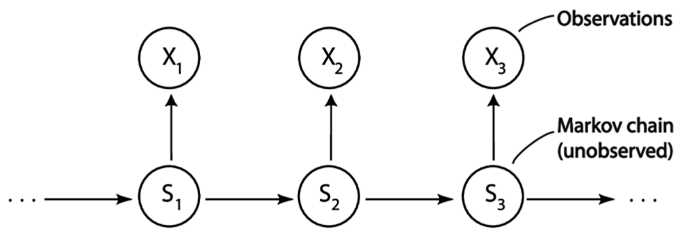


Figure 11. PCA

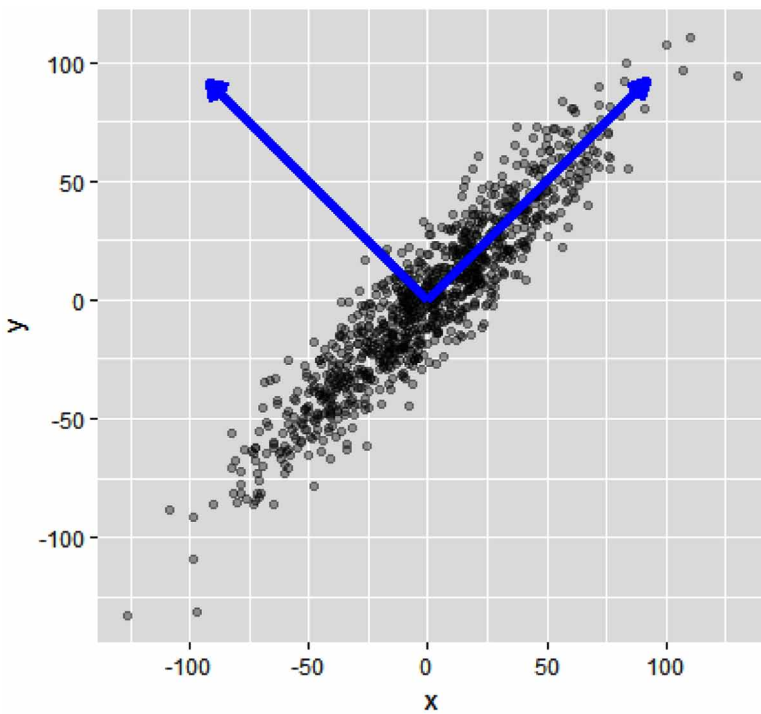
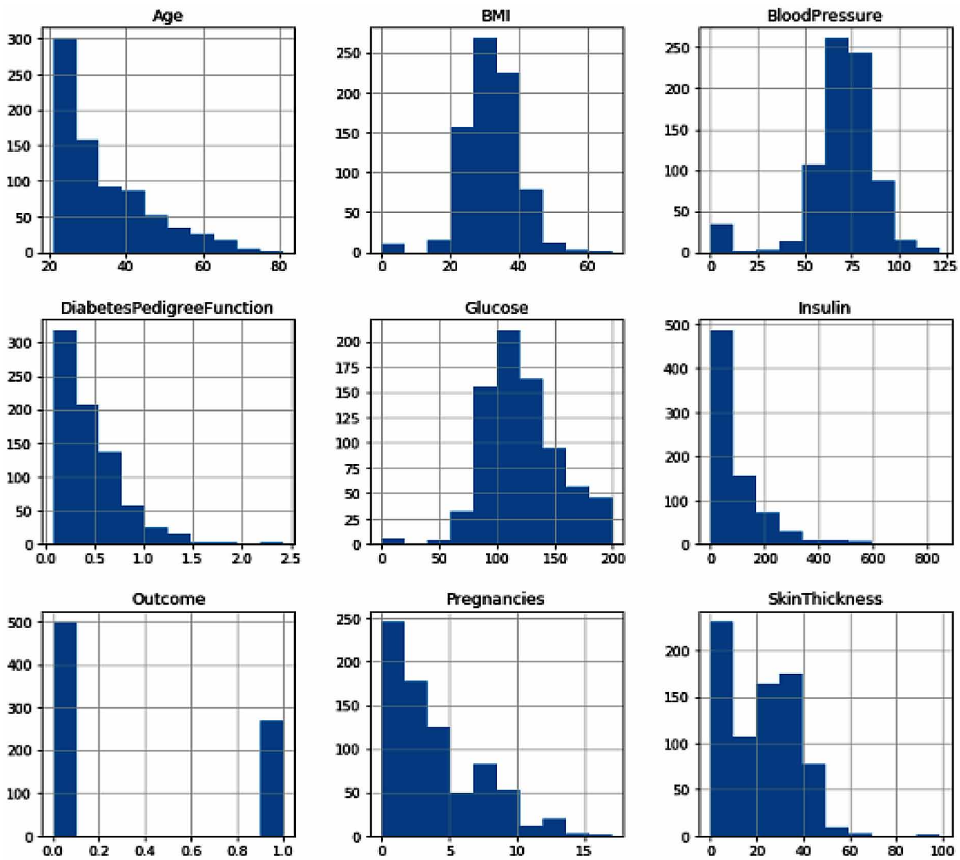


Figure 12. Summary about Datasets



SAMPLE CODING AND RESULTS

As a sample healthcare dataset we have consider PIMA Diabetic dataset here

```
import pandas as pd
from pandas.plotting import scatter_matrix
import matplotlib.pyplot as plt
from sklearn import model_selection
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
```

```
from sklearn.neighbors import KNeighborsClassifier
from sklearn.discriminant_analysis import
LinearDiscriminantAnalysis
from sklearn.naive_bayes import GaussianNB
from sklearn.svm import SVC
import seaborn as sns
url="D:/MYPAPER/diabetescsv/diabetes.csv"
names=['Pregnancies','GlucosePlasma','BloodPressureDiastolic','
SkinThicknessTriceps','Insulin','BMI','DiabetesPedigreeFunction
','Age','Outcome']
dataset=pd.read_csv(url,sep=',')
print(dataset.shape)
dataset.isnull()
dataset.dropna(inplace=True)
dataset.head(10)
```

Output

See Figure 13.

Figure 13. First 10 rows after data cleaning

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1
5	5	116	74	0	0	25.6	0.201	30	0
6	3	78	50	32	88	31.0	0.248	26	1
7	10	115	0	0	0	35.3	0.134	29	0
8	2	197	70	45	543	30.5	0.158	53	1
9	8	125	96	0	0	0.0	0.232	54	1

RESULTS FOR DECISION TREE

```
x=dataset.drop("Outcome",axis=1)
y=dataset["Outcome"]
```

```
from sklearn import model_selection
x_train,x_test,y_train,y_test=model_selection.train_test_
split(x,y,test_size=0.30,random_state=1)
from sklearn.tree import DecisionTreeClassifier
dtmodel=DecisionTreeClassifier()
dtmodel.fit(x_train,y_train)
predictions=dtmodel.predict(x_test)
from sklearn.metrics import classification_report
print(classification_report(y_test,predictions))
from sklearn.metrics import confusion_matrix
print(confusion_matrix(y_test,predictions))
Output:
precision    recall  f1-score   support
          0       0.63      1.00      0.77       146
          1       0.00      0.00      0.00        85
avg / total       0.40      0.63      0.49       231
[[146   0]
 [ 85   0]]
```

RESULTS FOR KNN CLASSIFIER

```
x=dataset.drop("Outcome",axis=1)
y=dataset["Outcome"]
from sklearn import model_selection
x_train,x_test,y_train,y_test=model_selection.train_test_
split(x,y,test_size=0.30,random_state=1)
from sklearn.neighbors import KNeighborsClassifier
knnmodel=KNeighborsClassifier()
knnmodel.fit(x_train,y_train)
predictions=knnmodel.predict(x_test)
from sklearn.metrics import classification_report
print(classification_report(y_test,predictions))
from sklearn.metrics import confusion_matrix
print(confusion_matrix(y_test,predictions))
Output:
precision    recall  f1-score   support
          0       0.80      0.88      0.84       146
          1       0.75      0.62      0.68        85
avg / total       0.78      0.78      0.78       231
```

```
[[128  18]
 [ 32  53]]
```

Results for SVM Classifier

```
x=dataset.drop("Outcome",axis=1)
y=dataset["Outcome"]
from sklearn import model_selection
x_train,x_test,y_train,y_test=model_selection.train_test_
split(x,y,test_size=0.30,random_state=1)
from sklearn.svm import SVC
svmmodel=SVC()
svmmodel.fit(x_train,y_train)
predictions=svmmodel.predict(x_test)
from sklearn.metrics import classification_report
print(classification_report(y_test,predictions))
from sklearn.metrics import confusion_matrix
print(confusion_matrix(y_test,predictions))
Output:
   precision    recall  f1-score   support
          0       0.63      1.00      0.77       146
          1       0.00      0.00      0.00        85
avg / total       0.40      0.63      0.49       231
[[146   0]
 [ 85   0]]
```

RESULTS FOR GAUSSIANNB CLASSIFIER

```
x=dataset.drop("Outcome",axis=1)
y=dataset["Outcome"]
from sklearn import model_selection
x_train,x_test,y_train,y_test=model_selection.train_test_
split(x,y,test_size=0.30,random_state=1)
from sklearn.naive_bayes import GaussianNB
gnbmodel=GaussianNB()
gnbmodel.fit(x_train,y_train)
predictions=gnbmodel.predict(x_test)
from sklearn.metrics import classification_report
```

```
print(classification_report(y_test,predictions))
from sklearn.metrics import confusion_matrix
print(confusion_matrix(y_test,predictions))
  Output:
           precision    recall  f1-score   support
          0       0.80      0.88      0.84       146
          1       0.75      0.62      0.68        85
avg / total       0.78      0.78      0.78       231
[[128  18]
 [ 32  53]]
```

Results for LDA Classifier

```
x=dataset.drop("Outcome",axis=1)
y=dataset["Outcome"]
from sklearn import model_selection
x_train,x_test,y_train,y_test=model_selection.train_test_
split(x,y,test_size=0.30,random_state=1)
from sklearn.discriminant_analysis import
LinearDiscriminantAnalysis
ldamodel=LinearDiscriminantAnalysis()
ldamodel.fit(x_train,y_train)
predictions=ldamodel.predict(x_test)
from sklearn.metrics import classification_report
print(classification_report(y_test,predictions))
from sklearn.metrics import confusion_matrix
print(confusion_matrix(y_test,predictions))
```

Output

```
             precision    recall  f1-score   support
          0       0.63      1.00      0.77       146
          1       0.00      0.00      0.00        85
avg / total       0.40      0.63      0.49       231
[[146   0]
 [ 85   0]]
```

K- FOLD VALIDATION

In summary, cross-validation combines (averages) measures of fitness in prediction to derive a more accurate estimate of model prediction performance.

In our dataset we have select 30% as validation size and rest as training dataset splits values with 5 (5 Fold Cross validation Procedure) with a random seed value 7.

```
from sklearn import model_selection
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
from sklearn.linear_model import LogisticRegression
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.discriminant_analysis import
LinearDiscriminantAnalysis
from sklearn.naive_bayes import GaussianNB
from sklearn.svm import SVC
x=dataset.drop("Outcome",axis=1)
y=dataset["Outcome"]
validation_size=0.30
seed = 7
x_train, x_validation, y_train, y_validation = model_selection.
train_test_split(x, y, test_size=validation_size, random_
state=seed)
seed = 7
scoring = 'accuracy'
models = []
models.append(('LR', LogisticRegression()))
models.append(('LDA', LinearDiscriminantAnalysis()))
models.append(('KNN', KNeighborsClassifier()))
models.append(('CART', DecisionTreeClassifier()))
models.append(('NB', GaussianNB()))
models.append(('SVM', SVC()))
results = []
names = []
for name, model in models:
        kfold = model_selection.KFold(n_splits=5, random_
state=seed)
        cv_results = model_selection.cross_val_score(model, x_
```

```
train, y_train, cv=kfold, scoring=scoring)
        results.append(cv_results)
        names.append(name)
        msg = "%s: %f (%f)" % (name, cv_results.mean(), cv_
results.std())
        print(msg)
 Output:
LR: 0.776625 (0.076214)
LDA: 0.772851 (0.072747)
KNN: 0.726171 (0.062883)
CART: 0.704053 (0.084624)
NB: 0.761740 (0.064313)
SVM: 0.657547 (0.077963)
```

RESULT ANALYSIS

```
fig = plt.figure()
fig.suptitle('Algorithm Comparison')
ax = fig.add_subplot(111)
plt.boxplot(results)
ax.set_xticklabels(names)
plt.show()
```

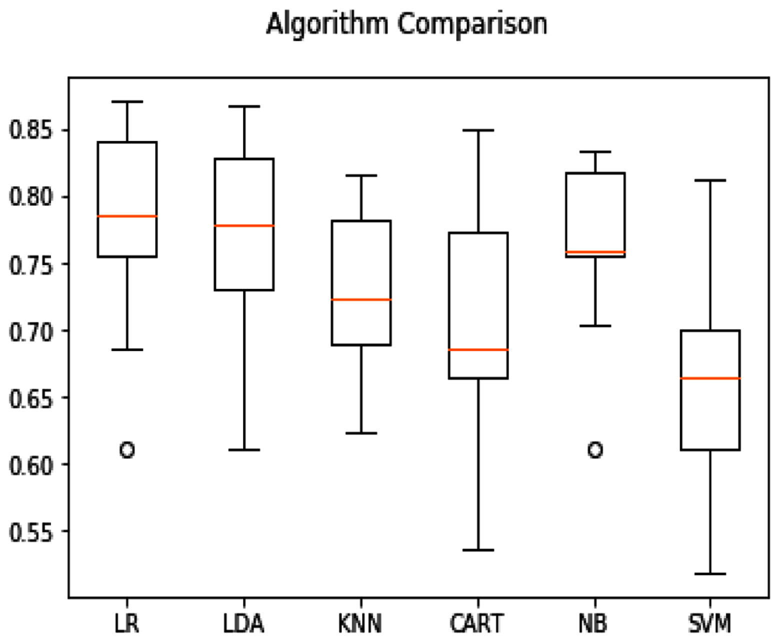
The above analysis shows the LDA and GNB will give the highest accuracy. Whereas LR and KNN are also providing a reasonably good accuracy, The CART and SVM both are not providing satisfactory accuracy. But K-Fold cross validation algorithm is also showing nearer accuracy . Here valur of k=5 and we have calculated accuracy as validation score and at the end we have calculated mean of accuracy.In case of K-Fold validation for CART and SVM accuracy level have been increased slightly and it cleared from above figure.

APPLICATIONS OF MACHINE LEARNING IN HEALTHCARE

Identifying Diseases and Medical Diagnosis and Analysis

One of the chief Machine Learning applications in healthcare is the identification and diagnosis of diseases .This can include anything from diabetes, heart diseases,

Figure 14. Algorithm Comparison



cancers which are tough to predict during the initial stages or to identify some genetic diseases.

Medical Imaging Diagnosis

Machine learning and deep learning are both responsible for the breakthrough technology called Computer Vision. E. g. InnerEye developed by Microsoft which works on image diagnostic tools for image analysis. process.

Smart Health Records

Machine Learning will provide best-predicted values related to the patients in their respected health condition and also it helps to analyze the previous health records. For that purpose we need to maintain a repository or in other words warehouse where we can maintain data related to the patients and their treatment. Regarding each hospital there should be health record system. These health records of the patients can be accessed only doctors or hospital by an identification number. This kind of approaches required an web application with prediction system. The records can be sequential or hierarchical. Admin can only change the order of data.

Clinical Trial and Research

Machine learning has immense potential in the field of clinical trials in pharmacy industry. It will decrease the clinical trial cost and save time. Applying ML-based predictive analytics to identify potential clinical trial results.

Crowdsourced Data Collection

Crowdsourcing is the new trends in market. Different kind of medical data of all the ages are collected. This live health data has great significant to researcher and scientists now a days. IBM recently partnered with Medtronic to predict through ML by available diabetes and insulin data in real time based on the crowdsourced information. With the advancements of IoT and bigdata with machine learning the healthcare industry is in booming condition. All over the world research has been faster than previous condition.

Better Radiotherapy

One of the most widely applications of machine learning in healthcare is in the field of Radiology. Through computer vision and medical image analysis we can model many tissue regions, cancer foci, etc by using complex equations. Since Machine Learning based algorithms learn from the trained dataset of different samples available globally and hence it becomes easier to diagnose and predict the factors responsible for cancer .As for example different classification based approaches for prediction of cancer stages. As for example Google's DeepMind Health is actively helping researchers in UCLH to develop algorithms which can detect the difference between healthy and cancerous tissue .

Outbreak Prediction

Machine learning is useful in monitoring and predicting epidemics around the world. There are huge data collected from satellites, social media updates, website information, etc through natural language processing. Artificial neural networks help to gather this information and predict everything from malaria, cholera, dengue outbreaks to severe diseases. Especially, this is helpful in third-world countries as they lack in medical infrastructure and educational systems due to huge population. As for example is the ProMED-mail which is an Internet based reporting system which monitors evolving diseases and emerging ones and provides outbreak reports in real-time.

Drug Discovery and Manufacturing

Manufacturing a new drug is very expensive and a long process because they are depended on variety of tests and their results.. With the advancements in ML machine learning can next-generation sequencing and precision medicine can be useful to help to cure many health diseases. Unsupervised machine learning algorithm can identify patterns in data without providing for any predictions.

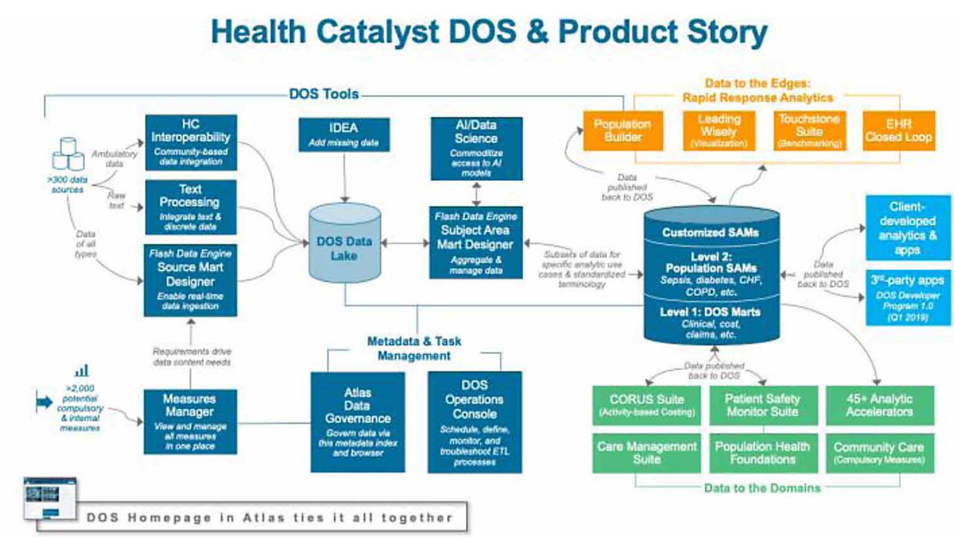
THE IDEAL SYSTEMS MACHINE LEARNING IN HEALTHCARE: CASE STUDY

As for example a cloud-based open system such as the Health Catalyst Data Operating System. Its aim to answer healthcares growing data needs by combining the features of data warehousing, clinical data repositories, and HIEs in a single, common-sense technology platform.

Following components are here

1. Cloud-Based Data Operating System
2. Health Information Exchange (HIE)
3. Hybrid Big data SQL Architecture-HTAP
4. Big Data Architecture

Figure 15. Health Catalyst Data Operating System



CONCLUSION

Many sectors like finance, education, agriculture are using machine learning and hence healthcare cannot stand behind. Google has developed an ML algorithm to identify cancerous tutors. Stanford is using it to identify skin cancer. People should stop thinking machine learning as a concept for future .Instead we should embrace the tools and make the use of all opportunities . These applications of machine learning are advancing the field of healthcare into a completely new arena of opportunities..

REFERENCES

Anthony, M., & Bartlet, P. (1999). *Neural Network Learning: Theoretical Foundations*. Cambridge University Press. doi:10.1017/CBO9780511624216

Langley, P. (1996). *Elements of Machine Learning*. Morgan Kaufmann.

Muller, A.C., & Guido, S. (n.d.). *Introduction to Machine Learning with Python.* O'Reilly.

Chapter 3
Detection of Tumor From Brain MRI Images Using Supervised and Unsupervised Methods

Kannan S.
Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

Anusuya S.
Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

ABSTRACT

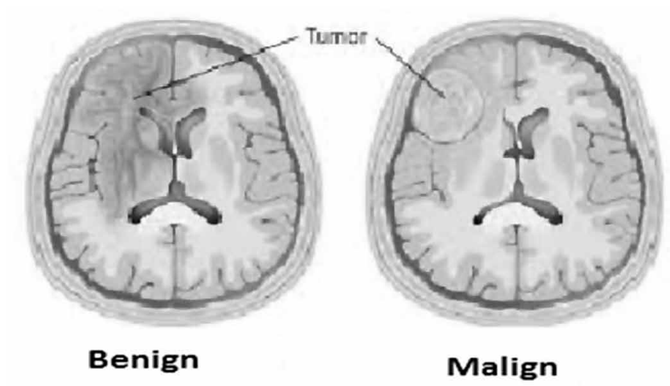
Brain tumor discovery and its segmentation from the magnetic resonance images (MRI) is a difficult task that has convoluted structures that make it hard to section the tumor with MR cerebrum images, different tissues, white issue, gray issue, and cerebrospinal liquid. A mechanized grouping for brain tumor location and division helps the patients for legitimate treatment. Additionally, the method improves the analysis and decreases the indicative time. In the separation of cerebrum tumor, MRI images would focus on the size, shape, area, and surface of MRI images. In this chapter, the authors have focused various supervised and unsupervised clustering techniques for identifying brain tumor and separating it using convolutional neural network (CNN), k-means clustering, fuzzy c-means grouping, and so on.

DOI: 10.4018/978-1-7998-3092-4.ch003

1. INTRODUCTION

Tumor refers to a mass of tissue which controls the development of growing of further tissue. It is an intracranial strong neoplasm and a gathering of anomalous cells develop inside, around the brain or cerebrum through uncontrolled cell division. Brain is the inward piece of the focal sensory system (Aggarwal & Kaur, 2012). Kindhearted tumor is portrayed by an ordinary shape, does not suddenly extend, not attack nonadjacent cells and neighboring sound tissues. Moles are a case of considerate tumors and premalignant tumor is a precancerous stage that can be viewed as a sickness, may prompt disease if not appropriately treated. Harm is the tumor type which develops like normal tissue that attacks solid neighboring tissues and can eventually bring about death. The term threatening is fundamentally a restorative term which alludes to an extreme advancing ailment and harmful tumor is utilized to depict malignant growth. In the figure.1 shows the benign and malignant tumors.

Figure 1. Benign and Malignant Tumor



The side effects for brain tumors can be perceived by spewing, queasiness, cerebral pain, sudden difference in character or conduct, deadness and shortcoming. Now, loss of sensation and memory can be experienced by the patient (Aggarwal & Kaur, 2012). The brain tumor division procedure contains preprocessing, extraction of highlights from MRI images, and division utilizing administered or solo strategies.

2. LITERATURE SURVEY

In the event that any of the evaluations, it must be recognized convenient and found precisely (Kaur & Rani, 2016)[5]. For example, Magnetic Imaging Resonance (MRI) and Computed Tomography (CT) are broadly used to identify the tumor. Among these restorative imaging modalities, MRI is most broadly utilized and exceedingly favored non-intrusive strategy in biomedical, radiology and therapeutic imaging fields because of its ability to identify and envision better subtleties in the inner structure of the body by creating three dimensional high goals point by point anatomical images without the utilization of any harming radiations(Pradhan, 2010)[10]. The brain MRI division into a few cerebrum tissues, for example, grey matter (GM), white matter (WM) and cerebrospinal liquid (CSL) is exceedingly fundamental for the conclusion of different ailments. This method is essentially used to distinguish the itemized contrasts in the tissues in non-obtrusive style which have not been analyzed by other imaging systems includes Computed Tomography (Sathies Kumar, 2017).

One of the serious issues in this entire procedure is isolating the anomalous cells from the remainder of the image content which is known as the procedure of division. The manual division is very testing just as tedious undertaking because of complex structure of the cerebrum and nonappearance of well-characterized limits among various brain tissues. In spite of the fact way toward it's partitioning the ideal area exceeding testing and confused however it has increased tremendous significance and a few examinations have been led in improving the exactness of this assignment (Freixenet et al., 2002)[8](Logeswari & Karnan, 2010).

The outcome from the diverse image division methods helpful in acquiring highlights of divided tumor locale (Sathies Kumar, 2017). Various research work has been done and few calculations are proposed with distinguishing position, limit of tumors consequently so they can do assist conclusion at their soonest. The examination exhibited in this work surveys the strategies and procedures of programmed division brain tumor from the MRI images. The remainder work and its areas are organized in the consequent way: In the second segment exhibits a conventional strategy received in the procedures executed for brain tumor division, trailed by a nitty gritty Literature overview in the third segment exhibits the general examination and assessment of the outcomes pursued by the last ends and future proposals in the closure segments.

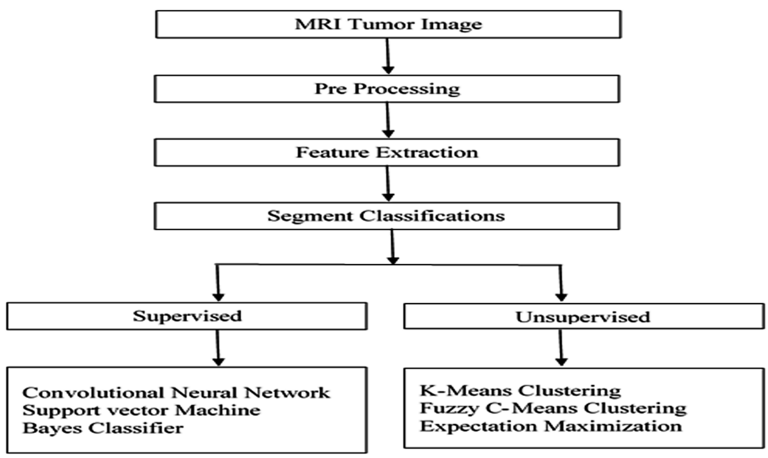
3. EXISTING METHODOLOGY

In the existing work, medical image processing to extract features and segment tumor from MRI images. It also dealt about Supervised, Unsupervised classification techniques. Figure 2, Elaborates the steps involved in segmentation of MRI image.

3.1 Pre Processing

Producing images from various therapeutic imaging procedures acquire pointless clamor into the image. A great deal of clamor accompanies in MRI, CT examine, Mammographic Image or and so on. This commotion is an obstacle when dividing tumor distinct from the given information image. To overcome this issue, first preprocess the image so as to expel undesirable exceptions, and then send for further preparing. In this preprocessing step deals procedures like clamor evacuation, channel application, image upgrade, standardization, and so on.

Figure 2. Brain Tumor Detection Systems



3.2 Feature Extraction

Post handling step is needed when the image has been separated into segments, so as to improve any edges and obscure any undesirable subtleties. This progression is called include extraction where highlights from the image are removed for examination that improved the tumor locale. In that most utilized component extraction steps includes Morphological activities, edge discovery strategies or histogram leveling.

3.3 Segmentation Classifications

Separating a image into various portions to further translation and division of image accomplished using numerous points of view. MR Images contain a high measure of information that makes the assignment of deciphering difficult and repetitive for a radiologist and clinical imaging authority. Similarly, the outcomes could diverse relying on the experience of the specific authority (Khadem, 2010).Additionally, various imaging frameworks present demand in the images, subsequently making it hard to fragment cerebrum tumor and give an adequate presentation. The significant division helps to compute the quantitative proportion of tumor in the cerebrum which is basic for treatment of patient and follow up of the infection.

An objective point of huge number of PC vision, image preparing and AI based applications recognized and separate the significant examples or indispensable highlights from the image information described by the machine for additionally gritty clarification (Tang et al., 2000)[7](Pradhan, 2010). The analysis of cerebrum tumors and complex illnesses from radiographs is one of the most significant testing because of high-time utilization and mutilation among these images help clinicians. The fundamental point of different research gatherings exhibit solid calculations that perform towards precise division. In this method, leads to develop a vigorous just as to guarantee a protected conclusion framework.

3.4 Supervised Learning Method

In the programmed cerebrum tumor location, prepared the model utilizing the regulated and solo AI strategies. As discussed above, many administered and unaided procedures are accessible for underlying trials as we individually managed solo AI methods utilizing CNN and K-Means. Additionally, the work utilized the morphological administrator for our examinations and these systems are portrayed underneath. The flow of brain tumor detection using supervised learning technique discussed in sections 3.4.1

3.4.1 Convolutional Neural Network (CNN)

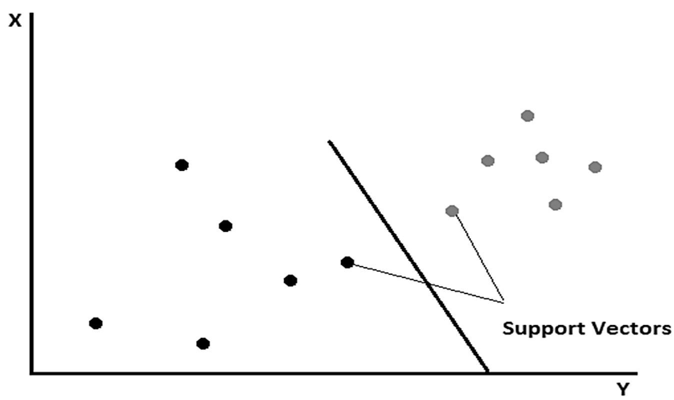
In computational sciences and in AI, the Convolutional neural systems (CNNs) increased the acknowledgment for assortment of image handling applications especially in programmed medicinal image division. As image segmentation isolates a image into various parts, the target of segmentation is to plan straightforward calculations as conceivable that process the images and examine the data in progressively powerful, significant, reasonable and advantageous way. Utilizing CNN designs as a model such significant data from medicinal images and radiographs promptly removed out with increased exactness and improved execution time.

The discriminative model CNN legitimately gains from clarified images with no earlier information (Liu, 2015). CNN based systems utilized the dataset preparation to educate a system and these prepared systems anticipate the class names as well as concentrate out the significant highlights includes designs, edges, lines. In that, further train the other arrangement of classifiers and patches of the data extracted out from the MRI images handled through convolution based channels. In that method acquired the intricate highlights and help to yield the area, size of the tumors dependent on their registered class scores. Besides, the CNN models have additionally bit of programmed learning the perplexing highlights which identified the solid tissues too strange tissues procured from MRI images.

3.4.2 Support Vector Machine (SVM)

SVM is one of the best predictable strategies for arranging the highlights. In SVM, the arrangement of images is fundamentally partitioned into two different resultant classes and the order is performed by finding the hyper-plane rule that separates the two classes as shown in figure 3. SVM builds a hyper plane receiving a part work (Sathies Kumar, 2017)[19] exhibited in the below figure3. In that, element vectors on the left half of the primary hyper plane place with the class - 1 whereas the component vectors assigned on the correct side of the fundamental hyper plane compares to the class +1.

Figure 3. Hyper Plane Classifications



The segmentation with SVM predominantly relies on the following stages (a) include extraction from preparing image (b) determination of SVM model (c) readiness of informational index (d) SVM preparing a classification plane (Mengqiao et al., 2017)[21]. A. Kumar et al. (Kumar, 2017) examination utilized Support vector machine integrated with K-implies grouping and Principle Component Analysis for extractions that order the tumor locale inside the cerebrum. In the given approach information including cerebrum outputs was prepared utilizing bolster vector machine whereas the tumor was sectioned utilizing k-means and PCA. The SVM classifier discovers the class of the tumor identified and had an exactness of 96% for tumor by the outcomes. Other than sectioning the tumor locale, the work additionally gave a definite data on K-means and PCA between them. Additionally, G. Gupta et al., SVM combined with Fuzzy C-Means to order the images and their methodology to apply FCM to segment the image and use of SVM to further group the images that gave progressively upgraded and better outcomes (Gupta & Singh, 2017).

3.4.3 Bayesian Approach

In this methodology, information is accepted to pursue a multivariate ordinary dispersion, where mean and covariance are evaluated from the preparation informational index (Jobin Christ et al., 2009). The technique consolidates a chart based calculation and Bayesian model portions the edema more over. Additionally it tends to be stretched out to vectorial factors to work on multi-methodology images. A Bayesian system is a model of compound likelihood circulation capacity of a lot of variable like coordinated non-cyclic diagram with a likelihood table for every hub. The hub in a Bayesian system relies on various factors in a space, and the circular segments among hubs to the reliance connections among the factors (Jobin Christ et al., 2009). The typical tissue classes are characterized by the enlisted spatial chart book to the patient images and the tumor spatial earlier is determined from the distinction image through histogram analysis (Corso et al., 2006). Also the likelihood conveyance of tumor and edema has been thought to be an ordinary appropriation that not right in the all cases (Corso et al., 2008). On account of edema, the creators have expected a small amount of white issue likelihood for edema (Jagath, 2001) incorporate the subsequent model-brain full affinities into the staggered division by weighted accumulation calculation and apply the system undertake the partitioning cerebrum tumor and edema in multichannel MR volumes.

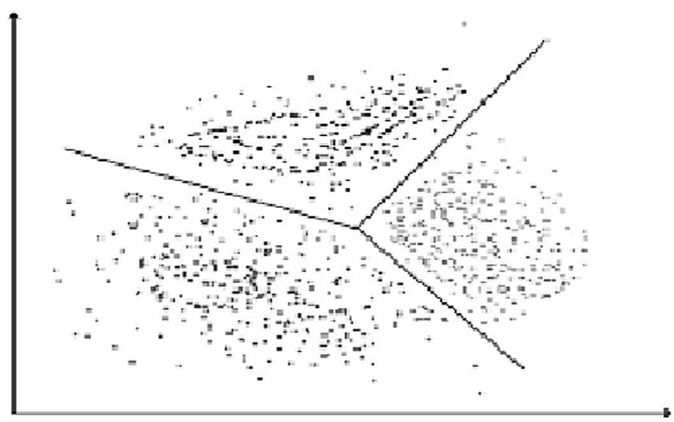
3.5 Unsupervised Learning Technique

In the learning method, the groups are framed with class labels. Here, discussed about K-implies grouping and morphological administrators that work as an unaided learning strategy for brain tumor discovery.

3.5.1 Fuzzy C-Means Clustering

Grouping approach has been broadly utilized in various computational spaces includes AI, PC vision and image handling. In addition, bunching method has late advanced into different biomedical and social insurance applications prevalently for the discovery of strange brain tissues (tumors) from radiographs procured through attractive reverberation imaging (MRI) methodology (Benson, 2016). Dunn et al., presented a bunching based Fuzzy C-Means (FCM) approach, which extraordinarily encouraged the division capability that bifurcates gathering of information into two or different distinctive relating groups as displayed in Figure 4

Figure 4. Clustering with FVM.



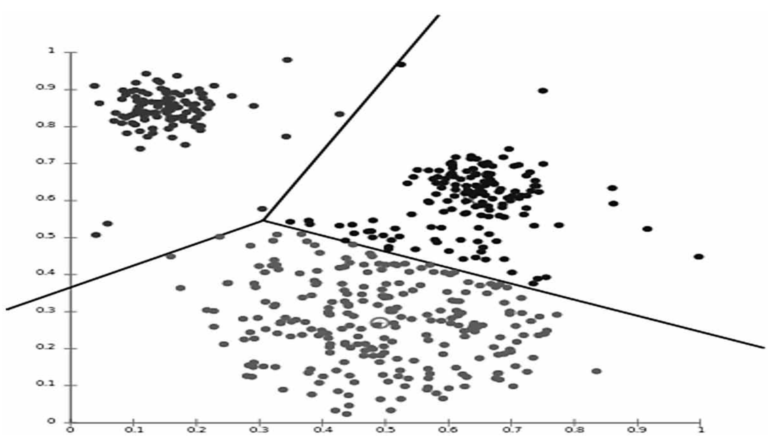
In this strategy, the examination of cerebrum tumor division recommended the method for FCM grouping calculation and the division of brain tumor bring dynamic cells, necrotic center and edema (Martin et al., 2001). In Suganya et al.,(2016) checked Fuzzy C Means calculation and its different applications in restorative imaging, design acknowledgment, bioinformatics and information mining (Suganya & Shanthi, 2012). Pham et. al., built up a novel grouping calculation by incorporating fluffy entropy bunching for division of brain tissues (tumors) from

MRI images (Pham et al., 2018). Kumar et. al., proposed a changed intuitionistic fluffy c-implies calculation (MIFCM) to systematically take care of the enhancement issue utilizing lag range strategy for dubious multipliers. The proposed MIFCM strategy concentrates cerebrum MRI information by defeating the confinements of commotion and loose estimation (Kumar, 2018). Shanmuga Priya et. al., foreseen FCM based staggered division by consolidating fluffy c-implies for distinguishing the tumor tissues and edema among brain MRI images. The bunching procedure improved by combining different bits dependent on the spatial data to perform effective division (Priya, 2018).

3.5.2 K-Means Clustering

In a standard K- Means calculation, there are four stages includes: introduction, grouping, computational and intermingling condition. The procedure introduces by partitioning a specific informational into stable positive K number of bunches with the goal that k centroids. The subsequent stage chooses a point which is in colleague to a given informational collection and changes it to the nearest centroid. The main gathering is finished when initial step is done which possibly happens when no any point exist of last. The result of initial step brings about K new centroids of the groups which ought to be recalculated toward the finish of the initial step. When new K centroids are initiated, there is a need of another association among similar informational collection focuses and the closest new centroid. In this way, a circle is made which aides in examining the changed area of K centroids in each stage except if all centroids come to static. As it were, centroids don't move any more appeared in Figure 5.

Figure 5. Clustering though K-means.



In Vijay et. al., (Vijay & Subhashini, 2013) contemplated the issue of marking in image segmentation uncommonly with regards to robotized brain tumor identification. They proposed a strategy that utilized morphological tasks for preprocessing and K-implies procedure with a slight change that was to diminish the quantity of cycles required for legitimate bunching by recommending the registered separation among a group focus and information point under assessment, which is put away in an information structure. This mix of K-implies bunching, and morphological tasks created 95% precise outcomes on an example space of 100 MRI images, as delineated in the outcomes. Dr. Patil et al. proposed a PC supported application to segmentation tumor from the given MRI filters (Patil, 2005). The division thought embraced for the investigation worked with an amalgamation of K-means bunching and Fuzzy C means based grouping draws near. Four distinct modalities of images were tried for investigations and the outcomes were produced dependent on parameters like, Mean Square Error (MSE), Contrast, Correlation, Max blunder, Area, and so on. The examination and results presumed that the technique proposed was strong, precise and efficient.

3.5.3 Expectation Maximization

In the model based tumor segmentation method utilized an Expectation Maximization (EM) separated the solid and the mild tissues. A lot of tumor attributes are exhibited which is exceptionally fundamental for precise division. However, the work examined the segment tumor area. EM steps are shown in the accompanying advances (Balafar et al., 2010):

Step1: Initialize mean and Covariance framework utilizing K-Means.
Step2: Calculate participation likelihood of each preparation information.
Step3: Compute mean and fluctuation of each Gaussian part utilizing enrollment capacity got in stage 2.

The stage 2 and 3 are rehashed until assembly. Gauss blend vector of each class is derived by EM preparing information for that class. The utilizations of the EM calculation brain MR image division accounted by (Wells et al., 1996) and (Leemput et al., 1999).

4. COMPARATIVE STUDY AND DISCUSSION

In the examination on different segmentation calculations, MATLAB tool used K-Means (Patil, 2005), Fuzzy C Means (Suganya & Shanthi, 2012) procedures whereas

Ibrahim et al. utilized CNN for order and exactness of the methods analyzed. As per the outcomes, K-Means and Fuzzy C Means calculations had a similar precision. G. Rao et al. (Kamnitsas et al., 2017) added to the space by differentiating Fuzzy C-Means and K-Means bunching methods. In this segmentation through FCM and K-Means was contrasted with Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Peak Time (PTime) and region estimation. Consequently investigation demonstrated that FCM had increased exactness of roughly 93%, alongside lower PTime in contrast with K-Means is about 76% precision.

Image Segmentation is one of the most central ideas in PC vision. Segmentation intends to change the representation of image to separate significant data from it. This work displays a complete survey of various division techniques for cerebrum tumor segmentation. Cross breed approaches integrated with various calculations of these techniques had exhibited. The different robotized and semi-mechanized procedures examined on constant usage contribute the greatness of PC innovation to aid the field of therapeutic science. The examination recommends the regulated learning technique have better exactness while inclination based strategies s are precise and require lesser assets.

5. CONCLUSION

In this book chapter, various cerebrum tumor segmentation methods are surveyed and also the steps for performing segmentation is explored and analytical results for the same are analyzed with respect to the size and shape of images. This sort of improvements in administered learning structures help in institutionalizing the present strategies that help in clinical acknowledgment.

REFERENCES

Aggarwal, R., & Kaur, A. (2012). Comparative Analysis of Different Algorithms For Brain Tumor Detection. *International Journal of Scientific Research*.

Balafar, M. A., Ramli, A. R., Saripan, M. I., & Mashohor, S. (2010). Review of brain MRI segmentation methods. *Artificial Intelligence Review*, *33*(3), 261–274. doi:10.100710462-010-9155-0

Benson, C. C. (2016). Brain Tumor Segmentation from MR Brain Images using Improved Fuzzy c-Means Clustering and Watershed Algorithm. *2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 10.1109/ICACCI.2016.7732045

Corso, J. J., Sharon, E., Brandt, A., & Yuille, A. (2006). Multilevel Segmentation and Integrated Bayesian Model Classification with an Application to Brain Tumor Segmentation. *MICCAI*, *4191*, 790–798. doi:10.1007/11866763_97 PMID:17354845

Corso, J. J., Sharon, E., Dube, S., El-Saden, S., Sinha, U., & Yuille, A. (2008, May). Efficient Multilevel Brain Tumor Segmentation with Integrated Bayesian Model Classification. *IEEE Transactions on Medical Imaging*, *27*(5), 629–640. doi:10.1109/TMI.2007.912817 PMID:18450536

Dvorak, P., & Menze, B. (2015). Structured prediction with convolutional neural networks for multimodal brain tumor segmentation. *Proceeding of the Multimodal Brain Tumor Image Segmentation Challenge*, 13-24.

Freixenet, J., Munoz, X., Raba, D., Marti, J., & Cufi, X. (2002). Yet another survey on image segmentation: Region and boundary information integration. *Proc. 7th Eur. Conf. Computer Vision Part III*, 408–422. 10.1007/3-540-47977-5_27

Girshick, R. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 10.1109/CVPR.2014.81

Gupta & Singh. (2017). Brain Tumor segmentation and classification using Fcm and support vector machine. *International Research Journal of Engineering and Technology, 4*(5).

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, *35*, 18–31. doi:10.1016/j.media.2016.05.004 PMID:27310171

Jagath, C. (2001, October). Bayesian Approach to Segmentation of Statistical Parametric Maps. *IEEE Transactions on Biomedical Engineering*, *48*(10).

Jobin Christ, M. C., Sasikumar, K., & Parwathy, R. M. S. (2009, July). Application of Bayesian Method in Medical Image Segmentation. *International Journal of Computing Science and Communication Technologies, VOL*, *2*(1).

Kamnitsas, K., Ledig, C., Newcombe, V. F. J., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., & Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, *36*, 61–78. doi:10.1016/j.media.2016.10.004 PMID:27865153

Kaur & Rani. (2016). MRI Brain Tumor Segmentation Methods- A Review. *International Journal of Current Engineering and Technology*.

Khadem. (2010). *MRI Brain image segmentation using graph cuts* (Master's thesis). Chalmers University of Technology, Goteborg, Sweden.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*.

Kumar, A. (2017). *A Novel Approach for Brain Tumor Detection Using Support Vector Machine*. K-Means and PCA Algorithm.

Kumar, D. (2018). A modified intuitionistic fuzzy c-means clustering approach to segment human brain MRI image. *Multimedia Tools and Applications*, 1–25.

Laddha, R. R. (2014). A Review on Brain Tumor Detection Using Segmentation And Threshold Operations. *International Journal of Computer Science and Information Technologies*, *5*(1), 607–611.

Leemput, K. V., Maes, F., Vandermeulen, D., & Suetens, P. (1999). Automated model-based tissue classification of MR images of brain. *IEEE Transactions on Medical Imaging*, *18*(10), 897–908. doi:10.1109/42.811270 PMID:10628949

Liu, Z. (2015). Semantic image segmentation via deep parsing network. *Proceedings of the IEEE International Conference on Computer Vision*. 10.1109/ICCV.2015.162

Logeswari & Karnan. (2010). An improved implementation of brain tumor detection using segmentation based on soft computing. *Journal of Cancer Research and Experimental Oncology, 2*(1).

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentationalgorithms and measuring ecological statistics. *Proc. 8th Int. Conf. Computer Vision*, 2, 416–423. 10.1109/ICCV.2001.937655

Mengqiao, W., Jie, Y., Yilei, C., & Hao, W. (2017). The multimodal brain tumor image segmentation based on convolutional neural networks. *2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA),* 336-339. 10.1109/CIAPP.2017.8167234

Moon, N., Bullitt, E., Leemput, K., & Gerig, G. (2002). Model based brain and tumor segmentation. *Int. Conf. on Pattern Recognition*, 528-531. 10.1109/ICPR.2002.1044787

Patil. (2005). *Pachpande, Automatic Brain Tumor Detection Using K-Means*. Academic Press.

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, *35*(5), 1240–1251. doi:10.1109/TMI.2016.2538465 PMID:26960222

Pham, T. X., Siarry, P., & Oulhadj, H. (2018). Integrating fuzzy entropy clustering with an improved PSO for MRI brain image segmentation. *Applied Soft Computing*, *65*, 230–242. doi:10.1016/j.asoc.2018.01.003

Pradhan, S. (2010). *Development of Unsupervised Image Segmentation Schemes for Brain MRI using HMRF model* (Master Thesis). Department of EE, NIT, Rourkela, India.

Priya. (2018). Efficient fuzzy c-means based multilevel image segmentation for brain tumor detection in MR images. *Design Automation for Embedded Systems*, 1–13.

Sathies Kumar, T. (2017). Brain Tumor Detection Using SVM Classifier. *IEEE 3rd International Conference on Sensing, Signal Processing and Security (ICSSS)*.

Suganya, R., & Shanthi, R. (2012). Fuzzy C- Means Algorithm. *RE:view*, *2*(11), 1–3.

Swapnil, R. T. (2016). Detection of brain tumor from MRI images by using segmentation & SVM. *World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*.

Tang, H., Wu, E. X., Ma, Q. Y., Gallagher, D., Perera, G. M., & Zhuang, T. (2000). MRI brain image segmentation by multi-resolution edge detection and region selection. *Computerized Medical Imaging and Graphics*, *24*(6), 349–357. doi:10.1016/S0895-6111(00)00037-9 PMID:11008183

Thuy, Hai, & Thai. (n.d.). *Image Classification using Support Vector Machine and Artificial Neural Network*. Academic Press.

Urban, G. (2014). Multi-modal brain tumor segmentation using deep convolutional neural networks. *MICCAI BraTS (Brain Tumor Segmentation) Challenge. Proceedings*, 31-35.

Vijay, J., & Subhashini, J. (2013). An Efficient Brain Tumor Detection Methodology Using K-Means Clustering Algorithm. *International conference on Communication and Signal Processing*.

Wang. (2017). The multimodal brain tumor image segmentation based on convolutional neural networks. *ICCIA*.

Wells, W. M., Grimson, W. E. L., Kikinis, R., & Jolesz, F. A. (1996). Adaptive segmentation of MRI data. *IEEE Transactions on Medical Imaging*, *15*(4), 429–442. doi:10.1109/42.511747 PMID:18215925

Yi, D. (2016). *3-D convolutional neural networks for glioblastoma segmentation*. arXiv preprint arXiv:1611.04534.

Zheng, S. (2015). Conditional random fields as recurrent neural networks. *Proceedings of the IEEE international conference on computer vision*. 10.1109/ICCV.2015.179

Zikic, D. (2014). Segmentation of brain tumor tissues with convolutional neural networks. *Proceedings MICCAI-BRATS*, 36-39.

Zulpe & Chowhan. (2011). Statical Approach For MRI Brain Tumor Quantification. *International Journal of Computer Applications, 35*(7).

Chapter 4
Breast Cancer Diagnosis in Mammograms Using Wavelet Analysis, Haralick Descriptors, and Autoencoder

Maira Araujo de Santana
Universidade Federal de Pernambuco, Brazil

Jessiane Mônica Silva Pereira
Universidade de Pernambuco, Brazil

Washington Wagner Azevedo da Silva
Universidade Federal de Pernambuco, Brazil

Wellington Pinheiro dos Santos
https://orcid.org/0000-0003-2558-6602
Universidade Federal de Pernambuco, Brazil

ABSTRACT

In this chapter, the authors used autoencoder in data preprocessing step in an attempt to improve image representation, consequently increasing classification performance. The authors applied autoencoder to the task of breast lesion classification in mammographic images. Image Retrieval in Medical Applications (IRMA) database was used. This database has a total of 2,796 ROI (regions of interest) images from mammograms. The images are from patients in one of the three conditions: with a benign lesion, a malignant lesion, or presenting healthy breast. In this study, images were from mostly fatty breasts and authors assessed different intelligent algorithms performance in grouping the images in their respective diagnosis.

DOI: 10.4018/978-1-7998-3092-4.ch004

INTRODUCTION

Cancer is a leading cause of death and, nowadays, is one of the largest public health issue worldwide. For decades, breast cancer has been the most common type of cancer among women around the world. The World Health Organization (WHO) estimates an occurrence of 1.7 million new cases per year (DeSantis et al., 2014). This disease is now placed on the top five causes of cancer death around the world (American Cancer Society, 2019). Survival rates for breast cancer can range from 80%, in high-income countries, to below 40%, in low-income countries (Coleman et al., 2008). The low survival rate in some countries is due to the lack of early detection programs. These programs have a major impact on the success of cancer treatment, since treatment becomes more difficult in later stages.

The gold standard method for breast cancer diagnosis is the digital mammography (Maitra, Nag & Bandyopadhyay, 2011). However, visual analysis of mammography can be a difficult task, even for specialists. Imaging diagnosis is a complex task due to the great variability of clinical cases (Ferreira, Oliveira & Martinez, 2011). Most of the cases observed in clinical practice do not match to classical images and theoretical descriptions (Juhl, Crummy, & Kuhlman, 2000). That is why Computer Aided Diagnosis (CAD) plays an important role in helping radiologists to improve diagnosis accuracy.

Many studies worldwide, are applying traditional image processing and analysis techniques to medical field. Therefore, the combination of professionals specialized knowledge and pattern recognition computational tools may improve diagnosis accuracy (Araujo et al., 2012; Azevedo et al., 2015; Bandyopadhyay, 2010; Commowick et al., 2018; Cordeiro, Bezerra & Santos, 2017; Cordeiro et al., 2012; Cordeiro, Santos & Silva-Filho, 2013; Cordeiro, Santos & Silva-Filho, 2016a; Cordeiro, Santos & Silva-Filho, 2016b; Cruz, Cruz e Santos, 2018; Fernandes & Santos, 2014; Lima, Silva-Filho & Santos, 2014; Mascaro et al., 2009; Santana et al., 2017; Santos, Assis, Souza & Santos Filho, 2009; Santos et al., 2008a; Santos et al., 2008b; Santos et al., 2009a; Santos et al., 2009b; Santos et al., 2010; Santos, Souza & Santos Filho, 2017). Intelligent systems may be used to assist these professionals in decision-making, thus improving the efficiency in identifying anatomical abnormalities (Araujo et al., 2012; Azevedo et al., 2015; Commowick et al., 2018; Cordeiro, Bezerra & Santos, 2017; Cordeiro et al., 2012; Cordeiro, Santos & Silva-Filho, 2013; Cordeiro, Santos & Silva-Filho, 2016a; Cordeiro, Santos & Silva-Filho, 2016b; Cruz, Cruz e Santos, 2018; Fernandes & Santos, 2014; Ferreira, Oliveira & Martinez, 2011; Lima, Silva-Filho & Santos, 2014; Mascaro et al., 2009; Santana et al., 2017; Santos, Assis, Souza & Santos Filho, 2009; Santos et al., 2008a; Santos et al., 2008b; Santos et al., 2009a; Santos et al., 2009b; Santos et al., 2010; Santos, Souza & Santos Filho, 2017; Wang, Yuan & Sheng, 2010; Ye, Zheng & Hao, 2010).

This chapter proposes the use of autoencoders to optimize images representation. As a case of study, authors applied the method to the task of detecting and classifying lesions in regions of interest of mammograms. They compared the results to previous approaches, also using Haralick descriptors, Wavelet transform and intelligent classifiers.

BACKGROUND

In this session, authors provide some related works and a broad definition of some topics they used along the experiments.

Related Works

In the study of Abdel-Zaher and Eldeib (2016) they developed a CAD approach for breast cancer detection. They used deep belief network unsupervised path followed by back-propagation supervised path. They proposed a neural back-propagation network with the Liebenberg Marquardt learning function. The weights are initialized from the deep belief network path (DBN-NN). They used the Wisconsin Breast Cancer Dataset (WBCD) to assess technique performance. The complex classifier achieved an accuracy of 99.68%, indicating promising results, when compared to previously published studies. The proposed system provides an effective classification model for breast cancer. In addition, we examined the architecture in several pieces of training-testing.

Bayramoglu, Kannala and Heikkila (2016) aimed to identify breast cancer using histopathological images, independent of their extensions using convolutional neural networks (CNNs). They proposed two different architectures: a single task CNN was used to predict malignancy and a multitasking CNN was used to simultaneously predict malignancy and the level of image enlargement. They used BreaKHis database to evaluate and compare the results to previous ones. The results of the experiments showed that the proposed approach improved the performance of the specific magnification model, regardless of magnification. Even though having an limited set of training data, the obtained results with the proposed model are comparable to previous results obtained by the state-of-the-art and results obtained by handmade resources. However, unlike previous methods, the proposed approach has the potential to directly benefit from additional training data. Such additional data can be captured at magnification levels equal to or different from previous data.

In the paper of Khuriwal and Mishra, they propose applying the convolutional neural network deep learning algorithm to the diagnosis of breast cancer. They used the mammography MIAS database. Their study shows how we can use deep learning technology to diagnose breast cancer using the MIAS Dataset.

In this work, when applying the Deep Learning technology in the database, an accuracy of 98% was achieved. MIAS database provides 200 images and 12 features in the data set. In this study, they used 12 features, which were extracted after preprocessing. However, before the training model, some preprocessing algorithms, such as Watershed Segmentation, Color based segmentation and Adaptive Mean Filters to staggered datasets were applied. After that, they applied the proposed model to perform classification. In this study, the Deep Learning algorithm is also compared to other machine learning algorithms. They found that the presented methodology achieves better results than other widely used intelligent algorithms.

In Jannesari et al. (2018), they applied pre-trained and adjusted Deep Learning networks. First, the authors tried to discriminate between different types of cancer. They used 6,402 tissue microarray samples (TMAs). Models, including ResNet V1 50, correctly predicted 99.8% of the four types of cancer, including breast, bladder, lung and lymphoma. In a second moment, they tried to assess method performance for the classification of breast cancer subtypes. To do so, they used 7,909 images of 82 patients, from the BreakHis database. ResNet V1 152 classified benign and malignant breast cancers with an accuracy of 98.7%. In addition, ResNet V150 and ResNet V1 152 categorized in benign (adenoses, fibroadenoma, lodia and tubular adenoma) or malignant (ductal carcinomas, lobular carcinomas, mucinous carcinomas and papillary carcinomas) subtypes with 94.8% and 96.4% accuracy, respectively. Confusion matrices revealed high sensitivity values of 1, 0.995 and 0.993 for cancers, as well as malignant and benign subgroups, respectively. The scores of the areas under the curve (AUC) were 0.996 for cancers, 0.973 for malign subtype and 0.996 for benign subtype. One of the most significant and impressive results to emerge from the data analysis was the insignificant false positive (FP) and false negative (FN). The optimal results indicate that FP is between 0 and 4 while FN is between 0 and 8 on which test data including 800, 900, 809, 1000 for four given classes.

In the studies pointed out by Xiao et al. (2018), they show a new method, integrating an unsupervised features extraction algorithm based on deep learning. They combined stacked autoencoders to a support vector machine, thus creating the SAE-SVM model. The approach was applied for breast cancer diagnosis. Stacked autoencoders with a fast pre-training layer and an improved momentum refresh training algorithm are applied to acquire essential information and extract relevant features from the original data. Next, they used a support vector machine to classify samples with new features into malignant and benign lesions. They tested the proposed method using the Wisconsin Diagnostic Breast Cancer database. Performance was assessed using various measures and compared to previously published results. The results shows that the proposed SAE-SVM method improves accuracy to 98.25% and outperforms other methods. The unsupervised features extraction based on deep learning significantly improves the classification performance and provides a promising approach to breast cancer diagnosis.

Features Extraction

As mentioned before, we used Haralick descriptors and Wavelet transform for features extraction. The first one extracts texture-based features from statistical calculations between neighboring pixels of the image (Haralick, Shanmugam & Dinstein, 1973). These features are obtained through a co-occurrence matrix. Co-occurrence matrix has the color occurrence values at a given image and represents the spatial distribution and dependence of gray levels within a local area. Haralick features are widely used as image descriptors, and had also been applied for breast lesion detection (Azevedo et al., 2015; Bhateja et al., 2018; Jenifer, Parasuraman & Kadirvel, 2014; Kishore et al., 2014; Santana et al., 2018; Yasiran, Salleh & Mahmud, 2016). From Haralick descriptor it is possible to differentiate textures that do not follow a certain pattern of repetition in the image (Haralick, Shanmugam & Dinstein, 1973).

Wavelets, however, are very effective tools for representing multi-resolution images. The wavelet transform relative to image processing can be implemented in a two-dimensional way. Mallat proposed a Discrete Wavelet Transform of a signal through the decomposition of an original image into a series of images generated by discrete high-pass and low-pass filters (Mallat, 1999). Such as Haralick descriptors, Wavelet transform is being successfully exploited to mammography representation, in order to detect breast lesions (Eltoukhy, Faye & Samir, 2009; Ganesan et al., 2014; Joseph & Balakrishnan, 2011; Roberts et al., 2017).

Autoencoder

Autoencoder is a neural network that is trained so that the input number is equal to the output number. Its main purpose is optimize the representation of the input data. It has an unsupervised training, so there is no need of labeled data. This training is based on optimizing a cost function. The cost function consists on the mean square error metric, which measures the error between the input, x, and its reconstruction in the output, y (Maria et al., 2016; Xu & Zhang, 2015).

The autoencoder architecture consists of an encoder and a decoder. The encoder structures the input data. Soon after, the decoder reverts the structuring to reconstruct the original input. The output layer has the same number of neurons as the input layer. It is done in order to reconstruct its own inputs, without predicting its outputs, by

Table 1. Autoencoder configuration

Parameter	Value
Hidden Size	10
Encoder Transfer Function	logsig
Decoder Transfer Function	logsig
Max Epochs	1000
L2 Weight Regularization	0.001
Loss Function	msesparse
Sparsity Proportion	0.05
Sparsity Regularization	1

using the unsupervised training (Vincent et al., 2008). Table 1 shows the parameters we set for the autoencoder to perform the experiments presented in this chapter.

The Hidden Size parameter matches the number of neurons in the hidden layer. Encoder Transfer Function and Decoder Transfer Function represent the transfer function for the encoder and decoder, respectively. In this application, the function used for both encoder and decoder, is the logsig, which is described by Equation 1.

$$f\left(z\right) = \frac{1}{1 + e^{-z}} \tag{1}$$

Max Epochs parameter denes the number of interaction. The coefficient L2 Weight Regularization is the regularizer. The Loss Function denes cost function and in this case it stands for to the mean square error. Sparsity Proportion controls the dispersion of the hidden layer output. Finally, the Sparsity Regularization parameter controls the dispersion smoothing impact on the cost function (MathWorks, 2019).

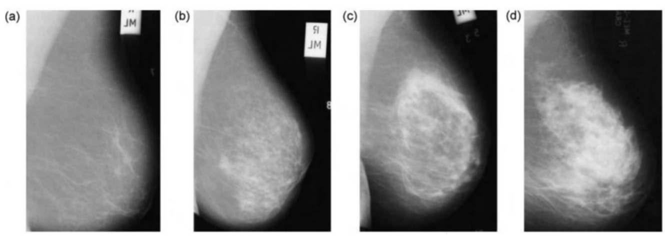
PROPOSED METHOD

In this study, we used Image Retrieval in Medical Applications (IRMA) database (Deserno et al., 2012a; Deserno et al., 2012b; Oliveira, 2010). This database has 2,796 ROI (Regions of Interest) images from mammograms, which were classified by radiologists and were resized to 128x128 pixels. IRMA was developed in Aachen University of Technology (RWTH Aachen), in Germany, and results from a combination of four open access repositories:

- 150 images from the Mini-MIAS database (Suckling et al., 1994);
- 2,576 images from DDSM (Digital Database for Screening Mammography) (Heath, Bowyer & Kopans, 2000);
- 1 image from the LLN database (Lawrence Livermore National Laboratory);
- 69 images from the Department of Radiology at Aachen University of Technology (RWTH), Germany.

The images from IRMA database have four types of tissue density, classified according to the BI-RADS classification (D'Orsi et al., 2013) into: adipose tissue (Type I), fibrous tissue (Type II), heterogeneously dense tissue (Type III), and extremely dense tissue (Type IV). Figure 1 shows examples of each of these classes. Moreover, IRMA has images of breasts with malignant lesion, benign lesion and of healthy breasts, in which there is no lesion.

Figure 1. Mammograms of different breast tissues: (a) adipose tissue, (b) fibrous tissue, (c) heterogeneously dense tissue and (d) extremely dense tissue.
Source: The authors

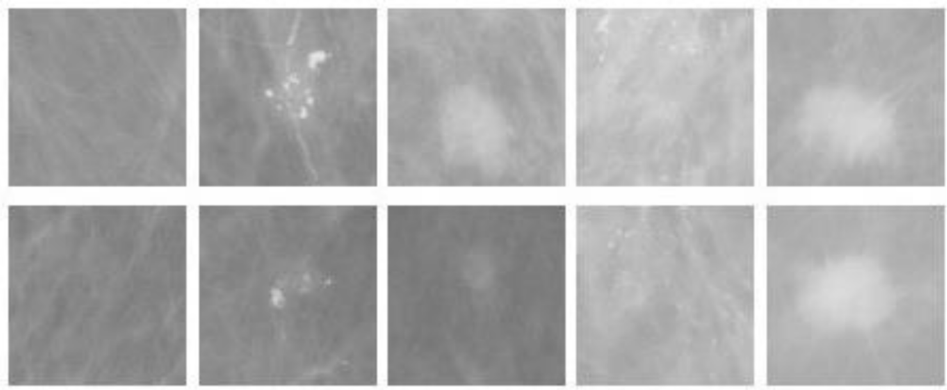


For this study, authors considered only the images of fatty breasts (Type I). They chose to use this class because most women who undergo mammography have this tissue composition in their breasts, since the amount of adipose tissue in the breasts tends to increase with age. Samples of these images can be seen in Figure 2.

Authors access to IRMA database was allowed upon an agreement between Federal University of Pernambuco, Brazil (UFPE) and the Department of Medical Informatics of Aachen University of Technology, Germany. This agreement vetoes the commercial use of partial or whole database.

In this chapter, the authors propose to use autoencoders to preprocess features from mammographic images, in order to optimize database representation. The features used in this study were extracted using both Haralick texture extractor and Wavelet transform. Our set of image features was submitted to the encoder and decoder processes, which worked as a kind of filter.

Figure 2. Example of ROI images from IRMA database classified as Type I.
Source: The authors



After features extraction and preprocessing, authors finally assessed classification performance of some widely used algorithms. The classification step was conducted using: Bayes Net, Naive Bayes, Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), J48, Random Tree, Random Forest, Extreme Learning Machine (ELM) and Morphological Extreme Learning Machine (mELM).

K-fold cross-validation method was used to perform all experiments. The authors also repeated each experiment 30 times, in order to acquire statistical information from them.

SOLUTIONS AND RECOMMENDATIONS

In this section, authors present the results obtained for the classification of breast images in one of the three possible diagnosis: benign lesion, malignant lesion or healthy breast.

Since the main goal in this study was to compare the quality of database representation, authors conducted all experiments in two different databases: the first one without using autoencoders and the second one using autoencoder to preprocess the features. Tables 2 and 3 show the mean and standard deviation (STD) for both accuracy and kappa statistic achieved for each classifier.

In Table 2 are the results for the database created without using autoencoder. From this table, you may see that mELM with erosion kernel outperformed the other methods in terms of both accuracy and kappa. mELM with erosion achieved an average accuracy of 93.28% and 0.92 of kappa. It was closely followed by SVM with linear kernel, mELM with dilatation kernel and ELM. MLP, Random Forest, J48 and SVM with polynomial kernel achieved less satisfying performance, all close to 80% of accuracy. Bayesian classifiers performed worse than the other methods, with accuracy around 60% and kappa around 0.45. Regarding to the standard deviation, mELMs and ELM achieved the best results, which were equal or really close to 0 (zero), meaning a very low data dispersion. Greater standard deviation values were associated to Naive Bayes and Random Tree classifiers, reaching a maximum value of 5.82.

Table 2. Classification results without using autoencoder

Classifier	Parameter	Accuracy (%)	Kappa Statistic
Bayes Net	-	64.08 ± 5.12	0.46 ± 0.08
Naive Bayes	-	61.66 ± 5.82	0.42 ± 0.09
MLP	-	84.20 ± 5.04	0.76 ± 0.08
SVM	Linear kernel	92.29 ± 3.05	0.88 ± 0.05
	Poly kernel (P=2)	78.27 ± 5.02	0.67 ± 0.08
J48	-	78.45 ± 4.89	0.68 ± 0.07
Random Tree	-	71.11 ± 5.61	0.57 ± 0.08
Random Forest	100 trees	81.46 ± 4.47	0.72 ± 0.07
ELM	Sigmoid kernel	91.86 ± 0.77	0.91 ± 0.01
mELM	Dilatation kernel	92.13 ± **0.00**	0.91 ± **0.00**
	Erosion kernel	**93.28** ± **0.00**	**0.92** ± **0.00**

Table 3 presents the results for the database in which we used autoencoder. These results shows that the use of autoencoder in the preprocessing step triggered a decrease on classifiers performance. While authors achieved an accuracy of 93.28% in the previous dataset, they found a maximum of 78.08% of accuracy for the dataset with autoencoder. The maximum accuracy, in this case (78.08%), was achieved by SVM classifier with polynomial kernel. However, this algorithm did not reached the best result for kappa statistic in this scenario. Again, the best kappa, of 0.73, was achieved by the mELMs with both kernels. Regarding to accuracy, SVM was followed by mELMs, Random Forest, MLP and ELM. One more time, Bayes Net and Naive Bayes were associated to the worst performances overall. Bayesian classifiers also presented the low values for kappa statistic, remaining close to 0.40. As to data dispersion, we observe a small decrease in the maximum value for standard deviation after using autoencoder. However, the methods that showed to minimize dispersion in the previous dataset (mELMs and ELM), presented an increase of up to 5.29 in standard deviation when using autoencoder.

Table 3. Classification results using autoencoder

Classifier	Parameter	Accuracy (%)	Kappa Statistic
Bayes Net	-	61.29 ± 5.03	0.42 ± 0.07
Naive Bayes	-	59.89 ± 5.70	0.40 ± 0.08
MLP	-	75.36 ± 5.07	0.63 ± 0.07
SVM	Linear kernel	76.44 ± 5.40	0.65 ± 0.08
	Poly kernel (P=2)	**78.08** ± 4.87	0.67 ± 0.07
J48	-	70.77 ± 5.34	0.56 ± 0.08
Random Tree	-	67.31 ± 5.41	0.51 ± 0.08
Random Forest	100 trees	75.86 ± 4.61	0.64 ± 0.06
ELM	Sigmoid kernel	75.01 ± 4.84	0.72 ± **0.05**
mELM	Dilatation kernel	76.24 ± 4.88	**0.73** ± **0.05**
	Erosion kernel	76.13 ± 5.29	**0.73** ± 0.06

CONCLUSION

This chapter proposed the application of autoencoder in mammographic images preprocessing step. The authors aimed to perform a kind of image filtering, in an attempt to build a better representation of the images and improve classification performance. Tools such as the autoencoder are being exploited to be used in optimization of pattern recognition systems. However, their findings did not meet expectations. Instead of improving classification, the use of autoencoders worsened the performance of the algorithms. They believe that the preprocessing method reduced the image details responsible for characterizing each breast lesions, worsening image representation.

It is important mentioning that the results shown in this chapter does not invalidates the use of autoencoders for identification of breast lesion in mammograms. The authors have no intention of generalizing the results presented here by stating that autoencoder should not be applied in situations like this. In fact, they actually aim to further analyze these results and possibly take a closer look on the parameters of this method. For future studies, the authors believe that they may improve results by modulating autoencoder parameters and characteristics. When it comes to diagnostic solutions, it is worth to keep on investing in optimization methods such as autoencoders, so that, one day, we may achieve the best possible performance.

ACKNOWLEDGMENT

We thank "Fundação de Amparo à Pesquisa do Estado de Pernambuco", FACEPE, Brazil, and "Conselho Nacional de Desenvolvimento Científico e Tecnológico", CNPq, Brazil, for the partial financial support for this research.

REFERENCES

Abdel-Zaher, A. M., & Eldeib, A. M. (2016). Breast cancer classification using deep belief networks. *Expert Systems with Applications*, *46*(1), 139–144. doi:10.1016/j.eswa.2015.10.015

American Cancer Society. (2019). *Cancer Facts & Figures 2019*. American Cancer Society.

Araujo, M., Queiroz, K., Pininga, M., Lima, R., & Santos, W. (2012). Uso de regiões elipsoidais como ferramenta de segmentação em termogramas de mama. In *XXIII Congresso Brasileiro de Engenharia Biomédica (CBEB 2012)*. Pernambuco: SBEB.

Azevedo, W. W., Lima, S. M., Fernandes, I. M., Rocha, A. D., Cordeiro, F. R., Silva-Filho, A. G., & Santos, W. P. (2015). Fuzzy Morphological Extreme Learning Machines to Detect and Classify Masses in Mammograms. In *2015 IEEE International Conference on Fuzzy Systems*. IEEE. 10.1109/FUZZ-IEEE.2015.7337975

Bandyopadhyay, S. K. (2010). Survey on Segmentation Methods for Locating Masses in a Mammogram Image. *International Journal of Computers and Applications*, *9*(11), 25–28. doi:10.5120/1429-1926

Bayramoglu, N., Kannala, J., & Heikkila, J. (2016). Deep Learning for Magnification Independent Breast Cancer Histopathology Image Classification. In *2016 23rd International Conference on Pattern Recognition (ICPR)*. Cancun: IEEE. 10.1109/ICPR.2016.7900002

Bhateja, V., Gautam, A., Tiwari, A., Bao, L. N., Satapathy, S. C., Nhu, N. G., & Le, D.-N. (2018). Haralick Features-Based Classification of Mammograms Using SVM. In *Information Systems Design and Intelligent Applications* (Vol. 672). Springer. doi:10.1007/978-981-10-7512-4_77

Coleman, M. P., Quaresma, M., Berrino, F., Lutz, J. M., De Angelis, R., Capocaccia, R., Baili, P., Rachet, B., Gatta, G., Hakulinen, T., Micheli, A., Sant, M., Weir, H. K., Elwood, J. M., Tsukuma, H., Koifman, S., & Silva, E. (2008). Cancer survival in five continents: A worldwide population-based study (CONCORD). *The Lancet. Oncology*, *9*(8), 730–756. doi:10.1016/S1470-2045(08)70179-7 PMID:18639491

Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., Simon, M., Pop, S. C., Girard, P., Ameli, R., Ferré, J.-C., Kerbrat, A., Tourdias, T., Cervenansky, F., Glatard, T., Beaumont, J., Doyle, S., Forbes, F., Knight, J., Khademi, A., ... Barillot, C. (2018). Objective Evaluation of Multiple Sclerosis Lesion Segmentation Using a Data Management and Processing Infrastructure. *Scientific Reports*, *8*(1), 13650. doi:10.103841598-018-31911-7 PMID:30209345

Cordeiro, F. R., Bezerra, K. F. P., & Santos, W. P. (2017). Random walker with fuzzy initialization applied to segment masses in mammography images. In *30th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE. 10.1109/CBMS.2017.40

Cordeiro, F. R., Lima, S. M., Silva-Filho, A. G., & Santos, W. P. (2012). Segmentation of mammography by applying extreme learning machine in tumor detection. In *International Conference of Intelligent Data Engineering and Automated Learning*. Berlin: Springer.

Cordeiro, F. R., Santos, W. P., & Silva-Filho, A. G. (2013). Segmentation of mammography by applying growcut for mass detection. *Studies in Health Technology and Informatics*, *192*, 87–91. PMID:23920521

Cordeiro, F. R., Santos, W. P., & Silva-Filho, A. G. (2016a). A semi-supervised fuzzy growcut algorithm to segment and classify regions of interest of mammographic images. *Expert Systems with Applications*, *65*, 116–126. doi:10.1016/j.eswa.2016.08.016

Cordeiro, F. R., Santos, W. P., & Silva-Filho, A. G. (2016b). An adaptive semi-supervised fuzzy growcut algorithm to segment masses of regions of interest of mammographic images. *Applied Soft Computing*, *46*, 613–628. doi:10.1016/j.asoc.2015.11.040

Cruz, T. N., Cruz, T. M., & Santos, W. P. (2018). Detection and classification of lesions in mammographies using neural networks and morphological wavelets. *IEEE Latin America Transactions*, *16*(3), 926–932. doi:10.1109/TLA.2018.8358675

D'Orsi, C. J., Sickles, E. A., Mendelson, E. B., & Morris, E. A. (2013). *Breast Imaging Reporting and Data System: ACR BI-RADS breast imaging atlas* (5th ed.). American College of Radiology.

DeSantis, C. E., Lin, C. C., Mariotto, A. B., Siegel, R. L., Stein, K. D., Kramer, J. L., Alteri, R., Robbins, A. S., & Jemal, A. (2014). Cancer treatment and survivorship statistics, 2014. *CA: a Cancer Journal for Clinicians*, *64*(4), 252–271. doi:10.3322/caac.21235 PMID:24890451

Deserno, T., Soiron, M., Oliveira, J., & Araújo, A. (2012a). Towards Computer-Aided Diagnostics of Screening Mammography Using Content-Based Image Retrieval. In *2011 24th SIBGRAPI Conference on Graphics, Patterns and Images*. Alagoas: IEEE.

Deserno, T. M., Soiron, M., Oliveira, J. E. E., & Araújo, A. A. (2012b). Computer-aided diagnostics of screening mammography using content-based image retrieval. In *Medical Imaging: Computer-Aided Diagnosis 2012*. SPIE. doi:10.1117/12.912392

Eltoukhy, M., Faye, I., & Samir, B. (2009). Breast Cancer Diagnosis in Mammograms using Multilevel Wavelet Analysis. *Proceeding of National Postgraduate Conference*.

Fernandes, I., & Santos, W. (2014). Classificação de mamografias utilizando extração de atributos de textura e redes neurais artificiais. In *Congresso Brasileiro de Engenharia Biomédica (CBEB 2014)*. SBEB.

Ferreira, J., Oliveira, H., & Martinez, M. (2011). Aplicação de uma metodologia computacional inteligente no diagnóstico de lesões cancerígenas. *Revista Brasileira de Inovação Tecnológica em Saúde*, *1*(2), 4-9.

Ganesan, K., Acharya, U. R., Chua, C. K., Min, L. C., & Abraham, T. K. (2014). Automated Diagnosis of Mammogram Images of Breast Cancer Using Discrete Wavelet Transform and Spherical Wavelet Transform Features: A Comparative Study. *Technology in Cancer Research & Treatment*, *13*(6), 605–615. doi:10.7785/tcrtexpress.2013.600262 PMID:24000991

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *3*(6), 610–621. doi:10.1109/TSMC.1973.4309314

Heath, M., Bowyer, K. W., & Kopans, D. (2000). The digital database for screening mammography. *Proceedings of the 5th International Workshop on Digital Mammography*

Jannesari, M., Habibzadeh, M., Aboulkheyr, H., Khosravi, P., Elemento, O., Totonchi, M., & Hajirasouliha, I. (2018). Breast Cancer Histopathological Image Classification: A Deep Learning Approach. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE. 10.1109/BIBM.2018.8621307

Jenifer, S., Parasuraman, S., & Kadirvel, A. (2014). An Efficient Biomedical Imaging Technique for Automatic Detection of Abnormalities in Digital Mammograms. *Journal of Medical Imaging and Health Informatics*, *4*(2), 291–296. doi:10.1166/jmihi.2014.1246

Joseph, S., & Balakrishnan, K. (2011). Local Binary Patterns, Haar Wavelet Features and Haralick Texture Features for Mammogram Image Classification using Artificial Neural Networks. In *International Conference on Advances in Computing and Information Technology*. Springer. 10.1007/978-3-642-22555-0_12

Juhl, J. H., Crummy, A. B., & Kuhlman, J. E. (2000). Paul & Juhl Interpretação Radiológica (7a ed.). Rio de Janeiro: Guanabara-Koogan.

Khuriwal, N., & Mishra, N. (2018). Breast Cancer Detection from Histopathological Images using Deep Learning. In *2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE)*. IEEE. 10.1109/ICRAIE.2018.8710426

Kishore, B., Arjunan, R. V., Saha, R., & Selvan, S. (2014). Using Haralick Features for the Distance Measure Classification of Digital Mammograms. *International Journal of Computers and Applications*, *6*(1), 17–21.

Lima, S. M., Silva-Filho, A. G., & Santos, W. P. (2014). A methodology for classification of lesions in mammographies using Zernike moments, ELM and SVM neural networks in a multi-kernel approach. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE. 10.1109/SMC.2014.6974041

Maitra, I. K., Nag, S., & Bandyopadhyay, S. K. (2011). Identification of Abnormal Masses in Digital Mammography Images. *International Journal of Computer Graphics*, *2*(1).

Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Academic.

Maria, J., Amaro, J., Falcao, G., & Alexandre, L. A. (2016). Stacked Autoencoders Using Low-Power Accelerated Architectures for Object Recognition in Autonomous Systems. *Neural Processing Letters*, *43*(2), 445–458. doi:10.100711063-015-9430-9

Mascaro, A. A., Mello, C. A., Santos, W. P., & Cavalcanti, G. D. (2009). Mammographic images segmentation using texture descriptors. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 10.1109/IEMBS.2009.5333696

MathWorks. (2019). *Deep Learning Toolbox ™ Reference*. Author.

Oliveira, J. E., Machado, A. M., Chavez, G. C., Lopes, A. P., Deserno, T. M., & Araújo, A. A. (2010). MammoSys: A content-based image retrieval system using breast density patterns. *Computer Methods and Programs in Biomedicine*, *99*(3), 289–297. doi:10.1016/j.cmpb.2010.01.005 PMID:20207441

Roberts, T., Newell, M., Auffermann, W., & Vidakovic, B. (2017). Wavelet-based scaling indices for breast cancer diagnostics. *Statistics in Medicine*, *36*(12), 1989–2000. doi:10.1002im.7264 PMID:28226399

Santana, M., Pereira, J., Lima, N., Sousa, F., Lima, R., & Santos, W. (2017). Classificação de lesões em imagens frontais de termografia de mama a partir de sistema inteligente de suporte ao diagnóstico. In *Anais do I Simpósio de Inovação em Engenharia Biomédica*. SABIO.

Santana, M. A., Pereira, J. M. S., Silva, F. L., Lima, N. M., Sousa, F. N., Arruda, G. M. S., Lima, R. C. F., Silva, W. W. A., & Santos, W. P. (2018). Breast cancer diagnosis based on mammary thermography and extreme learning machines. *Research on Biomedical Engineering*, *34*(1), 45–53. doi:10.1590/2446-4740.05217

Santos, W. P., Assis, F., Souza, R., Santos Filho, P. B., & Neto, F. L. (2009b). Dialectical Multispectral Classification of Diffusion-weighted Magnetic Resonance Images as an Alternative to Apparent Diffusion Coefficients Maps to Perform Anatomical Analysis. *Computerized Medical Imaging and Graphics*, *33*(6), 442–460. doi:10.1016/j.compmedimag.2009.04.004 PMID:19446434

Santos, W. P., Assis, F. M., Souza, R. E., Mendes, P. B., Monteiro, H. S. S., & Alves, H. D. (2009a). A Dialectical Method to Classify Alzheimer's Magnetic Resonance Images. In *Evolutionary Computation*. IntechOpen. doi:10.5772/9609

Santos, W. P., Assis, F. M., Souza, R. E., Mendes, P. B., Monteiro, H. S. S., & Alves, H. D. (2010). Fuzzy-based Dialectical Non-supervised Image Classification and Clustering. *International Journal of Hybrid Intelligent Systems*, *7*(2), 115–124. doi:10.3233/HIS-2010-0108

Santos, W. P., Assis, F. M., Souza, R. E., & Santos Filho, P. B. (2008a). *Evaluation of Alzheimer's Disease by Analysis of MR Images using Objective Dialectical Classifiers as an Alternative to ADC Maps. In 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE.

Santos, W. P., Assis, F. M., Souza, R. E., & Santos Filho, P. B. (2009). Dialectical Classification of MR Images for the Evaluation of Alzheimer's Disease. In *Recent Advances in Biomedical Engineering*. IntechOpen. doi:10.5772/7475

Santos, W. P., Souza, R. E., & Santos Filho, P. B. (2017). Evaluation of Alzheimer's Disease by Analysis of MR Images using Multilayer Perceptrons and Kohonen SOM Classifiers as an Alternative to the ADC Maps. In *2017 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE.

Santos, W. P., Souza, R. E., Santos Filho, P. B., Neto, F. B. L., & Assis, F. M. (2008b). A Dialectical Approach for Classification of DW-MR Alzheimer's Images. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on COmputational Intelligence)*. Hong Kong: IEEE. 10.1109/CEC.2008.4631023

Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S., Taylor, P., Betal, D., & Savage, J. (1994). The mammographic image analysis society digital mammogram database. In *2nd International Workshop on Digital Mammography*. Excerpta Medica.

Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P.-A. (2008). Extracting and Composing Robust Features with Denoising Autoencoders. In *25th International Conference on Machine Learning*. New York: ACM. 10.1145/1390156.1390294

Wang, D., Yuan, F., & Sheng, H. (2010). An Algorithm for Medical Imaging Identification based on Edge Detection and Seed Filling. In *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*. Taiyuan: IEEE.

Xiao, Y., Wu, J., Lin, Z., & Zhao, X. (2018). Breast Cancer Diagnosis Using an Unsupervised Feature Extraction Algorithm Based on Deep Learning. In *2018 37th Chinese Control Conference (CCC)*. IEEE. 10.23919/ChiCC.2018.8483140

Xu, Q., & Zhang, L. (2015). The Effect of Different Hidden Unit Number of Sparse Autoencoder. In *The 27th Chinese Control and Decision Conference (2015 CCDC)*. IEEE. 10.1109/CCDC.2015.7162335

Yasiran, S. S., Salleh, S., & Mahmud, R. (2016). Haralick texture and invariant moments features for breast cancer classification. In AIP Conference Proceedings. AIP Publishing. doi:10.1063/1.4954535

Ye, S., Zheng, S., & Hao, W. (2010). Medical image edge detection method based on adaptive facet model. In *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*. Taiyuan: IEEE.

Chapter 5
Feature Selection Using Random Forest Algorithm to Diagnose Tuberculosis From Lung CT Images

Beaulah Jeyavathana Rajendran
Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

Kanimozhi K. V.
Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

ABSTRACT

Tuberculosis is one of the hazardous infectious diseases that can be categorized by the evolution of tubercles in the tissues. This disease mainly affects the lungs and also the other parts of the body. The disease can be easily diagnosed by the radiologists. The main objective of this chapter is to get best solution selected by means of modified particle swarm optimization is regarded as optimal feature descriptor. Five stages are being used to detect tuberculosis disease. They are pre-processing an image, segmenting the lungs and extracting the feature, feature selection and classification. These stages that are used in medical image processing to identify the tuberculosis. In the feature extraction, the GLCM approach is used to extract the features and from the extracted feature sets the optimal features are selected by random forest. Finally, support vector machine classifier method is used for image classification. The experimentation is done, and intermediate results are obtained. The proposed system accuracy results are better than the existing method in classification.

DOI: 10.4018/978-1-7998-3092-4.ch005

INTRODUCTION

Tuberculosis is one of the communicable bacterial diseases that is caused by the bacterium called Mycobacterium tuberculosis and that may affect any tissues of the body but it primarily disturbs the lungs. TB is one of the airborne pathogens that can binge through air or by coughing or sneezing from one person to another. TB disturbs all age groups in all parts of the world. Tuberculosis bacteria are present in sputum trials and it is identified under a microscope.

In 2015, around 11 million people fell ill with TB and 2 million people were died from the disease. Over 95% of the deaths in TB occur in low and middle- revenue countries.

X-ray is not easily predicting the early stage of tuberculosis. Hence, because of this wrong prediction of tuberculosis, an automated detection of tuberculosis is used. To overcome the problems in existing methods, CT lung images are used for diagnosis of tuberculosis.

In image processing Feature extraction is an important step, which is a special form of dimensionality reduction. When the input data is too large to be processed and alleged to be redundant then the data is transformed into a reduced set of feature representations. Feature contains the information that is related to colour, shape, texture and context. Modified Random Forest Algorithm technique is based on optimization searching technique and it is used to find the optimal solutions. It is used for selecting the best features after the feature extraction process. This will continue until a needed solution is obtained. Classifying the images whether it is normal or abnormal by SVM classifier.

RELATED WORKS

Les Folio (2014), presented the automated approach for detecting tuberculosis in conventional poster anterior chest radiographs. For the extracted region, set of texture features and shape features are computed, which enable the X-rays to be classified as normal or abnormal using a binary classifier. The pre- processing techniques is used to remove the noises and the feature extraction are done to extract the useful features in given image and the feature selection technique will optimize the top ranking features that are relevant for the image and the classifiers are employed to classify the images and the performance measures are found for the same (Sun et al., 2015). Laurens Hogeweg, Clara I. The performance is evaluated on a TB screening and a TB suspect database using both an external and a radiological reference standard. The systems to detect different types of TB related abnormalities and their combination is described. Yan Kang (2015), Wenbo Li, using a new adaptive

VOI selection method. The improved GA algorithm to select the optimal feature combination from the feature pool to establish SVM classifier (Omisore, 2014). G. Vijaya, A. Suhasini identified the cancer tumor from lung CT images using edge detection and boundary tracing. To classify the lung cancer, by using the data mining, classification techniques like SMO (Sequential Minimal Optimization), J48 decision tree, Naive Bayes.

Once the classification is performed, we have to compare the experimental results of the above classification techniques, and determine which one gives accurate and correct answers (Girisha et al., 2013). Mumini Olatunji Omisore (2014) proposed the genetic neuro-fuzzy inferential model for the diagnosis of tuberculosis. Finally, SVM is used in the classification stage (Linguraru et al., n.d.). A. Zabidi, L.Y. Khuan IEEE International conference(2011) proposed the Binary Particle Swarm Optimization For Feature selection in Detection of Infants with Hypothyroidism. In this, he investigates the effect of feature selection with Binary PSO on performance of Multilayer perceptron classifier in discriminating between the healthy infants and infants with hypothyroidism from their cry signals. The performance was examined by varying the number of coefficients.

PROPOSED WORK

Material and Methods

In the study, dataset containing lung CT images comprising abnormal lung and normal lung are taken from several patients was utilized. The lung diseases are categorized by the radiologist from the CT Image. Images are collected from male and female patients whose ages are ranging from 15 to 78 years.

Pre-Processing

Pre-processing is done to remove unwanted noise and it gives quality to the images at this stage where filtering is done to remove noise. In our proposed system we have used wiener filter to remove noise. Wiener filter preserves the edges and fine details of lungs. It is low pass-filter. The filter size of 5*5 is selected to avoid over smoothing of the image. 2D Wiener filter is used for lessening of additive gaussian white noise in images.

SEGMENTATION

K-Means Clustering

K-Means is possibly the most well-known clustering algorithm. It's easy to understand and implement in code.

1. To begin, we first hand-picked a number of classes/groups to practice and erratically set their respective center points. To number out the number of classes to use, it's good to take a rapid look at the data and attempt to recognize any distinct groupings. The center points are vectors of the same length as each data point vector.
2. Each data point is classified by calculating the distance between that point and each group center, and then classifying the point to be in the group whose center is closest to it.
3. Based on these classified points, we recompute the group center by taking the mean of all the vectors in the group.
4. Repeat these steps for a set number of iterations or until the group centers don't change much between iterations. We can also opt to randomly initialize the group centers a few times, and then select the run that looks like it provided the best results.

ROI EXTRACTION

ROI's are taken out using the radiologists and thus it is authorized to attain the clinical relevance which progresses the performance of the system. Extract the defected tissues from the lung as ROI's and then find the intensity level of the pixels and using the range of pixel intensity values discriminate the defected tissues and other lung tissues. If there is no defected tissues are present, then the slice is considered to be Normal. Then obtain the class labels for each ROI's from the experts. Finally, ROI's are extracted and also the class label information is obtained.

FEATURE EXTRACTION

GLCM Approach

The feature extraction based on Texture feature is carried out. GLCM approach is used for extracting the features in given image such as entropy, energy, contrast, correlation, variance, sum average, homogeneity cluster shade and etc.., are considered for feature selection. Extract these twenty-two features for each ROI in four orientations 0^o, 45^o, 90^o, 135^o using GLCM also called the Grey Tone Spatial Dependency Matrix. GLCM contains the information about the positions of pixel having similar grey level values.

CLASSIFICATION SUBSYSTEM

Naïve Bayes Classifier

It is not a single algorithm but a family of algorithms where all of them stake a communal principle, i.e. every pair of features being classified is self-determining of each other. Naïve Bayes classifiers are highly ascendable, necessitating a number of parameters linear in the number of variables (features/predictors) in a learning problem.

FEATURE SELECTION

The term Feature selection deals with selecting a subset of features, among the entire features, that shows the best performance in classification accuracy. Optimization searching process is done by Modified Random Forest Algorithm.

Modified Random Forest Algorithm

Random forests are an unification of tree predictors such that each tree depends on the values of a random vector sampled self-reliantly. In year 2001, Leo Breiman, a statistician recognizes the difficulties in prevailing machine learning techniques. In former tree approach of machine learning data set is not consistently disseminated lead to imbalance of data. Imbalanced data set performance is underprivileged with the classification, this lead to miss classification and error in the training phase. He recommended data set were collected and then divided into two or more subset of data, where one or more data set used as learner and residual is used for test purpose. Many researchers got fascinated towards Random Forest approach of handling data set and started employed on different attributes of Random Forest like features, concepts, analysis and modification of the proposed model of Random Forest algorithm. Research works going on in the field of Random Forest can be approximately classified into three categories:

- Research in Random Forest Improving accuracy
- Improving performance
- Investigation of new domain for application of Random Forest.

Number of trees engendered in the Random Forest is the task for the Researchers because it devours extra space in memory and also increases run time of the algorithm. Rudimentary problem with this method is that it takes only time complexity under contemplation leaving the space complexity. Because of this pruning approach is required. Redeemable more time when assessment to static approach but tactlessly it is hard to contrivance because of this researchers are also not showing interest for this approach. Research work completed under static pruning approach fall in to three majority categories:

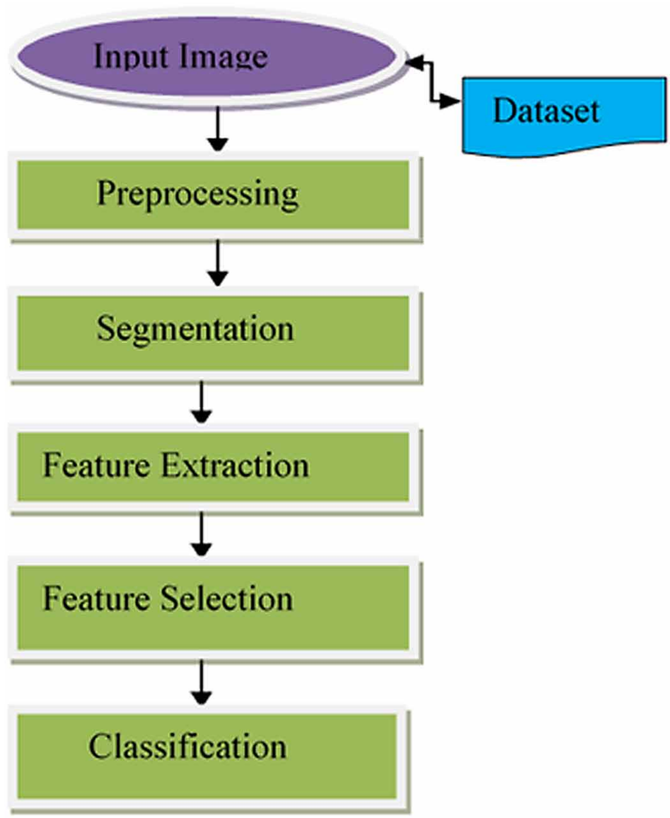
- Weighted voting method
- Ranking Based method
- Search Based method

In one forerunner work of static pruning, genetic algorithm is used to choice most optimal candidate from pool of Decision Tree. Other work uses elimination of similar Decision Tree if their output class and accuracy are same then keep single copy of Tree eliminating others. Dynamic pruning necessitates help of statistics and probability along with nature inspired algorithm to get recovering results. In one approach authors have tried to model dynamic pruning approach with the help of eight degree mathematical equation.

RESULTS AND DISCUSSION

The CT images used for testing and training purpose for classification were collected from AARTHI SCANS & LABS at TIRUNELVELI. We have several CT images, but we use 197 images for my work out of which 94 images have tuberculosis and the remaining images do not have tuberculosis. The segmentation of the image takes place, in which K-Means Clustering is done. The set of Tuberculosis (TB) CT images and non-Tuberculosis CT images are tested to give an accurate result. Thus, the technique deals with the accurate detection of tuberculosis.

Figure 1.



Processing Time Analysis

In our study, to implement our proposed algorithm, we used MATLAB software (R2016a) on a laptop, Intel Core i3 (2.0 GHZ) and 4GB memory. The resolution of images in our database was 512 x 512.

To evaluate our proposed algorithm efficiently, we analyzed each step of our algorithm based on processing time. Table 1 presents average of the processing time of each module of proposed algorithm.

Table 1.

Proposed algorithm.	Processing time of each Module
Module	Processing time (s)
Pre-processing	1.1456 seconds
K-Means segmentation	20.0196 seconds
Feature extraction	0.24534 seconds

Table 2.

Parameter	**Using PSO**	**Using MPSO**	**Using Modified Random Forest**
TP	**43**	**46**	**49**
TN	**41**	**45**	**48**
FP	**09**	**05**	**02**
FN	**07**	**04**	**01**
Accuracy	**84**	**91**	**98**
Sensitivity	**86**	**92**	**96**
Specificity	**82**	**90**	**94**

Table 3.

Classifier	**Performance Parameters**	**Percentage**
Bayes classifier	**Accuracy**	**92.30%**
	Sensitivity	**96%**
	Specificity	**49%**

CONCLUSION

In this work the preprocessing of the images is done, then the segmentation are done by K-Means Clustering algorithm. It is one of the distinctive clustering algorithms. Furthermore several algorithms are developed based on K-Means. During the implementation of this algorithm, find some points that can be further improvement in the future using some advanced clustering to achieve more accuracy and GLCM-based feature extraction technique was described. The texture features are served as the input to classify the image accurately. Effective use of these multiple features and the selection of suitable classification method is significant for improving accuracy.

REFERENCES

Bhuvaneswari, Aruna, & Loganathan. (2014). Classification of Lung Diseases by Image Processing Techniques Using Computed Tomography Images. *International Journal of Advanced Computer Research, 4*(1).

Candemir, Jaeger, Palaniappan, & Musco. (2014). Lung Segmentation in Chest Radiographs Using Anatomical Atlases With Non-rigid Registration. *IEEE Transactions on Medical Imaging, 33*(2).

Dai, S., Lu, K., & Dong, J. (2015). Lung segmentation with improved graph cuts on chest CT images. *3rd IAPR Asian Conference on Pattern Recognition*. 10.1109/ACPR.2015.7486502

Girisha, Chandrashekhar, & Kurian. (2013). Texture Feature Extraction of Video Frames Using GLCM. *International Journal of Engineering Trends and Technology, 4*(6).

Hogeweg, Sánchez, Maduskar, Philipsen, & Story. (2015). Automatic Detection of Tuberculosis in Chest Radiographs Using a Combination of Textural, Focal and Shape Abnormality Analysis. *IEEE Transactions on Medical Imaging, 34*(12).

Jaeger, Karargyris, Candemir, Folio, & Siegelman. (2014). Automatic Tuberculosis Screening Using Chest Radiographs. *IEEE Transactions on Medical Imaging, 33*(2).

Linguraru, Richbourg, Liu, & Watt. (n.d.). *Tumor Burden Analysis on Computed Tomography by Automated Liver and Tumor Segmentation*. IEEE.

Omisore. (2014). Proposed the genetic neuro-fuzzy inferential model for the diagnosis of tuberculosis. *IEEE Transactions*.

Rendon-Gonzalez & Ponomaryov. (2016). Automatic Lung Nodule Segmentation and Classification in CT Images Based on SVM. *International conferences IEEE.*

Sun, S., Li, W., & Kang, Y. (2015). Lung Nodule Detection Based on GA and SVM. *8th International Conference on Bio Medical Engineering and Informatics (BMEI 2015).*

Vijaya, Suhasini, & Priya. (n.d.). Automatic detection of lung cancer in CT images. *IJRET: International Journal of Research in Engineering and Technology.*

Chapter 6
An Ensemble Feature Subset Selection for Women Breast Cancer Classification

A. Kalaivani
Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

ABSTRACT

Breast cancer leads to fatal diseases both in India and America and takes the lives of thousands of women in the world every year. The patients can be easily treated if the signs and symptoms are identified at the early stages. But the symptoms identified at the final stage spreads in the human body, and most of the time, the cancer is identified at the final stage. Breast cancer detected at the early stage is treated easily rather than at the advanced stage. Computer-aided diagnosis came into existence from 2000 with high expectations to improve true positive diagnosis and reduce false positive marks. Artificial intelligence revolved in computing drives the attention of deep learning for an automated breast cancer detection and diagnosis in digital mammography. The chapter focuses on automatic feature selection algorithm for diagnosis of women breast cancer from digital mammographic images achieved through multi-layer perceptron techniques.

1. INTRODUCTION

Breast cancer (BC) is the tumor that originates in the cells of women breast and grows into breast cancer. Breast Cancer tumor has a nature to spread to different parts of the body (Y.S. Hotko, 2013). Breast Cancer is a universal disease which harms the

DOI: 10.4018/978-1-7998-3092-4.ch006

lives of women in the age group of 25– 50. There is a potential rise in the number of BC cases in India and America. During the past five years, the survival rates of BC patients are about 90% in the USA and whereas in India the figure reports approximately 60%. Breast Cancer projection for India suspect to reach higher rates may be two millions (S. Malvia, 2017).

Medical World identified hormonal, life style and environmental factors are the root cause for development of Breast Cancer. Around 5%–6% of breast cancer patients are due to gene mutations that went through the ages of the family. The most common factors due to which breast cancer caused are Obesity, increasing age, postmenopausal hormonal imbalances. The only mechanism to diagnose breast cancer The early detection of breast cancer can reduce the costs of the treatment as there is no prevention mechanism for breast cancer. But the early detection is difficult since most of the times it is unusual to show cancer symptoms. It is indispensable for the patients to test using digital mammograms or self-breast tests to detect any early irregularities in the breast and also to get the tumor advanced (Shallu, Rajesh Mehra, 2018).

Medical Experts only deals with the diagnosis of disease purely based on the various tests performed upon the patient. The important factors in diagnosis is based on the data evaluation of patients data and experts knowledge. The medical diagnosis focused on this paper leads to the early diagnosis of women breast cancer from digital mammographic images predicts the malignant cases in a timely manner and which increased life span of patients from 56 to 86%.

Breast Cancer shows four signs of liaisons which are micro-calcification, mass, architectural distortion, and breast asymmetries(Hazlina H,et.al., 2004). The medical modalities supported for breast cancer diagnosis are positron emission tomography (PET), magnetic resonance imaging (MRI), CT scan, X-ray, digital mammography, ultrasound, tomography of photo-acoustic, optical, electrical impedance, opto-acoustic imaging(Sulochana Wadhwani et.al., 2013). The results obtained from these methods are used to recognize the patterns, which help medical experts to classify breast cancer into malignant or benign cases.

Digital Mammography System used for early stage breast cancer replaces X-ray film by electronics produces mammographic pictures of the breast enables better picture with a lower radiation dose. The breast images are transferred to a computer for review by the radiologist and can also be used for long term storage of patient record.

As per World Health Organization report, women breast cancer found to be the common women diagnosed cancer disease which also leads to death mortality among women worldwide. On an average, a woman is diagnosed with breast cancer every two minutes and one woman dies of it every 13 minutes worldwide. Survey statistics in 2019, says an estimated 2,68,600 new cases of invasive breast cancer

Figure 1. Digital mammography system
(from (Bronzino, 2000)).

are diagnosed in U.S women and 62,960 new cases of noninvasive breast cancer and mortality rate reached 41,760.

iThe recent introduction of slide scanners that digitize the biopsy into multi-resolution images, along with advances in deep learning methods are used for computer-aided diagnosis of breast cancer. The intermediate steps from tissue localization, image enhancement, segmentation, annotation made us to make diagnosis accurate, reliable, efficient and cost-effective.

The introduction of deep learning convolutional neural networks (CNNs) in medical image analysis has brought forth a potential revolution in computer-based interpretation of Digital Mammography. Deep learning convolutional neural networks involve the processing of an image by multiple sequential stages of maximum convolutions and down sampling operators which combine the spatially correlated information contained in images. During this multiple-stage process, this information is broken down into different representations, and the analysis are more abstract, and the ability of the network to recognize the image was made accurate.

The paper introduces the discussed the related work by the researchers for women breast cancer using computer aided detection and diagnosis system in in section 2. A detailed view of how an Artificial Neural Network system can play a vital role in CAD diagnosis and the proposed system methodology is explained in section 3. Materials and Methods of the proposed technology are explained in Section 5 and experimental results and discussion are done in Section 6 and finally paper is ended in the conclusion section.

2. RELATED WORK

The Computer Aided Diagnosis system detects the suspicious regions with high sensitivity and presents the results to the radiologist with a focus to reduce false positives. The preprocessing algorithm reduces the noise acquired in the image.

A segmentation process will identifies different region of interest to detect high suspicion of some signs of cancer. Features are extracted from the segmented region and the classification of positives or negatives prediction of Breast Cancer is obtained through an artificial neural network.

A brief breakup of different types of artificial neural network architecture, which is done by a number of researchers are shown in table 2. Computer Aided diagnosis reduce the workloads on clinicians by detecting artifacts and providing decision support for better performance. Table I provides information about different artificial neural network algorithms studied by various researchers to diagnose women breast cancer.

Researchers have carried out the research work on the artificial neural network techniques and applied various algorithms to diagnose women breast cancer. List of the researchers works are listed and explained in the table 2.

3. ARTIFICIAL NEURAL NETWORKS

An artificial neural network (ANN) is a computational model based on the basis of biological neural networks similar representation of human neural system. ANN consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. Artificial Neural Network architectures are designed to model complex relationships between inputs and outputs or used to find patterns in data.

Table 1. Research work on Medical Images based on Artifineural Neural Network

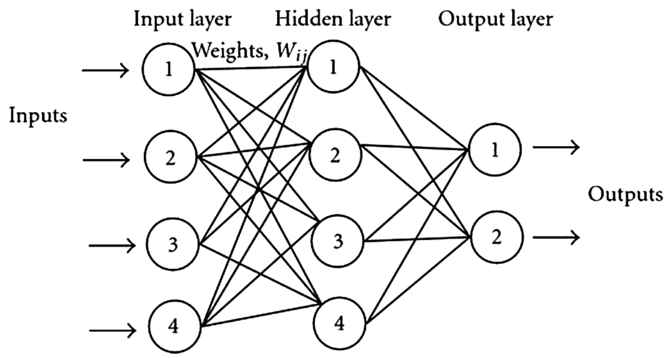
S.Nos.	ANN Types	Research Outcome	Image Modality
1	Cellular Neural Network	Detect Boundary/ area	X-ray
2	GA and CNN	Detect nodular shadows	X-ray
3	Hybrid Neural Digital CAD	Classify 3-15 mm size nodules	X-ray
4	ANN Feed Forward	Increase sensitivity & accuracy	X-ray
5	Artificial CNN & application	Detect False Positive & increase sensitivity	X-ray
6	Convolution Neural Network	Decrease False & Increase True Positive	X-ray
7	Two - level Convolution Neural Network	Reduce False Positive	X-ray
8	NN Ensembles	Reduce False Positive	X-ray
9	J-net	Improve sensitivity & accuracy	CT Image
10	Massive Training ANN	Enhancement of lung nodules	CT Image

Table 2. Existing methods of neural network algorithms for Women Breast Cancer Diagnosis

S.Nos.	Authors	Technique	Algorithm	Results
1	Sulochana Wadhwani, A.K Wadhwani, Monika Saraswat.	Artificial Neural Network	Back propagation Algorithm	Classification of Breast cancer into malignant or benign with the accuracies of 94.11% and 100% (Sulochana Wadhwani et. al., 2013)
2	Pankaj Sapra, Rupinderpal Singh, Shivani Khurana.	Computer Aided Detection System and Probabilistic Neural Network	Competitive Learning Algorithm	Detection of Brain Tumor, obtained 100% accuracy (Pankaj Sapra et.al, 2013)
3	Yongjun WU, Na Wang, Hongsheng ZHANG, Lijuan Qin, Zhen YAN, Yiming WU.	Artificial Neural Network	Back propagation Algorithm Diagnosis of lung cancer.	Provides accuracy of 96.6%. (Yongjun WU et. al., 2010)
4	Ayoub Arafi, Youssef Safi, Rkia Fajr and Abdelaziz Bouroumi.	Image processing and Artificial Neural Network	Multilayer Perceptron Training Algorithm	Classification of mammographic images of breast cancer. Accuracy obtained is 95.49%. (Ayoub Arafi et. al., 2013)
5	Seema Singh, Sunita Saini, Mandeep Singh.	Artificial Neural Network	Adaptive Resonance Theory Detection of cancer using ART	Obtained accuracy 82.64% (Seema Singh et. al., 2012)
6	Ali Raad, Ali Kalakech, Mohammad Ayache.	. Artificial Neural Network	Back propagation Algorithm Breast cancer detection and classification using ANN.	provided an accuracy of 94% (Ali Raad et. al., 2012)
7	Yuehui Chen, Yan Wang, Bo Yang.	Artificial neural network	Hierarchical Radial Basis Function	Breast cancer detection using hierarchical RBF with the accuracy of 97.09%. (Yuehui Chen et. al., 2006)

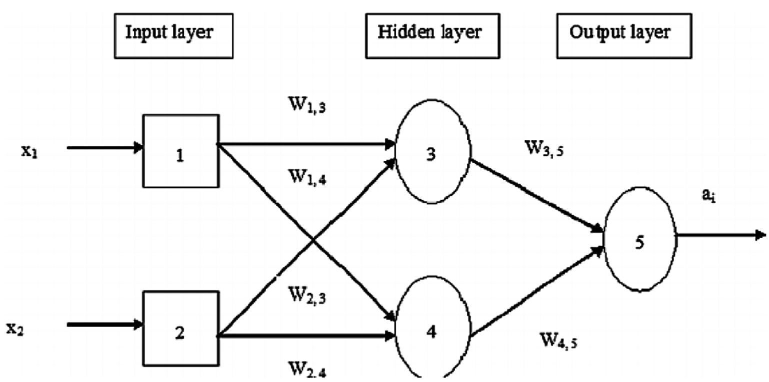
Artificial Neural Network are composed of multiple nodes present in input layer, hidden layer and output layer similar to nerve system of human body. The nodes in the input layer, hidden layer and output layer interact with each other which can take input data and perform operations on the input nodes. The result of these operations is passed to other nodes in the next layer and the output at each node are activated to the next layer through an activation function. Weights are associated in the node for learning the network and final output is obtained in the output layer. The artificial neural network architecture is of two broad categories: feed forward neural network and feed backward neural network.

Figure 2. Artificial Neural Network



A feed forward neural network is an artificial neural network connections between the nodes are forwarded from input layer to hidden layer and forwarded from hidden layer into output layer. The feed forward neural network was the first and simplest type of artificial neural network devised which will not form loops. The information in feed forward neural network moves in single forward direction, moving from the input nodes, to the hidden nodes and finally move to the output nodes. There are no cycles or loops in the network. f'(x)=f(x)(1-f(x))}.The Feed forward network can be formed using a single layer perceptron and multi-layer perceptron.

Figure 3. Feed Forward Neural Network



The most popular form of artificial neural network architecture focused by the researchers for effective medical diagnosis is the multilayer perceptron (MLP). A multilayer perceptron artificial neural network has any number of inputs, which can have one or more hidden layers with any number of units and takes forward sigmoid activation functions in the hidden layers. MLP is highly suitable to approximate a classification function which is determined by the vector attribute values based on one or more classes. MLP trained with forward or backward propagation algorithm can be effectively used for medical diagnosis.

In our proposed research work we have used WEKA tool for the feature selection and training the model using multi-layer perceptron. WEKA is an open source software stands for Waikato Environment for Knowledge Learning which was developed by the University of Waikato, New Zealand. WEKA tool supports for data mining and machine learning algorithms tasks such as data preprocessing, classification, clustering, regression, feature selection and visualize the data. The data formats to be handled by WEKA are ARFF, CSV, C4.5 and binary and can also port the data from URL or an SQL database. After loading the data, preprocessing filters could be used for adding/removing features, attributes, discretization, sampling, randomizing etc.

Classifier model can be implemented using training and testing model building using multilayer perceptron neural network. For model building, the entire data women breast cancer data set are split into 70% training and 30% testing groups. The preprocessed feature subset datasets are applied into this training model and the data can be again built using Cross Validation 0f 10-fold. The training data sets includes 9 folds and the test set includes a single fold. The Multilayer Perceptron neural network is build and the classifier performance are analyzed. The datasets for different feature subsets generated are observed and the final results are analyzed for to fix the model for further medical analysis. Unnecessary features are removed from the data set using filter ranking method and MLP neural network training model is built which gives a better result than simple train & test method. For the most fare evaluation of the classifier model result of K-fold cross validation method of MLP training was used.

Multi-Layer Perceptron neural network classifier creates two sets training and test set which are used for MLP neural network training in WEKA. Training sets and Test sets are used for learning which fits the weight parameters of the classifier. Test set are the examples used only to assess the performance of a fully-specified trained classifier. The algorithm for MLP neural network classifier model to train and test is given in table 3.

The learning rate and momentum are adjusted to get a better training result. Main parameters for building the classifier with a number of hidden layers, learning rate, momentum, trainingTime. Paramenter setting function isweka.classifiers.functions. MultilayerPerceptron

For better classifier model building and testing, the 10-fold Cross Validation is used in which datasets are randomly divided into 10 folds in which any one fold is used as a testing set and the remaining 9 subsets are used as training sets.

Table 3. MLP Neural Network Classifier

Multi-Layer Perceptron Neural Network Classifier Model
Initialize all weights to small random numbers Until satisfied DO
For each training example, do
1. Input the training example to the network and compute the training outputs
2. For each output unit k
$\delta_k \leftarrow o_k \left(1 - o_k\right)\left(t_k - o_k\right)$
3. For each hidden unit h
$\delta_h \leftarrow o_h \left(1 - o_h\right) \sum_{k \in outputs} w_{h,k} \delta_k$
4. Update each network weight
$w_{i,j} \leftarrow w_{i,j} + \Delta w_{i,j}$ where
$\Delta w_{i,j} = \overset{\overset{Learning\,rate}{\uparrow}}{\eta} \delta_j x_{i.j} \quad \Delta w_{i,j}\left(n\right) = \eta \delta_j x_{i.j} + \overset{\overset{Momentum}{\uparrow}}{\alpha} \Delta w_{i,j}\left(n - 1\right)$

4. MATERIALS AND METHODS

Computer Aided Diagnosis systems perform automatic assessments of patient images and present to radiologist towards the appearance of an abnormality. Computer Aided Diagnosis is an effective and efficient solution for implementing double reading, which provides double perception and interpretation. In a Real Time Scenario, technicians can be attracted by some features and can miss a lesion which can be used to identify disease. The CAD systems read images faster without reducing accuracy than radiologists, but medical community needs to be confidant in the results.

The automatic computer aided diagnosis of breast cancer is an important real-world medical problem and an active research field in the detection of subtle abnormalities in mammograms to improve the accuracy and efficiency of radiologists. Mammograms are the images but difficult to interpret, and a mammographic images are preprocessed to improve the quality of the images and make the feature extraction phase more reliable. The image enhancement technique is applied to enhance the quality of the images. After preprocessing phase, features relevant to the classification are extracted from the preprocessed images in which the most relevant and non-redundant features are selected for the classification model to develop.

The proposed work data sets are taken from UCI machine repository. It consists of 569 Fine Needle Aspirate biopsy samples of human breast tissues. There are 32 attributes computed for each cell samples which includes radius, perimeter, texture, area, smoothness, compactness, concavity, concave points, symmetry and fractal dimension are the 10 most important features which have been used as the only inputs to the network as these are sufficient to obtain good results. This makes the network more concise and less complex.

We used the WEKA toolkit to experiment the breast cancer dataset captured through digital mammography used to evaluate the performance and effectiveness of the breast cancer prediction models. The features of the given women breast cancer data sets features radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, fractal dimensions of mean data. The features includes radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, fractal dimensions of standard error data. The worst data features of radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, fractal dimensions and finally includes class label for classification. The various Feature Selection Methods chosen such as fssubset evaluation, filtered subset evaluation, Gain Ratio, Chisquare, SVM Attribute, Relief Attribute with their corresponding search methods such as Best first and Ranker methods. The features selected are using Gain Ratio, InfoGain, Chisquare, Filtered Attribute Evaluation, One R Attribute Evaluation, Relief Attribute Evaluator, Symmetrical Uncertain Evaluator using Ranking Methods are given below in the table 4.

Features subsets are selected based on the filtered methods and list of the features are chosen based on 65%, 70% and 75% of feature subset. All the feature subsets results are obtained and the detailed feature subsets are given in the below table 5. The total features for the given data set is 30 and the subset of features based on ranking method are shown in below table.

The features are analyzed in each ranking method and union set of all features are obtained and the results are shown in the table 6 out of 30 features excluding class label.

Initially the performance of the multi-layer perceptron on all the features without pre-processing are done at the various training split of 66%, 70%,, 72% and 75% respectively. The performance metrics are given in the table 7.

The performance of a trained classifier based on MLP neural network was evaluated using four performance measures: correctly classified instances (CCI), incorrectly classified instances(ICCI), precision, recall, F-Score and ROC. These measures are defined by four decisions: true positive (TP), true negative (TN), false positive (FN), and false negative (FN). TP decision occurs when malignant instances are predicted rightly. TN decision benign instances are predicted rightly. FP decision occurs when benign instances are predicted as malignant. FN decision occurs when malignant instances are predicted as benign.

Table 4. Filter Features Selection Attributes

InfoGain
Selected attributes:
23,24,21,28,8,3,4,1,7,14,27,11,13,26,6,17,18,22,2,29,16,25,9,5,30,20,19,10,12,15: 30
GainRatio
Selected attributes:
23,21,24,28,8,7,27,3,4,1,14,6,11,13,26,17,2,19,18,25,22,29,5,16,30,9,20,12,10,15: 30
Chisquare
Selected attributes:
23,21,24,28,8,3,4,1,7,14,27,11,13,26,6,17,18,22,2,29,25,16,9,5,30,20,19,10,12,15: 30
Filtered AttributeEval
Selected attributes:
23,24,21,28,8,3,4,1,7,14,27,11,13,26,6,17,18,22,2,29,16,25,9,5,30,20,19,10,12,15: 30
One R Attribute Evaluator
Selected attributes:
8,21,28,23,3,24,7,1,4,14,27,13,11,6,26,25,22,18,2,29,17,5,30,9,15,16,20,10,12,19: 30
Relief Attribute Evaluator
Selected attributes:
21,28,23,22,1,3,8,24,4,7,2,27,25,11,26,14,10,13,6,5,29,12,19,18,15,30,16,17,9,20: 30
SVM Attribute Evaluator
Selected attributes:
21,28,23,22,8,24,29,1,25,4,11,2,3,7,16,13,10,27,14,9,6,5,15,20,30,12,17,18,19,26: 30
Symmetrical Uncert Attribute Eval
Selected attributes:
23,21,24,28,8,3,7,4,1,27,14,11,13,6,26,17,2,18,22,25,29,16,5,30,9,19,20,10,12,15: 30

Table 5. Filter Features Subset Selection

S.Nos	Ranking	65% of Features	70% of Features	75% of Features
1	InfoGain	20	21	23
2	GainRatio	20	21	23
3	Chisquare	20	21	23
4	Filtered AttributeEval	20	21	23
5	One R Attribute Evaluator	20	21	23
6	Relief Attribute Evaluator	20	21	23
7	Symmetrical Uncert Attribute Eval	20	21	23

Table 6. Final Subset Features

S.Nos.	Ranking Features Subset	Total Features Subset	Final Ensemble Features Subset
1	Ensemble Features Subset 1 -65%	20	24
2	Ensemble Features Subset 2 -70% Set	21	25
3	Ensemble Features Subset 2 -75%	23	26

Table 7. MLP Classifier with all features (30)

S.Nos.	Training Split	CCI	ICCI	Precision	Recall	F-Score	ROC
1	Default - 66%	96.37	3.63	0.964	0.964	0.964	0.987
2	70%	97.66	2.34	0.977	0.977	0.977	0.988
3	72%	96.23	3.77	0.962	0.962	0.962	0.987
4	75%	95.77	4.22	0.958	0.958	0.958	0.975

Table 8. MLP Classifier based on 10-Fold Cross Validation with all features (30)

S.Nos.	Training Split	CCI	ICCI	Precision	Recall	F-Score	ROC
1	Cross Validation – All Features	95.66	4.33	0.95	0.95	0.95	0.98
2	Cross Validation – Features Subset 1	96.31	3.69	0.96	0.96	0.96	0.99
3	Cross Validation – Features Subset 2	96.30	3.69	0.96	0.96	0.96	0.99
4	Cross Validation – Features Subset 3	96.48	3.51	0.97	0.97	0.97	0.99
5	Training Split – 70% – Features Subset 3	97.07	2.92	0.97	0.97	0.97	0.98

To evaluate the performance of above methods of neural network training different parameters are accuracy, precision, recall, F-Measure, kappa score etc. For our proposed system the metrics used are accuracy, precision, recall and ROC are considered. Accuracy gives a measured value in par with the actual (true) value and retrieves the percentage of correctly classified instances.

6. CONCLUSION

In this paper, an automated computer aided diagnosis system has been devised for diagnosis of breast cancer. The performance are measured and computed with supervised neural network classifier model. The produced classification results are very much promising with 97% accuracy of correct classification and F-Score is of 93% with reduced error measures is of 4% by including all features. The proposed algorithm gives up an ensemble feature subsets and this is applied to artificial neural network with a better accuracy for breast cancer diagnosis. The proposed method may provide an adequate support to the radiologists in differentiating between normal and abnormal breast cancer identification with high accuracy and of low error measures. The research can be focused further to develop better preprocessing, enhancement and segmentation techniques. The proposed work can be further expanded to design enhanced feature extraction and selection and also appropriate classification algorithms can be used to reduce both false positives and false negatives by employing high resolution mammograms and investigating 3D mammograms.

REFERENCES

Arafi, A., Safi, Y., Fajr, R., & Bouroumi, A. (2013). Classification of Mammographic Images using Artificial Neural Network. *Applied Mathematical Sciences*, *7*(89), 4415–4423. doi:10.12988/ams.2013.35293

Chang, R. F., Wu, W. J., Moon, W. K., & Chen, D.-R. (2005). Dr Chen, “Automatic Ultrasound Segmentation and Morphology based Diagnosis of Solid Breast Tumors. *Breast Cancer Research and Treatment*, *89*(2), 179–185. doi:10.100710549-004-2043-z PMID:15692761

Chen, Y., Wang, Y., & Yang, B. (2006). Evolving Hierarchical RBF Neural Networks for Breast Cancer Detection. LNCS, 4234, 137-144. doi:10.1007/11893295_16

Hazlina, H., & Sameem, A. K. (2004). Back Propagation Neural Network for the Prognosis of Breast Cancer: Comparison on Different Training Algorithms. *Proceedings Second International Conference on Artificial Intelligence in Engineering & Technology*, 445-449.

Horsch, K., Giger, M. L., Venkata, L. A., & Vybomya, C. J. (2001). Automatic Segmentation of Breast Lesions on Ultrasound. *Medical Physics*, *28*(8), 1652–1659. doi:10.1118/1.1386426 PMID:11548934

Hotko, Y. S. (2013). Male breast cancer: Clinical presentation, diagnosis, treatment. *Experimental Oncology*, *35*(4), 303–310. PMID:24382442

Kiyan, T., & Yildrim, T. (2004). Breast Cancer Diagnosis using Statistical Neural Networks. *Journal of Electrical and Electronics Engineering (Oradea)*, *4*(2), 1149–1153.

Madabhushi, A., & Metaxas, D. (2003). Combining low-, high-level and Empirical Domain Knowledge for Automated Segmentation of Ultrasonic Breast Lesions. *IEEE Transactions on Medical Imaging*, *22*(2), 155–169. doi:10.1109/TMI.2002.808364 PMID:12715992

Malvia, S., Bagadi, S. A., Dubey, U. S., & Saxena, S. (2017). Epidemiology ofbreast cancer in Indian women. *Asia Pacific Journal of Clinical Oncology*, *13*(4), 289–295. doi:10.1111/ajco.12661 PMID:28181405

Raad, A., Kalakech, A., & Ayache, M. (2012). Breast Cancer Classification using Neural Network Approach: MLP AND RBF. *The 13th international Arab conference on information technology,* 10 – 13.

Sahiner, B., Heang-Ping, C., Patrick, N., Wei, D. M. A., Helie, D., Adler, D., & Goodsitt, M. M. (1996). Classification of Mass and Normal Breast Tissue: A Convolution Neural Network Classifier with Spatial Domain and Texture Images. *IEEE Transactions on Medical Imaging*, *15*(5), 598–610. doi:10.1109/42.538937 PMID:18215941

Sapra, P., Singh, R., & Khurana, S. (2013). Brain Tumor Detection using Neural Network. *International Journal of Science and Modern Engineering*, *1*(9).

Shallu, R. M., & Mehra, R. (2018). Breast cancer histology images classification:Training from scratch or transfer learning? *ICT Express*, *4*(4), 247–254. doi:10.1016/j.icte.2018.10.007

Singh, S., Saini, S., & Singh, M. (2012). Cancer Detection using Adaptive Neural Network. *International Journal of Advancements in Research and Technology*, *1*(4).

Wadhwani & Saraswat. (2009). Classification of breast cancer using artificial neural network. *Current Research in Engineering. Science and Technology Journals*.

Wu, Y., Wang, N., Zhang, H., Qin, L., Yan, Z., & Wu, Y. (2010). Application of Neural Networks in the Diagnosis of Lung Cancer by Computed Tomography. *Sixth International Conference on Natural Computation*. 10.1109/ICNC.2010.5583316

Chapter 7
A Content-Based Approach to Medical Image Retrieval

Anitha K.
Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

Naresh K.
VIT University, India

Rukmani Devi D.
RMD Engineering College, India

ABSTRACT

Medical images stored in distributed and centralized servers are referred to for knowledge, teaching, information, and diagnosis. Content-based image retrieval (CBIR) is used to locate images in vast databases. Images are indexed and retrieved with a set of features. The CBIR model on receipt of query extracts same set of features of query, matches with indexed features index, and retrieves similar images from database. Thus, the system performance mainly depends on the features adopted for indexing. Features selected must require lesser storage, retrieval time, cost of retrieval model, and must support different classifier algorithms. Feature set adopted should support to improve the performance of the system. The chapter briefs on the strength of local binary patterns (LBP) and its variants for indexing medical images. Efficacy of the LBP is verified using medical images from OASIS. The results presented in the chapter are obtained by direct method without the aid of any classification techniques like SVM, neural networks, etc. The results prove good prospects of LBP and its variants.

DOI: 10.4018/978-1-7998-3092-4.ch007

INTRODUCTION

Due to the enormous size of medical image data repository, CBIR can be used for medical image retrieval. This chapter is envisioned to propagate the knowledge of the CBIR approach to deal with the applications of medical image management and to pull in more prominent enthusiasm from various research groups to rapidly propel research in this field.

The image is presumably a standout amongst the most essential tools in medicine since it provides a method for diagnosis, monitoring drug treatment responses and disease management of patients with the advantage of being a very fast non-invasive procedure, having very few side effects and with an excellent cost-effect relationship.

Hard-copy image formats used to support for medical images are not utilized these days. The expense and resource involved in maintenance, storage room and the amount of material to display images in this format contributed for its disuse. Nowadays digital images, that doesn't face problems mentioned for hard copy formats are used. **Table 1** gives a review of digital images per exam in medical imaging. This transition from hard-copy to soft-copy images is still the center of an interesting debate related with human perception and understanding issues during exam analysis. Elizabeth (2000) have tended to the significance of observation in medical imaging.

Table 1. Types and sizes of some commonly used digital medical images from Huang (2004)

Image Type	One Image(bits)	No. of Images/ Exam	One Examination
Nuclear medicine (**NM**)	128X128X12	30-60	1-2 MB
Magnetic resonance imaging (**MRI**)	256X256X12	60-3000	8 MB
Ultrasound (**US**)*	512X512X8	20-240	5-60 MB
Digital subtraction angiography (**DS**)	512X512X8	15-40	4-10 MB
Digital microscopy	512X512X8	1	0.25 MB
Digital color microscopy	512X512X24	1	0.75 MB
Color light images	512X512X24	4-20	3-15 MB
Computed tomography (**CT**)	512X512X24	40-3000	20 MB
Computed/digital radiography (**CR/DR**)	2048X2048X12	2	16 MB
Digitized X-rays	2048X2048X12	2	16 MB
Digital mammography	4000X5000X12	4	160 MB

*Doppler US with 24 bit color images

Increase of medicinal information in digital libraries makes tougher to perform analysis on search related tasks. Since textual information retrieval is as of now a developed discipline, an approach to overcome this issue is to utilize metadata for image indexing. Key description, patient identification, kind of exam and its technical details or even a small text comment concerning clinical relevant information can be utilized to represent the image in its index. With these annotations, text-matching techniques that assess the similarity between the search statement and the metadata can be applied for retrieving images. This is called text-based or concept-based image retrieval.

LIMITATIONS OF CONCEPT-BASED RETRIEVAL AND THE NEED FOR CBIR

Generally to archive an image with a general vocabulary of medical terms devours numerous resources and requests broad collaboration efforts that is difficult to incorporate. It is sensible to utilize inductive methodologies by beginning with more particular standards and endeavor generalization later. This strategy is utilized by (Dean Bidgood 1998) in the composite SNOMED-DICOM micro-glossary. Annotation of images by human is a time consuming and cumbersome task and also lead to unrecoverable errors. (Mark et al 2002) has undergone the study of medical images using DICOM headers and revealed that 15% of annotation errors occur from both human and machine origin. Another significant obstacle in concept-based image retrieval systems is the fact that the query, does not allow the user to switch and/or combine interaction paradigms (Shi-kuo and Arding 1992) during text transactions. The ideal framework would relieve the human from the annotation task, by doing it automatically, and allowing image retrieval by its content instead of textual description. This framework is known as Content Based Image Retrieval (CBIR).

CBIR IN MEDICAL APPLICATIONS

The CBIR in the medical field presents a developing trend in publications. Comparative analysis of CBIR implementations in medical imaging are presented by (Long et al 2009 and Hussain et al 2020). A study proposed by (Ogul et al 2020) implemented a new method of PD detection from gait signals, using artificial neural networks and a novel technique framework called Neighborhood Representation Local Binary Pattern (NR-LBP). Vertical Ground Reaction Force (VGRF) readings are preprocessed and transformed using several methods within the proposed framework. Despite the growth of CBIR frameworks in medical imaging, the utility of the frameworks in number of medical applications is still very constrained. There exist only a couple of systems with relative success. The Cervigram Finder system (Zhiyun et al 2008) was developed to study the uterine cervix cancer. It is a computer assisted framework where local features from a user-defined region in an image are computed and, using similarity measures, similar images are retrieved from a database. The Image Retrieval for Medical Applications (IRMA) framework (Thomas et al 2008) is a web-based x-ray retrieval system. It permits the user to retrieve images from a database given an x-ray image as query. CBIR for medical applications can be found in (Lehmann et al 2004 and Muller et al 2007). Medical applications are one of the priority areas where CBIR can meet more accomplishment due to population aging in developed countries.

PROBLEM DEFINITION

Generally to archive an image with a general vocabulary of medical terms devours numerous resources and requests broad collaboration efforts that is difficult to incorporate. It is sensible to utilize inductive methodologies by beginning with more particular standards and endeavor generalization later. Another significant obstacle in concept-based image retrieval systems is the fact that the query, does not allow the user to switch and/or combine interaction paradigms during text transactions. The ideal framework would relieve the human from the annotation task, by doing it automatically, and allowing image retrieval by its content instead of textual description. This framework is known as retrieving images with the help of the image content.

Increase of medicinal information in digital libraries and availability of advanced technologies to store and manage shall be utilized effectively in diagnosis and treating patients. Searching and retrieving similar cases from the huge database plays a vital role in developing an efficient architecture to perform the task. The information stored shall be in the form of medical images and text records. Medical Image Retrieval Systems (MIRS) with determined feature vector to index the images and adopting advanced Classifying methods shall be more efficient to cluster and locate similar medical images or records stored in vast databases.

The proposed work concentrates on classification of medical images provided by OASIS. The work aims at the following goals:

- Implement and analysis a simple, effective and efficient medical image retrieval system
- To use Local Binary Pattern (LBP operator) to represent medical images and develop a retrieval system that match and retrieve images with the support of the operator.

- To derive hybrid descriptors combining two LBP operators with different sizes of neighborhood pixels and study the impact on the retrieval results
- To point new directions and considerations where future work can be developed

IMAGE DOMAIN

The digital image is an approximation of a two-dimensional image by set of values called pixels. Each pixel is described in terms of its color, intensity/luminance or value. In order to retrieve images according to a given query we need to enhance its relevant elements while reducing the remaining aspects. This is the goal of image processing. Generically, the image is processed using an operator, **g**, over the full spatial domain of the image, $I(X,Y)$, an interest point, $I(x_n, y_m)$ to generate a feature space containing the information needed to identify the objects in the following way:

$$f(r) = g \circ I(x,y) \tag{1}$$

$$f(r_{(x,y)}) = g \circ I(x,y) \tag{2}$$

$$f(r_{(x_n,y_m)}) = g \circ I(x_n, y_m) \tag{3}$$

Where $I(X,Y)$ is the full image; $I(x,y)$ is an image patch, i.e., a connected subset of Cartesian points with $(x,y) \in (X,Y), \forall x,y \in R$ and $I(x_n, y_m)$ is at interest point (x_n, y_m) where $n \in \{1,2,3,.....,N_x\}$ and $m \in \{1,2,3,.....,M_x\}$.

Image Properties

The relationships between image properties like color, shape, texture and interest points are the fundamental characteristics of an image. Thus the properties used to represent images mathematically are named as image descriptors. The similarity between images is also estimated using these descriptors.

Image Descriptors

It is very difficult to ascertain which image properties are fundamental to characterize a specific image. It depends on the context of the problem to be solved and the knowledge within the image itself. If there is a need to recognize specific objects in a scene, probably the shape property is more relevant than the others. However if such objects have a distinct color then the relevance of this property is higher than the rest. If there is a need to detect light bulbs in a night scenario we rely on an interest point detector. Sometimes the image can be quite complex and all properties are essential for its characterization. Eq. (1) to (3) enables the quantification of the image properties. This quantification bridges the gap between human perception of image properties and mathematical measure(s) taken from the image. This is the aim of an image descriptor.

Image descriptors can be global, like Eq. (1), or local, like Eq. (2) and Eq. (3). The feature database is constructed using global descriptor that represents the visual features of the whole image as well as the local descriptor that represents the visual features of regions or objects to describe the image. Similarity measurement is done by many methods like Euclidean distance (L2), L1 distance etc. The selection of feature descriptors and similarity metrics greatly affects the retrieval performance.

The description of the features available in the literature for biomedical imaging is briefed as follows. (Hersh et al 2009) proposed Image CLEF medical image retrieval task to improve understanding and system capability in search for medical images. They described the development and use of a medical image test collection design to facilitate research with image retrieval systems and their users. (Manjunath et al 2007) presented bit plane histogram and hierarchical bit plane histogram along with cumulative distribution function (CDF) for CT and MRI image retrieval. (Fahimeh et al 2010) has done the classification of benign and malignant breast masses based on shape and texture features in sonography images. The blood cell image retrieval using color histogram and wavelet transform can be seen in the work proposed by (Woo and Seyed 2009).

Medical image retrieval system detailed in this chapter uses Local binary patterns (LBP) proposed by (Ojala et al 1996) to represent images in database. LBP can show better performance as well as less computational complexity for image classification. LBP have been used in many research areas such as texture classification (Guo et al 2010, Xueming et al 2011, Loris et al 2012), face recognition (Zhang et al 2010, Tan et al 2010, Papakostas et al 2013), and image retrieval(Cheng-Hao et al 2003, Sorensen et al 2010).

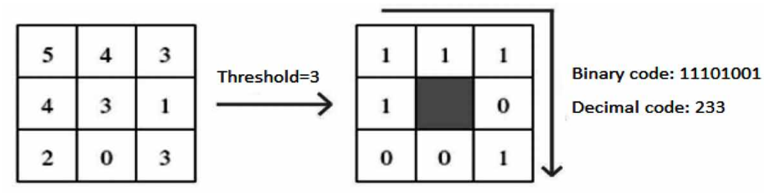
LBP METHODOLOGY

The image descriptors developed for representing images mathematically shows more emphasis on effectiveness and less on efficiency. Effectiveness is the closeness of the retrieved images to the query. Efficiency is the utilization of optimum resources to retrieve more similar images from the datasets. The need is an effective and efficient system.

Focuses of the chapter is to brief an analysis and implementation of an image descriptor to retrieve medical images that speed up the process by reducing the dimension of image descriptor. To get real time processing speeds, less dimension of the image descriptor is an important property of the methods which take less time to build the feature and less time to retrieve the matched images. However, lower dimension usually has less distinctiveness than higher dimension. Therefore, we must need to take care of this tradeoff while building the descriptor to represent the image. Among the image description methods, local binary pattern (LBP) has received considerable attentions in many computer vision applications, such as face recognition, image retrieval and motion analysis, for its efficiency and simple computational complexity to build the image descriptor. LBP method had been proposed for texture analysis. This operator is defined as a monotonic illumination invariant texture measure, derived from local neighborhood. For each pixel in an image, a binary code is produced by thresholding center intensity value with the intensity value of the neighbor pixel. Then, A histogram, created to collect the occurrences of different binary patterns, is used to represent the image. The basic LBP operator considers only the eight neighbors of a center pixel. However, the definition has been extended to include any number of circular neighborhoods by using the interpolation technique

The LBP operator was initially presented as a complementary measure for local image contrast (Ojala et al 1996). An approach that opted to adopt Local Binary Patterns and k-means clustering for precise identification of lesions boundaries (Pedro et al 2020), particularly the melanocytic has been derived. A blind detection based uniform local binary patterns (ULBP) is proposed (Zhang et al 2020) to detect seam-carved image. The gentle boost decision trees are trained(Gogic et al 2020) to extract highly discriminative feature vectors (local binary features) for each basic facial expression around distinct facial landmark points for faster fascial recognition. The operator works with the eight-neighbors of a pixel, using the center pixel value as a threshold. LBP code for a neighborhood was generated by multiplying the threshold values with weights assigned to the corresponding pixels, and summing up the result Figure 6.1. Since the LBP was, invariant to monotonic changes in gray scale, it was supplemented by an orthogonal measure of local contrast.

Figure 1. Basic Version of LBP operator



Derivation

The texture T in a local neighborhood of a grayscale image is defined as the joint distribution of the gray levels of P + 1 (P > 0) image pixels:

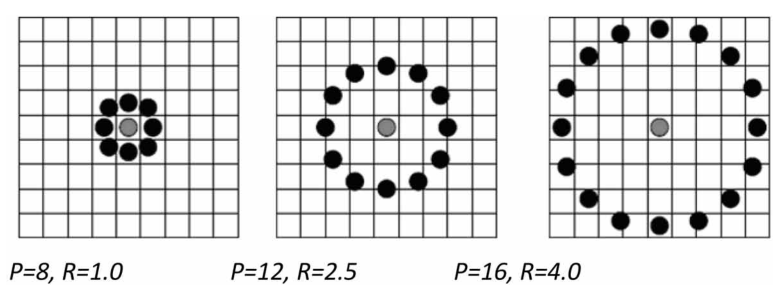
$$T = t(g_c, g_0......., g_{p-1}) \tag{4}$$

where g_c corresponds to the gray value of the center pixel of a local neighborhood. $g_p(p = 0,1,2,3......, P-1)$ correspond to the gray values of P pixels equally spaced on a circle of radius R (R > 0) that form a circularly symmetric set of neighbors. This set of $P+1$ pixels is later denoted by G_P. In a digital image domain, the coordinates of the g_P neighbors are given by

$$(x_c + R\cos(2\Pi p / P), y_c - R\sin(2\Pi p / P))$$

where (x_c, y_c) are the coordinates of the center pixel. Figure 2 illustrates three circularly symmetric neighbor sets for different values of P and R.

Figure 2. Circularly symmetric neighbor sets. Samples that do not exactly match the pixel grid are obtained via interpolation.



The values of neighbors that do not fall exactly on pixels are estimated by bilinear interpolation and the operator can be expressed using

$$T = t(g_c, g_0 - g_c\ldots\ldots, g_{p-1} - g_c) \tag{5}$$

Assuming that the differences are independent of g_c, the distribution can be factorized:

$$T \approx t(g_c), t(g_0 - g_c\ldots\ldots, g_{p-1} - g_c) \tag{6}$$

Since $t(g_c)$ describes the overall luminance of an image, which is unrelated to local image texture, it does not provide useful information for texture analysis Eq. 6 simplifies to

$$T \approx t(g_0 - g_c\ldots\ldots, g_{p-1} - g_c) \tag{7}$$

The P−dimensional difference distribution records the occurrences of different texture patterns in the neighborhood of each pixel. For constant or slowly varying regions, the differences cluster near zero. On a spot, all differences are relatively large. On an edge, differences in some directions are larger than the others. Although invariant against gray scale shifts, the differences are affected by scaling. To achieve invariance with respect to any monotonic transformation of the gray scale, only the signs of the differences are considered:

$$T \approx t(s(g_0 - g_c)\ldots\ldots, s(g_{p-1} - g_c)) \tag{8}$$

where

$$s(x) = \{^{1 \quad s \geq 0}_{0 \quad s < 0} \tag{9}$$

Now, a binomial weight 2^p is assigned to each sign $s(g_p - g_c)$, transforming the differences in a neighborhood into a unique LBP code. The code characterizes the local image texture around (x_c, y_c)

$$LBP_{P,R}(x_c, y_c) = \sum_{p}^{P-1} s(g_p - g_c)2^p \tag{10}$$

In practice, Eq. 6.8 means that the signs of the differences in a neighborhood are interpreted as a P-bit binary number, resulting in 2^P distinct values for the LBP code. The local gray-scale distribution, i.e. texture, can thus be approximately described with a 2^P bin discrete distribution of LBP codes:

$$T \approx t(LBP_{P,R}(x_c, y_c)) \tag{11}$$

Let us assume we are given an NXM image sample

$(x_c \in \{0,1,2,....,N-1\}, y_c \in \{0,1,2....,M-1\})$.

In calculating the $LBP_{P,R}$ distribution (feature vector) for this image, the central part is only considered because a sufficiently large neighborhood cannot be used on the borders. The LBP code is calculated for each pixel in the cropped portion of the image, and the distribution of the codes is used as a feature vector, denoted by S

$$S = t(LBR_{P,R}(x, y)) \tag{12}$$

where

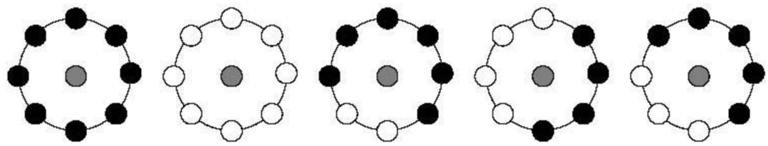
$x \in \{[R],......,N-1-[R]\}, y \in \{[R],......,M-1-[R]\}$.

Combination of the Structural and Stochastic Approaches by LBP

The LBP operator to represent images is a unifying approach and can be considered as a micro-texton. The texture primitive that best matches the nearby neighborhood is used to label every pixel. Spots, edges, curves, flat areas, edge ends etc. are detected by local primitives of LBP.

Figure 3 provides few of such illustrations. White and black circles in the figure, represents ones and zeros respectively.

Figure 3. Different texture primitives detected by the LBP



Statistical and structural methods have been normally used separately to detect the textures. The LBP technique has both of these properties, texture primitives and placement rules. Thus the operator is a better option to distinguish/match a variety of texture images

Rotation Invariance

Circular sampling of neighborhoods to form the operator supports to make LBP operator invariant to the rotation of an image. The below assumptions are considered to derive the operator.

- Rotation invariance here does not count for textural differences caused by changes in the relative positions of a light source and the target object.
- The effects caused due to digitizing effects are neglected.
- Each pixel is considered a rotation center

The gray values g_p in a neighbor set of the circle move along the perimeter centered at g_c due to the rotation of the image. $LBP_{P,R}$ Value will be different for a rotated image from that of the actual. Patterns comprising of only zeros or ones will remain unchanged with rotation. Rotating the LBP operator back to a reference position will eliminate the effect of rotation and makes versions of a binary code the same. This transformation can be defined as follows:

$$LBP_{P,R}^{ri} = \min\{ROR(LBP_{P,R}, i)\forall i = 0,1,....,P-1\} \tag{13}$$

where the superscript ri stands for rotation invariant. The function $ROR(x,i)$ circularly shifts the P-bit binary number x by i times to the right $(|i| < p)$. That is, given a binary number x :

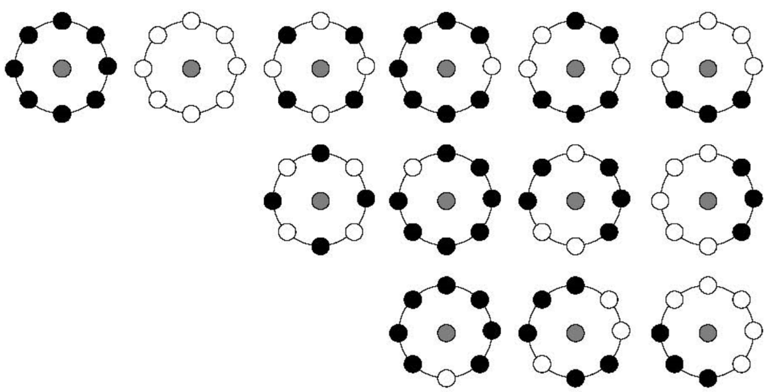
$$x = \sum_{k=0}^{P-1} 2^k a_k \forall a_k = \{0,1\} \tag{14}$$

the ROR operation is defined as:

$$ROR(x,i) = \begin{cases} \sum_{k=i}^{P-1} 2^{k-i} a_k + \sum_{k=0}^{i-1} 2^{P-i+k} a_k \; i > 0 \\ x \; i = 0 \\ ROR(x, P+i) \; i < 0 \end{cases} \tag{15}$$

In short, the rotation invariant code is produced by circularly rotating the original code until its minimum value is attained. Figure 6.4 illustrates six rotation invariant codes in the top row. Below these, examples of rotated neighborhoods that result in the same rotation invariant code are shown. In total, there are 36 different 8-bit rotation invariant codes. Therefore, $LBR_{8,R}^{ri}$, produces 36-bin histograms.

Figure 4. Neighborhoods rotated to their minimum value (top row) and that produce the same rotation invariant LBP codes.



The LBP codes shown in Figure 3 are all uniform. Examples of non-uniform codes can be seen in Figure 4, in the third and fifth columns. To formally define the uniformity of a neighborhood G, a uniformity measure U is needed:

$$U(G) = \left| s(g_{P-1} - g_c) - s(g_0 - g_c) \right| + \sum_{p=1}^{P-1} \left| s(g_p - g_c) - s(g_{p-1} - g_c) \right| \tag{16}$$

Patterns with a U value of less than or equal to two are designated as uniform. For a P-bit binary number, the U value can be calculated efficiently as follows:

$$U(x) = \sum_{p=0}^{P-1} F_b(x \quad xor - ROR(x,1), p) \tag{17}$$

where b is a binary number. The function $F_b(x,i)$ extracts b circularly successive bits from a binary number x

$$F_b(x,i) = ROR(x,i)\, and\, (2^b - 1) \tag{18}$$

where i is the index of the least significant bit of the bit sequence. The operators "and" and "xor" denote bitwise logical operations.

The total number of patterns with U(GP)≤2 is P(P−1)+2. When uniform codes are rotated to their minimum values, the total number of patterns becomes P + 1. The rotation invariant uniform (riu2) pattern code for any uniform pattern is calculated by simply counting ones in the binary number. All other patterns are labeled "miscellaneous" and collapsed into one value

$$LBP_{P,R}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) U(G_P) < 2 \\ P + 1\, otherwise \end{cases} \tag{19}$$

In practice, $LBP_{P,R}^{riu2}$ is best implemented by creating a look-up table that converts the "basic" LBP codes into their $LBP_{P,R}^{riu2}$ correspondents.

Contrast and Texture Patterns

The LBP operator ignores the amount of gray level divergences. But the magnitude of gray level that provides the contrast is a property of texture and is important for our vision system to arrive a result. An operator that is not influenced by gray scale may waste useful information obtained from applications that have a reasonable control on illumination accurately. The accuracy of the operator can be enhanced by including information about gray-scale. Texture is identified with two properties spatial structure and contrast. Spatial structure is independent on gray scale and affected by rotation whereas contrast is dependent on gray scale but not affected by rotation. A joint distribution of LBP operator and local contrast measure (LBP/C) as a texture descriptor has been implemented by (Ojala et al 1996).

$VAR_{P,R}$ does not change due to variations in the gray scale. It can be estimated in circular sets as similar to LBP.

$$
\begin{aligned}
&VAR_{P,R} = \frac{1}{P}\sum_{p=0}^{P-1}(g_p - \mu)^2 \\
&where\, \mu = \frac{1}{P}\sum_{p=0}^{P-1} g_p
\end{aligned}
\tag{20}
$$

An operator denoted as $LBP_{P1,R1}^{riu2}/VAR_{P2,R2}$ formed with joint distribution of LBP and local variance can be formed. The texture descriptor thus formed will not vary with rotation.

EXPERIMENTAL RESULTS

OASIS (Open Access Series of Imaging Studies) is a series of magnetic resonance imaging (MRI) dataset that is publicly available for study and evaluation. This dataset proposed by (Marcus et al 2007) comprises cross-sectional collection of 421 subjects. The images are grouped into four categories (Group-1: 124, Group-2:102, Group-3: 89, and Group-4: 106 images) and used for experimentation of proposed CBIR system. Grouping is done based on the shape of ventricular in the images. For training the classifier 30% of images in each category are utilized and 30% and 40% of images are used for validation and testing respectively.

The purpose of experiments is to estimate and prove the capability of the LBP operator to represent the medical images mathematically. Experiments are performed separately to test the capacity of the operator to classify the images and to retrieve the images. A basic retrieval system that classifies and retrieves images using LBP (without help of trained classifier) is tested and the results are shown.

Similarity Measure

The objective of any CBIR system is to archive n best images from an image database (N number of images) that resemble the query image. The selection of n images that best matches is selected by measuring the distance between query image and *N* images in the database. In literature it is found four types of similarity distance metrics have been used for the purpose.

L_1 or Manhattan distance measure

$$D(Q, DB_j) = \sum_{i=1}^{lf} \left| \left(f_{DBji} - f_{Qi} \right) \right| \tag{21}$$

L_2 Euclidean distance measure:

$$D(Q, DB_j) = \sqrt{\sum_{i=1}^{lf} \left(f_{DBji} - f_{Qi} \right)^2} \tag{22}$$

Canberra distance measure:

$$D(Q, DB_j) = \sum_{i=1}^{lf} \frac{\left| \left(f_{DBji} - f_{Qi} \right) \right|}{\left| f_{DBji} \right| + \left| f_{Qi} \right|} \tag{23}$$

d1 distance measure:

$$D(Q, DB_j) = \sum_{i=1}^{lf} \left| \frac{\left(f_{DBji} - f_{Qi} \right)}{1 + f_{DBji} + f_{Qi}} \right| \tag{24}$$

Where f_{DBji} is the jth (length of feature vector is lf) feature of ith image in Data base *DB* of *N* images.

Evaluation Metrics

The performance of the proposed system is evaluated by the parameters average retrieval precision (ARP) and average retrieval rate (ARR). The evaluation parameters are calculated using Eqs. (26)–(29). The equations also define the precision (*P*) and recall (*R*) for the query image $I_{q.}$

$$Precision\ P(I_q) = \frac{Number\ of\ relevant\ images\ retrieved}{Number\ of\ images\ retrieved} \tag{25}$$

$$ARP = \frac{1}{N} \sum_{i=1}^{N} P(I_i) \Big|_{n \leq 10} \tag{26}$$

$$Recall\ R(I_q) = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ relevant\ images\ in\ data\ base} \tag{27}$$

$$ARR = \frac{1}{N}\sum_{i=1}^{N} R(I_i)\Bigg|_{n \geq 10} \tag{28}$$

In order to analyze the performance of our algorithm biomedical image retrieval is performed on two different medical databases. Results obtained are discussed in the following subsections. In all experiments, each image in the database is used as the query image. For each query, the system collects n database images with the shortest image matching distance is given by Eq. (26-28). If the retrieved image belongs to the same category of the query image, we say the system has correctly matched the desired.

Classification and Retrieval Using Basic Retrieval System

Results in Table 2 illustrate the group wise percentage of correctly retrieved images from OASIS-MRI database. In this chapter, experiments are performed with individual LBP and VAR features of images and by combination of both to develop a hybrid system. In addition to the improvements achieved with hybrid features, the combined features significantly found robust under illumination variations and invariance to rotation. It has been observed from detailed experimental results that $LBP_{16,2}^{riu2}$ and $VAR_{8,1}$ supersede other variants considered individually or if combined for various images groups. The accuracy of the retrieval system is improved further by integrating variants of LBP and VAR. Several experiments are performed to analyze and confirm the superiority of the proposed hybrid approach. $LBP_{16,2}^{riu2}/VAR_{8,1}$ based hybrid approach also improves the image retrieval rate but slightly less than $LBP_{8,1}^{riu2}/LBP_{24,3}^{riu2}$. It has also been observed that hybrid approach is more effective for extracting features offline and for online retrieval in comparison to other hybrid methods. Therefore, all the experiments performed in this chapter lead to proving LBP based indexing, the proposed hybrid approach is more superior.

From **Table 2,** the following inference is drawn. $LBP_{16,2}^{riu2}$ and $LBP_{24,3}^{riu2}$ clearly outperformed their simpler counterpart $LBP_{8,1}^{riu2}$, which had difficulties in discriminating strongly oriented texture. In nearly 176 test cases, the system using $LBP_{8,1}^{riu2}$ identified true class of 74 test samples, $LBP_{16,2}^{riu2}$ identified 79 test samples $LBP_{24,3}^{riu2}$ identified 78 test samples correctly. If group wise classification of images is considered, $LBP_{16,2}^{riu2}$ did much better, classifying more than 45% samples correctly. Combining the $LBP_{P,R}^{riu2}$ operator with the $VAR_{P,R}$ operator, improved the performance. It was observed that $LBP_{16,2}^{riu2}/VAR_{16,2}$ provided more comparable results. It is noticed from the classification results the combined features of LBP and VAR aids to improve the search and retrieval process. From the evaluations it is evident that LBP and VAR are redundant and provide excellent results under illumination variations and texture images

Table 2. Group wise percentage of correctly retrieved images from OASIS-MRI database

Method	Category				
	1	2	3	4	Avg
$LBP_{8,1}^{riu2}$	54.10	36.50	38.89	41.86	42.04
$LBP_{16,2}^{riu2}$	58.00	39.02	36.11	55.81	46.59
$LBP_{24,3}^{riu2}$	48.00	46.34	38.89	48.83	44.31
$VAR_{8,1}$	58.00	41.46	41.66	44.18	45.45
$VAR_{16,2}$	54.00	36.58	36.11	37.20	40.34
$VAR_{24,3}$	52.50	36.58	38.88	34.88	39.77
$LBP_{8,1}^{riu2}/LBP_{16,2}^{riu2}$	68.00	53.66	58.33	53.49	**56.82**
$LBP_{8,1}^{riu2}/LBP_{24,3}^{riu2}$	72.00	63.41	63.89	55.14	**61.93**
$LBP_{16,2}^{riu2}/LBP_{24,3}^{riu2}$	68.00	60.98	61.11	51.16	**58.52**
$VAR_{8,1}/VAR_{16,2}$	62.00	46.34	47.22	48.84	50.00
$VAR_{8,1}/VAR_{24,3}$	54.00	43.90	44.44	46.51	46.02
$VAR_{16,2}/VAR_{24,3}$	62.00	58.54	58.33	51.16	55.68

continues on following page

Table 2. Continued

Method	Category				
	1	2	3	4	Avg
$LBP_{8,1}^{riu2}/VAR_{8,1}$	55.35	48.78	52.776	48.83	51.70
$LBP_{8,1}^{riu2}/VAR_{16,2}$	62.00	51.21	50.00	44.18	50.00
$LBP_{8,1}^{riu2}/VAR_{24,3}$	60.00	48.78	55.00	44.18	49.43
$LBP_{16,2}^{riu2}/VAR_{8,1}$	70.00	63.41	63.89	51.16	**60.23**
$LBP_{16,2}^{riu2}/VAR_{16,2}$	64.00	51.21	58.33	51.16	54.44
$LBP_{16,2}^{riu2}/VAR_{24,3}$	58.00	48.78	47.22	44.18	48.86
$LBP_{24,3}^{riu2}/VAR_{8,1}$	68.00	58.54	63.89	48.84	**57.95**
$LBP_{24,3}^{riu2}/VAR_{16,2}$	62.00	43.90	44.44	55.81	50.57
$LBP_{24,3}^{riu2}/VAR_{24,3}$	60.00	48.78	52.77	48.83	51.13

Table 3. Percentage of classification accuracy achieved using LBP operator with different distance measures on OASIS database

Method	Distance Measure			
	L_1	L_2	Canberra	d_1
$LBP_{8,1}^{riu2}/LBP_{16,2}^{riu2}$	56.82	55.11	57.39	59.09
$LBP_{8,1}^{riu2}/LBP_{24,3}^{riu2}$	61.93	59.09	63.64	65.91
$LBP_{16,2}^{riu2}/LBP_{24,3}^{riu2}$	58.52	55.68	59.66	61.36
$LBP_{16,2}^{riu2}/VAR_{8,1}$	60.23	57.95	61.93	63.64
$LBP_{24,3}^{riu2}/VAR_{8,1}$	57.95	56.25	60.80	62.50

Table 3 illustrates the percentage of classification accuracy achieved using LBP operator with different distance measures. It is clear that d1 similarity measure shows better performance as compared to other measures.

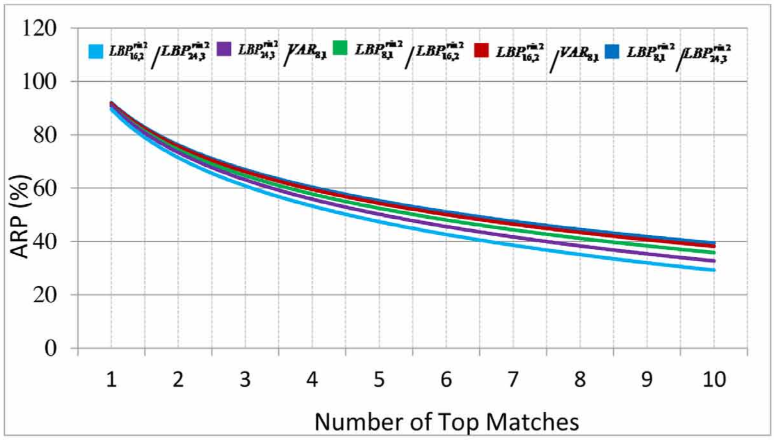
Figure 5 shows the retrieval performance of CBIR system with LBP/VAR operators in terms of ARP and ARR as a function of number of top matches and Figure 6 and 7 provides the snap shots of MIRS. The system performs better, if the required number of similar images are lesser. That is if the database has 25 images similar to query image and the retrieval system is requested for 10 similar images, then the system can retrieve 8-10 similar images. In the other hand if the system is asked for retrieving 25 images, then 13-15 images are retrieved. The system performance declines with increase in requirement. The system performance is little less consistent. The same can be improved by combining global features with LBP. Since our interest is to evaluate the performance of LBP features, the investigations of combining LBP with global features shall be considered as a separate study. The retrieval performance of hybrid approaches is determined in terms of precision and recall (ARP − ARR) curves. The comparison is performed in terms of ARP − ARR curves. The analysis is performed for OASIS-MRI dataset in which images of the same class show various variations. It is observed that among the hybrid LBP features, the superior performance is given by $LBP_{8,1}^{riu2} / LBP_{24,3}^{riu2}$. While observing the performance of combined LBP, we see that the proposed approach effectively overpowers other local and achieves a very high retrieval rate. It is apparent from the ARP − ARR curves.

SUMMARY

The work presented a theoretically and computationally simple multi resolution approach to gray scale and rotation invariant texture classification. A generalized gray scale and rotation invariant operator $LBP_{P,R}^{riu2}$ for detecting 'uniform' patterns in circular neighborhoods is employed to represent images. To experiment the rotation invariance, the classifier was trained at one particular rotation angle and tested with samples from other rotation angles.

Retrieval results proved the proposed approach is very robust in terms of gray scale variations caused e.g. by changes in illumination intensity, since the $LBP_{P,R}^{riu2}$ operator is by definition invariant against any monotonic transformation of the gray scale. This should make it very attractive in situations where non uniform illumination conditions are a concern, e.g. in visual inspection. Gray scale invariance is also necessary if the gray scale properties of the training and testing data are different. The performance can be further enhanced by multi resolution analysis, combining Global features such as WM, ART, PCTs etc. with LB and adopting advanced classification techniques involving Machine learning algorithms.

Figure 5. Performance of CBIR system with LBP/VAR operators in terms of ARP and ARR as a function of number of top matches



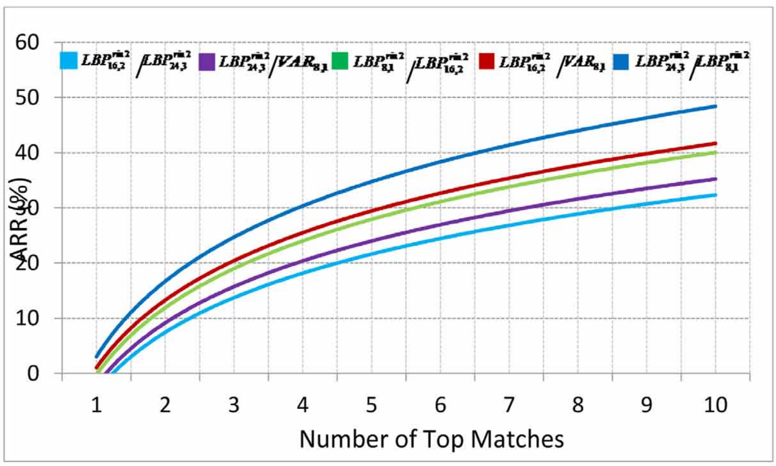


Figure 6. Medical Image Retrieval System with LBP/VAR operators for images from OASIS datasets

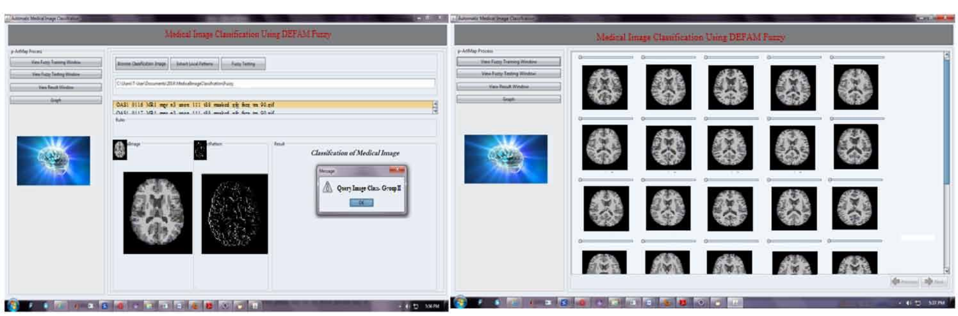
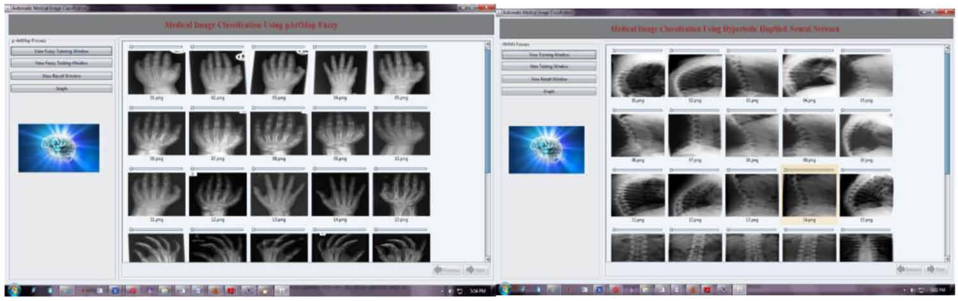


Figure 7. Snap shots for retrieval results with different classifiers on database of human parts



REFERENCES

Chang, S.-K., & Hsu, A. (1992). Image information systems: Where do we go from here? *IEEE Transactions on Knowledge and Data Engineering*, *4*(5), 431–442. doi:10.1109/69.166986

Dean Bidgood, W. Jr. (1998). The SNOMED DICOM Microglossary: Controlled terminology resource for Idata interchange in biomedical imaging. *Methods of Information in Medicine*, *37*(4/5), 404–414. doi:10.1055-0038-1634557 PMID:9865038

Gogic, I., Manhart, M., Pandzic, I. S., & Ahlberg, J. (2020). Fast facial expression recognition using local binary features and shallow neural networks. *The Visual Computer*, *36*(1), 97–112. doi:10.100700371-018-1585-8

Guld, Kohnen, Keysers, Schubert, Wein, Bredno, & Lehmann. (2002). Quality of DICOM header information for image categorization. *SER*, *Proc. SPIE*, 4685, 280-287. 10.1117/12.467017

Guo, Z., Zhang, L., & Zhang, D. (2010). Rotation invariant texture classification using LBP variance with global matching. *Pattern Recognition*, *43*(3), 706–716. doi:10.1016/j.patcog.2009.08.017

Hersh, W., Muller, H., & Kalpathy. (2009). The imageCLEFmed medical image retrieval task test collection. *J. Digital Imaging, 22*(6), 648-655.

Huang, H. K. (2004). PACS and imaging informatics: basic principles and applications. John Wiley & Sons Inc. doi:10.1002/0471654787

Hussain, C. A., Rao, D. V., & Mastani, S. A. (2020). *RetrieveNet: a novel deep network for medical image retrieval*. Evol. Intel. doi:10.100712065-020-00401-z

Krupinski. (2000). The Importance of Perception Research in Medical Imaging. *Radiation Medicine, 8*(6), 329-334.

Lehmann, T. M., Guld, M. O., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H., & Wein, B. B. (2004). Content-based image retrieval in medical applications. *Methods of Information in Medicine*, *43*(4), 354–361. doi:10.1055-0038-1633877 PMID:15472746

Long, L. R., Antani, S., Deserno, T. M., & Thoma, G. R. (2009). Content-Based Image Retrieval in Medicine: Retrospective Assessment, State of the Art, and Future Directions. *International Journal of Healthcare Information Systems and Informatics*, *4*(1), 1–16. doi:10.4018/jhisi.2009010101 PMID:20523757

Manjunath, K. N., Renuka, A., & Niranjan, U. C. (2007). Linear models of cumulative distribution function for content-based medical image retrieval. *Journal of Medical Systems*, *31*(6), 433–443. doi:10.100710916-007-9075-y PMID:18041275

Muller, Michoux, Bandon, & Geissbuhler. (2007). A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. *Intl. Journal of Medical Informatics, 73*(1), 1–23.

Nanni, L., Brahnam, S., & Lumini, A. (2012). A simple method for improving local binary patterns by considering non-uniform patterns. *Pattern Recognition*, *45*(10), 3844–3852. doi:10.1016/j.patcog.2012.04.007

Ojala, T., Pietikainen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, *29*(1), 51–59. doi:10.1016/0031-3203(95)00067-4

Papakostas, G. A., Koulouriotis, D. E., Karakasis, E. G., & Tourassis, V. D. (2013). Moment- based local binary patterns: A novel descriptor for invariant pattern recognition applications. *Neurocomputing*, *99*, 358–371. doi:10.1016/j.neucom.2012.06.031

Pereira, Fonseca-Pinto, Paiva, Tavora, Assuncao, & Faria (2020). *Accurate segmentation of desmoscopic images based on local binary pattern clustering*. International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia.

Seng & Mirisaee. (2009). Evaluation of a content-based retrieval system for blood cell images with automated methods. *Journal of Medical Systems*, *35*, 571–578.

Sorensen, Shaker, & Bruijne. (2010). Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Trans. Med. Imaging, 29*(2), 559-569.

Tan, X., & Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, *19*(6), 1635–1650. doi:10.1109/TIP.2010.2042645 PMID:20172829

Xue, Long, & Antani, Jeronimo, & Thoma. (2008). A Web-accessible content-based cervicographic image retrieval system. *Proceedings of the Society for Photo-Instrumentation Engineers*, 6919.

Xueming, Hua, Chen, & Liangjun. (2011). PLBP: An effective local binary patterns texture descriptor with pyramid representation. *Pattern Recognition*, *44*, 2502–2515.

Yao, C.-H., & Chen, S.-Y. (2003). Retrieval of translated, rotated and scaled color textures. *Pattern Recognition*, *36*(4), 913–929. doi:10.1016/S0031-3203(02)00124-3

Yurdakul, Subathra, & Georgec. (2020). Detection of Parkinson's Disease from gait using Neighborhood Representation Local Binary Patterns. *Biomedical Signal Processing and Control, 62*.

Zakeri, F. S., Behnam, H., & Ahmadinejad, N. (2010). Classification of benign and malignant breast masses based on shape and texture features in sonography images. *Journal of Medical Systems*, *36*(3), 1621–1627. doi:10.100710916-010-9624-7 PMID:21082222

Zhang, B., Gao, Y., Zhao, S., & Liu, J. (2010). Local derivative pattern versus local binary pattern: Face recognition with higher-order local pattern descriptor. *IEEE Transactions on Image Processing*, *19*(2), 533–544.

Zhang, D., Yang, G., & Li, F. (2020). Detecting seam carved images using uniform local binary patterns. *Multimedia Tools and Applications*, *79*, 8415–8430. doi:10.100711042-018-6470

Chapter 8
Correlation and Analysis of Overlapping Leukocytes in Blood Cell Images Using Intracellular Markers and Colocalization Operation

Balanagireddy G.
Rajiv Gandhi University of Knowledge Technologies, India & Dr. A. P. J. Abdul Kalam Technical University, Ongole, India

Ananthajothi K.
https://orcid.org/0000-0002-6390-2082
Misrimal Navajee Munoth Jain Engineering College, India

Ganesh Babu T. R.
Muthayammal Engineering College, India

Sudha V.
Sona College of Technology, India

ABSTRACT

This chapter contributes to the study of uncertainty of signal dimensions within a microscopic image of blood sample. Appropriate colocalization indicator classifies the leukocytes in the region of interest having ragged boundaries. Signal transduction has been interpreted using correlation function determined fluorescence intensity in proposed work using just another colocalization plugin (JaCoP). Dependence between

DOI: 10.4018/978-1-7998-3092-4.ch008

the channels in the colocalization region is being analysed in a linear fashion using Pearson correlation coefficient. Manders split, which gives intensity, is represented in a channel by co-localizing pixels. Overlap coefficients are also being analysed to analyse coefficient of each channel. Li's intensity correlation coefficient is being used in specific cases to interpret the impact of staining.

1. INTRODUCTION

Blood sample image usually contains the following: erythrocytes (RBC), leukocytes (WBC) and platelets. The major classification on "White blood cells" are denoted as follows, "Neutrophils", "Eosinophils", "Monocytes" and "Lymphocytes". Each of these subtype cells contributes to the usefulness in body defence. Hence each subtype of cells are taken and classified as imaging dimensions in accordance to its shape. The main limitation behind the four subtypes of leukocytes is that, if they are clumped together they may reduce the accuracy of classification. The purpose of this research is to interpret the biological relevance between specific classes of leukocytes using colocalization procedures. Spatial point characteristics are visually evaluated to provide protrusion of cells associated with the region of interest. The chapter is organised as follows section 2 deals with the previous research of processing blood samples using image processing. Section 3, discusses algorithm for segmenting the class of leukocytes with JaCoP. Section 4 discusses result of medical image segmentation using colocalization method. Section 5 concludes the overall work.

2. LITERATURE SURVEY

"Immunohistochemical" slide image registration accuracy depends on the field of view of few cells. Registration accuracy is achieved with biomarker colocalization using an elastix framework based on dynamic resolution levels (Moles Lopez et al., 2015). The classification of leukocytes based on their shape and lobes of nucleus is given as follows. It can be mononuclear which includes "Monocytes" and "Lymphocytes". The other contains granules named as granulocytes, which includes Neutrophils and Eosinophils. The extraction of WBC cells from blood samples is followed by separation of cytoplasm and nucleus thereby further enhanced classification has been done in (Putzu et al., 2014). Marker controlled watershed has been used for segmentation based on cell nucleus. Subsequently, classification has been done to separate WBC and RBC (Miao & Xiao, 2018). However, misclassification may result in improper movement of WBC leading to cell adhesion.

A review of colocalization techniques has been discussed with "Manders co-occurrence" (MOC) which considered pixel intensity (Manders et al., 1993). The limitation of MOC is that, it can be affected by useless signal. Similarly, Pearsons Correlation Coefficient (PCC) based on interdependency may result that depended on threshold values (Aaron et al., 2018). Localization using colour transformation methods and subsequent segmentation of cytoplasm via region growing, watershed and finally classification is done using support vector machine has been done using bone marrow images (Liu et al., 2019). Leukocyte nuclei segmentation using channel splitting with blue and green channel can reduce the time incurred to analyse the blood cell image (Wang & Cao, 2019). Both spatial and spectral features for segmenting the nucleus and leucktocyte region was discussed. Spatial feature used morphological operations, whereas spectral features used support vector machine (Duan et al., 2019). Feature classifying benign and malignant cells incorporating convolutional neural network and statistical property has been discussed with Salp Swarm Algorithm (Sahlol et al., 2020). But the work included, most features which are relevant and excludes noise. Arbitrary shape of biomedical image has been done using colocalization in a cell of interest with dilated radius(CIRCOAST) (Corliss et al., 2019). Molecular interaction of irregular shape has been discussed using Geo-coPositioning system. It combines an object based method to intensity based method. Object based method provides resilience against noise and content of fluorescence signal is estimated using intensity based method (Lavancier et al., 2020). Biological processing with "Digital Lensless holographic microscopy" states that ImageJ plugin has been used for interoperability for calculating numerical simulation (Trujillo et al., 2020). Hence this work used ImageJ with JaCoP plugins for analysis.

Semantic segmentation has been used initially after pre-processing the blood cell images the pixel level features are extracted using deep convolutional encoder and decoder. The accuracy in classifying the RBC, WBC and platelets are better using Intersection of Union and Boundary Score (Shahzad et al., 2020).

2.2 Problem Definition

Fluorescence spectra are not separated well in image acquisition. Misalignment of signal with imperfect representation of leukocytes cells with nucleus and cytoplasm along with its overlap can result in various quantitative values or volumetric changes leading to wrong diagnosis. Thus colocalization of individual images has to be done manually, to interpret the intracellular markers.

3. PROPOSED SYSTEM

3.1 Assumptions

The assumption is made such that the background is excluded and the region of interest in an overlapping area, marked in leukocytes is taken for colocalization.

3.2 Algorithm

Step 1. Convert the given image into its corresponding colour across the red, green and blue channel.

Step 2. Depending on the hue values through visual inspection the corresponding green or red channel is taken for processing the leukocyte region.

Step 3. Segment the leukocyte region of interest manually by adjusting the wand tool using appropriate threshold values. Crop this region of interest and display as separate image.

Step 4. Segment different level of nuclei and cytoplasm inside the leukocyte region from the cropped image.

Step 5. Correlation coefficients for interpreting the region of interest in the images are Pearson's correlation coefficient, Manderco localization coefficient, Overlap coefficient. In addition,

Li's intensity correlation coefficient is used in certain images.

Step 6. Coste's mask (Costes et al., 2004) is being used with channel intensity as in equation 1 and equation 2.

$$I_1 = C + ROI_1 \tag{1}$$

$$I_2 = \alpha + C + ROI_2 \tag{2}$$

Notations used are denoted as follows, I_1 and I_2 denote the channel intensity of the image 1 and image 2 respectively. Colocalized component is denoted by C, random component in image 1 is ROI_1 and random component in image 2 is ROI_2 and α is "stoichiometry coefficient".

4. RESULTS AND DISCUSSION

The Blood cell image has been obtained from (Mooney, 2018). Correlation functions of images for colocalization has been analysed using "overlap coefficient", "Mander coefficient", "Pearson coefficient" and "Li's intensity correlation coefficient"(Bolte & Cordelières, 2006). The acquired images are in JPEG format but are converted into dtif for colocalization procedures. The darkest purple colour represents the nucleus surrounded by light purple denoting the cytoplasm. Erythrocytes are being represented in pink colour.

Table 1. Description of leukocytes cells used in this work.

Type of WBC Cells	Number of Images	Image Size
Neutrophils	Figure 1a, Input image is taken from (Mooney, 2018).	Width 640 Height 480 Depth 3
Eosinophils	Figure 9a, Input image taken from (Mooney, 2018).	Width 640 Height 480 Depth 3
Monocytes	Figure 13a, Input image taken from (Mooney, 2018).	Width 640 Height 480 Depth 3
Lymphocytes	Figure 16a, Input image taken from (Mooney, 2018).	Width 640 Height 480 Depth 3

A high amount of saturationis present in green channel in the image used for analysis in the first image work. Figure 2 (a) shows the whole image of a leukocyte region and the (b), (c) and (d) represents nucleus area.

The whole region of the cell is being correlated with its sub-region nucleus. There is no colocalization that exhibits between the cells. The threshold valueis taken for first ROI and second ROI is being set to116.

The figure 3 (a), shows the whole cell where nucleus is in graycolor and cytoplasm in red color. Thefigure 3 (b), sub region of interest where graycoloris nucleus and cytoplasm in red color with yellow boundaries.

The whole region of the cell is being correlated with its sub-region of nucleus. There is no colocalization that exhibits between these cells. The threshold value taken for First ROI is 121 and second ROI is set as 116.

The figure 4 shows Coste's automatic threshold using whole area threshold as 160 and threshold (c) as 123.

Figure 1. (a) Neutrophil identified as purple colourimage in RGB format. (b) Neutrophil image with green channel.(c) Neutrophil image with red channel.(b) Neutrophil image with blue channel.

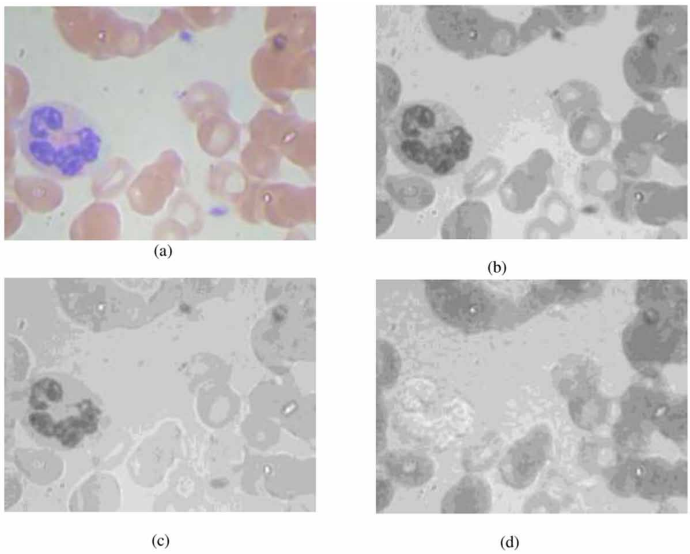


Figure 2. (a) First Region of Interest (ROI) withwhole cell within Neutrophil image ingreen channel. (b) Second Region of Interest (ROI) Neutrophil image with green channel.(c)Third Region of Interest (ROI) Neutrophil image with green channel. (d) Forth Region of Interest (ROI) Neutrophil image with green channel.

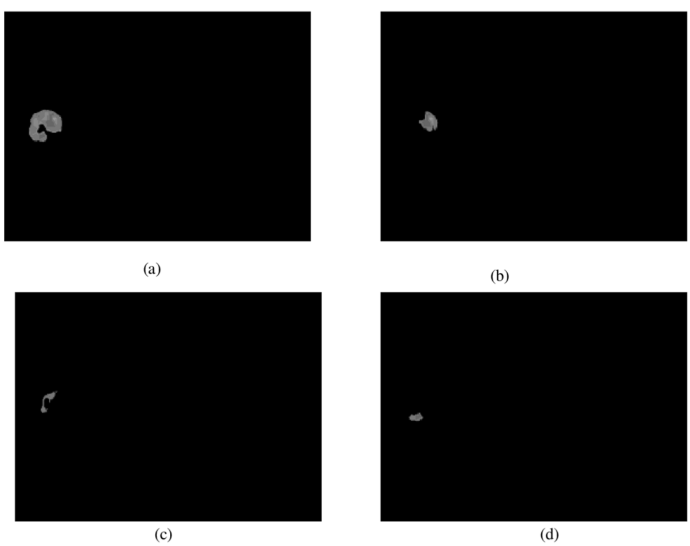
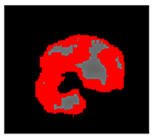


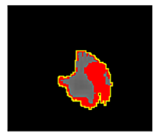
Table 2. Denotes the correlation coefficient used for figure 2.

Pearson coefficient	0.549
Overlap coefficient	0.551
Mander coefficient (Image 2 (b) overlapping with 2 (a) fraction	0.318
Mander coefficient (Image 2 (a) overlapping with 2(b) fraction using threshold values	0.226

Figure 3.



(a)



(b)

Table 3. Denotes the correlation coefficient used for figure 2(a) and 2 (c).

Pearson coefficient	0.368
Overlap coefficient	0.37
Mander coefficient (orginal)	0.143
Mander coefficient (Image 2 (a) overlapping with 2(c) fraction using threshold value)	0.051
Mander coefficient (Image 2 (c) overlapping with 2(a) fraction using threshold value)	0.292

Figure 4. Coste's mask for ROI-3 over whole cell area.

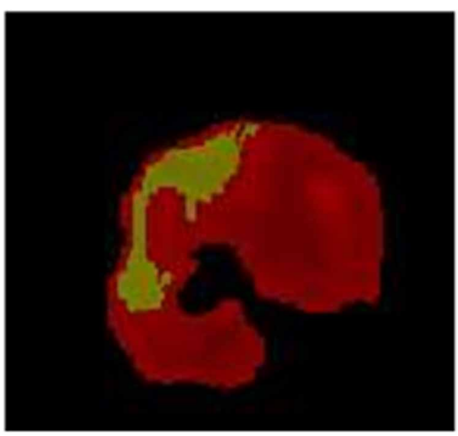


Table 4. Denotes the correlation coefficient used for figure 2 (a) and 2(d).

Pearson coefficient	0.318
Overlap coefficient	0.32
Mander coefficient (original)	0.108
Mander coefficient (Image 2 (a) overlapping with 2(d) fraction using threshold value)	0.057
Mander coefficient (Image 2 (d) overlapping with 2(a) fraction using threshold value)	0.481

The figure 5, shows Coste's automatic threshold using whole area threshold as 160 and threshold as 125.

The whole region of the cell is being correlated with its sub-region of nucleus. There is one colocalization that exhibits between these cells. The threshold value taken for First ROI is 118 and second ROI is set as 104.

The figure 7 shows Costes automatic threshold using whole area threshold as 143 for image 6(a) and 120 threshold for image 6 (c).

The whole region of the cell is being correlated with its sub-region of nucleus. There is one colocalization that exhibits between these cells. The threshold value taken for First ROI is 118 and second ROI is set as 104.

The figure 8 shows Costes automatic threshold using whole area threshold as 143 for image 6(a) and 128 threshold for image 6 (b).

The whole region of the cell is being correlated with its sub-region of nucleus.

The whole region of the cell is being correlated with its sub-region of nucleus. There is minimal colocalization that exhibits between these cells. The threshold value taken for First ROI is94 and second ROI is set as 85.

The figure 11 shows Costes automatic threshold using whole area threshold as 94 for image 10(a) and 85 threshold for image 10(b).

The whole region of the cell is being correlated with its sub-region of nucleus. There is minimal colocalization that exhibits between these cells. The threshold value taken for First ROI is116 and second ROI is set as 97.

The figure 12 shows Costes automatic threshold using whole area threshold as 116 for image 10(a) and 97 threshold for image 10(c).

The figure 15 shows Costes automatic threshold using whole area threshold as 111 for image 14(a) and 102 threshold for image 14(b).

The whole region of the cell is being correlated with its sub-region of nucleus. There is minimal colocalization that exhibits between these cells. The threshold value taken for First ROI is111 and second ROI is set as 102.

Figure 5. Costes mask for ROI-4 over whole cell area.

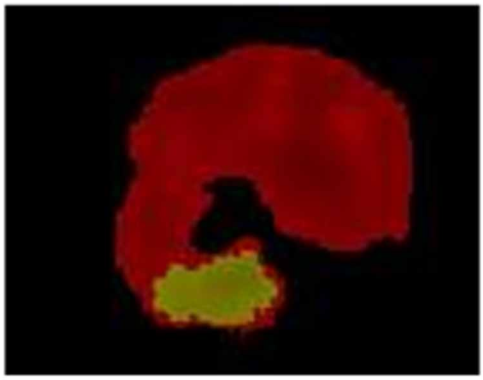
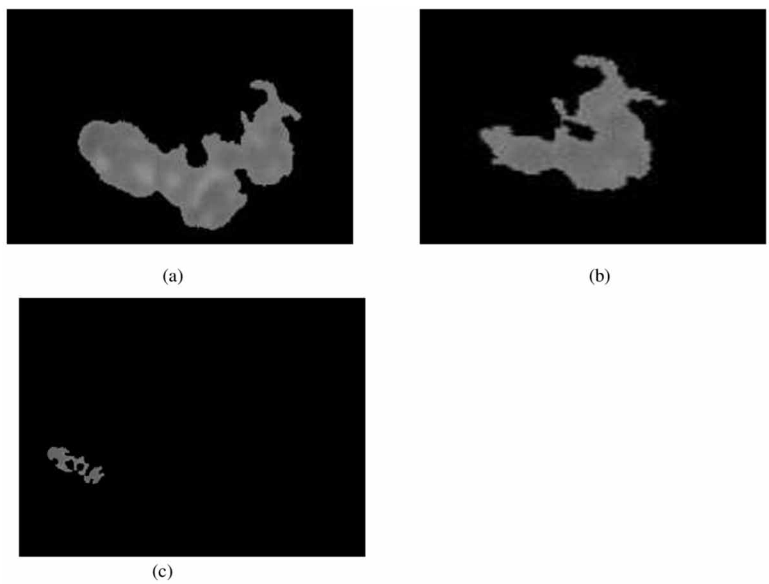


Figure 6. (a) First Region of Interest (ROI) withwhole cell within Neutrophil image in green channel. (b) Second Region of Interest (ROI) Neutrophil image ingreen channel.(c) Third Region of Interest (ROI) Neutrophil image ingreen channel.



The whole region of the cell is being correlated with its sub-region of nucleus. There is one colocalization that exhibits between these cells. The threshold value taken for First ROI is 98 and second ROI is set as 72.The correlation coefficient values used for figure 17(a) and 17(b) is denoted in table 10.

Table 5. Denotes the correlation coefficient used for figure 6 (a) and 6(c).

Pearson coefficient	0.61
Overlap coefficient	0.614
Mander coefficient (orginal)	0.397
Mander coefficient (Image 6 (a) overlapping with 6(c) fraction using threshold value)	0.063
Mander coefficient (Image 6 (c) overlapping with 6(a) fraction using threshold value)	0.081

Figure 7. Costes mask for image 6(a) with ROI-(a) and ROI (c) over whole cell area.

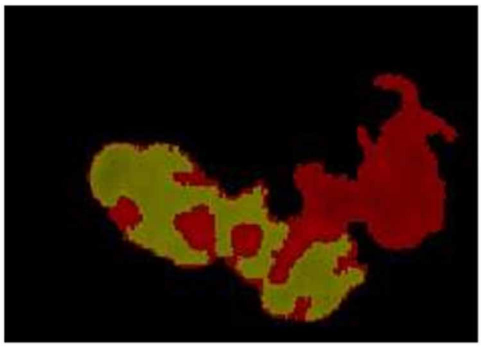


Table 6. Denotes the correlation coefficient used for figure 6 (a) and 6(b).

Pearson coefficient	0.548
Overlap coefficient	0.551
Mander coefficient (original)	0.307
Mander coefficient (Image 6 (a) overlapping with 6(b) fraction using threshold value)	0.305
Mander coefficient (Image 6 (b) overlapping with 6(a) fraction using threshold value)	0.91

Figure 8. Costes mask for image 6(a) with ROI-(a) and ROI (b) over whole cell area.

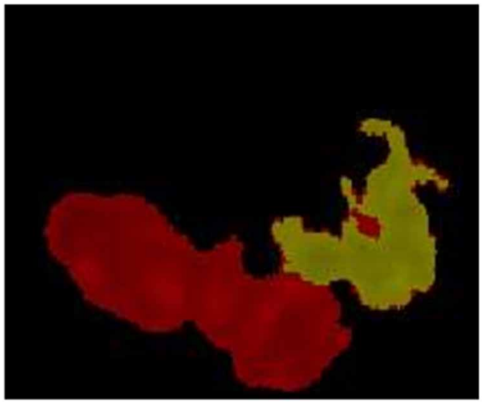


Figure 9. (a) Eosinophils identified in purple color representation in RGB image. (b) Eosinophils image with green channel.(c) Eosinophils image with red channel. (d) Eosinophils image with blue channel.

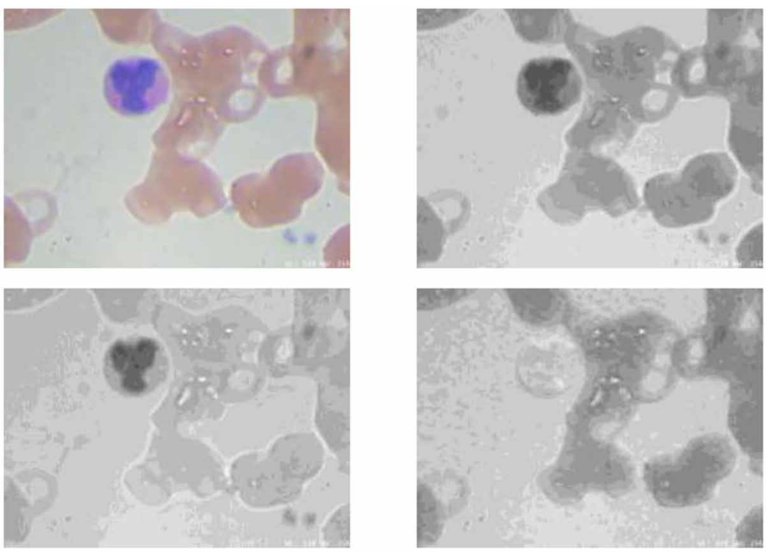


Figure 10. (a) First Region of Interest (ROI) withwhole cell within Eosinophil image withgreen channel. (b) Second Region of Interest (ROI) Eosinophil image with green channel.(c) Third Region of Interest (ROI) Eosinophil image with green channel.

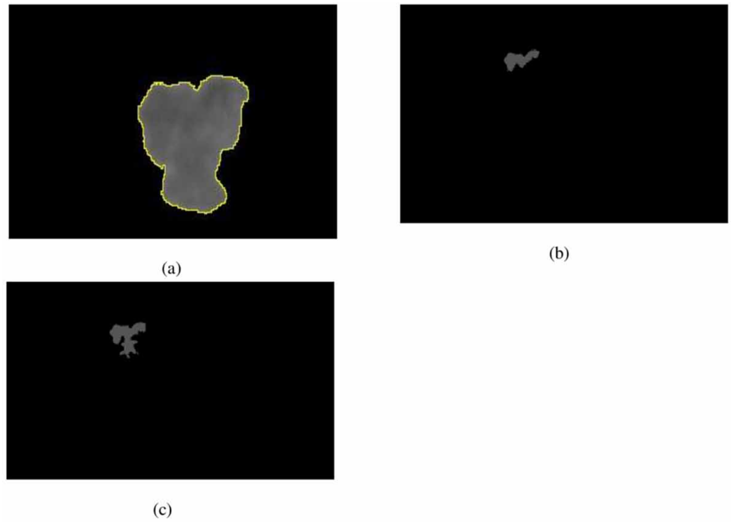


Table 7. Denotes the correlation coefficient used for figure 10(a) and 10(b).

Pearson coefficient	0.548
Overlap coefficient	0.551
Mander coefficient (original)	0.307
Li's intensity correlation coefficient	0.487

Figure 11. a, Costes mask for image 10 with ROI-(a)over whole cell area and ROI (b) with nucleus.

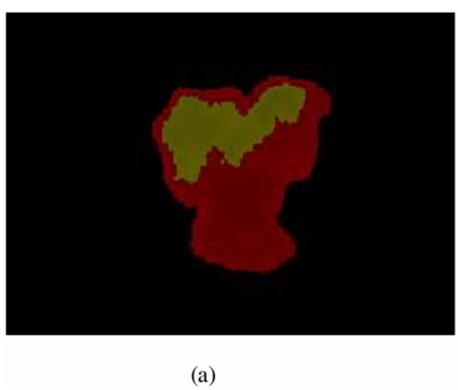


Table 8. Denotes the correlation coefficient used for figure 10(a) and 10(c).

Pearson coefficient	0.671
Overlap coefficient	0.675
Mander coefficient (original)	0.49
Li's intensity correlation coefficient	0.491

Figure 12. a, Costes mask for image 10 with ROI-(a)over whole cell area and ROI (c) with nucleus.

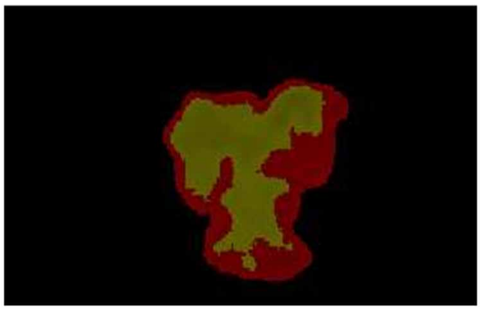
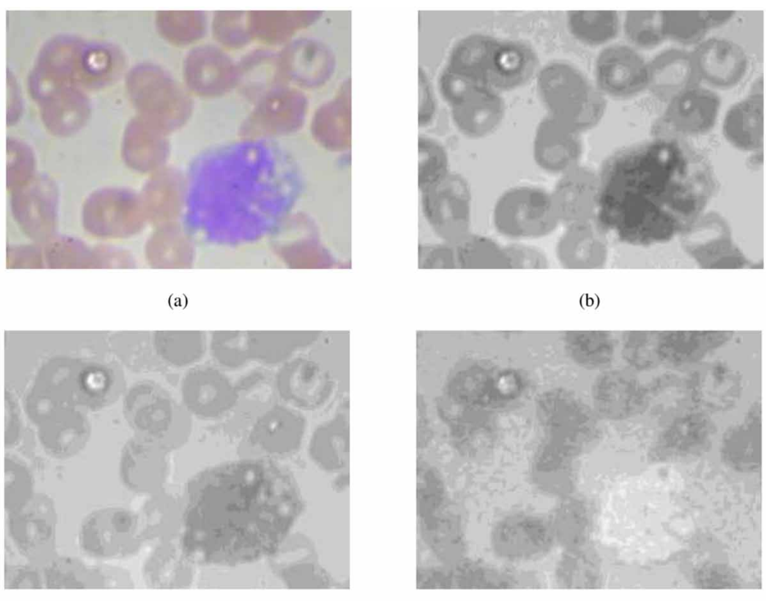


Figure 13. (a) Monocytesimagein RGB image. (b) Monocytes image with green channel.(c) Monocytes image with red channel.(d) Monocytes image with blue channel.



(a) (b) (c) (d)

Figure 14. (a) First Region of Interest (ROI) withwhole cell of Monocytesimage in green channel. (b) Second Region of Interest (ROI) sub cell of Monocytesimage in green channel.

(a) (b)

Figure 15. a, Costes mask for image 14 with ROI-(a)over whole cell area and ROI (b) with nucleus.

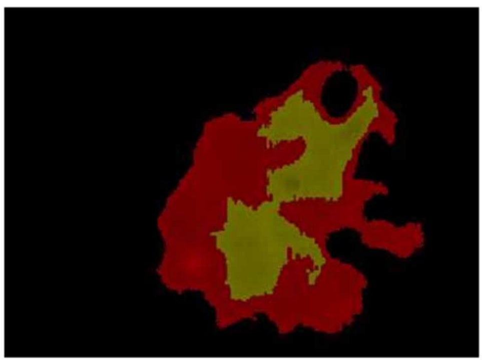


Table 9. Denotes the correlation coefficient used for figure 14(a) and 14(b).

Pearson coefficient	0.507
Overlap coefficient	0.52
Mander coefficient (original)	0.293
Li's intensity correlation coefficient	0.460

Figure 16. (a) Lymphocyteidentified as whole image in RGB format. (b) Lymphocyteimage with green channel.(c) Lymphocyteimage with red channel.(d) Lymphocyteimage with blue channel.

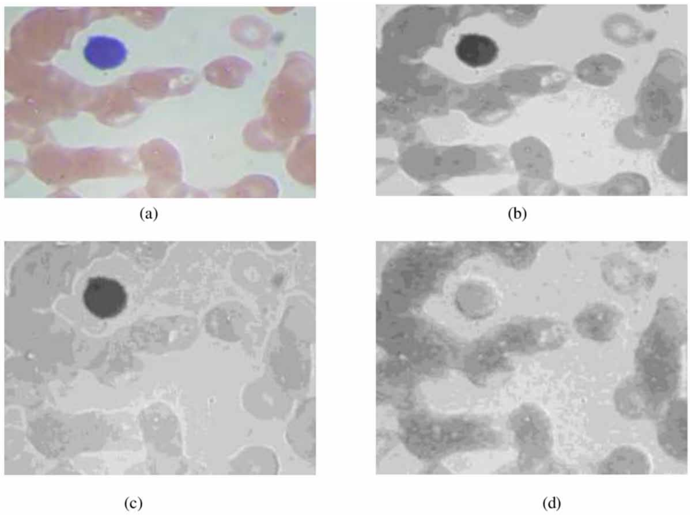


Figure 17. (a) First Region of Interest (ROI) Lymphocyteidentified as whole imagein green channel. (b) Second Region of Interest (ROI) sub cell of Lymphocyteimage in green channel.

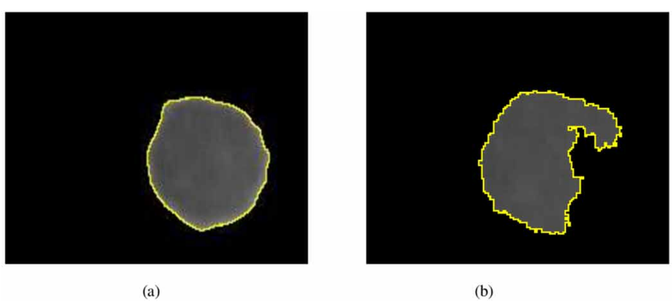
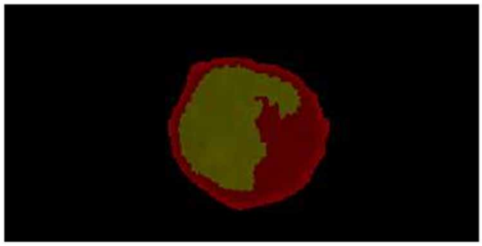


Table 10. Denotes the correlation coefficient used for figure 17(a) and 17(b).

Pearson coefficient	0.625
Overlap coefficient	0.629
Mander coefficient (original)	0.45
Li's intensity correlation coefficient	0.491

Figure 18. a, Costes mask for image 17 with ROI-(a) over whole cell area and ROI (b) with nucleus.



5. CONCLUSION

This research contributes to the colocalization of overlapping leukocytes and its calibration with correlation function. Thus individual leukocytes are analysed with co-occurrence .and clustered leukocytes and are visually inspected and their impact of correlation is being analysed. Larger cells of leukocytes and ragged boundaries which navigate through complex tissues are being examined visually via correlating coefficients. The limitation behind this approach taken depends on the brightness of computer screen which cannot be taken as reliable marker for medical diagnosis. Future work will deal considering the impact of bio-molecules using colocalizationtechniques using over restoration of nucleus in WBC cells.

REFERENCES

Aaron, J. S., Taylor, A. B., & Chew, T. L. (2018). Image co-localization–co-occurrence versus correlation. *Journal of Cell Science*, *131*(3), jcs211847. doi:10.1242/jcs.211847 PMID:29439158

Bolte, S., & Cordelières, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy*, *224*(3), 213–232. doi:10.1111/j.1365-2818.2006.01706.x PMID:17210054

Corliss, B. A., Ray, H. C., Patrie, J. T., Mansour, J., Kesting, S., Park, J. H., & Peirce, S. M. (2019). CIRCOAST: A statistical hypothesis test for cellular colocalization with network structures. *Bioinformatics (Oxford, England)*, *35*(3), 506–514. doi:10.1093/bioinformatics/bty638 PMID:30032263

Costes, S. V., Daelemans, D., Cho, E. H., Dobbin, Z., Pavlakis, G., & Lockett, S. (2004). Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical Journal*, *86*(6), 3993–4003. doi:10.1529/biophysj.103.038422 PMID:15189895

Duan, Y., Wang, J., Hu, M., Zhou, M., Li, Q., Sun, L., & Wang, Y. (2019). Leukocyte classification based on spatial and spectral features of microscopic hyperspectral images. *Optics & Laser Technology*, *112*, 530–538. doi:10.1016/j.optlastec.2018.11.057

Lavancier, F., Pécot, T., Zengzhen, L., & Kervrann, C. (2020). Testing independence between two random sets for the analysis of colocalization in bioimaging. *Biometrics*, *76*(1), 36–46. doi:10.1111/biom.13115 PMID:31271216

Liu, H., Cao, H., & Song, E. (2019). Bone marrow cells detection: A technique for the microscopic image analysis. *Journal of Medical Systems*, *43*(4), 82. doi:10.100710916-019-1185-9 PMID:30798374

Manders, E. M. M., Verbeek, F. J., & Aten, J. A. (1993). Measurement of co-localization of objects in dual colour confocal images. *Journal of Microscopy*, *169*(3), 375–382. doi:10.1111/j.1365-2818.1993.tb03313.x

Miao, H., & Xiao, C. (2018). Simultaneous segmentation of leukocyte and erythrocyte in microscopic images using a marker-controlled watershed algorithm. *Computational and Mathematical Methods in Medicine*, *2018*, 2018. doi:10.1155/2018/7235795 PMID:29681997

Moles Lopez, X., Barbot, P., Van Eycke, Y. R., Verset, L., Trépant, A. L., Larbanoix, L., & Decaestecker, C. (2015). Registration of whole immunohistochemical slide images: An efficient way to characterize biomarker colocalization. *Journal of the American Medical Informatics Association: JAMIA*, *22*(1), 86–99. doi:10.1136/amiajnl-2014-002710 PMID:25125687

Mooney. (2018 April). *Blood cell images Version 6*. Retrieved May 23 2020 from https://www.kaggle.com/paultimothymooney/blood-cells

Putzu, L., Caocci, G., & Di Ruberto, C. (2014). Leucocyte classification for leukaemia detection using image processing techniques. *Artificial Intelligence in Medicine*, *62*(3), 179–191. doi:10.1016/j.artmed.2014.09.002 PMID:25241903

Sahlol, A. T., Kollmannsberger, P., & Ewees, A. A. (2020). Efficient classification of white blood cell leukemia with improved Swarm optimization of deep features. *Scientific Reports*, *10*(1), 1–11. doi:10.103841598-020-59215-9 PMID:32054876

Shahzad, M., Umar, A. I., Khan, M. A., Shirazi, S. H., Khan, Z., & Yousaf, W. (2020). Robust Method for Semantic Segmentation of Whole-Slide Blood Cell Microscopic Images. *Computational and Mathematical Methods in Medicine*, *2020*, 2020. doi:10.1155/2020/4015323 PMID:32411282

Trujillo, C., Piedrahita-Quintero, P., & Garcia-Sucerquia, J. (2020). Digital lensless holographic microscopy: Numerical simulation and reconstruction with ImageJ. *Applied Optics*, *59*(19), 5788–5795. doi:10.1364/AO.395672 PMID:32609706

Wang, Y., & Cao, Y. (2019). Quick leukocyte nucleus segmentation in leukocyte counting. *Computational and Mathematical Methods in Medicine*, *2019*, 2019. doi:10.1155/2019/3072498 PMID:31308855

Chapter 9
Enchodroma Tumor Detection From MRI Images Using SVM Classifier

G. Durgadevi
New Prince Shri Bhavani College of Engineering and Technology, India

K. Sujatha
Dr. M. G. R. Educational and Research Institute, India

K.S. Thivya
Dr. M.G.R. Educational and Research Institute, India

S. Elakkiya
Dr. M.G.R. Educational and Research Institute, India

M. Anand
Dr. M.G.R. Educational and Research Institute, India

S. Shobana
New Prince Shri Bhavani College of Engineering and Technology, India

ABSTRACT

Magnetic resonance imaging is a standard modality used in medicine for bone diagnosis and treatment. It offers the advantage to be a non-invasive technique that enables the analysis of bone tissues. The early detection of tumor in the bone leads on saving the patients' life through proper care. The accurate detection of tumor in the MRI scans are very easy to perform. Furthermore, the tumor detection in an image is useful not only for medical experts, but also for other purposes like segmentation and 3D reconstruction. The manual delineation and visual inspection will be limited to avoid time consumption by medical doctors. The bone tumor tissue detection allows localizing a mass of abnormal cells in a slice of magnetic resonance (MR).

DOI: 10.4018/978-1-7998-3092-4.ch009

INTRODUCTION

Medical image processing is an important field of research as its outcomes are used for the betterment of health issues. A tumor is an abnormal growth of tissues in any part of the body. As the tumor grows, the abnormal tissue displaces healthy tissue. There is a large class of bone tumor types which have different characteristics. There are two types of bone tumors, Noncancerous (Benign) and Cancerous (Malignant). The benign tumor grows very large and press on nearby tissues, once removed by surgery, they don't usually reoccur. Malignant tumor has a larger nucleus that looks different from a normal cell's nucleus and can also reoccur after they are removed. Hence care as to be taken in order to completely avoid tumors. There are different image modalities like X-ray, MRI, CT, PET SCANS has shown in figure 1.1. The MRI imaging technique is the best because it has a higher resolution. Magnetic resonance imaging (MRI) is a non-invasive medical system used to show 2D images of the body. This technique is based on a process that uses highly charged magnetic fields and radio waves to make images of the inside the body. It is an unharmed method of obtaining images of the human body. Its data are most relevant and it helps in early detection of tumors and precise estimation of tumor boundaries. Magnetic resonance (MR) sequences such as T1-weighted, T2-weighted, contrast-enhanced T1W and T2W, STIR (Short T1 inversion recovery), PD-Weighted series provide different information. Thus MRI scans have a best non-invasive medical systems used to show 2D images of the body. This technique is based on a process that used highly charged magnetic fields to make images of the body. Hence MRI has more than one methodology to classify images. These are atlas methods, shape methods, fuzzy methods, and variations methods. New technology MRI are T1 weighted, T2 weighted and proton density weighted images.

The rest of the paper includes section 2 gives the brief glimpse of the relevant work that was carried out all in the various fields of research. Section 3explains segmentation process -thresholding and morphological operations. Section 4 includes the proposed method with results and experimental results. Section 5 includes the conclusions followed by future enhancements.

REVIEW OF LITERATURE

Sinan Onal *et* al. (Onal et al., 2014) proposed a method of automatic localization of multiple pelvic bone structure on MRI, and they have used an SVM classification and nonlinear regression model with global and local information, and they are presented to automatically localize multiple pelvic bone Durgadevi *et* al. (Durgadevi & Shekhar, 2015) proposed a method of Identification of tumour using k-means algorithm.

The identification of breast cancer from the MRI images is made automatic using K-means clustering and wavelet transform. The human perception at many times may lead to erroneous diagnosis. Variation in diagnosis may produce adverse effect on the patients. Hence to improve the accuracy this system is made automatic using machine vision algorithms. Alan Jose *et* al. (Emran et al., 2015) described Brain Tumour Segmentation Using K-Means Clustering and Fuzzy C-Means Algorithms and Its Area Calculation; they have given Simple Algorithm for detection of range and shape of tumour in brain MR Images. Normally the anatomy of the Brain can be viewed by the MRI. MRI scanned image is used for the entire process. The MRI scan is more comfortable than any other scans for diagnosis Deepak *et* al. (Jose et al., 2014) discussed Comparative Study of Tumour Detection Techniques with their Suitability for Brain MRI Images The canny edge detection technique defines edges of the MRI image by using many parameter like thresholding, thinning etc. canny with morphological operation like dilation, erosion etc., where simply applied on it for getting better results, and fuzzy c-means method gives best results for segmentation of Brain tumour in MRI images. M. Koch, *et* al,(Bagahel & Kiran, 2015) described automatically segmenting the wrist bones in the arthritis patients using the k-means clustering process.

Figure 1. MRI SCAN

PROPOSED SYSTEM

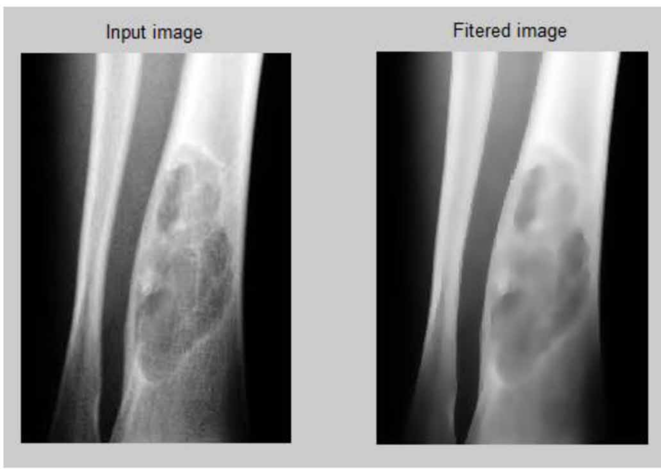
Pre-processing is a basic step, which aims to improve the quality of an image by removing the noise artefacts and increasing contrast. some pre-processing techniques are image acquisition, histogram equalization filtering. Filtering is the basic step among all. Filtering reduces the noise, saves the edges and smooth improper images produced by the MRI. Figures 1.2 shows images with enchondroma.

We chose the bilateral filter. The bilateral filter is a filter for sharping and for using nonlinear images. Since we use SVM classifier along with the bi lateral used to detect the nonlinear images the bi lateral filter is also used to reduce the noise over the images and also gives the contrast to the image. Figure 3 shows the bilateral output.

Figure 2. Enchondroma images

The input image is taken as a noisy one and then pre-processing is carried out in previous step then the boundaries have to be calculated with the help of the thresholding and morphological operation. Segmentation is the process of dividing the images into regions. We are using two segmentation process thresholding and morphological operations. Thresholding is the most common process, which divides the images into a binary images.it is very productivity with images that have high resolution. Figure 4 and 5 shows the results of thresholding and morphological operations. Thresholding alone is not enough because in some cases it may case false segments. That is false images. So, dilation and erosion is used. Which means shrink and expand. Both the operations are helpful to detect the tumor area.

Figure 3. preprocessing technique



In Proposed, SVM Classifier used to segment the cancer detected portion. To segment the portion, first have to filter out the acquired image based upon the masking methodology. The Morphological function including dilation and erosion method will be applied extracted throughout the filtered image. By the method of morphological bounding box will be drawn over the affected portion. Then, the region enclosed by bounding box will be splitted out separately with the SVM Classifier.

The block diagram in figure 6 explains the various blocks to execute the segmentation process. In thatInput Bone tumor image is given as input in this module. That is taking the affected tumor part from the MRI images and processing In this stage we used two method of segmentation, they are morphological operation and thresholding segmentation. Thresholding alone is not enough because in most cases the images have artefacts or false segments.so erosion and dilation operations are carried out to expand and shrink the region of interest

EXPERIMENTAL RESULTS

The following screenshot explains the classification done by SVM classifier and the tumor region is separated fig 1.7shows the SVM classifier is used to classify the segmented region, where as figure 1.8 shows the classified tumor detection.

Figure 4. Thresholding segmentation

CONCLUSION

The detection of Bone Tumour from the MRI images is taken away, and the images that do not have a tumour or an unrelated images requires two main steps, namely pre-processing and segmentation. In the pre-processing step, bilateral filters smooth the images and remove the noise. We combine two segmentation along with it namely morphological as well thresholding process. With various process. SVM classifier is used to detect the tumor portions and finally segmentation is carried out. With the SVM classifier the tumor regions are separated from the non-tumor sections. The detection of enchondroma tumor from the MRI images using SVM classifier is carried out along with the pre-processing techniques and segmentations process. We develop an application to assess the performance of the proposed method via MATLAB 2015R and in advanced.

Figure 5. Morphological operation

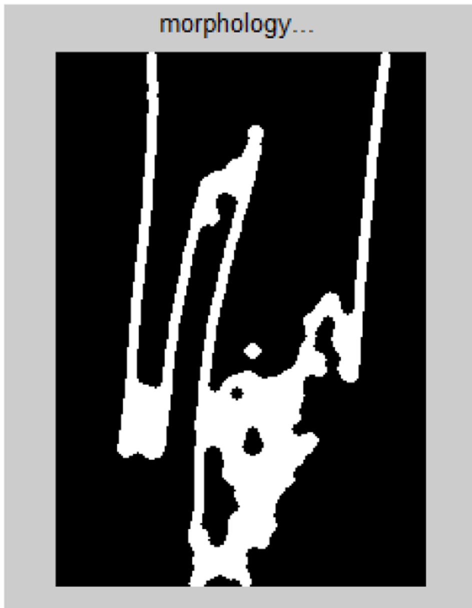


Figure 6. Proposed architecture

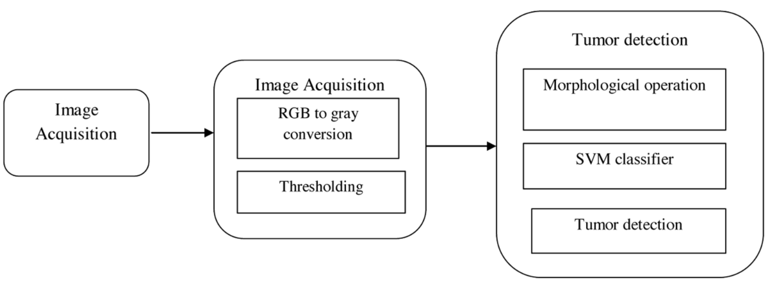


Figure 7. Classification

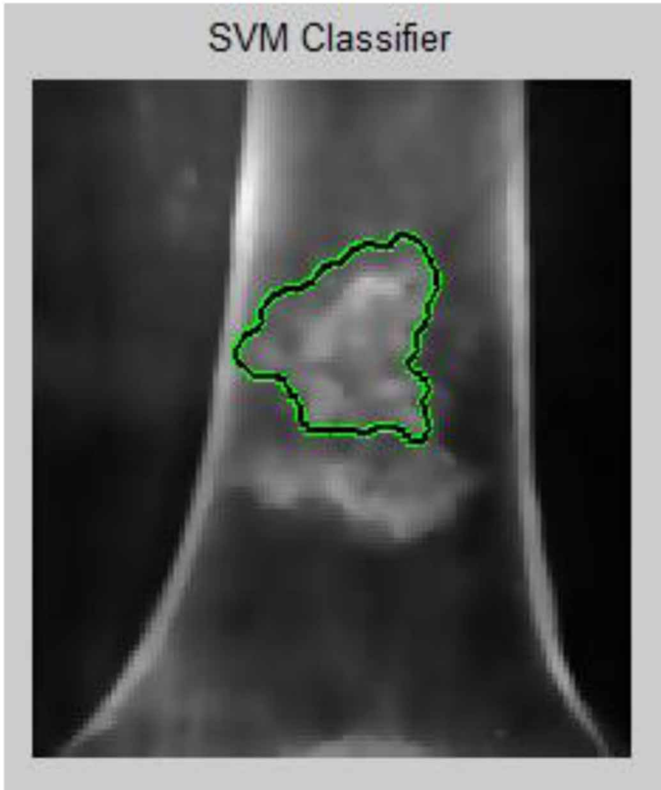
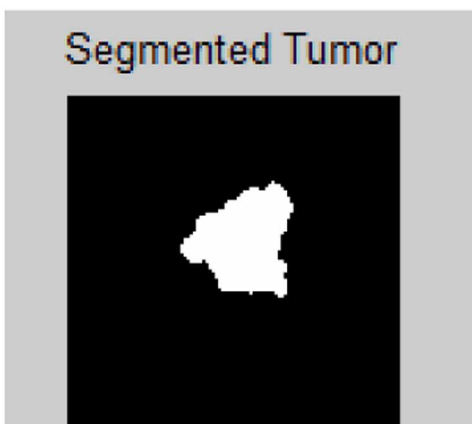


Figure 8. Segmented Tumor



REFERENCES

Bagahel, D., & Kiran, K.G. (2015). Comparative study of tumor detection techniques with their suitability for brain MRI images. Inlt. Jrl., 127(13).

Durgadevi & Shekhar. (2015). Identification of tumor using K-means algorithm. *Intl. Jrl. Adv. Res. Inn. Id. Edu, 1,* 227-231.

Emran, Abtin, & David. (2015). Automatic segmentation of wrist bones in CT using a statistical wrist shape pose model. *Inlt. Jrl.*

Jose, Ravi, & Sampath. (2014). Brain tumor segmentation a performance analysis using K-means, fuzzy c-means and region growing algorithm. *Inlt. Jrl., 2*(3).

Onal, Susana, Paul, & Alferedo. (2014). Automated localization of multiple pelvic bone structure in MRI. *Intl Jrl*.

Chapter 10
An Approach to Cloud Computing for Medical Image Analysis

M. P. Chitra
Panimalar Institute of Technology, India

R. S. Ponmagal
SRM Institute of Science and Technology, India

N. P. G. Bhavani
Meenakshi College of Engineering, India

V. Srividhya
Meenakshi College of Engineering, India

ABSTRACT

Cloud computing has become popular among users in organizations and companies. Security and efficiency are the two major problems facing cloud service providers and their customers. Cloud data allocation facilities that allow groups of users to work together to access the shared data are the most standard and effective working styles in the enterprises. So, in spite of having advantages of scalability and flexibility, cloud storage service comes with confidential and security concerns. A direct method to defend the user data is to encrypt the data stored at the cloud. In this research work, a secure cloud model (SCM) that contains user authentication and data scheduling approach is scheduled. An innovative digital signature with chaotic secure hashing (DS-CS) is used for user authentication, followed by an enhanced work scheduling based on improved genetic algorithm to reduce the execution cost.

DOI: 10.4018/978-1-7998-3092-4.ch010

INTRODUCTION TO CLOUD COMPUTING

Cloud computing is a computing model, where a substantial pool of computing frameworks are associated in the private or open systems, to give progressively versatile foundation to the execution of Personal Computer (PC) application and information stockpiling. With the approach of this innovation, the expense of the computational procedure, application facilitating, content stockpiling and conveyance is diminished altogether. The virtual pictures of the physical machines in the data centers are provided to the clients. Virtualization is one of the principle ideas of the cloud computing framework as it basically assembles the reflection over the physical framework. Cloud computing is a prototype for enabling appropriate and on-demand network access to a shared pool of computing resources that can be rapidly stipulated and released with the cloud service provider interaction or minimal management effort. The service provider provides different types of services such as Software as a Service (SaaS), Platform as a Service (PaaS) or Infrastructure as a Service (IaaS) to the customers across the world through the Internet. The cloud resizes the virtualized hardware automatically. Cloud computing system provides the technologies and tools to compute intensive parallel applications with affordable prices when compared to the existing parallel computing techniques. The architecture of cloud computing is shown in the Figure 1. The benefits of cloud computing system are depicted in Figure 2. The advantages of the cloud computing system are described below Ankita Yadav et.al., (2016).

- Cost effective
- On-demand services
- Remote access
- High efficiency and scalability
- Improved flexibility and reliability
- Maximum resilience without redundancy

TYPES OF CLOUD DEPLOYMENT AND SERVICE MODELS

Cloud Deployment Models

Cloud deployment models can be divided into four types: public, private, community and hybrid cloud. Different types of deployment models are described below

- Public Cloud
- Private Cloud

- Community Cloud
- Hybrid Cloud

Public Cloud

The public cloud allows the systems and services to be easily accessible to the general users. Public cloud might be less secure because of its receptiveness, e.g. email. Public cloud is a completely virtualized condition. Furthermore, suppliers have a multi-inhabitant engineering that empowers clients or occupants to share the figuring assets, for example, PCs, virtual machines, servers, applications and capacity Yarlagadda et.al.,(2011). Information of each tenant is stored in the public cloud. Notwithstanding, it stays segregated from different inhabitants. Public cloud likewise depends on the high-transmission capacity arrange network to quickly transmit information Masdari et.al., 92017). Public cloud stockpiling is commonly excess, utilizing various data centers and cautious replication of document renditions. Figure 3 demonstrates the public cloud.

Figure 1. Cloud Computing Architecture

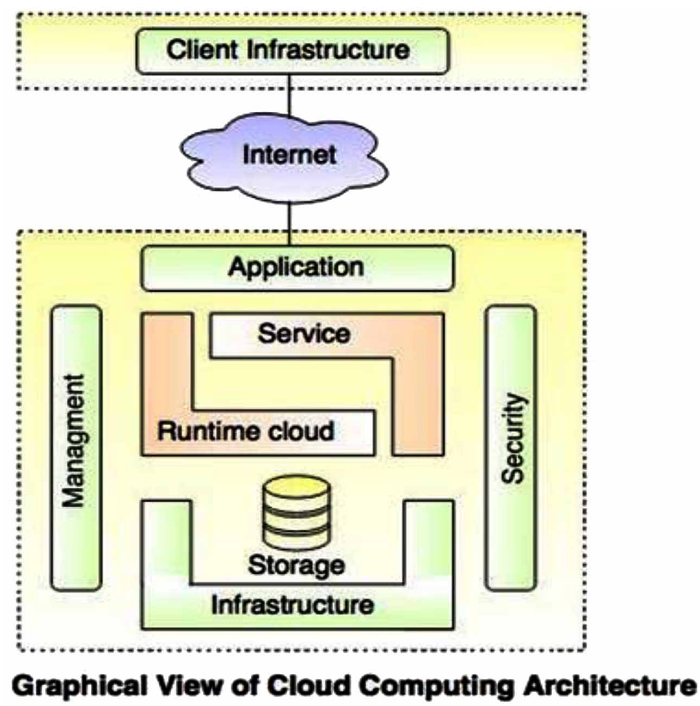
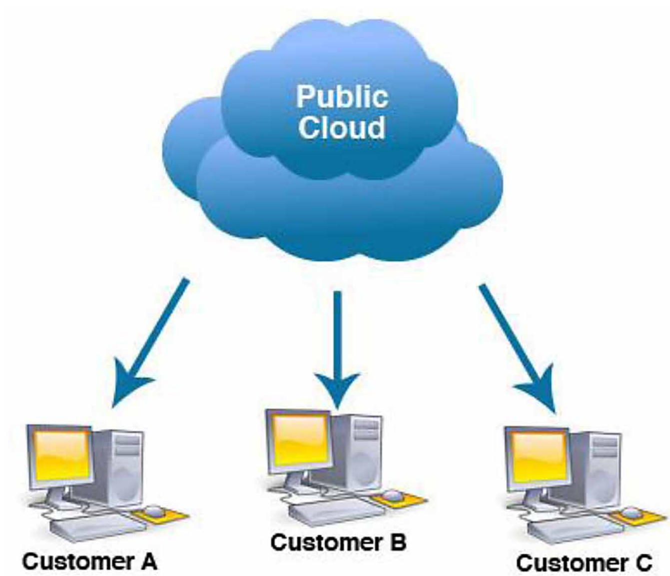


Figure 2. Benefits of cloud computing system

The main benefits of a public cloud service are:

- It prevents the need for the business organizations to invest and maintain their own on-premises resources.
- High scalability to meet workload and user demands.
- Fewer wasted resources as the customers should only pay for the resources they use.

Figure 3. Public cloud

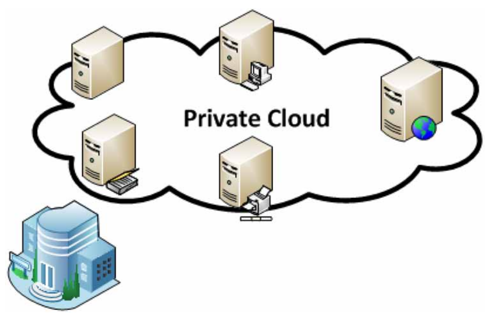


Private Cloud

The private cloud allows the computing systems and services to be accessible within an organization or business enterprise. It offers expanded security due to its private nature. Private cloud is a sort of cloud computing that conveys comparable points of interest to the general population cloud, including adaptability and self-benefit, yet through an exclusive engineering. In contrast to open clouds, which convey administrations to numerous associations, a private cloud is committed to the necessities and objectives of a solitary association Mehmood et.al., (2013).

Private clouds are sent inside firewalls and offer strong Information Technology (IT) security for the association. On the off chance that a server farm foundation is as of now accessible with the association the private cloud can be executed in-house. For having in-house private clouds the association needs to put vigorously in running and keeping up the framework which can result in critical capital consumption. Figure 4 shows the private cloud Liao et.al.,(2013).

Figure 4. Private cloud



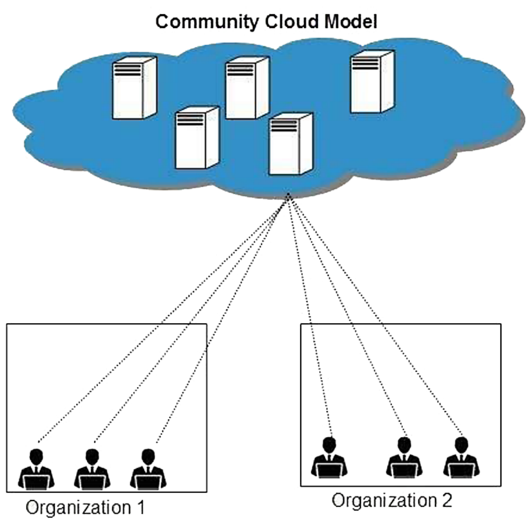
Advantages

- Better data control for the users and cloud service providers.
- As the cloud belongs to a single client, there will be high levels of security.
- As they are deployed inside the firewall of the organization's intranet, it ensures high efficiency and good network performance.
- Easy automation of the hardware and other resources by the company.

Community Cloud

The community cloud allows the systems and services to be accessible by group of organizations. A community cloud is a cloud benefit show that gives a cloud registering solution for a predetermined number of people or associations that is represented, overseen and anchored regularly by all the taking an interest associations or an outsider overseen specialist organization Bace et.al., (2011). Community clouds are a half and half type of private clouds assembled and worked particularly for a focused on gathering. These people group have comparable cloud prerequisites and their definitive objective is to cooperate to accomplish their business targets. Community clouds are intended for organizations and associations chipping away at joint tasks, applications, or research, which requires a focal cloud office for building, overseeing and executing such undertakings, paying little mind to the arrangement leased Bashir et.al.,(2014).

Figure 5. Community cloud



Advantages

- Highly secure private multi-tenant cloud computing
- Improved scalability
- Low cost by eliminating the infrastructure and software licenses needs
- Faster implementation, minimum development time and easy access to the test environments

- Better application, hosting, storage and backup
- Access to the distributed information and advanced analytics solutions

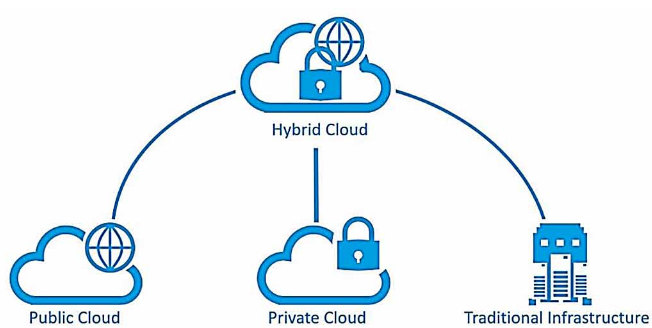
Hybrid Cloud

The Hybrid Cloud is a combination of the public and private cloud. Nonetheless, the basic exercises are performed utilizing the private cloud, while the non-basic exercises are performed utilizing the public cloud. A hybrid cloud is a model in which a private cloud associates with public cloud framework, enabling an association to arrange workloads across the two environments. A hybrid cloud organization requires an abnormal state of similarity between the hidden programming and administrations utilized by both the public and private clouds. Figure 6 depicts the hybrid cloud Bharati et.al.,(2017), Werlinger et.al.,(2008) .

Advantages

- High scalability
- Improved Security
- Enhanced organizational operability
- Greater data accessibility
- Low cost requirement

Figure 6. Hybrid cloud

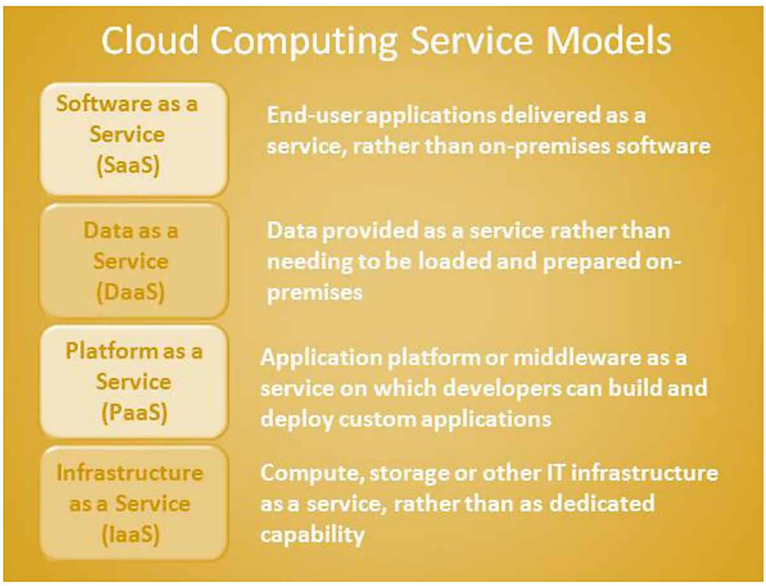


Service Models

Figure 7 depicts the cloud service models. Different types of service models are described below

- Infrastructure as a Service (IaaS)
    - Software IaaS (For example: Google Compute Engine, IBM Smart Cloud, AWS)
    - Hardware IaaS (For example: Rackspace Cloud, AWS, Google Compute Engine)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)
- Data as a Service (DaaS)

Figure 7. Cloud Service Models



IaaS

Figure8 illustrates the IaaS model. IaaS provides access to fundamental resources such as physical machines, virtual machines, virtual storage, etc. In an IaaS model, a cloud specialist organization has the foundation parts generally present in on-premises server farm, including servers, stockpiling and networking equipment, and the virtualization or hypervisor layer. IaaS supplier likewise supplies a scope of administrations to go with those framework parts. IaaS clients get to assets and administrations through a Wide Area Network (WAN) for example, the web, and can utilize the cloud supplier's administrations to introduce the rest of the components of an application stack. Associations pick IaaS since usually less demanding, quicker and more cost-productive to work an outstanding task at the hand without purchasing, oversee and bolster the hidden foundation. With IaaS, a business can basically lease or rent that framework from another business Mell et.al., (2011), Modi et.al.,(2013).

PaaS

PaaS is a cloud computing model in which a third-party provider delivers hardware and software tools usually those needed for application development to the users over the internet. PaaS provides the runtime environment for applications, development & deployment tools, etc. A PaaS provider builds and supplies a resilient and optimized environment on which users can install applications and data sets. Users can focus on creating and running applications rather than constructing and maintaining the underlying infrastructure and services Modi et.al., (2012), Oktay et.al., (2013). Instead of obtaining the hardware downright, users pay for the IaaS on demand. Infrastructure is scalable depending on processing and storage needs. Figure 9 shows the PaaS.

Figure 8. IaaS

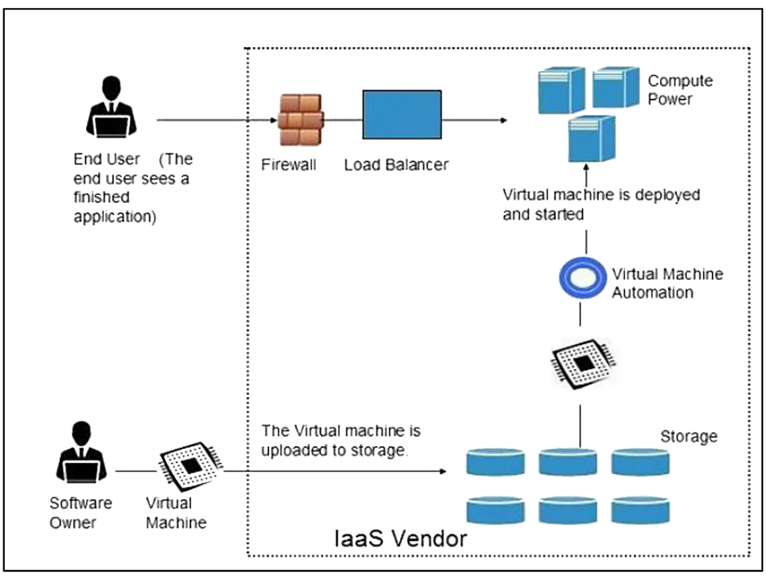
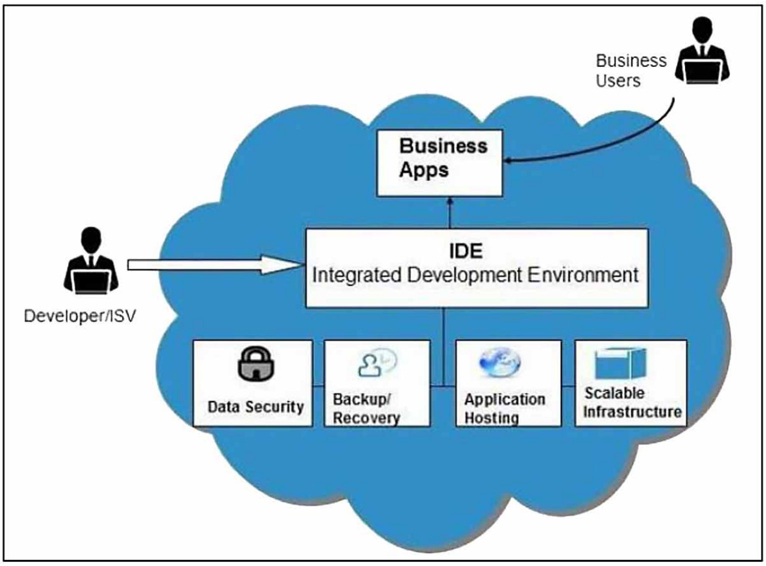


Figure 9. PaaS



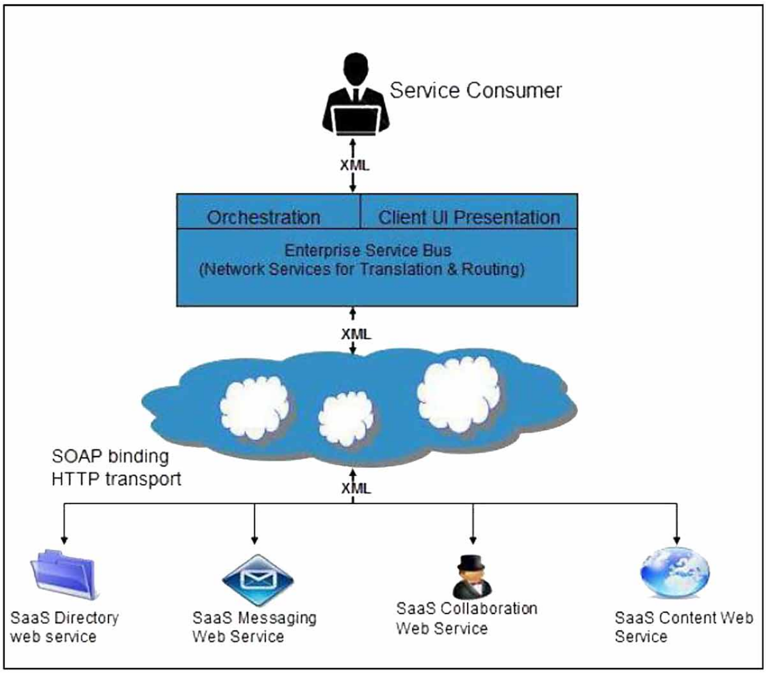
Advantages of PaaS

PaaS offers additional features such as middleware, development tools and other business tools offer more advantages like cut coding time, add development capabilities without adding staff, develop for multiple platforms including mobile, use sophisticated tools affordably, support geographically distributed development teams and efficiently manage the application lifecycle Mehmood et.al., (2013), Dhage et.al.,(2011).

SaaS

SaaS model allows usage of software applications by the end users. Figure 10 shows the SaaS model. SaaS provides a complete software solution on a pay-as-per-go basis from a cloud service provider. SaaS is a cloud computing offering that provides users with access to the cloud-based software of vendor. Users do not install applications on their local devices. Instead, the applications reside on a remote cloud network accessed through the web or an Application Programming Interface (API). Through the application, users can store and analyze data and collaborate on projects. SaaS vendors provide users with software and applications via a subscription model. SaaS providers manage, install or upgrade software without requiring any manual intervention Shelke et.al., (2012), Modi et.al.,(2012).

Figure 10. SaaS



DaaS

DaaS is an information provision and distribution model in which data files including text, images, sounds, and videos are made available to customers over a network, typically the Internet. DaaS offers convenient and cost-effective solutions for customer- and client-oriented enterprises Patel et.al.,(2013), Lee et.al., (2011).

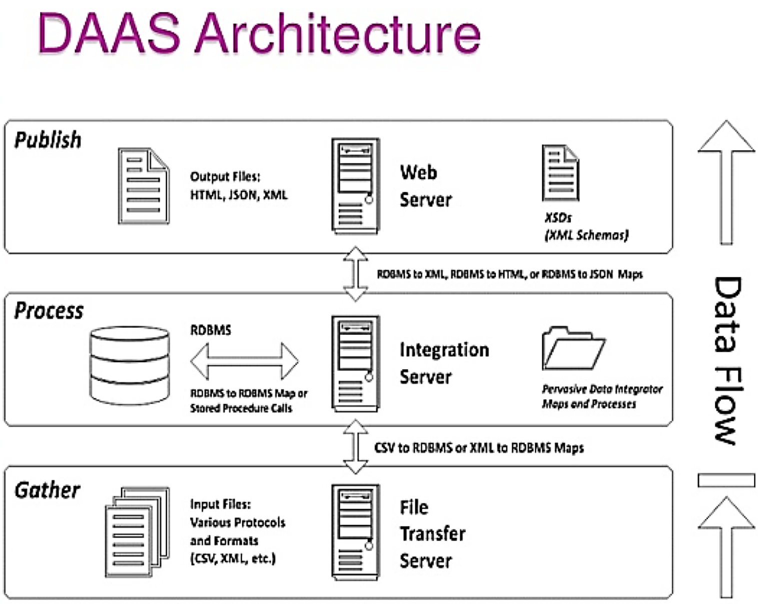
DaaS allows for the separation of data cost and usage from software or platform cost and usage. DaaS is expected to facilitate new and more effective ways of distributing and processing data. DaaS is closely related to Storage as a Service and (SaaS) and may be integrated with one or both of these provision models. Figure 11 shows the layout diagram of DaaS Alharkan et.al., (2012), Xing et.al.,(2013).

The characteristics of cloud are

- On demand self-services
- Wide Network Access
- Resource Pooling
- Rapid Elasticity
- Dynamic Computing Infrastructure

- IT Service-centric Approach
- Minimally or Self-managed Platform
- Consumption-based Billing
- Multi Tenancy
- Cost-effective
- Scalable

Figure 11. DaaS



SECURITY IN CLOUD COMPUTING

Security is a major concern in the cloud computing system. Companies such as Amazon, Google and Microsoft are enhancing the services provided for their users. The Cloud Service Providers (CSPs) are concerned about the non-adequate security measures and aspects such as data integrity, control, audit, confidentiality and availability. In the cloud computing environment, the consumers can access the computational resources online at any time through Internet without managing the original resources issues such as physical and technical management. Protection of the private information of the user is a major concern in the cloud computing environment. When deciding whether or not to move into the cloud, the cloud users would take into account factors such as service availability, security, system performance, etc. The cloud users may lose physical control over their applications and data. In cloud environment, network perimeters will no longer exist from the perspective of cloud users, which renders traditional security protection mechanisms such as firewalls not applicable to the cloud applications Modi et.al.,(2012) Kholidy et.al.,(2012).

In the healthcare applications, cloud service providers and/or system administrators may not be allowed to access sensitive data when providing improved data security protection according to the corresponding compliances. It requires that cloud service providers are able to provide necessary security services to meet security requirements of the individual cloud user while abiding to the regulations and/or compliances. Data protection is a crucial security issue for most of the organizations. Before moving into the cloud, cloud users need to clearly identify data objects to be protected and classify data based on their implication on security, and then define the security policy for data protection as well as the policy enforcement mechanisms. For most applications, data objects would include not only bulky data at rest in cloud servers, but also data in transit between the cloud and the users can be transmitted over the Internet or via mobile media. Data objects may also include user identity information created by the user management model, service audit data produced by the auditing model, service profile information used to describe the service instances, temporary runtime data generated by the instances, and other application data Chung et.al., (2013).

Different types of data would be of different value and hence have different security implication to cloud users. User identity information can contain Personally Identifiable Information (PII) and has impact on user privacy. Therefore, just authorized users should be allowed to access user identity information. Service audit data provide the evidences related to compliances and the fulfillment of Service Level Agreement (SLA), and should not be maliciously manipulated. Service profile information could help attackers locate and identify the service instances and should be well protected. Temporary runtime data may contain critical data related to user business and should be segregated during runtime and securely destroyed after runtime Modi et.al., (2012).

Security Services

The basic security services for information security include assurance of data Confidentiality, Integrity, and Availability (CIA). In the Cloud Computing system, the issue of data security becomes more complicated because of the intrinsic cloud characteristics. Potential cloud users were able to safely move their applications/ data to the cloud. A suit of security services are described as follows.

Data Confidentiality Assurance

This service protects data from being disclosed to illegitimate parties. In Cloud Computing, data confidentiality is a basic security service to be in place. Although different applications may have different requirements in terms of what kind of data need confidentiality protection, this security service could be applicable to all the data objects Bace et.al., (2001).

Data Integrity Protection

This service protects data from malicious modification. When having outsourced their data to remote cloud servers, cloud users must have a way to check whether or not their data at rest or in transit are intact. Such a security service would be of the core value to cloud users. When auditing cloud services, it is also critical to guarantee that all the audit data are authentic since these data would be of legal concerns Werlinger et.al.,(2008).

Guarantee of Data Availability

This service assures that data stored in the cloud are available on each user retrieval request. This service is particularly important for data at rest in cloud servers and related to the fulfillment of SLA. For long-term data storage services, data availability assurance is of more importance because of the increasing possibility of data damage or loss over the time Mell et.al., (2011).

Secure Data Access

This security service is to limit the disclosure of data content to authorized users. In practical applications, disclosing application data to unauthorized users may threat the cloud user's business goal. In the critical applications, inappropriate disclosure of sensitive data can have juristic concerns. For better protection on sensitive data, cloud users may need fine-grained data access control in the sense that different users may have access to different set of data.

Regulations and Compliances

In practical application scenarios, storage and access of sensitive data should obey specific compliance. In addition to this, the geographic location of data would frequently be of concern due to export-law violation issues. Cloud users should thoroughly review these regulation and compliance issues before moving their data into the cloud.

Service Audition

This service provides a way for cloud users to monitor data access and is critical for compliance enforcement. In the case of local storage, it is not hard to audit the system. In Cloud Computing, it requires the CSP to support trustworthy transparency of data access Yarlagadda et.al., (2011).

1ATTACKER MODEL

In the Cloud Computing system, the cloud users move applications from within their enterprise/organization boundary into the open cloud. The cloud users lose physical control over their data. In an open environment, the cloud users may confront all kinds of attacks. Though there might be various categorization methods for the attacks, it is useful to identify where these attackers come from and what kind of attacks they can launch. Based on this criterion, the attackers are classified as insiders and outsiders in the cloud computing environment Masdari et.al., (2017).

Insiders

The insiders refer to the subjects within the system. They could be malicious employees with authorized access privileges inside the cloud user's organization, malicious employees at the side of CSP, and even the CSP itself. In practice, an employee, at both the cloud user side and the CSP side, could become malicious for reasons such as economic benefits. These insider attackers can launch serious attacks such as learning other cloud users' passwords or authentication information, obtaining control of the virtual machines, logging all the communication of other cloud users, and even abusing their access privilege to help unauthorized users gain access to sensitive information. In practical deployments, the cloud users may have to establish trust relationship with the CSPs.

The misbehavior of the cloud server can be anyone or the combination of the following:

1. Potentially decide to hide data corruptions caused by server hacks or Byzantine failures to maintain reputation.
2. Neglect to keep or deliberately delete some rarely accessed data files to save resources.

3. Acquire data information by eavesdropping and monitoring the network traffic.
4. Even collude with a small number of malicious users for the purpose of harvesting the data file contents when it is highly beneficial.

Cloud users should thoroughly review all the potential vulnerabilities and protect their assets on any intentional or in advertent security breaches. More specifically, the users should be aware about the security services offered by the service providers and implementation of these security services. Verification mechanisms should be available to cloud users for verifying the security services provided by the CSPs. For valuable and/or sensitive data, the cloud users may also have to implement their own security protection mechanisms, e.g., strong cryptographic protection, in addition to the security service offered by the cloud service providers Liao et.al., 2013).

Types of Insiders

Companies implement a sophisticated technology to monitor their employees but it's not always easy for them to distinguish between an insider and an outside attack. Those who target and plan attacks from the outside might create strategies for obtaining insider knowledge and access by either resorting to an existing employee, or by making one of their own an insider.

Compromised Actors

Insiders with the access credentials or computing devices have been compromised by an outside threat actor. These insiders are more challenging to address since the real attack is coming from outside, posing a much lower risk of being identified.

Unintentional Actors

Insiders who expose data accidentally, such as an employee who accesses company data through public Wireless Fidelity (Wi-Fi) without the unsafe knowledge are unintentional actors. A large number of data breach incidents result from employee negligence towards security measures, policies, and practices.

Emotional Attackers

Insiders who steal data or destroy company networks intentionally, such as a former employee who injects malware or logic bomb in corporate computers on his last day at work.

Tech Savvy Actors

Insiders who react to challenges are tech savvy actors. They use their knowledge of weaknesses and vulnerabilities to breach clearance and access sensitive information. Tech savvy actors can pose some of the most dangerous insider threats, and are likely to sell confidential information to external parties or black market bidders.

Outsiders

The percentage of external threats to an organization is very high. It includes well-funded hackers, organized crime groups, and government entities. Attacks can be either active or passive. An active attack generates packets or participates in the network while a passive attack is dropping the network or tracking users. By moving data into the cloud, users will lose their conventional network perimeters and expose their data in an open system. Just like any other open systems, Cloud Computing could be vulnerable to malicious attacks through the Internet. This is because Cloud Computing usually does not limit the type of user when providing services. Malicious attackers can easily log into the cloud and launch attacks. More specifically, outsider attackers can launch both passive attacks such as eavesdropping the network traffic, and active attacks like phishing legitimate users' credential, manipulating network traffic and probing the cloud structure. For some cloud services, outsider attackers can launch very severe attacks by taking advantage of the system flaw. By launching the cross virtual machine attacks, the attackers are able to monitor VMs from their co-resident VMs and threaten their security. By subverting hypervisors, attackers are even able to control the whole system stack above the hypervisor.

To address outsider attacks, cloud service providers have the responsibility to secure their cloud infrastructure, isolate user application in the cloud, patch system flaws timely, and notify the cloud users with any discovered security risks. Cloud users should strictly abide to the security guidance when using cloud services for the purpose of reducing the possibility of security breach. Cloud users need to negotiate recovery and backup mechanism with service providers for better security protection. Table 1 shows different types of attacks Masdari et.al.,(2017).

INTRUSION DETECTION IN CLOUD

Intrusion Detection System (IDS) is an essential component of defensive measures protecting computer systems and network against harm abuse. It becomes a crucial section in the Cloud computing environment. The main aim of IDS is to detect computer attacks and provide the proper response. An IDS is defined as the technique that is used to detect and respond to intrusion activities from malicious host or network. There are mainly two categories of IDSs, network based and host based. In addition, the IDS can be defined as a defense system, which detects hostile activities in a network. The key is to detect and possibly prevent activities that may compromise system security, or some hacking attempt in progress including reconnaissance/data collection phases that involve for example, port scans. One key feature of the IDS is their ability to provide a view of unusual activity and to issue alerts notifying administrators and/or blocking a suspected connection. Intrusion detection is defined as the process of identifying and responding to malicious activity targeted at computing and networking resources. In addition, IDS tools are capable of distinguishing between insider attacks and external attacks.

Table 1. Attacks in Cloud Computing

Attacks	Description
Password guessing attack	This includes multiple attacks, including brute force, common passwords and dictionary attacks, which aim to obtain password of the user. The attacker can try to guess a specific user's password, try common passwords to all users or use an already made list of passwords to match against the password file, in their attempt to find a valid password.
Replay attack	The attacker tracks the authentication packet and replays this information to get an unauthorized access to the server.
Man-in-the-middle attack	The attacker passively puts himself between the user and the verifier during the authentication process. Then, the attacker attempts to authenticate by pretending to be as the user to the verifier and the verifier to the user.
Masquerade attack	The attacker pretends to be the verifier to the user to obtain authentication keys or data that may be used to authenticate fallaciously to the verifier.
Insider assisted attack	The systems managers intentionally compromise the authentication system or thieve authentication keys or relevant data of users.
Phishing attack	It is a web based attack in which the attacker redirects the user to the fake website to get passwords/ PIN of the user. Social engineering attacks that use fake emails, web pages and other electronic communications to encourage the user to disclose their password and other susceptible information to the attacker.
Shoulder-surfing attack	The attacker spies the user's movements to get his/her password. In this type of attack the attacker observes the user; how he enters the password i.e. what keys of keyboard the user has pressed.
Denial of Service (DoS) attacks	In DoS attack, an attacker overloads the target cloud system with service requests so that it stop responding to any new requests and hence made resources unavailable to its users.
Cloud Malware Injection Attack	An attacker tries to inject malicious service or virtual machine into the cloud. In this type of attack, the attacker creates own malicious service implementation module (SaaS or PaaS) or virtual machine (IaaS), and add it to the Cloud system.

continues on following page

Table 1. Continued

Attacks	Description
Side channel attack	An attacker attempts to compromise the cloud system by placing a malicious virtual machine in close proximity to a target cloud server system and then launching a side channel attack.
Brute Force Attacks	In this type of attack, all possible combinations of password apply to break the password. The brute force attack is generally applied to crack the encrypted passwords that are saved in the form of encrypted text.
Dictionary Attack	This attack tries to match the password with most occurring words or words of daily life usage.
Key Loggers	The key loggers are the software programs which monitors the user activities by recording each and every key pressed by the user.
Wrapping attack	In wrapping attack, the attacker tries to insert the malicious element in the Simple Object Access Protocol (SOAP) message structure in Transport Layer Service (TLS) and copies the fake content of the message into the server to interrupt the server operation.
Flooding attack	In flooding attack, an adversary can easily create fake data and whenever the server is overloaded, it allocates the job to the nearest server.
Data stealing attack	The attacker steals the information of user account and password. In this attack, confidential information user is lost by the activity of the challenger.
Eavesdropping attack	If an attacker can read the transmitted keys, an eavesdropping will happen.
Spoofing attack	An attacker makes an interrupt by changing routing information and keys.
Privileged insider attack	When the server needs to retain the password for later authentication, the keys are probably being stolen by the adversary because the server can find out the new password.
Server impersonation attack	An attacker masquerades as a legitimate user. To succeed the user impersonation attack, an attacker has to generate a valid login message.
Stolen-verifier attack	In this attack, an adversary which is plotted inside the member can modify the passwords or the verification tables stored in the server's database.
Parallel session attack	In this attack, an adversary applies messages in another authentication process to replace the messages in the authentication operation.
Perfect forward secrecy	This attack happens when an adversary is able to acquire the patient password or a secret key, and it will still be able to compute previous session keys.
Resistance to server spoofing attack	This type of attack can be completely solved by providing the mutual authentication between user and server.
Identity guessing attack	An adversary can reveal the identity through offline exhaustive guessing. The user's identity is usually short and has a certain format. Hence, an adversary may find the identity (ID) within multinomial time by executing complete guessing.
Reflection attacks	When a challenge-response authentication system is used as the same protocol in both directions by each side to authenticate the other side, a reflection attack will happen.

Host-Based IDS (HIDS)

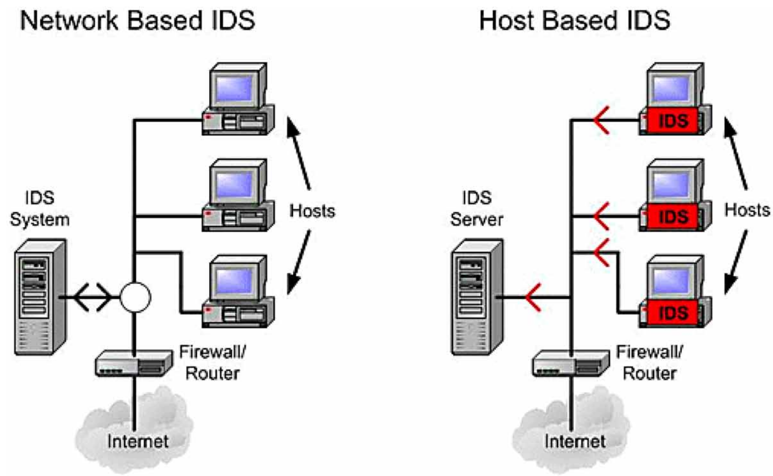
Host-based IDS involves software or agent components run on the server, router, switch or network appliance. But, the agent versions must report to a console or can be run together on the same host. Basically, HIDS provides poor real-time response and cannot effectively defend against one-time catastrophic events. In fact, the HIDSs are much better in detecting and responding to long term attacks such as data theft. HIDS collect information from a particular host and analyze to detect intrusive events. The information may be system logs or audit trails of operating system.

HIDS analyzes the information and if there is any change in the behavior of system or program, it reports to the0 network manager that the system is under attack. The effectiveness of HIDS can be improved by specifying the features that provide it more information for detection. However, it requires more storage for information to be analyzed. In the case of cloud computing network, it is possible to deploy HIDS on hypervisor, VM or host to analyze the system logs, user login information or access control policies and detect intrusion events. HIDS is capable of analyzing encrypted traffic however; it is susceptible to DoS attack and can even be disabled. HIDS are commonly used to protect the integrity of the software Mehmood et.al., (2013).

Network-Based IDS (NIDS)

This type of IDS captures network traffic packets such as TCP, UDP and IPX/SPX and analyzes the content against a set of rules or signatures to determine if a possible event took place. False positives are common when an IDS system is not configured or tuned to the environment traffic it is trying to analyze. Networks based IDS (NIDS) capture the traffic of entire network and analyze it to detect possible intrusions like port scanning, DoS attacks etc. NIDS usually performs intrusion detection by processing the IP and transport layer headers of captured network packets. It utilizes the anomaly based and signature based detection methods to identify intrusions. NIDS collects the network packets and looks for their correlation with signatures of known attacks or compares the user current behavior with their already known profiles in real-time. Multiple hosts in the network can be secured from attackers by utilizing a few properly deployed NIDSs. If run in the stealth mode, the location of NIDS can be hidden from attacker. The NIDS is unable to perform analysis if traffic is encrypted. In cloud environment, the attacks on hypervisor or VMs are detected by positioning NIDS at the cloud server that interacts with external network. However, it cannot detect attacks inside a virtual network contained by hypervisor. Cloud provider is responsible for installing NIDS in the cloud Mehmood et.al.,(2013). Figure 12 shows the host-based IDS and network based IDS.

Figure 12. Host-based and Network-based IDS



Virtual Machine Monitor (VMM)/Hypervisor Based IDS

Hypervisor provides a platform for communication among VMs. Hypervisor based IDSs is deployed at the hypervisor layer. It helps in analysis of available information to detect the anomalous activities. The information is based on communication at various levels like communication between VM and hypervisor, between VMs and communication within the hypervisor based virtual network Mehmood et.al.,(2013).

Distributed IDS (DIDS)

DIDS comprises numerous IDSs that are deployed across a large network to monitor the traffic for intrusive behavior. The participant IDSs can communicate with each other or with a centralized server. Each of these individual IDSs has its own two function components: detection component and correlation manager. Detection component monitors the system or subnet and transmits the collected information in a standard format to the correlation manager. Correlation manager combines information from multiple IDS and generates high level alerts corresponding to an attack. Analysis phase makes use of signature and anomaly based detection methods hence DIDS can detect known and unknown attacks. In case of cloud, DIDS can be located at any of two positions: at processing server or host machine Mehmood et.al., (2013).

DETECTION TECHNIQUES USED BY IDS

Signature Based IDS

Signature based detection is performed by comparing the information collected from a network or system against a database of signatures. A signature is a predefined set of rules or patterns that correspond to a known attack. This technique is also known as misuse detection Liao et.ai.,(2013). It can efficiently detect known attacks with negligible false alarms. Signature based method helps network managers with average security expertise to identify intrusions accurately. It is a flexible approach since new signatures can be added to database without modifying existing ones. However, it is unable to detect unknown attacks Bace et.al., (2001). In the Cloud environment, signature based intrusion detection method can be utilized at front-end of the cloud for detection of the known attacks from external network. It can also detect both internal and external intrusions if deployed at back end of the cloud.

Anomaly Based Detection

Anomaly based detection compares current user activities against preloaded profiles of users or networks to detect abnormal behavior that may be intrusions. The profiles may be dynamic or static and correspond to the expected or benign behavior of users. To build a profile, regular activities of users, network connections, or hosts are monitored for a specific period of time called as training period. Profiles are developed using various features like failed login attempts, number of times a file is accessed by a particular user over a particular time duration, CPU usage etc. Anomaly based detection is effective against unknown attacks. An attack detected by anomaly based technique can be used as a signature in signature based detection. However it produces a large number of false alarms due to irregular network and user behavior. Moreover, it also requires large data sets to train the system for the normal user profiles.

Soft computing techniques used for intrusion detection are described below

1. Fuzzy Logic: It is based on probability, uses values ranging between 0 and 1, and is used to define degree of anomaly in intrusion detection.
2. Artificial Neural Networks (ANN): In intrusion detection, ANN can be used for generalization of data from imperfect data. It is also used to categorize data as being normal or anomalous.
3. Support Vector Machines (SVM): SVM can be an effective way to detect intrusive events in case of limited data samples, where data dimensions will not change the accuracy.

4. Association rules: This technique helps in creation of new signatures which can be used to detect intrusions. Such intrusions consist of some known attacks or variation of known attacks.
5. Genetic Algorithm (GA): The network features selected by using GAs can be applied in other techniques which improves the detection accuracy of IDS.

Hybrid Detection

The efficiency of IDS can be significantly improved by combining signature based and anomaly based techniques that are called as Hybrid detection technique. The motivation behind this combination is the ability to detect both known and unknown attacks using signature based and anomaly based detection techniques. Hybrid IDSs can be divided into two categories: 1) sequence-based, in which either anomaly detection or misuse detection is applied first, and the other one is applied next; 2) parallel-based, in which multiple detectors are applied in parallel, and the final decision is made based on multiple output sources. The most common type of hybrid system is to combine misuse detection and anomaly detection together. In such a hybrid system, the signature detection technique detects known attacks, and the anomaly detection technique detects new or unknown attacks.

PROBLEM STATEMENT

IDS technology has not reached a level where it does not require human intervention. Latest IDS technology offers some automation like notifying the administrator in case of detection of a malicious activity, shunning the malicious connection for a configurable period of time, dynamically modifying a router's access control list in order to stop a malicious connection etc. But it is still very important to monitor the IDS logs regularly to stay on top of the occurrence of events. Monitoring the logs on a daily basis is required to analyze the kind of malicious activities detected by the IDS over a period of time. Today's IDS has not yet reached the level where it can give historical analysis of the intrusions detected over a period of time. This is still a manual activity. The IDS technology works on the attack signatures. Attack signatures are attack patterns of previous attacks. The signature database needs to be updated whenever a different kind of attack is detected and the fix for the same is available. The frequency of signature update varies from vendor to vendor.

The successful growth of the artificial intelligence techniques has placed a great challenge of incorporating this new field in IDS. Use of neural networks can also be effective in IDS. Their capability to process huge data and derive meaning and patterns from it can be applied to find attacks. Gradually, it keeps on learning keeping track of previous penetrations and analyzing data for newer ones. As cloud sellers utilize virtual machine innovation Host and system interruption attacks on remote hypervisors are a genuine security challenge. DOS and Distributed (DDOS) attacks are hurled to refuse assistance accessibility to complete clients. Protection of data from the outsider inspector is another stress of cloud security. Cloud reviewing might be a hard assignment to look at consistence of all the security approaches by the vendor Bharati et.al.,(2017). An IDS is complex and provide many challenges for security practitioners. IDS research has focused largely on improving the accuracy of these systems and on providing support to practitioners during the ongoing task of monitoring alerts and analyzing potential security incidents. The installation and the initial configuration of the IDS can be challenging that they can serve as a barrier to use Werlinger et.al.,(2008).

RESEARCH CONTRIBUTION

Detection of the intrusions and attacks through unauthorized users is one of the major challenges for both cloud service providers and cloud users. The first phase of the research work proposes a new IDS based on the combination of One-Class Support Vector Machine (OC-SVM) network, and Artificial Bee Colony (ABC) to detect anomalies in complex dataset. The hybrid OC-SVM algorithm is substandard because of it is not able to effect the representation based learning in the middle hidden layer. This approach was implemented for different datasets such as NSL-KDD, KDD-CUP datasets. The experimental results showed improved accuracy in intrusion detecting attacks by the unauthorized access.

Cloud data allocation facility allows a group of user to work together to access and the shared data is one of the most standard and effective working styles in the enterprises. Despite of the scalability and flexibility benefits, the cloud storage service comes with the data confidentiality and the security concerns. A direct method to defend the user data is to encrypt the data stored at the cloud. In this research work, a Secured Cloud Model (SCM) that contains user authentication, and data scheduling is suggested. An innovative Digital Signature with Chaotic Secure Hashing (DS-CSH) is applied for user authentication, followed by an enhanced work scheduling based on the improved genetic algorithm to reduce the execution cost. The proposed SCM architecture yields better throughput, schedule success rate, lower normalized schedule cost, end-to-end delay and packet loss rate. Thus, the proposed SCM provides a secure environment with a higher QoS that can support more users.

In recent years, outlier detection is a well-investigated topic in data science Bashiret. al.,(2014), Bharati et.al.,(2017). Learning rule to classify normal and anomalous data without prior label is called unsupervised anomaly detection. One-class SVM

(OC-SVM) is a most popular approach of unsupervised anomaly detection, in which smooth boundary is constructed around the majority of probability mass of data Werlinger et.al.,(2008). Nowadays, numerous feature extraction and feature selection methods have been implemented by use with OC-SVM for high dimensional dataset Ankita Yadav et.al.,(2016). Because of extraordinary success of deep auto encoder networks used as a feature extractors in speech anomaly detection Yarlagadda et.al., (2011), several hybrid models uses combine method of OC-SVM with deep learning for their applications Liao et.al., (2013). This hybrid model was proved that in terms on accuracy and recognition rate for anomaly detection with maximum benefits using two publically available CNN models ImageNet-MatConvNet-VGG-F (VGG-F) and ImageNetMatConvNet-VGG-M (VGG-M) Bace et.al.,2001).

Robust Deep Autoencoders for Anomaly Detection

Along with the OC-SVM with deep learning hybrid approaches, a new novel method for anomaly detection is to use deep autoencoders. Motivated by RPCA Bashir et.al.,(2014) robust deep autoencoder method is introduced for anomaly detection in unsupervised manner Bashir et.al., (2017). In which, the input data is decomposed by two parts in Robust Deep Autoencoder (RDA) or Robust Deep Convolutional Autoencoder (RCAE) as $X = L_D + S$, where L_D denotes the representation the hidden layer of the autoencoder. The matrix S captures unwanted and outlier data which are hard to estimate as shown in Equation 1. The decomposition is improved by optimizing the objective function shown below:

$$\min_{\theta,s} + L_D - D_\theta\left(E_\theta\left(L_D\right)\right)_2 + \lambda S^T{}_{2,1} \tag{2.1}$$

Such that $X - L_D - S = 0$

The mentioned optimization issue is carried out by a hybrid method of back propagation and Alternating Direct Method of Multipliers (ADMM) approach.

One Class SVM for Anomaly Detection

Apart from many algorithms, One-Class SVM (OC-SVM) is most commonly used approach for outlier detection in an unsupervised manner Chung et.al.,(2013). This approach is special case of SVM, which separate all the data by learning hyper-plane from origin in a Reproducing Kernel Hilbert Space (RKHS) and maximize the space between hyper-plane to the origin. Automatically in OC-SVM all the data points are denoted as positive labeled data and the origin as the only negative labeled data.

More exactly, consider the data set X without class information, and $\Phi\left(X\right)$ denote as RKHS map function from the input space to the feature space F and a hyper plane $f\left(X_n\right)$ is constructed by

$$f\left(X_n\right) = w^T\Phi\left(X_n\right) - r \tag{1}$$

This function is used to separate as many as mapped vectors $\Phi\left(X\right)_n, n:1,2,\ldots,$ from the origin.

Where $w-$ Norm perpendicular to the hyperplane and $r-$ bias Factor of the hyperplane.

By solving the below equation, the value of $r \text{ and } w$ can be obtained.

$$\min_{w,r} \frac{1}{2} w_2^2 + \frac{1}{v} \cdot \frac{1}{N} \sum_{n=1}^{N} \max\left(0, r - w, \Phi\left(X_{n:}\right)\right) - r \tag{2}$$

Where $v \in \left(0,1\right),$ is a parameter that control the hyper plane distance from origin and number of points cross over the hyper-plane.

The proposed SCM-cloud framework is estimated with the novel algorithms implemented for security. This section comprises two sub-sections: simulation setup, and comparative analysis.

Simulation Setup

The proposed novel SCM-cloud architecture is implemented in the Java simulation environment. The proposed SCM architecture is compared with OMC-RP, SecSDN-Cloud, Opensec, AuthFlow, RPL-based SDN, FlowDefender, IC-PCP, BHEFT and CWTS-GA. Table 2 defines the most important parameters used in the proposed algorithm simulations. The number of users, Cloud Service Provider (CSP) and difference time are varied to analyze the network performance. However, the parameters are not limited to this list.

Performance Evaluation

The QoS efficiency is defined in terms of three important throughput parameters as well as end-to-end delay and packet loss. Network throughput is defined as the amount of data transmitted from user to cloud in a specified time period. The throughput T_p, is calculated using the following formula:

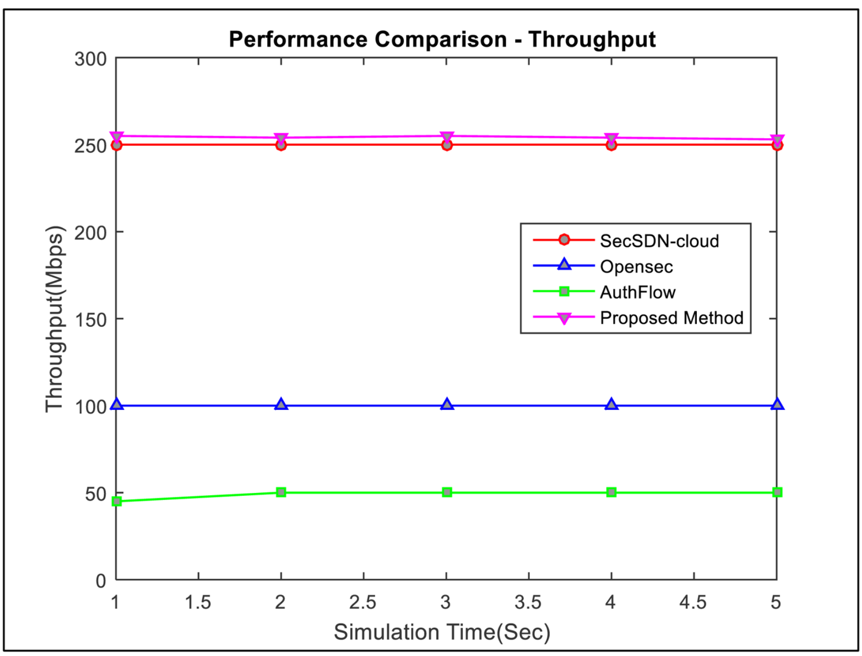
Table 2. Simulation Parameters

Number of users	10
Number of cloud providers	3
Queue type	Drop Tail
Buffer capacity	3
Data Rate	100 Mbps
Transmission Interval	2 seconds
Simulation time	30 seconds

$$T_p = \frac{NP_s}{T} \tag{3}$$

Here, T represents the time interval of transmission, NP_s denotes the number of packet to be transferred. Figure 13 shows that the proposed SCM model show the gradual improvement of the throughput when simulation time is increased. The throughput of the SCM model is maximum of about 1.186%, 60.47% and 80.24% than the SecSDN-Cloud, Opensec and AuthFlow approaches, respectively.

Figure 13. Performance Comparison of Throughput



CONCLUSION

The overview of cloud computing for medical image analysis gives a brief outline of the existing intrusion detection approaches for the cloud computing environment in the field of medical image processing. The new IDS based on the combination of OC-SVM network, and ABC to detect anomalies in complex dataset. A new intrusion detection system (IDS) is proposed based on a combination of a one-class Support Vector Machine (OC-SVM) network, and artificial bee colony (ABC) to detect anomalies in complex dataset.The experimental results showed improved accuracy in intrusion detecting attacks by unauthorized access. A new Secure Cloud Model (SCM) framework for medical image analysis using digital signature with chaotic secure hashing and Work scheduling based on improved Genetic algorithm. An innovative digital signature with chaotic secure hashing (DS-CS) is used for user authentication, followed by an enhanced work scheduling based on improved genetic algorithm to reduce the execution cost.

REFERENCES

Aggarwal, C. C. (2015). Outlier analysis. Data Mining, 237-263. doi:10.1007/978-3-319-14142-8_8

Alharkan, T., & Martin, P. (2012). IDSaaS: Intrusion detection system as a service in public clouds. *Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (ccgrid 2012)*, 686-687. 10.1109/CCGrid.2012.81

Bace & Mell. (2001). *NIST special publication on intrusion detection systems*. Booz-Allen and Hamilton Inc.

Bashir, U., & Chachoo, M. (2014). Intrusion detection and prevention system: Challenges & opportunities. *Computing for Sustainable Global Development (INDIACom), 2014 International Conference on*, 806-809. 10.1109/IndiaCom.2014.6828073

Bharati, M., & Tamane, S. (2017). Intrusion detection systems (IDS) & future challenges in cloud based environment. Intelligent Systems and Information Management (ICISIM), 240-250. doi:10.1109/ICISIM.2017.8122180

Chung, C.-J., Khatkar, P., Xing, T., Lee, J., & Huang, D. (2013). NICE: Network Intrusion Detection and Countermeasure Selection in Virtual Network Systems. *IEEE Transactions on Dependable and Secure Computing*, *10*(4), 198–211. doi:10.1109/TDSC.2013.8

Dhage, S. N., Meshram, B., Rawat, R., Padawe, S., Paingaokar, M., & Misra, A. (2011). Intrusion detection system in cloud computing environment. *Proceedings of the International Conference & Workshop on Emerging Trends in Technology*, 235-239. 10.1145/1980022.1980076

Kholidy, H. A., & Baiardi, F. (2012). CIDS: A framework for intrusion detection in cloud systems. *Information Technology: New Generations (ITNG), Ninth International Conference on*, 379-385. 10.1109/ITNG.2012.94

Lee, J.-H., Park, M.-W., Eom, J.-H., & Chung, T.-M. (2011). Multi-level intrusion detection system and log management in cloud computing. *Advanced Communication Technology (ICACT), 13th International Conference on*, 552-555.

Liao, H.-J., Lin, C.-H. R., Lin, Y.-C., & Tung, K.-Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, *36*(1), 16–24. doi:10.1016/j.jnca.2012.09.004

Masdari, M., & Ahmadzadeh, S. (2017). A survey and taxonomy of the authentication schemes in Telecare Medicine Information Systems. *Journal of Network and Computer Applications*, *87*, 1–19. doi:10.1016/j.jnca.2017.03.003

Mehmood, Y., Shibli, M. A., Habiba, U., & Masood, R. (2013). Intrusion detection system in cloud computing: challenges and opportunities. *2013 2nd National Conference on Information Assurance (NCIA)*, 59-66. 10.1109/NCIA.2013.6725325

Mehmood, Y., Shibli, M. A., Habiba, U., & Masood, R. (2013). Intrusion detection system in cloud computing: challenges and opportunities. *2nd National Conference on Information Assurance (NCIA)*, 59-66. 10.1109/NCIA.2013.6725325

Mell & Grance. (2011). *The NIST definition of cloud computing*. Academic Press.

Modi, C., Patel, D., Borisaniya, B., Patel, H., Patel, A., & Rajarajan, M. (2013). A survey of intrusion detection techniques in cloud. *Journal of Network and Computer Applications*, *36*(1), 42–57. doi:10.1016/j.jnca.2012.05.003

Modi, C. N., Patel, D. R., Patel, A., & Muttukrishnan, R. (2012). Bayesian Classifier and Snort based network intrusion detection system in cloud computing. In *Computing Communication & Networking Technologies* (pp. 1–7). ICCCNT. doi:10.1109/ICCCNT.2012.6396086

Modi, C. N., Patel, D. R., Patel, A., & Muttukrishnan, R. (2012). Bayesian Classifier and Snort based network intrusion detection system in cloud computing. *Third International Conference on Computing Communication & Networking Technologies (ICCCNT)*, 1-7.

Modi, C. N., Patel, D. R., Patel, A., & Rajarajan, M. (2012). Integrating signature apriori based network intrusion detection system (NIDS) in cloud computing. *Procedia Technology*, *6*, 905–912. doi:10.1016/j.protcy.2012.10.110

Nikolai, J., & Wang, Y. (2014). Hypervisor-based cloud intrusion detection system. In *Computing* (pp. 989–993). Networking and Communications.

Oktay, U., & Sahingoz, O. K. (2013). Proxy network intrusion detection system for cloud computing. *International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE)*, 98-104. 10.1109/TAEECE.2013.6557203

Oktay, U., & Sahingoz, O. K. (2013). Proxy network intrusion detection system for cloud computing. *Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE) International Conference on*, 98-104. 10.1109/TAEECE.2013.6557203

Patel, A., Taghavi, M., Bakhtiyari, K., & Celestino Júnior, J. (2013). An intrusion detection and prevention system in cloud computing: A systematic review. *Journal of Network and Computer Applications*, *36*(1), 25–41. doi:10.1016/j.jnca.2012.08.007

Shelke, M. P. K., Sontakke, M. S., & Gawande, A. (2012). Intrusion detection system for cloud computing. *International Journal of Scientific & Technology Research*, *1*, 67–71.

Werlinger, R., Hawkey, K., Muldner, K., Jaferian, P., & Beznosov, K. (2008). The challenges of using an intrusion detection system: is it worth the effort? *Proceedings of the 4th symposium on Usable privacy and security*, 107-118. 10.1145/1408664.1408679

Xing, T., Huang, D., Xu, L., Chung, C.-J., & Khatkar, P. (2013). Snortflow: A openflow-based intrusion prevention system in cloud environment. *Research and Educational Experiment Workshop (GREE)*, 89-92. 10.1109/GREE.2013.25

Yadav, A., & Kumar, N. (2016). A Survey of Authentication Methods in Cloud Computing. *International Journal of Innovative Research in Computer and Communication Engineering*, *4*, 19529–19533.

Yarlagadda, V. K., & Ramanujam, S. (2011). Data security in cloud computing. *Journal of Computer and Mathematical Sciences*, *2*, 1–169.

Chapter 11

Segmentation of Spine Tumour Using K-Means and Active Contour and Feature Extraction Using GLCM

Malathi M.
Rajalakshmi Institute of Technology, India

Sujatha Kesavan
Dr. M. G. R. Educational Research Institute of Technology, India

Praveen K.
Chennai Institute of Technology, India

ABSTRACT

MRI imaging technique is used to detect spine tumours. After getting the spine image through MRI scans calculation of area, size, and position of the spine tumour are important to give treatment for the patient. The earlier the tumour portion of the spine is detected using manual labeling. This is a challenging task for the radiologist, and also it is a time-consuming process. Manual labeling of the tumour is a tiring, tedious process for the radiologist. Accurate detection of tumour is important for the doctor because by knowing the position and the stage of the tumour, the doctor can decide the type of treatment for the patient. Next, important consideration in the detection of a tumour is earlier diagnosis of a tumour; this will improve the lifetime of the patient. Hence, a method which helps to segment the tumour region automatically is proposed. Most of the research work uses clustering techniques for segmentation. The research work used k-means clustering and active contour segmentation to find the tumour portion.

DOI: 10.4018/978-1-7998-3092-4.ch011

INTRODUCTION

A spine tumour is abandoned growth of cells, which was found in the spinal cord. It grows uncontrollably. It may cancerous (Malignant) or non-cancerous (Benign). It may cause neurologic problems and in some cases it produces paralysis. It may occur in the region of the spine. The two types of tumour are primary and secondary. The primary tumour starts in the spinal cord and secondary spreads to another part of the spine. Based on the region of the spine it may occur in cervical, thoracic lumber and sacrum. Based on the location of the spine it may be classified into three types like Intradural-extramedullary, intramedullary and extradural. (Yezzi et al., 1997) The clear and accurate visual arrangements of an internal organ of our body have been generated using various medical imaging modalities like CT (Computed Tomography), MRI (Magnetic Resonance Imaging). This can be used to provide the internal organizations of bone and skin. The various diseases of the human body should be identified with the help of Medical imaging techniques. It can be used to generate an actual structure of the human body to detect the abnormalities. MRI, CT, Ultrasound, Positron Emission Tomography (PET), etc. were the different kinds of medical imaging techniques.

(Hai, S, Fuyong Xing & Lin Yang 2016) Demonstrates the various brain imaging techniques. For the CT images the tomography is the word, which originates from two Greek words like tomos and graphia. The word tomos represents slice or section and graphia represents the picture. From this, it will understand that CT provides the detailed structure of internal organ of the human body. CT utilizes X-rays to reproduce the internal organization of the human body. After CT imaging, the reconstruction of an image depends on the X-ray absorption profile.

One of the dynamic and flexible radio imaging technique was MRI. The technique uses electromagnetic radiation to acquire the internal structure of the human body. The abnormalities in the soft region were found by the invasive MRI imaging methods. The technology helps the physicians to find the abnormalities in chest, lungs, bones etc. Unlike X–ray MRI does not uses harmful radiation. During MRI imaging the human body aligns the hydrogen atoms of the body.

X-rays are electromagnetic waves which are used to provide useful information about the human body. The X-ray absorption profile will differ for every tissue. Dense tissue seems like white on CT film and soft tissue looks like gray. During CT imaging techniques the appearance of the Lungs is black because the hollow space within the lungs is filled with air. Unlike X-ray CT does not use the dangerous radiation, it affects the human body. CT is one of the best medical imaging technique and it helps to diagnose the diseases in various human body parts like Brain, Pelvis, Liver, Chest, Abdomen, and Spine etc. Hence the suggested method utilizes MRI Imaging Techniques to find tumour in the spinal cord.

RELATED STUDIES

Manual segmentation refers the human operator or physician performs segmentation and labeling of an image by hand. The separation is performed in a slice by slice method on a 3-D volumetric image. Depending on the artifacts present in the medical image makes the segmentation is an easy or difficult process. But manual segmentation requires a long time to complete the task. Recent days the CT and MRI imaging technique is mostly used in the diagnosis, treatment planning and clinical studies require computers to assist the radiologist experts. In order to perform the segmentation of a large amount of images with the same accuracy, the computer aided diagnosis was implemented

(Shan Shen et al., 2005) The author proposed technique in which noises are present in MRI brain images due to poor operating performance, disturbances, surroundings and poor equipment maintenance. This noise affects the accuracy of the segmentation process. Many conventional algorithms are available for tumour segmentation. Clustering is grouping of pixels depends on the intensity value of a pixel. Conventional clustering method mainly based on the intensity value of a pixel, which leads to poor segmentation. Hence segmentation of the tumour is improved by using a neighborhood attraction method, which will provide comparative position and structures of neighbouring pixels. The proposed methodology does not use intensity as a single parameter to perform segmentation of tumour, but the intensity of neighbouring pixels was also taken into account for segmentation of tumour.

(Li Hong Juang& Ming Ni Wu 2010) The author used conventional clustering method for MRI brain tumour segmentation. The author describes the segmentation process on a colour brain image. Most of the segmentation process is executed on gray image. But the proposed work by the author is performed on RGB image. The segmentation process on colour image provides a good accuracy and also reduces the iteration time or time required.

(Madhukumar, S &Santhiyakumari, N 2015) The author used histogram guided initialization to do qualitative comparison between Fuzzy C-means (FCM) and k-Means segmentation. The accuracy of the above mentioned methods depends on its ability to distinguish dissimilar tissue classes, independently. The research discloses that FCM detects the vasogenic edema and the white matter as a only tissue class and correspondingly gray matter and necrotic focus, besides this FCM is better to designate these areas when compared with K means clustering.

(Georges et al., 1999) The author discussed and introduced a Deformable model in order to obtain the boundary information of lesions. It may be calculated effectively with the help of snake model in active contour segmentation. It helps the radiologist to find the spatial relationship between every pixel.

(Andrew et al., 2020) The author uses deep learning technology to perform segmentation from MRI images. The manual segmentation of affected region from an input image is tedious process. Hence the research work uses the automatic segmentation in order to improve the accuracy and overall efficiency of the process. Recent years CNN is used for automatic segmentation because it has the ability to handle large data set and provides better accuracy by comparing with the existing segmentation methods. The CNN performance is evaluated using state of art results.

(Hyunseok Seo &Masoud Badiei Khuzani, 2020) The author discussed various machine learning algorithms for segmentation of medical and natural images. The author compared different machine learning algorithms like Markov random field, K- means clustering, random forest etc., the conventional segmentation yields less accurate then compared to deep learning techniques. But it is simple to implement and have less complex structures. The research work also compares the various learning architectures like ANN, CNN and RNN. Recent deep learning techniques need a several hyper parameter tuning. Small changes in hyper planeparameters yieldsmany changes in the network output

MATERIALS AND METHODS

Existing Method

MRI Brain-imaging technique is used to detect the brain tumour. After getting the Brain image through MRI scans calculation of area, size and position of the brain tumour are important to give the treatment for the patient. Earlier days the tumour portion of the brain detected using manual labeling. This is a challenging task for the radiologist and also it is a time- consuming process. Manual labeling of the tumour is fatigue, tedious process for the radiologist. Accurate detection of brain tumour is important for the doctor because by knowing the position and the stage of the tumour, the doctor can decide the type of treatment for the patient. Next important consideration in the detection of a tumour is earlier diagnosis of a tumour; this will improve the lifetime of the patient. From this motive while analyzing the spine tumour there is no good segmentation methods. Hence the article is initiated to start with the conventional segmentation process like k means clustering followed by active contour segmentation. But this method provides false segmentation in the presence of noise and also it is difficult to assign the K – value. The proposed work overcomes the drawback with suitable filters to remove nose and false segmentation is avoided by active contour segmentation in order to partition the affected portion. Further the image underdone for the feature extraction.

Proposed Method

Thereare many types of segmentation methods used in medical images are thresholding method, edge detection method based technology, region based technology, clustering based technology. The proposed research work uses clustering based segmentation due to its simplicity, accuracy and easy implementation. Recent years many types of machine learning algorithms were developed like Artificial neural Networks (ANN's), Convolutional Neural networks (CNN's), Recurrent neural networks(RNN's) and Deep neural networks (DNN's). The future research work focused on developing the segmentation using CNN and DNN. Since it has the ability to process large amount of data. Before learning the recent machine learning algorithms it is necessary to know the fundamental concepts on segmentation. Hence we proposed spine tumour segmentation using conventional K means and active contour segmentation. In the presence of noise K means algorithms provides false or over segmentation. The traditional clustering based segmentation provides acceptable efficiency; it may be improved by using CNN, DNN and ANN in future research work.

The segmentation of abnormal portions was calculated for many organs like brain, breast, lungs etc., All segmentation process uses clustering methods. From this motivation we applied automatic Spine tumour segmentation by implementing K-means clustering and active contour segmentation process.

(Eman et al., 2015) finally the performance of the segmentation techniques is measured by the number of GLCM features. The extracted features helps to classify the abnormal portion is malignant or benign. The flow diagram is shown in figure 1.

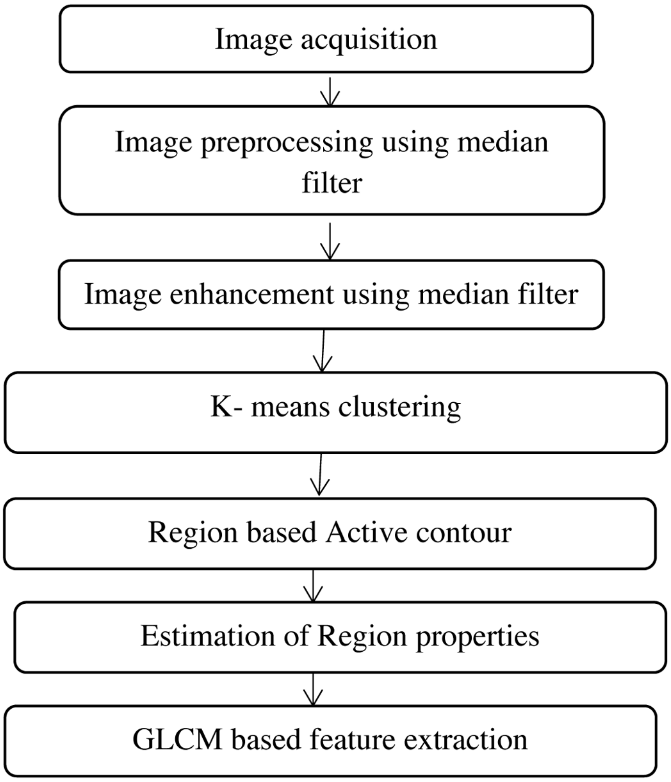
Image Preprocessing

The After getting MRI Image preprocessing is the first step of segmentation process, which helps to reduce noise and artifacts from the acquired image. The 0.02 amount of salt and pepper noise is added to the acquired image. It can be replaced by median filter. Since we have many types of filters to remove noise, but median filter was mostly used due to the following reasons.

- The edge information should be preserved during preprocessing
- The median value is estimated by considering the value of an all pixels which is available in the image.
- Nonlinear digital filtering method is used.

The following steps should be followed to calculate the median value. Consider the particular pixel, for which the median pixel is calculated by organizing all the pixels which is near to that. For an even number of pixels median value is calculated by taking the average of two center pixel values is calculated to find median value. Similarly for Odd value the center value of the pixels are median. Finally the superiority of the MRI, spinal image is enhanced by conserving the edges.

Figure 1. Flow diagram of proposed model



K- Means Clustering

The flow diagram for K-means algorithm are denoted in the following figure

The algorithm is presented by Macqueen in the year 1997. It is an unsupervised algorithm. It begins by randomly allocating the K total number of cluster center. The cluster is calculated and it is named as centroid. Compare the every pixel with neighboring pixels.

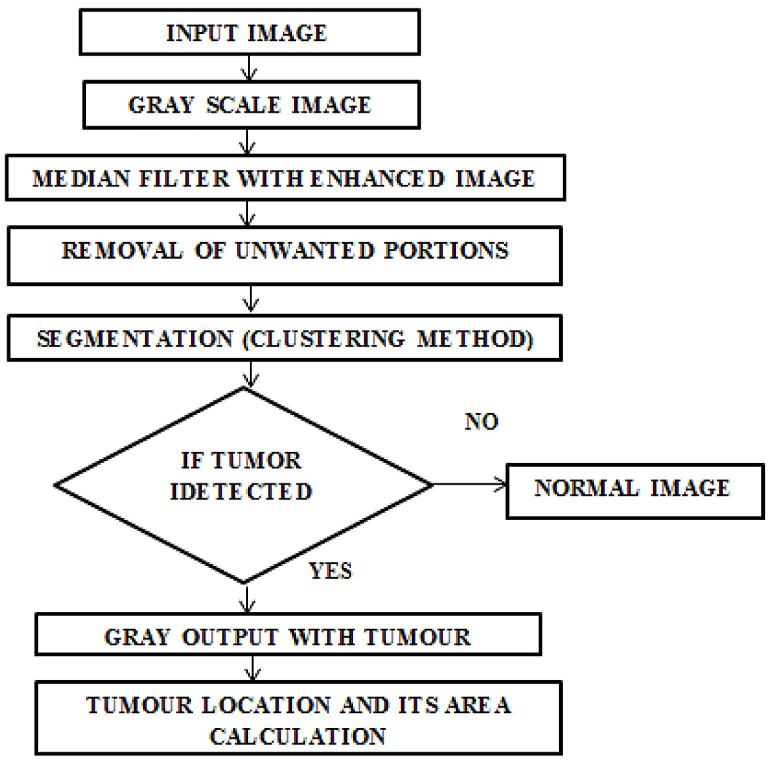
(Bjoern 2016 et al.,) Further, each pixel has shifted to nearest cluster center, which has the shortest distance among all. This procedure is continuing till the center converges. The several steps of the algorithms are described as follows.

Step1: Arbitrarily choose the C cluster center.
Step 2: Evaluate Euclidean distance for all pixel to cluster centre.
Step 3: Each pixel is assigned to particular cluster, which has the smallest distance
Step 4: The algorithm helps to diminish the squared error,

$X_i - V_i$ is the Euclidean distance between X_i, V_i.
C is the number of clusters.
C_i is the number of data points in the i^{th}cluster. Next, calculate the cluster centre by using the following formula

$$C = \frac{1}{C_i}\sum_{j=1}^{C_i} X_i$$

Figure 2. Flow chart for K means algorithm



The advantages are

- Implementation is easy.
- Easy to understand

The disadvantages are

- Care should be taken to choose the Proper K value.
- It is immune to noise and artifacts.

Clustering finds numerous applications in a variety of areas like image processing, data mining, Image retrieval, pattern recognition, Image segmentation etc.

REGION BASED ACTIVE CONTOUR SEGEMENTATION

Active Contours

Active contouris one of the segmentation methods; it uses energy forces and constraints for separation of the pixels of interest from the image for further processing. It uses snake model, gradient vector flow snake model, balloon model and geometric contours. Active contour model performs the segmentation on MRI, CT, PET and SPECT images. The early diagnosis and detection of abnormalities of affected region can be evaluated with the help of active contour models in 3-D imaging. This model provides accurate results for 3-D CT and MRI images when comparing with other methods. Segmentation of fine structures from the affected object in an image is possible with the help of active contour models.

Region based segmentation check for the similarity of pixels based on the properties like intensity, color and texture. These variation findings can be used by active contour segmentation. For object detection, active contour model uses 33 types of curve evolution. Active contour is also named deformable model. It is presented by Kass et al in 2-D space, and it can be upgraded for 3 –D space by Terzopoulos et al. An Active contour or snakes used in 2D space, and balloons used in 3 D space. In the presence of external force, the parametric curve transfers inside the image to detect the boundaries of the object. (Georges et al., 1999) Discussed the geometry; physics related information related of each pixels can be obtained by deformable model. The above mentioned properties provide the variation of shape over space and time.

Step 1: Place the Active contour or snake neighboring to the abnormal region.
Step 2: In the presence of inside and outside forces produced in the image the snake is relocated immediately to the target by an iterative process.
Step 3: Estimate the energy function of the forces.
Step 4: The main purpose of this technique is to reduce the energy function. The data can be smoothened by internal forces and the contours are shifted to next to the region of interest.

Region Properties

Eman et al. (2015) shows that the particular portion of an image has several properties like perimeter, area, boundary box, major axis, centroid, eccentricity, convex, area, etc. The property of the certain region represents the mathematical feature of the specific portion of an image.

A region in an image can have many properties like area, perimeter, boundary box, major axis length, minor axis length, centroid, eccentricity, filled image, orientation, convex area, convex image, pixel list, solidity, Euler number, filled area, extrema, subarray. Fundamentally the Region Properties indicate the mathematical features of a particular region of image.

Area

It is the total amount of pixels in this specified portion. It is returned as a scalar

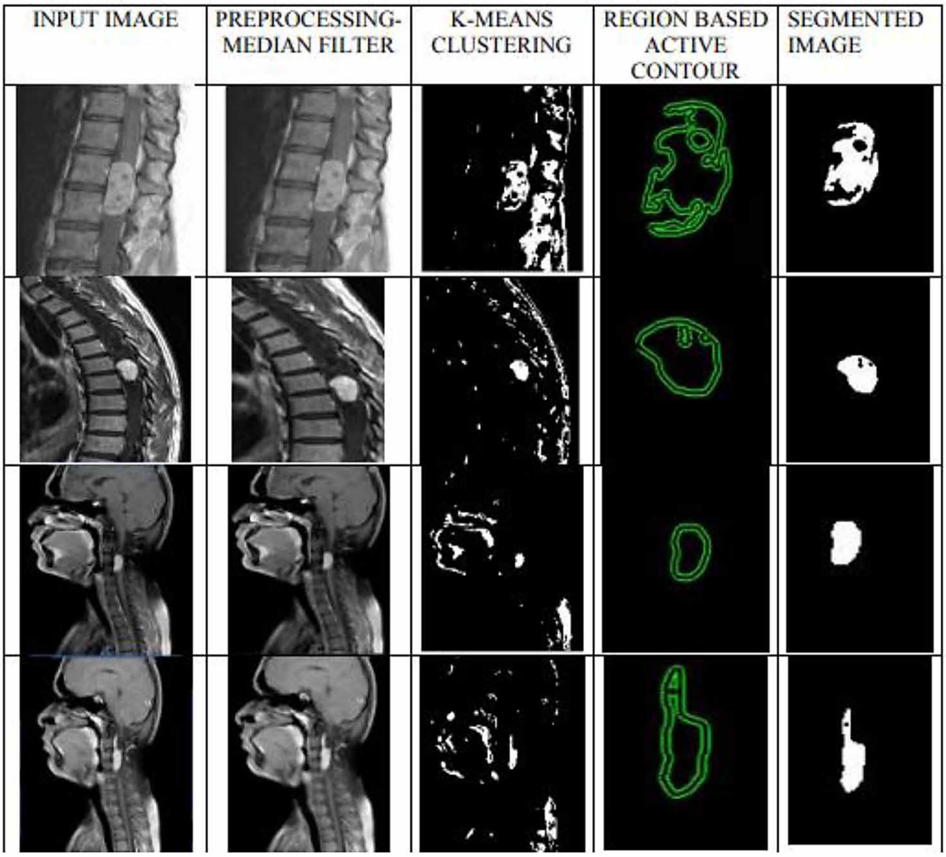
Perimeter

It is defined as the distance around boundary of the specific area, defined as a scalar. The perimeter is estimated by measuring the distance among the all adjacent pair of pixels around the boundary of the specific portion. In vase of discontinuous region the command provides unexpected value.

FEATURE EXTRACTION USING GLCM

(Hai, S et al., 2016) The organization of pattern is made to be easy with the help of feature extraction. It provides the information which is related to the shape of an image. It reduces the number of resources required to classify the pattern. The proposed work uses GLCM feature (Gray Level Co-occurrence Matrix) extraction.

Figure 3. Results of clustering and active contour segmentation



GLCM matrix provides a statistical method which examines the texture of pixel by considering a spatial relationship. (Mohammad Fathy et al., 2019) It characterizes the quality of an image by estimating how frequently the pair of pixels with a particular value occurs in the specified spatial relationship of an image.

A GLCM is a matrix, in which has equal number of rows and columns for gray level of an image. The matrix element P (i, j | Δx, Δy), in which the two pixels has relative frequency, separated by a distance (Δx, Δy) with the neighbourhood one with intensity 'j'.

For variations in the gray levels of 'i' and 'j' at a certain displacement distance d and at a specific angle ɵ m the matrix element P (i, j | d, ɵ) provides the second order statistical probability values.

GLCM is very sensitive to dimensions. The proposed extract some of the feature using GLCM. It can be listed as follows

Contrast

The difference in intensity contrast between the pixel and to adjacent pixels over the entire image is represented by contrast.

$$C = \left|i - j\right|^2 p\left(i, j\right) \tag{1}$$

Correlation

The relationship between the pixels and its neighbourhood pixel is represented by correlation. The range of the value lies between (-1, 1). The correlation value is 1 for the positively correlated image, -1 for the negatively correlated image.

$$corr = \sum \frac{\left(i - \mu_i\right)\left(j - \mu_j\right)\left(p\left(i, j\right)\right)}{\sigma_i \sigma_j} \tag{2}$$

$\mu_{i,}$ μ_j - Mean, $\sigma_{i\,,}\,\sigma_j$ - Standard deviation
Pi, Pj- Partial probability function

Energy

Energy is defined as a sum of squared elements in the GLCM (Gray Level Co-occurrence Matrix) Energy is also mentioned as the quantity of uniformity. The range of the value is between [0 1], for a constant, the energy value is 1

$$E = \sum_{i,j} P\left(i, j\right)^2 \tag{3}$$

Entropy

It is used to represented as the uncertainty of the textural image.

$$h = -\sum_{k=0}^{L-1} p_{lk(\log_2 p_{lk})} \tag{4}$$

p_{lk}= probability of the K^{th}level

The extracted features can be used by different classifiers like in such a way to help the neural network system to classify the abnormality of tumour affected portion.

Table 1. Feature extraction table

Input Images	Contrast	Correlation	Energy	Entropy
Image 1	0.205	0.742	0.8422	0.153
Image 2	0.901	0.756	0.924	0.531
Image 3	0.029	0.688	0.893	0.026
Image 4	0.291	0.7801	0.883	0.368

RESULTS AND DISCUSSION

Many of the brain tumour segmentation methods uses K means, Fuzzy C means for separating tumour regions from the affected regions. While doing segmentation using above two segmentation there is a possibility of over segmentation and false segmentation happening in the presence of noise. Hence from the detailed review hybrid clustering helps to overcome the disadvantages along with the BPN or SVM classifier. It provides the efficiency of 93.28%. Further to perform the segmentation on 3-D images we used active contour segmentation for our proposed work. The accuracy and overall performance of the segmentation was improved with the help of CNN. From this motivation the research work uses K means and Active contour segmentation methods for spine tumour segmentation.

CONCLUSION

Initially the proposed work starts with a conventional segmentation algorithm called K-means clustering. Since much of the research work for segmentation of tumour portion is designed using clustering techniques. Because the method is easy to understand and implement. But it finds many disadvantages in the presence of artifacts, which occurs during MRI scan. From the motivation we have started the work to find the tumor in the spinal cord is done by K means clustering. The disadvantage of this technology is overcome through the active contour model. Since it is compulsory to find the exact boundary of affected portions for feature extraction.

REFERENCES

Bjoern, H. M., & Koen, V. L. (2016). A Generative Probabilistic Model and Discriminative Extensions for Brain Lesion Segmentation With Application to Tumor and Stroke. *IEEE Transactions on Medical Imaging*, *35*(4), 933–946. doi:10.1109/TMI.2015.2502596 PMID:26599702

Eman, A. M., Mohammed, E., & Rashid, A. L. (2015). Brain tumor segmentation based on a hybrid clustering technique. *Egyptian Informatics Journal*, *16*(1), 71–81. doi:10.1016/j.eij.2015.01.003

Fathy, M., Keshk, M., & El Sherif, A. (2019). Surgical management and outcome of intramedullary spinal cord tumour. *Egyptian Journal of Neurosurgery.*, *34*(2), 2–7. doi:10.118641984-019-0028-9

Georges, B. A. (1999). Model Creation and Deformation for the Automatic Segmentation of the Brain in MR Images. *IEEE Transactions on Biomedical Engineering*, *46*(11), 1346–1356. doi:10.1109/10.797995 PMID:10582420

Hai, S., Xing, F., & Yang, L. (2016). Robust Cell Detection of Histopathological Brain Tumor Images Using Sparse Reconstruction and Adaptive Dictionary Selection. *IEEE Transactions on Medical Imaging*, *35*(6), 1575–1586. doi:10.1109/TMI.2016.2520502 PMID:26812706

Hai, S., Xing, F., & Yang, L. (2016). Robust Cell Detection of Histopathological Brain Tumor Images Using Sparse Reconstruction and Adaptive Dictionary Selection. *IEEE Transactions on Medical Imaging*, *35*(6), 1575–1586. doi:10.1109/TMI.2016.2520502 PMID:26812706

Juang, L. H., & Ming, N. W. (2010). MRI brain lesion image detection based on color-converted K-means. *Science Direct Measurement*, *43*(7), 941–949.

Madhukumar, S., & Santhiyakumari, N. (2015). Evaluation of k-Means and fuzzy C-means segmentation on MR images of brain. *Egyptian Society of Radiology and Nuclear Medicine*, *46*(2), 475–479. doi:10.1016/j.ejrnm.2015.02.008

Murathoti Varshini, Barjo, & Tigga. (2020). Spine Magnetic Resonance Image Segmentation Using Deep Learning Techniques. *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS*).

Seo, H., & Khuzani, M. B. (2020). Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. *American Association of Physicists in Medicine*, *45*(5), 148–167. PMID:32418337

Shen, S., Sandham, W., Granat, M., & Sterr, A. (2005). MRI Fuzzy Segmentation of Brain Tissue using Neighborhood Attraction with Neural-Network Optimization. *IEEE Transactions on Information Technology in Biomedicine*, *9*(3), 459–497. doi:10.1109/TITB.2005.847500 PMID:16167700

Yezzi, A. J., Kichenassamy, S., Kumar, A., Olver, P., & Tannenbaum, A. (1997). A geometric snake model for segmentation of medical imagery. *IEEE Transactions on Medical Imaging*, *16*(2), 199–209. doi:10.1109/42.563665 PMID:9101329

Chapter 12
A Survey on Early Detection of Women's Breast Cancer Using IoT

P. Malathi
Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

A. Kalaivani
Saveetha School of Engineering, India & Saveetha Institute of Medical and Technical Sciences, Chennai, India

ABSTRACT

The internet of things is probably one of the most challenging and disruptive concepts raised in recent years. Recent development in innovation and availability have prompted the rise of internet of things (IoT). IoT technology is used in a wide scope of certified application circumstances. Internet of things has witnessed the transition in life for the last few years which provides a way to analyze both the real-time data and past data by the emerging role. The current state-of-the-art method does not effectively diagnose breast cancer in the early stages. Thus, the early detection of breast cancer poses a great challenge for medical experts and researchers. This chapter alleviates this by developing a novel software to detect breast cancer at a much earlier stage than traditional methods or self-examination.

DOI: 10.4018/978-1-7998-3092-4.ch012

INTRODUCTION

Breast cancer the second leading cause of death for women. Breast cancer is cancer that develops in breast cells. Typically, the cancer forms in either the lobules or the ducts of the breast. Lobules are the glands that produce milk, and ducts are the pathways that bring the milk from the glands to the nipple. Cancer can also occur in the fatty tissue or the fibrous connective tissue within your breast. Among women, breast cancer is the most second most common cancer diagnosed, after skin cancer, and the second leading cause of cancer deaths, after lung cancer. On average, 1 in 8 women will develop breast cancer in their lifetimes. About two-thirds of women with breast cancer are 55 or older. Most of the rest are between 35 and 54. Fortunately, breast cancer is very treatable if you spot it early. Localized cancer (meaning it hasn't spread outside your breast) can usually be treated before it spreads. Once the cancer begins to spread, treatment becomes more complicated. It can often control the disease for years.

Symptoms of Breast Cancer

In Early stages, breast cancer may not cause any symptoms. In many cases, a tumor may be too small to be felt, but an abnormality can still be seen on a mammogram. If a tumor can be felt, the first sign is usually a new lump in the breast that was not there before. However, not all lumps are cancer. Each type of breast cancer can cause a variety of symptoms.

- A breast lump or tissue thickening that feels different than surrounding tissue and has developed recently
- Breast pain
- Red, pitted skin over your entire breast
- Swelling in all or part of your breast
- A nipple discharge other than breast milk
- Bloody discharge from your nipple
- Peeling, scaling, or flaking of skin on your nipple or breast
- A sudden, unexplained change in the shape or size of your breast
- Inverted nipple
- Changes to the appearance of the skin on your breasts
- A lump or swelling under your arm

Treatment of Breast Cancer

Breast cancer is treated in several ways. It depends on the kind of breast cancer and how far it has spread. People with breast cancer often get more than one kind of treatment.

- *Surgery*: An operation where doctors cut out cancer tissue.
- *Chemotherapy*: Using special medicines to shrink or kill the cancer cells. The drugs can be pills you take or medicines given in your veins, or sometimes both.
- *Hormonal therapy*: Blocks cancer cells from getting the hormones they need to grow.
- *Biological therapy*: Works with your body's immune system to help it fight cancer cells or to control side effects from other cancer treatments.
- *Radiation therapy*: Using high-energy rays (similar to X-rays) to kill the cancer cells.

BREAST CANCER SCREENING TEST

Digital Mammogram

A mammogram is an X-ray of the breast. Mammograms are the best way to find breast cancer early, when it is easier to treat and before it is big enough to feel or cause symptoms. Having regular mammograms can lower the risk of dying from breast cancer. At this time, a mammogram is the best way to find breast cancer for most women.

Magnetic Resonance Imaging (MRI)

A breast MRI uses magnets and radio waves to take pictures of the breast. MRI is used along with mammograms to screen women who are at high risk for getting breast cancer. Because breast MRIs may appear abnormal even when there is no cancer, they are not used for women at average risk.

Clinical Breast Exam

A *clinical breast exam* is an examination by a doctor or nurse, who uses his or her hands to feel for lumps or other changes. The benefit of screening is finding cancer early, when it's easier to treat. Harms can include **false positive test results,** when a doctor sees something that looks like cancer but is not. This can lead to more tests, which can be expensive, invasive, time-consuming, and may cause anxiety.

Early Detection of Women Breast Cancer Techniques

Newer diagnostic techniques such as sestamibi scans, optical imaging and molecular diagnostic techniques look promising, but need more investigation into their use. Their roles will appear clearer in coming years, and they may prove to be of help in further investigating lesions that are indeterminate on standard imaging. Other upcoming techniques are contrast-enhanced mammography and tomosynthesis. These may give additional information in indeterminate lesions, and when used in screening they aid in reducing recall rates, as shown in recent studies. Tomography has a role in detecting local disease recurrence and distant metastasis in breast cancer patients. Computer-aided detection (CAD) is a software technology that has become widespread in radiology practices, particularly in breast cancer screening for improving detection rates at earlier stages. Many studies have investigated the diagnostic accuracy of CAD. The current level of performance for the CAD systems is encouraging but not enough to make CAD systems standalone detection and diagnose clinical systems. Unless the performance of CAD systems enhanced dramatically from its current level by enhancing the existing methods. Traditionally, doctors have relied on mammograms to detect changes in breast tissue that could indicate the growth of cancerous tissue. Unfortunately, mammograms are not always accurate, particularly for women with dense breast tissue. However, an innovative new "Internet of Things" (IoT). In the future, that home may be filled with IoT-enabled devices that could transmit patient-generated health data to their doctors. The information obtained from these devices could include vitals such as heart rate, pulse ox, and respiratory rate.

In addition, IoT-enabled pillboxes, appliances, and even toothbrushes could also generate a plethora of useful data. Yet other devices will detect time in bed, falls, and even gait. All of this information will give clinicians (and family members) a better idea of how patients are faring at home. For example, if IoT devices detect that the patient hasn't left their bed in a number of days nor opened their pill box in a week, the system could alert their physician to take appropriate measures to check on their patient.

Women Breast Cancer Detection Using Wearable Technology

In an innovative IoT wearable technology application, Cyrcadia Health is developing a bra that could alert women to the early signs of breast cancer. A method to detect breast cancer early. The accurate and efficient diagnosis of breast cancer is extremely necessary for recovery and treatment in early stages in IoT healthcare environment. However, that is not where they're spending the majority of their day. That place is their home. An iTBra and a vision of using wearable technology to drastically improve for less money as well. In the future, that home may be filled with IoT-enabled devices that could transmit patient-generated health data to their doctors. The information obtained from these devices could include vitals such as heart rate, pulse ox, and respiratory rate. In addition, IoT-enabled pillboxes, appliances, and even toothbrushes could also generate a plethora of useful data. Yet other devices will detect time in bed, falls, and even gait.

All of this information will give clinicians (and family members) a better idea of how patients are faring at home. For example, if IoT devices detect that the patient hasn't left their bed in a number of days nor opened their pill box in a week, the system could alert their physician to take appropriate measures to check on their patient.

Figure 1. Workflow for IoT bra

RELATED WORK

Internet of Things World Forum, Now-a-days hearing a lot about the transformational value of the Internet of Things (IoT) across many industries manufacturing, transportation, agriculture, smart cities, retail, and finance. So many new solutions are on display that help organizations either save or make money. But in healthcare, IoT can actually do more than that, it has the potential to save lives.

Lucia Arcarisi et. al.,(2019) in their research study, a non-invasive wearable device designed to mimic the process of breast self-examination. It uses pressure sensing textiles and thus increase the confidence and self-awareness of women. Combined with other screening methods, the device can increase the odds of early detection for better prognosis. The research work demonstrates that it can detect nodules in much the same way as does the human hand during breast self-examination.

Muhammad Hammad Memon et. al., (2019) proposed machine learning-based diagnostic system which effectively classifies the malignant and benign people in the environment of IoT. The proposed system performance is excellent due to the selection of recursive feature selection algorithm. The implementation of the proposed system is very reliable in all aspects of IoT healthcare for breast cancer.

Md. Milon Islam et.al.,(2020) proposed a smart healthcare system in IoT environment that can monitor a patient's basic health signs and room condition of the patients in real-time. In this system, five sensors are used to capture the data from hospital environment named heart beat sensor, body temperature sensor, room temperature sensor, CO sensor, and CO_2 sensor. The condition of the patients is conveyed via a portal to medical staff, where they can process and analyze the current situation of the patients. The developed prototype is well suited for healthcare monitoring that is proved by the effectiveness of the system.

Javier Andreu-Perez et. al., (2015) present the milestones and recent developments of different generations of pervasive sensing applications for health monitoring. The opportunities of pervasive health monitoring through data linkages with other health informatics systems including the mining of health records, clinical trial databases, multi-omics data integration and social media. Technical advances have supported the evolution of the pervasive health paradigm towards preventative, predictive, personalised and participatory medicine.

M. Sung et. al., (2005) describe LiveNet, a flexible wearable platform intended for long-term Ambulatory health monitoring with real-time data streaming and context classification. LiveNet is a stable, accessible system that combines inexpensive, commodity hardware;, a flexible sensor interconnection bus and a powerful, light-weight distributed sensing, classification, and inter-process communications software architecture to facilitate the development of distributed real-time multi-modal and context-aware applications. The paper demonstrate the power and functionality of this platform by describing a number of health monitoring applications using the LiveNet system in a variety of clinical studies.

Kim Gau Ng et. al.,(2012) proposed Cadi ThermoSENSOR skin-contact thermometer measures body temperature continuously and transmits readings wirelessly to a central server. This study evaluated the ThermoSENSOR against ear temperatures (ETs) measured by a Braun ThermoScan ear thermometer and axillary temperatures (ATs) measured by a Terumo digital clinical thermometer. These results suggest that the TTs were comparable to the ETs and ATs.

This study employed a recognized protocol of activities both pre-operatively, and at regular intervals up to twenty-four weeks post-total knee arthroplasty. The results suggest that a wearable miniaturised ear-worn sensor is potentially useful in monitoring post-operative recovery, and in identifying patients who fail to improve as expected, thus facilitating early clinical review and intervention.

Direct interfacing of nanosensors onto biomaterials could impact health quality monitoring and adaptive threat detection. The paper demonstrates the integration onto a tooth for remote monitoring of respiration and bacteria detection in saliva. Overall, this strategy of interfacing graphene nanosensors with biomaterials represents a versatile approach for ubiquitous detection of biochemical targets.

M.-Z. Po et. al., (2010) evaluated the choice of electrode material by comparing conductive fabric with Ag / AgCl electrodes and discuss the limitations found. The evidence given in this work is a viable alternative to the traditional palmar sites for EDA measurements. Our device offers the unprecedented ability to perform comfortable, long-term, and in situ assessment of EDA. This paper opens up opportunities for future investigations that were previously not feasible, and could have far-reaching implications for diagnosis and understanding of psychological or neurological conditions.

S.Patel (2009) presents the results of a pilot study to assess the feasibility of using accelerometer data to estimate the severity of symptoms and motor complications in patients with Parkinson's disease. A support vector machine (SVM) classifier was implemented to estimate the severity from accelerometer data features. SVM-based estimates were compared with clinical scores derived via visual inspection of video recordings taken while patients performed a series of standardized motor tasks. The outcome of the proposed work is a thin, comfortable device technology that can softly laminate on to the surface of the skin to enable advanced, multifunctional operation for physiological monitoring in a wireless mode.

H. Zhou et. al., (2008) presents a new human motion tracking system using two wearable inertial sensors that are placed near the wrist and elbow joints of the upper limb. Each inertial sensor consists of a tri-axial accelerometer, a tri-axial gyroscope and a tri-axial magnetometer. The turning rates of the gyroscope are utilised for localizing the wrist and elbow joints on the assumption that the two upper limb segment lengths. Experimental results demonstrate that this new system, compared to an optical motion tracker, has RMS position errors that are normally less than 0.01 m, and RMS angle errors that are 2.5-4.8 degrees.

E. S. Sazonov (2011) presents a novel wearable sensor capable of very accurate recognition of common postures and activities. The shoe sensor was tested in nine adults performing sitting and standing postures and while walking, running, stair ascent/descent and cycling. Support vector machines (SVMs) were used for classification. A fourfold validation of a six-class subject-independent group model showed 95.2% average accuracy of posture/activity classification on full sensor set and over 98% on optimized sensor set. The proposed methods are implemented by using signals from sensor-embedded shoes called smart shoes. Each smart shoe has four GCF sensors installed between the cushion pad and the sole. The GCF sensor applies an air pressure sensor connected to an air bladder. A gait monitoring system that integrates the proposed methods is shown in this paper and verified for both normal and abnormal gaits.

Flynn (2013) specified Rheumatoid Arthritis (RA) is a disease which attacks the synovial tissue lubricating skeletal joints. This systemic condition affects the musculoskeletal system, including bones, joints, muscles and tendons that contribute to loss of function and Range of Motion (ROM). Traditional measurement of arthritis requires labour intensive personal examination by medical staff which through their objective measures may hinder the enactment and analysis of arthritis rehabilitation.

IOT Wearable for Breast

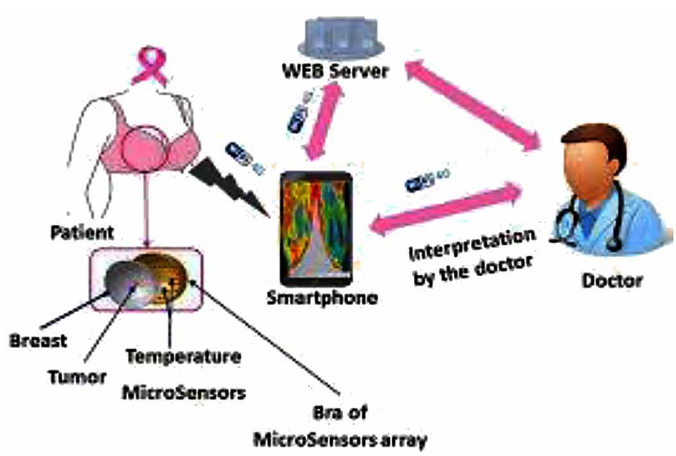
In an innovative IoT wearable technology application, Cyrcadia Health is developing a bra that could alert women to the early signs of breast cancer. Cancer cells exhibit abnormal temperature patterns, because the increased activity of the cancer cells generates heat. The bra has built-in sensors that read cell temperatures and transmit the data in real-time to a patient database. This Wi-Fi enabled garment contains 16 sensors that can detect changes in the wearer's breasts. After wearing the bra for two hours, the data is transmitted directly to the patient's physician. This data is then paired with a predictive algorithm that analyzes the information for known risk factors. While it's no replacement for an annual mammogram, this IoT bra could alert patients and physicians to changes between their yearly appointments.

An abnormal reading triggers an alert which is sent via smartphone to the patient and their doctor. Currently undergoing medical trials, to date the iT Bra has been tested with 500 patients and has demonstrated an 87 percent correlation to a verified, clinical diagnosis of breast cancer. In order for the cell temperature readings to be taken, women wear the iT Bra for 12 hours and it is being presented as an alternative to a monthly breast examination. Early detection and treatment has been shown to improve breast cancer survival rates and Cyrcadia believes its technology could reduce the number of unnecessary breast biopsies by up to 50 percent, especially among women who have dense breast tissue, for whom mammography often doesn't work.

A key feature of our proposed detection concept is the simultaneous sensing of tissue property changes to the two female breasts since the right and left healthy breasts are morphologically and materially identical. It's developed a breast health system made up of two parts. The first is a sensor device that's placed in a bra to measure cell temperature changes created over time by new blood vessel growth associated with developing tumors. The second is proprietary software that uses pattern recognition, chronology and artificial intelligence to look for changes in breast tissue that could indicate the presence of a tumor. Doctors and researchers say the size of breast cancer and how far it has spread are some of the most important factors in determining the prognosis of the more than 225,000 women alone who are diagnosed with invasive breast cancer each year. Recent advancements in breast imaging like 3D mammography and thermography may help in detection and diagnosis, but there are still flaws in the gold standard of mammography (which have stirred up some debate about its value as an annual screening method for women over 40).

The sensors are contained in a patch that attaches with an adhesive to the patient's skin. The patient wears the iTBra for 2 hours, and the data collected is sent directly to her physician for analysis. It's an alternative to the discomfort of a mammogram, and it's especially helpful for women with dense breast tissue, such as Royea's own wife, Kelli Royea, who is featured in the documentary.

Figure 2. Self Examination using IOT Sensors



Mammography is the current gold standard diagnostic tool for breast cancer screening. But screening mammography has an important limitation: its results are significantly less accurate in women with dense breast tissue. Breast tissue density is a recognized medical condition which affects more than 40% of women worldwide. Dense breast tissue is comprised of less fat and more connective/fibrous and glandular tissue, and ranges in severity from Level A (fatty) to Level D (extremely dense). As the density of a breast increases, the ability of the mammogram to reveal cancer decreases. Because both dense breast tissue and breast cancer appear white on mammography images, finding cancer in these dense tissue breasts is akin to looking for a distinct snowflake in a snowstorm. The cancer risk in women with extremely dense breasts is up to 6 times higher compared to normal/fatty tissue, and shows a much more rapid acceleration of the condition. Still, 70% of breast biopsies that are conducted as a result of suspicious findings on a mammogram are performed on non-cancerous tissue. The staggering numbers of such unnecessary biopsies could be reduced with improved diagnostic screening in women with dense breast tissue.

One recent development, however, has really captured my attention because of its potential to help with early detection of breast cancer. That's the idea behind the connected bra—dubbed the" iTBra" by its inventor, Rob Royea, CEO of Cyrcadia Health. With embedded temperature sensors, this new kind of wearable technology tracks changes in temperature in breast tissue over time. It uses machine learning and predictive analytics to identify and classify abnormal patterns that could indicate early stage breast cancer. And women will be glad to know that they just need to wear the iTBra for 2 to 12 hours once a month as they go about their daily activities—there's no painful squashing or prodding or radiation involved. Internet of Things (IoT) across many industries – manufacturing, transportation, agriculture, smart cities, retail, and finance. So many new solutions are on display that help organizations either save or make money. But in healthcare, IoT can actually do more than that, it has the potential to save lives. The Internet of Things (IoT) also has the potential to bring a huge transformation to the Healthcare industry. For the starters, it promises to reduce the emergency room wait time, track patient data accurately, and manage healthcare inventory. All these functions will improve the efficiency of the healthcare sector to an unimaginable extent.

The iTBra is comprised of two wearable breast sensor patches; flexible and intelligent, they detect dynamic circadian temperature changes within breast tissue. In contrast with mammography and ultrasound, which rely on specular reflective capabilities of varying tissue densities, the Cyrcadia Health solution is tissue agnostic and is able to detect early circadian cellular changes in all tissue types and varied age groups. The abnormalities in the circadian rhythm-based temperature variances of cell cycles are present at the earliest stages of abnormal cellular growth and proliferation, and serve as the early indicators of breast cancer. Exclusively patented predictive analytics technology identifies these abnormal cellular changes and reports results to the health care provider, used to assist in their clinical decision process.

CONCLUSION

These methods are impractical to be used as personal monitoring device due to its high cost and uncomfortable procedures on the patient. Therefore, this study proposes the use of multiple sensors positioned on brassiere cloth covering all four quadrants to provide continuous monitoring of temperature changes on the breasts. To test the reliability of the developed device, breast phantom and heater were used to mimic women breasts and the tumor respectively. Camera was used to verify the changes of surface temperature on breast phantom. Result obtained shows that the reading of sensors on each quadrant with a tumor to have higher temperature compared to the rest. Its indicates that this low-cost wearable device can be potentially used as breast cancer monitoring tool.

REFERENCES

Andreu-Perez, J., Leff, D. R., Ip, H. M. D., & Yang, G.-Z. (2015). From Wearable Sensors to Smart Implants – Towards Pervasive and Personalised Healthcare. *IEEE Transactions on Biomedical Engineering*, *62*(12), 2750–2762. doi:10.1109/TBME.2015.2422751 PMID:25879838

Arcarisi & Di Pietro. (2019). Palpreast—A New Wearable Device for Breast Self-Examination. *Appl. Sci.*, *9*(3).

Atallah, L. (2011). *Observing recovery from knee-replacement surgery by using wearable sensors*. IEEE Digital Xplore.

Kong, K., & Tomizuka, M. (2009). A gait monitoring system based on air pressure sensors embedded in a shoe. *IEEE/ASME Transactions on Mechatronics*, *14*(3), 358–370. doi:10.1109/TMECH.2008.2008803

Liana, D. D., Raguse, B., Gooding, J. J., & Chow, E. (2012). Recent Advances in Paper-Based Sensors. *Sensors (Basel)*, *12*(9), 11505–11526. doi:10.3390120911505 PMID:23112667

Mannoor, M. S., Tao, H., Clayton, J. D., Sengupta, A., Kaplan, D. L., Naik, R. R., Verma, N., Omenetto, F. G., & McAlpine, M. C. (2012). Graphene-based wireless bacteria detection on tooth enamel. *Nature Communications*, *3*(1), 1–8. doi:10.1038/ncomms1767 PMID:22453836

Memon & Li. (2019). Breast Cancer Detection in the IOT Health Environment Using Modified Recursive Feature Selection. *Wireless Communications and Mobile Computing*, 1-19.

Milon Islam, Md., & Rashedul Islam, Md. (2020). Development of Smart Healthcare Monitoring System in IoT Environment. *SN Computer Science*, *185*, 1–11.

Ng, K.-G., Wong, S.-T., Lim, S.-M., & Goh, Z. (2010). Evaluation of the cadi thermosensor wireless skin-contact thermometer against ear and axillary temperatures in children. *Journal of Pediatric Nursing-Nursing Care of Children & Families*, *25*(3), 176–186. doi:10.1016/j.pedn.2008.12.002 PMID:20430278

O'Flynn, J. (2013). Novel smart sensor glove for arthritis rehabilitation. *BSN*, 1-6.

Patel, S., Lorincz, K., Hughes, R., Huggins, N., Growdon, J., Standaert, D., Akay, M., Dy, J., Welsh, M., & Bonato, P. (2009). Monitoring motor fluctuations in patients with parkinson's disease using wearable sensors. *IEEE Transactions on Information Technology in Biomedicine*, *13*(6), 864–873. doi:10.1109/TITB.2009.2033471 PMID:19846382

Po, M.-Z., Swenson, C., & Rosalind, W. (2010). A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity. *IEEE Transactions on Biomedical Engineering*, *57*(5), 1243–1252. doi:10.1109/TBME.2009.2038487 PMID:20172811

Sazonov, E. S., Fulk, G., Hill, J., Schutz, Y., & Browning, R. (2011). Monitoring of posture allocations and activities by a shoe-based wearable sensor. *IEEE Transactions on Biomedical Engineering*, *58*(4), 983–990. doi:10.1109/TBME.2010.2046738 PMID:20403783

Sung, M., Marci, C., & Pentland, A. (2005). Carl Marci, Alex Pentland, "Wearable feedback systems for rehabilitation. *Journal of Neuroengineering and Rehabilitation*, *2*(1), 1–12. doi:10.1186/1743-0003-2-17

Xu, S., Zhang, Y., Jia, L., & Kyle, E. (2014). Soft microfluidic assemblies of sensors, circuits, and radios for the Skin. *Science*, *344*(6179), 70–74. doi:10.1126cience.1250169 PMID:24700852

Zhou, H., Stone, T., Hu, H., & Harris, N. (2008). Use of multiple wearable inertial sensors in upper limb motion tracking. *Medical Engineering & Physics*, *30*(1), 123–133. doi:10.1016/j.medengphy.2006.11.010 PMID:17251049

Compilation of References

Aaron, J. S., Taylor, A. B., & Chew, T. L. (2018). Image co-localization–co-occurrence versus correlation. *Journal of Cell Science*, *131*(3), jcs211847. doi:10.1242/jcs.211847 PMID:29439158

Abdel-Zaher, A. M., & Eldeib, A. M. (2016). Breast cancer classification using deep belief networks. *Expert Systems with Applications*, *46*(1), 139–144. doi:10.1016/j.eswa.2015.10.015

Aggarwal, C. C. (2015). Outlier analysis. Data Mining, 237-263. doi:10.1007/978-3-319-14142-8_8

Aggarwal, M. K., & Khare, V. (2015). Automatic localization and contour detection of Optic disc. *2015 International Conference on Signal Processing and Communication (ICSC)*. 10.1109/ICSPCom.2015.7150686

Aggarwal, R., & Kaur, A. (2012). Comparative Analysis of Different Algorithms For Brain Tumor Detection. *International Journal of Scientific Research*.

Alharkan, T., & Martin, P. (2012). IDSaaS: Intrusion detection system as a service in public clouds. *Proceedings of the 2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (ccgrid 2012)*, 686-687. 10.1109/CCGrid.2012.81

Aloudat & Faezipour. (2016). Determination for Glaucoma Disease Based on Red Area Percentage. *2016 IEEE Long Island Systems*, *Applications and Technology Conference (LISAT)*.

American Cancer Society. (2019). *Cancer Facts & Figures 2019*. American Cancer Society.

Andreu-Perez, J., Leff, D. R., Ip, H. M. D., & Yang, G.-Z. (2015). From Wearable Sensors to Smart Implants – Towards Pervasive and Personalised Healthcare. *IEEE Transactions on Biomedical Engineering*, *62*(12), 2750–2762. doi:10.1109/TBME.2015.2422751 PMID:25879838

Anthony, M., & Bartlet, P. (1999). *Neural Network Learning: Theoretical Foundations*. Cambridge University Press. doi:10.1017/CBO9780511624216

Arafi, A., Safi, Y., Fajr, R., & Bouroumi, A. (2013). Classification of Mammographic Images using Artificial Neural Network. *Applied Mathematical Sciences*, *7*(89), 4415–4423. doi:10.12988/ams.2013.35293

Araujo, M., Queiroz, K., Pininga, M., Lima, R., & Santos, W. (2012). Uso de regiões elipsoidais como ferramenta de segmentação em termogramas de mama. In *XXIII Congresso Brasileiro de Engenharia Biomédica (CBEB 2012)*. Pernambuco: SBEB.

Arcarisi & Di Pietro. (2019). Palpreast—A New Wearable Device for Breast Self-Examination. *Appl. Sci.*, *9*(3).

Aslam, T. M., Tan, S. Z., & Dhillon, B. (2009). Iris recognition in the presence of ocular disease. *Journal of the Royal Society, Interface*, *6*(34), 2009. doi:10.1098/rsif.2008.0530 PMID:19324690

Atallah, L. (2011). *Observing recovery from knee-replacement surgery by using wearable sensors*. IEEE Digital Xplore.

Azevedo, W. W., Lima, S. M., Fernandes, I. M., Rocha, A. D., Cordeiro, F. R., Silva-Filho, A. G., & Santos, W. P. (2015). Fuzzy Morphological Extreme Learning Machines to Detect and Classify Masses in Mammograms. In *2015 IEEE International Conference on Fuzzy Systems*. IEEE. 10.1109/FUZZ-IEEE.2015.7337975

Bace & Mell. (2001). *NIST special publication on intrusion detection systems*. Booz-Allen and Hamilton Inc.

Bagahel, D., & Kiran, K.G. (2015). Comparative study of tumor detection techniques with their suitability for brain MRI images. Inlt. Jrl., 127(13).

Balafar, M. A., Ramli, A. R., Saripan, M. I., & Mashohor, S. (2010). Review of brain MRI segmentation methods. *Artificial Intelligence Review*, *33*(3), 261–274. doi:10.100710462-010-9155-0

Bandyopadhyay, S. K. (2010). Survey on Segmentation Methods for Locating Masses in a Mammogram Image. *International Journal of Computers and Applications*, *9*(11), 25–28. doi:10.5120/1429-1926

Bashir, U., & Chachoo, M. (2014). Intrusion detection and prevention system: Challenges & opportunities. *Computing for Sustainable Global Development (INDIACom), 2014 International Conference on*, 806-809. 10.1109/IndiaCom.2014.6828073

Bayramoglu, N., Kannala, J., & Heikkila, J. (2016). Deep Learning for Magnification Independent Breast Cancer Histopathology Image Classification. In *2016 23rd International Conference on Pattern Recognition (ICPR)*. Cancun: IEEE. 10.1109/ICPR.2016.7900002

Benson, C. C. (2016). Brain Tumor Segmentation from MR Brain Images using Improved Fuzzy c-Means Clustering and Watershed Algorithm. *2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 10.1109/ICACCI.2016.7732045

Bharati, M., & Tamane, S. (2017). Intrusion detection systems (IDS) & future challenges in cloud based environment. Intelligent Systems and Information Management (ICISIM), 240-250. doi:10.1109/ICISIM.2017.8122180

Bhateja, V., Gautam, A., Tiwari, A., Bao, L. N., Satapathy, S. C., Nhu, N. G., & Le, D.-N. (2018). Haralick Features-Based Classification of Mammograms Using SVM. In *Information Systems Design and Intelligent Applications* (Vol. 672). Springer. doi:10.1007/978-981-10-7512-4_77

Bhuvaneswari, Aruna, & Loganathan. (2014). Classification of Lung Diseases by Image Processing Techniques Using Computed Tomography Images. *International Journal of Advanced Computer Research, 4*(1).

Bjoern, H. M., & Koen, V. L. (2016). A Generative Probabilistic Model and Discriminative Extensions for Brain Lesion Segmentation With Application to Tumor and Stroke. *IEEE Transactions on Medical Imaging*, *35*(4), 933–946. doi:10.1109/TMI.2015.2502596 PMID:26599702

Bolte, S., & Cordelières, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *Journal of Microscopy*, *224*(3), 213–232. doi:10.1111/j.1365-2818.2006.01706.x PMID:17210054

Borgen, H., Bours, P., & Wolthusen, S. D. (2009). Simulating the Influences of Aging and Ocular Disease on Biometric Recognition Performance. *International Conference on Biometrics 2009, 8*(8), 857–867. 10.1007/978-3-642-01793-3_87

Budai, A., Bock, R., Maier, A., Hornegger, J., & Michelson, G. (2013). Robust Vessel Segmentation in Fundus Images. *International Journal of Biomedical Imaging*.

Canadian Border Services Agency. (2015). *CANPASS Air*. Available: http: //www.cbsa-asfc.gc.ca/prog/canpass/canpassair-eng.html

Candemir, Jaeger, Palaniappan, & Musco. (2014). Lung Segmentation in Chest Radiographs Using Anatomical Atlases With Non-rigid Registration. *IEEE Transactions on Medical Imaging, 33*(2).

Chang, R. F., Wu, W. J., Moon, W. K., & Chen, D.-R. (2005). Dr Chen, "Automatic Ultrasound Segmentation and Morphology based Diagnosis of Solid Breast Tumors. *Breast Cancer Research and Treatment*, *89*(2), 179–185. doi:10.100710549-004-2043-z PMID:15692761

Chang, S.-K., & Hsu, A. (1992). Image information systems: Where do we go from here? *IEEE Transactions on Knowledge and Data Engineering*, *4*(5), 431–442. doi:10.1109/69.166986

Chen, Y., Wang, Y., & Yang, B. (2006). Evolving Hierarchical RBF Neural Networks for Breast Cancer Detection. LNCS, 4234, 137-144. doi:10.1007/11893295_16

Chung, C.-J., Khatkar, P., Xing, T., Lee, J., & Huang, D. (2013). NICE: Network Intrusion Detection and Countermeasure Selection in Virtual Network Systems. *IEEE Transactions on Dependable and Secure Computing*, *10*(4), 198–211. doi:10.1109/TDSC.2013.8

Coleman, M. P., Quaresma, M., Berrino, F., Lutz, J. M., De Angelis, R., Capocaccia, R., Baili, P., Rachet, B., Gatta, G., Hakulinen, T., Micheli, A., Sant, M., Weir, H. K., Elwood, J. M., Tsukuma, H., Koifman, S., & Silva, E. (2008). Cancer survival in five continents: A worldwide population-based study (CONCORD). *The Lancet. Oncology*, *9*(8), 730–756. doi:10.1016/S1470-2045(08)70179-7 PMID:18639491

Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., Simon, M., Pop, S. C., Girard, P., Ameli, R., Ferré, J.-C., Kerbrat, A., Tourdias, T., Cervenansky, F., Glatard, T., Beaumont, J., Doyle, S., Forbes, F., Knight, J., Khademi, A., ... Barillot, C. (2018). Objective Evaluation of Multiple Sclerosis Lesion Segmentation Using a Data Management and Processing Infrastructure. *Scientific Reports*, *8*(1), 13650. doi:10.103841598-018-31911-7 PMID:30209345

Cordeiro, F. R., Bezerra, K. F. P., & Santos, W. P. (2017). Random walker with fuzzy initialization applied to segment masses in mammography images. In *30th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE. 10.1109/CBMS.2017.40

Cordeiro, F. R., Lima, S. M., Silva-Filho, A. G., & Santos, W. P. (2012). Segmentation of mammography by applying extreme learning machine in tumor detection. In *International Conference of Intelligent Data Engineering and Automated Learning*. Berlin: Springer.

Cordeiro, F. R., Santos, W. P., & Silva-Filho, A. G. (2013). Segmentation of mammography by applying growcut for mass detection. *Studies in Health Technology and Informatics*, *192*, 87–91. PMID:23920521

Cordeiro, F. R., Santos, W. P., & Silva-Filho, A. G. (2016a). A semi-supervised fuzzy growcut algorithm to segment and classify regions of interest of mammographic images. *Expert Systems with Applications*, *65*, 116–126. doi:10.1016/j.eswa.2016.08.016

Cordeiro, F. R., Santos, W. P., & Silva-Filho, A. G. (2016b). An adaptive semi-supervised fuzzy growcut algorithm to segment masses of regions of interest of mammographic images. *Applied Soft Computing*, *46*, 613–628. doi:10.1016/j.asoc.2015.11.040

Corliss, B. A., Ray, H. C., Patrie, J. T., Mansour, J., Kesting, S., Park, J. H., & Peirce, S. M. (2019). CIRCOAST: A statistical hypothesis test for cellular colocalization with network structures. *Bioinformatics (Oxford, England)*, *35*(3), 506–514. doi:10.1093/bioinformatics/bty638 PMID:30032263

Corso, J. J., Sharon, E., Brandt, A., & Yuille, A. (2006). Multilevel Segmentation and Integrated Bayesian Model Classification with an Application to Brain Tumor Segmentation. *MICCAI*, *4191*, 790–798. doi:10.1007/11866763_97 PMID:17354845

Corso, J. J., Sharon, E., Dube, S., El-Saden, S., Sinha, U., & Yuille, A. (2008, May). Efficient Multilevel Brain Tumor Segmentation with Integrated Bayesian Model Classification. *IEEE Transactions on Medical Imaging*, *27*(5), 629–640. doi:10.1109/TMI.2007.912817 PMID:18450536

Costes, S. V., Daelemans, D., Cho, E. H., Dobbin, Z., Pavlakis, G., & Lockett, S. (2004). Automatic and quantitative measurement of protein-protein colocalization in live cells. *Biophysical Journal*, *86*(6), 3993–4003. doi:10.1529/biophysj.103.038422 PMID:15189895

Cruz, T. N., Cruz, T. M., & Santos, W. P. (2018). Detection and classification of lesions in mammographies using neural networks and morphological wavelets. *IEEE Latin America Transactions*, *16*(3), 926–932. doi:10.1109/TLA.2018.8358675

D'Orsi, C. J., Sickles, E. A., Mendelson, E. B., & Morris, E. A. (2013). *Breast Imaging Reporting and Data System: ACR BI-RADS breast imaging atlas* (5th ed.). American College of Radiology.

Dai, S., Lu, K., & Dong, J. (2015). Lung segmentation with improved graph cuts on chest CT images. *3rd IAPR Asian Conference on Pattern Recognition*. 10.1109/ACPR.2015.7486502

Dean Bidgood, W. Jr. (1998). The SNOMED DICOM Microglossary: Controlled terminology resource for Idata interchange in biomedical imaging. *Methods of Information in Medicine*, *37*(4/5), 404–414. doi:10.1055-0038-1634557 PMID:9865038

DeSantis, C. E., Lin, C. C., Mariotto, A. B., Siegel, R. L., Stein, K. D., Kramer, J. L., Alteri, R., Robbins, A. S., & Jemal, A. (2014). Cancer treatment and survivorship statistics, 2014. *CA: a Cancer Journal for Clinicians*, *64*(4), 252–271. doi:10.3322/caac.21235 PMID:24890451

Deserno, T., Soiron, M., Oliveira, J., & Araújo, A. (2012a). Towards Computer-Aided Diagnostics of Screening Mammography Using Content-Based Image Retrieval. In *2011 24th SIBGRAPI Conference on Graphics, Patterns and Images*. Alagoas: IEEE.

Deserno, T. M., Soiron, M., Oliveira, J. E. E., & Araújo, A. A. (2012b). Computer-aided diagnostics of screening mammography using content-based image retrieval. In *Medical Imaging: Computer-Aided Diagnosis 2012*. SPIE. doi:10.1117/12.912392

Dhage, S. N., Meshram, B., Rawat, R., Padawe, S., Paingaokar, M., & Misra, A. (2011). Intrusion detection system in cloud computing environment. *Proceedings of the International Conference & Workshop on Emerging Trends in Technology*, 235-239. 10.1145/1980022.1980076

Dhir, L., Habib, N. E., Monro, D. M., & Rakshit, S. (2010). Effect of cataract surgery and pupil dilation on iris pattern recognition for personal authentication. *Eye (London, England)*, *24*(6), 1006–1010. doi:10.1038/eye.2009.275 PMID:19911017

Dhooge, M., & de Laey, J. J. (1989). The ocular ischemic syndrome. *Bulletin de la Société Belge d'Ophtalmologie*, *231*, 1–13. PMID:2488440

Duan, Y., Wang, J., Hu, M., Zhou, M., Li, Q., Sun, L., & Wang, Y. (2019). Leukocyte classification based on spatial and spectral features of microscopic hyperspectral images. *Optics & Laser Technology*, *112*, 530–538. doi:10.1016/j.optlastec.2018.11.057

Durgadevi & Shekhar. (2015). Identification of tumor using K-means algorithm. *Intl. Jrl. Adv. Res. Inn. Id. Edu, 1,* 227-231.

Dvorak, P., & Menze, B. (2015). Structured prediction with convolutional neural networks for multimodal brain tumor segmentation. *Proceeding of the Multimodal Brain Tumor Image Segmentation Challenge*, 13-24.

Elbalaoui, Fakir, Taifi, & Merbohua. (2016). Automatic Detection of Blood Vessel in Retinal Images. *13th International Conference Computer Graphics, Imaging and Visualization*.

Eltoukhy, M., Faye, I., & Samir, B. (2009). Breast Cancer Diagnosis in Mammograms using Multilevel Wavelet Analysis. *Proceeding of National Postgraduate Conference*.

Eman, A. M., Mohammed, E., & Rashid, A. L. (2015). Brain tumor segmentation based on a hybrid clustering technique. *Egyptian Informatics Journal*, *16*(1), 71–81. doi:10.1016/j.eij.2015.01.003

Emran, Abtin, & David. (2015). Automatic segmentation of wrist bones in CT using a statistical wrist shape pose model. *Inlt. Jrl.*

Fathy, M., Keshk, M., & El Sherif, A. (2019). Surgical management and outcome of intramedullary spinal cord tumour. *Egyptian Journal of Neurosurgery.*, *34*(2), 2–7. doi:10.118641984-019-0028-9

Fernandes, I., & Santos, W. (2014). Classificação de mamografias utilizando extração de atributos de textura e redes neurais artificiais. In *Congresso Brasileiro de Engenharia Biomédica (CBEB 2014)*. SBEB.

Ferreira, J., Oliveira, H., & Martinez, M. (2011). Aplicação de uma metodologia computacional inteligente no diagnóstico de lesões cancerígenas. *Revista Brasileira de Inovação Tecnológica em Saúde*, *1*(2), 4-9.

Freixenet, J., Munoz, X., Raba, D., Marti, J., & Cufi, X. (2002). Yet another survey on image segmentation: Region and boundary information integration. *Proc. 7th Eur. Conf. Computer Vision Part III*, 408–422. 10.1007/3-540-47977-5_27

Fuadah, Setiawan, & Mengko. (2015). Mobile Cataract Detection using Optimal Combination of Statistical Texture Analysis. *4th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME).*

Fuadah, Setiawan, Mengko, & Budiman. (2015). *A computer aided healthcare system for cataract classification and grading based on fundus image analysis*. Elsevier Science Publishers B. V.

Ganesan, K., Acharya, U. R., Chua, C. K., Min, L. C., & Abraham, T. K. (2014). Automated Diagnosis of Mammogram Images of Breast Cancer Using Discrete Wavelet Transform and Spherical Wavelet Transform Features: A Comparative Study. *Technology in Cancer Research & Treatment*, *13*(6), 605–615. doi:10.7785/tcrtexpress.2013.600262 PMID:24000991

Georges, B. A. (1999). Model Creation and Deformation for the Automatic Segmentation of the Brain in MR Images. *IEEE Transactions on Biomedical Engineering*, *46*(11), 1346–1356. doi:10.1109/10.797995 PMID:10582420

Girisha, Chandrashekhar, & Kurian. (2013). Texture Feature Extraction of Video Frames Using GLCM. *International Journal of Engineering Trends and Technology, 4*(6).

Girshick, R. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 10.1109/CVPR.2014.81

Gogic, I., Manhart, M., Pandzic, I. S., & Ahlberg, J. (2020). Fast facial expression recognition using local binary features and shallow neural networks. *The Visual Computer*, *36*(1), 97–112. doi:10.100700371-018-1585-8

Guld, Kohnen, Keysers, Schubert, Wein, Bredno, & Lehmann. (2002). Quality of DICOM header information for image categorization. *SER*, *Proc. SPIE*, 4685, 280-287. 10.1117/12.467017

Guo, Z., Zhang, L., & Zhang, D. (2010). Rotation invariant texture classification using LBP variance with global matching. *Pattern Recognition*, *43*(3), 706–716. doi:10.1016/j.patcog.2009.08.017

Gupta & Singh. (2017). Brain Tumor segmentation and classification using Fcm and support vector machine. *International Research Journal of Engineering and Technology, 4*(5).

Hai, S., Xing, F., & Yang, L. (2016). Robust Cell Detection of Histopathological Brain Tumor Images Using Sparse Reconstruction and Adaptive Dictionary Selection. *IEEE Transactions on Medical Imaging*, *35*(6), 1575–1586. doi:10.1109/TMI.2016.2520502 PMID:26812706

Haleem, M. S., Han, L., van Hemert, J., & Fleming, A. (2015). Glaucoma Classification using Regional Wavelet Features of the ONH and its Surroundinga. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).*

Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, *3*(6), 610–621. doi:10.1109/TSMC.1973.4309314

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, *35*, 18–31. doi:10.1016/j.media.2016.05.004 PMID:27310171

Hazlina, H., & Sameem, A. K. (2004). Back Propagation Neural Network for the Prognosis of Breast Cancer: Comparison on Different Training Algorithms. *Proceedings Second International Conference on Artificial Intelligence in Engineering & Technology*, 445-449.

Heath, M., Bowyer, K. W., & Kopans, D. (2000). The digital database for screening mammography. *Proceedings of the 5th International Workshop on Digital Mammography*

Hersh, W., Muller, H., & Kalpathy. (2009). The imageCLEFmed medical image retrieval task test collection. *J. Digital Imaging, 22*(6), 648-655.

Hogeweg, Sánchez, Maduskar, Philipsen, & Story. (2015). Automatic Detection of Tuberculosis in Chest Radiographs Using a Combination of Textural, Focal and Shape Abnormality Analysis. *IEEE Transactions on Medical Imaging, 34*(12).

Horsch, K., Giger, M. L., Venkata, L. A., & Vybomya, C. J. (2001). Automatic Segmentation of Breast Lesions on Ultrasound. *Medical Physics*, *28*(8), 1652–1659. doi:10.1118/1.1386426 PMID:11548934

Hotko, Y. S. (2013). Male breast cancer: Clinical presentation, diagnosis, treatment. *Experimental Oncology*, *35*(4), 303–310. PMID:24382442

Huang, H. K. (2004). PACS and imaging informatics: basic principles and applications. John Wiley & Sons Inc. doi:10.1002/0471654787

Hussain, C. A., Rao, D. V., & Mastani, S. A. (2020). *RetrieveNet: a novel deep network for medical image retrieval*. Evol. Intel. doi:10.100712065-020-00401-z

ISO/IEC 19794-6:2011. (2011). Information technology – Biometric data interchange formats – Part 6: Iris image data.

Jaeger, Karargyris, Candemir, Folio, & Siegelman. (2014). Automatic Tuberculosis Screening Using Chest Radiographs. *IEEE Transactions on Medical Imaging, 33*(2).

Jagath, C. (2001, October). Bayesian Approach to Segmentation of Statistical Parametric Maps. *IEEE Transactions on Biomedical Engineering*, *48*(10).

Jannesari, M., Habibzadeh, M., Aboulkheyr, H., Khosravi, P., Elemento, O., Totonchi, M., & Hajirasouliha, I. (2018). Breast Cancer Histopathological Image Classification: A Deep Learning Approach. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE. 10.1109/BIBM.2018.8621307

Jenifer, S., Parasuraman, S., & Kadirvel, A. (2014). An Efficient Biomedical Imaging Technique for Automatic Detection of Abnormalities in Digital Mammograms. *Journal of Medical Imaging and Health Informatics*, *4*(2), 291–296. doi:10.1166/jmihi.2014.1246

Jobin Christ, M. C., Sasikumar, K., & Parwathy, R. M. S. (2009, July). Application of Bayesian Method in Medical Image Segmentation. *International Journal of Computing Science and Communication Technologies, VOL*, *2*(1).

Jose, Ravi, & Sampath. (2014). Brain tumor segmentation a performance analysis using K-means, fuzzy c-means and region growing algorithm. *Inlt. Jrl., 2*(3).

Joseph, S., & Balakrishnan, K. (2011). Local Binary Patterns, Haar Wavelet Features and Haralick Texture Features for Mammogram Image Classification using Artificial Neural Networks. In *International Conference on Advances in Computing and Information Technology*. Springer. 10.1007/978-3-642-22555-0_12

Juang, L. H., & Ming, N. W. (2010). MRI brain lesion image detection based on color-converted K-means. *Science Direct Measurement*, *43*(7), 941–949.

Juhl, J. H., Crummy, A. B., & Kuhlman, J. E. (2000). Paul & Juhl Interpretação Radiológica (7a ed.). Rio de Janeiro: Guanabara-Koogan.

Kamnitsas, K., Ledig, C., Newcombe, V. F. J., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., & Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis*, *36*, 61–78. doi:10.1016/j.media.2016.10.004 PMID:27865153

Kaur & Rani. (2016). MRI Brain Tumor Segmentation Methods- A Review. *International Journal of Current Engineering and Technology*.

Khadem. (2010). *MRI Brain image segmentation using graph cuts* (Master's thesis). Chalmers University of Technology, Goteborg, Sweden.

Kholidy, H. A., & Baiardi, F. (2012). CIDS: A framework for intrusion detection in cloud systems. *Information Technology: New Generations (ITNG), Ninth International Conference on*, 379-385. 10.1109/ITNG.2012.94

Khuriwal, N., & Mishra, N. (2018). Breast Cancer Detection from Histopathological Images using Deep Learning. In *2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE)*. IEEE. 10.1109/ICRAIE.2018.8710426

Kishore, B., Arjunan, R. V., Saha, R., & Selvan, S. (2014). Using Haralick Features for the Distance Measure Classification of Digital Mammograms. *International Journal of Computers and Applications*, *6*(1), 17–21.

Kiyan, T., & Yildrim, T. (2004). Breast Cancer Diagnosis using Statistical Neural Networks. *Journal of Electrical and Electronics Engineering (Oradea)*, *4*(2), 1149–1153.

Kong, K., & Tomizuka, M. (2009). A gait monitoring system based on air pressure sensors embedded in a shoe. *IEEE/ASME Transactions on Mechatronics*, *14*(3), 358–370. doi:10.1109/TMECH.2008.2008803

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*.

Krupinski. (2000). The Importance of Perception Research in Medical Imaging. *Radiation Medicine, 8*(6), 329-334.

Kumar, Manjunathand, & Sheshadri. (2015). Feature extraction from the fundus images for the diagnosis of diabetic retinopathy. *International Conference on Emerging Research in Electronics. Computer Science and Technology*.

Kumar, A. (2017). *A Novel Approach for Brain Tumor Detection Using Support Vector Machine*. K-Means and PCA Algorithm.

Kumar, D. (2018). A modified intuitionistic fuzzy c-means clustering approach to segment human brain MRI image. *Multimedia Tools and Applications*, 1–25.

Laddha, R. R. (2014). A Review on Brain Tumor Detection Using Segmentation And Threshold Operations. *International Journal of Computer Science and Information Technologies*, *5*(1), 607–611.

Langley, P. (1996). *Elements of Machine Learning*. Morgan Kaufmann.

Lavancier, F., Pécot, T., Zengzhen, L., & Kervrann, C. (2020). Testing independence between two random sets for the analysis of colocalization in bioimaging. *Biometrics*, *76*(1), 36–46. doi:10.1111/biom.13115 PMID:31271216

Lee, J.-H., Park, M.-W., Eom, J.-H., & Chung, T.-M. (2011). Multi-level intrusion detection system and log management in cloud computing. *Advanced Communication Technology (ICACT), 13th International Conference on*, 552-555.

Leemput, K. V., Maes, F., Vandermeulen, D., & Suetens, P. (1999). Automated model-based tissue classification of MR images of brain. *IEEE Transactions on Medical Imaging*, *18*(10), 897–908. doi:10.1109/42.811270 PMID:10628949

Lehmann, T. M., Guld, M. O., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H., & Wein, B. B. (2004). Content-based image retrieval in medical applications. *Methods of Information in Medicine*, *43*(4), 354–361. doi:10.1055-0038-1633877 PMID:15472746

Liana, D. D., Raguse, B., Gooding, J. J., & Chow, E. (2012). Recent Advances in Paper-Based Sensors. *Sensors (Basel)*, *12*(9), 11505–11526. doi:10.3390120911505 PMID:23112667

Liao, H.-J., Lin, C.-H. R., Lin, Y.-C., & Tung, K.-Y. (2013). Intrusion detection system: A comprehensive review. *Journal of Network and Computer Applications*, *36*(1), 16–24. doi:10.1016/j.jnca.2012.09.004

Lima, S. M., Silva-Filho, A. G., & Santos, W. P. (2014). A methodology for classification of lesions in mammographies using Zernike moments, ELM and SVM neural networks in a multi-kernel approach. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE. 10.1109/SMC.2014.6974041

Linguraru, Richbourg, Liu, & Watt. (n.d.). *Tumor Burden Analysis on Computed Tomography by Automated Liver and Tumor Segmentation*. IEEE.

Liu, H., Cao, H., & Song, E. (2019). Bone marrow cells detection: A technique for the microscopic image analysis. *Journal of Medical Systems*, *43*(4), 82. doi:10.100710916-019-1185-9 PMID:30798374

Liu, Z. (2015). Semantic image segmentation via deep parsing network. *Proceedings of the IEEE International Conference on Computer Vision*. 10.1109/ICCV.2015.162

Logeswari & Karnan. (2010). An improved implementation of brain tumor detection using segmentation based on soft computing. *Journal of Cancer Research and Experimental Oncology, 2*(1).

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Long, L. R., Antani, S., Deserno, T. M., & Thoma, G. R. (2009). Content-Based Image Retrieval in Medicine: Retrospective Assessment, State of the Art, and Future Directions. *International Journal of Healthcare Information Systems and Informatics*, *4*(1), 1–16. doi:10.4018/jhisi.2009010101 PMID:20523757

Lotankar, M., Noronha, K., & Koti, J. (2015). Detection of Optic Disc and Cup from Color Retinal Images for Automated Diagnosis of Glaucoma. *IEEE UP Section Conference on Electrical Computer and Electronics (UPCON).*

Madabhushi, A., & Metaxas, D. (2003). Combining low-, high-level and Empirical Domain Knowledge for Automated Segmentation of Ultrasonic Breast Lesions. *IEEE Transactions on Medical Imaging*, *22*(2), 155–169. doi:10.1109/TMI.2002.808364 PMID:12715992

Madhukumar, S., & Santhiyakumari, N. (2015). Evaluation of k-Means and fuzzy C-means segmentation on MR images of brain. *Egyptian Society of Radiology and Nuclear Medicine*, *46*(2), 475–479. doi:10.1016/j.ejrnm.2015.02.008

Maitra, I. K., Nag, S., & Bandyopadhyay, S. K. (2011). Identification of Abnormal Masses in Digital Mammography Images. *International Journal of Computer Graphics*, *2*(1).

Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Academic.

Malvia, S., Bagadi, S. A., Dubey, U. S., & Saxena, S. (2017). Epidemiology ofbreast cancer in Indian women. *Asia Pacific Journal of Clinical Oncology*, *13*(4), 289–295. doi:10.1111/ajco.12661 PMID:28181405

Manders, E. M. M., Verbeek, F. J., & Aten, J. A. (1993). Measurement of co-localization of objects in dual colour confocal images. *Journal of Microscopy*, *169*(3), 375–382. doi:10.1111/j.1365-2818.1993.tb03313.x

Manjunath, K. N., Renuka, A., & Niranjan, U. C. (2007). Linear models of cumulative distribution function for content-based medical image retrieval. *Journal of Medical Systems*, *31*(6), 433–443. doi:10.100710916-007-9075-y PMID:18041275

Mannoor, M. S., Tao, H., Clayton, J. D., Sengupta, A., Kaplan, D. L., Naik, R. R., Verma, N., Omenetto, F. G., & McAlpine, M. C. (2012). Graphene-based wireless bacteria detection on tooth enamel. *Nature Communications*, *3*(1), 1–8. doi:10.1038/ncomms1767 PMID:22453836

Maria, J., Amaro, J., Falcao, G., & Alexandre, L. A. (2016). Stacked Autoencoders Using Low-Power Accelerated Architectures for Object Recognition in Autonomous Systems. *Neural Processing Letters*, *43*(2), 445–458. doi:10.100711063-015-9430-9

Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentationalgorithms and measuring ecological statistics. *Proc. 8th Int. Conf. Computer Vision*, 2, 416–423. 10.1109/ICCV.2001.937655

Mascaro, A. A., Mello, C. A., Santos, W. P., & Cavalcanti, G. D. (2009). Mammographic images segmentation using texture descriptors. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE. 10.1109/IEMBS.2009.5333696

Masdari, M., & Ahmadzadeh, S. (2017). A survey and taxonomy of the authentication schemes in Telecare Medicine Information Systems. *Journal of Network and Computer Applications*, *87*, 1–19. doi:10.1016/j.jnca.2017.03.003

MathWorks. (2019). *Deep Learning Toolbox ™ Reference*. Author.

McConnon, G., Deravi, F., Hoque, S., Sirlantzis, K., & Howells, G. (2012). Impact of Common Ophthalmic Disorders on Iris Recognition. *2012 5th IAPR International Conference on Biometrics Compendium*, 277–282.

Mehmood, Y., Shibli, M. A., Habiba, U., & Masood, R. (2013). Intrusion detection system in cloud computing: challenges and opportunities. *2013 2nd National Conference on Information Assurance (NCIA)*, 59-66. 10.1109/NCIA.2013.6725325

Mell & Grance. (2011). *The NIST definition of cloud computing*. Academic Press.

Memon & Li. (2019). Breast Cancer Detection in the IOT Health Environment Using Modified Recursive Feature Selection. *Wireless Communications and Mobile Computing*, 1-19.

Mengqiao, W., Jie, Y., Yilei, C., & Hao, W. (2017). The multimodal brain tumor image segmentation based on convolutional neural networks. *2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA),* 336-339. 10.1109/CIAPP.2017.8167234

Miao, H., & Xiao, C. (2018). Simultaneous segmentation of leukocyte and erythrocyte in microscopic images using a marker-controlled watershed algorithm. *Computational and Mathematical Methods in Medicine*, *2018*, 2018. doi:10.1155/2018/7235795 PMID:29681997

Milon Islam, Md., & Rashedul Islam, Md. (2020). Development of Smart Healthcare Monitoring System in IoT Environment. *SN Computer Science*, *185*, 1–11.

Modi, C. N., Patel, D. R., Patel, A., & Muttukrishnan, R. (2012). Bayesian Classifier and Snort based network intrusion detection system in cloud computing. In *Computing Communication & Networking Technologies* (pp. 1–7). ICCCNT. doi:10.1109/ICCCNT.2012.6396086

Modi, C. N., Patel, D. R., Patel, A., & Muttukrishnan, R. (2012). Bayesian Classifier and Snort based network intrusion detection system in cloud computing. *Third International Conference on Computing Communication & Networking Technologies (ICCCNT)*, 1-7.

Modi, C. N., Patel, D. R., Patel, A., & Rajarajan, M. (2012). Integrating signature apriori based network intrusion detection system (NIDS) in cloud computing. *Procedia Technology*, *6*, 905–912. doi:10.1016/j.protcy.2012.10.110

Modi, C., Patel, D., Borisaniya, B., Patel, H., Patel, A., & Rajarajan, M. (2013). A survey of intrusion detection techniques in cloud. *Journal of Network and Computer Applications*, *36*(1), 42–57. doi:10.1016/j.jnca.2012.05.003

Moles Lopez, X., Barbot, P., Van Eycke, Y. R., Verset, L., Trépant, A. L., Larbanoix, L., & Decaestecker, C. (2015). Registration of whole immunohistochemical slide images: An efficient way to characterize biomarker colocalization. *Journal of the American Medical Informatics Association: JAMIA*, *22*(1), 86–99. doi:10.1136/amiajnl-2014-002710 PMID:25125687

Monro, D. M., Rakshit, S., & Zhang, D. (2009). DCT-Based Iris Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(4), 586–595. doi:10.1109/TPAMI.2007.1002 PMID:17299216

Moon, N., Bullitt, E., Leemput, K., & Gerig, G. (2002). Model based brain and tumor segmentation. *Int. Conf. on Pattern Recognition*, 528-531. 10.1109/ICPR.2002.1044787

Mooney. (2018 April). *Blood cell images Version 6*. Retrieved May 23 2020 from https://www.kaggle.com/paultimothymooney/blood-cells

Muller, A.C., & Guido, S. (n.d.). *Introduction to Machine Learning with Python.* O'Reilly.

Muller, Michoux, Bandon, & Geissbuhler. (2007). A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. *Intl. Journal of Medical Informatics, 73*(1), 1–23.

Murathoti Varshini, Barjo, & Tigga. (2020). Spine Magnetic Resonance Image Segmentation Using Deep Learning Techniques. *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS*).

Nanni, L., Brahnam, S., & Lumini, A. (2012). A simple method for improving local binary patterns by considering non-uniform patterns. *Pattern Recognition*, *45*(10), 3844–3852. doi:10.1016/j.patcog.2012.04.007

Naveen Kumar, B., Chauhan, R. P., & Dahiya, N. (2016). Detection of Glaucoma using Image processing techniques: A Review. *2016 International Conference on Microelectronics, Computing and Communications (MicroCom)*.

Neurotechnology. (2012). *VeriEye SDK, v. 4.3*. Available: https://www.neurotechnology.com/verieye.html

Ng, K.-G., Wong, S.-T., Lim, S.-M., & Goh, Z. (2010). Evaluation of the cadi thermosensor wireless skin-contact thermometer against ear and axillary temperatures in children. *Journal of Pediatric Nursing-Nursing Care of Children & Families*, *25*(3), 176–186. doi:10.1016/j.pedn.2008.12.002 PMID:20430278

Nikolai, J., & Wang, Y. (2014). Hypervisor-based cloud intrusion detection system. In *Computing* (pp. 989–993). Networking and Communications.

Niwas, Lin, Kwoh, Kuo, Sng, Aquino, & Chew. (2016). Cross-examination for Angle-Closure Glaucoma Feature Detection. *IEEE Journal of Biomedical and Health Informatics*.

Odstrcilik, J., Budai, A., Kolar, R., & Hornegger, J. (2013, June). Retinal vessel segmentation by improved matched filtering: Evaluation on a new high-resolution fundus image database. *IET Image Processing*, *7*(4), 373–383. doi:10.1049/iet-ipr.2012.0455

O'Flynn, J. (2013). Novel smart sensor glove for arthritis rehabilitation. *BSN*, 1-6.

Ojala, T., Pietikainen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, *29*(1), 51–59. doi:10.1016/0031-3203(95)00067-4

Oktay, U., & Sahingoz, O. K. (2013). Proxy network intrusion detection system for cloud computing. *International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE)*, 98-104. 10.1109/TAEECE.2013.6557203

Oliveira, J. E., Machado, A. M., Chavez, G. C., Lopes, A. P., Deserno, T. M., & Araújo, A. A. (2010). MammoSys: A content-based image retrieval system using breast density patterns. *Computer Methods and Programs in Biomedicine*, *99*(3), 289–297. doi:10.1016/j.cmpb.2010.01.005 PMID:20207441

Omisore. (2014). Proposed the genetic neuro-fuzzy inferential model for the diagnosis of tuberculosis. *IEEE Transactions*.

Onal, Susana, Paul, & Alferedo. (2014). Automated localization of multiple pelvic bone structure in MRI. *Intl Jrl*.

Panse, N. D., Ghorpade, T., & Jethani, V. (2015). Retinal Fundus Diseases Diagnosis using Image Mining. *IEEE International Conference on Computer, Communication and Control (IC4-2015).* 10.1109/IC4.2015.7375721

Papakostas, G. A., Koulouriotis, D. E., Karakasis, E. G., & Tourassis, V. D. (2013). Moment-based local binary patterns: A novel descriptor for invariant pattern recognition applications. *Neurocomputing*, *99*, 358–371. doi:10.1016/j.neucom.2012.06.031

Patel, A., Taghavi, M., Bakhtiyari, K., & Celestino Júnior, J. (2013). An intrusion detection and prevention system in cloud computing: A systematic review. *Journal of Network and Computer Applications*, *36*(1), 25–41. doi:10.1016/j.jnca.2012.08.007

Patel, S., Lorincz, K., Hughes, R., Huggins, N., Growdon, J., Standaert, D., Akay, M., Dy, J., Welsh, M., & Bonato, P. (2009). Monitoring motor fluctuations in patients with parkinson's disease using wearable sensors. *IEEE Transactions on Information Technology in Biomedicine*, *13*(6), 864–873. doi:10.1109/TITB.2009.2033471 PMID:19846382

Patil. (2005). *Pachpande, Automatic Brain Tumor Detection Using K-Means*. Academic Press.

Pereira, Fonseca-Pinto, Paiva, Tavora, Assuncao, & Faria (2020). *Accurate segmentation of desmoscopic images based on local binary pattern clustering*. International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia.

Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, *35*(5), 1240–1251. doi:10.1109/TMI.2016.2538465 PMID:26960222

Pham, T. X., Siarry, P., & Oulhadj, H. (2018). Integrating fuzzy entropy clustering with an improved PSO for MRI brain image segmentation. *Applied Soft Computing*, *65*, 230–242. doi:10.1016/j.asoc.2018.01.003

Po, M.-Z., Swenson, C., & Rosalind, W. (2010). A Wearable Sensor for Unobtrusive, Long-Term Assessment of Electrodermal Activity. *IEEE Transactions on Biomedical Engineering*, *57*(5), 1243–1252. doi:10.1109/TBME.2009.2038487 PMID:20172811

Pradhan, S. (2010). *Development of Unsupervised Image Segmentation Schemes for Brain MRI using HMRF model* (Master Thesis). Department of EE, NIT, Rourkela, India.

Priya. (2018). Efficient fuzzy c-means based multilevel image segmentation for brain tumor detection in MR images. *Design Automation for Embedded Systems*, 1–13.

Putzu, L., Caocci, G., & Di Ruberto, C. (2014). Leucocyte classification for leukaemia detection using image processing techniques. *Artificial Intelligence in Medicine*, *62*(3), 179–191. doi:10.1016/j.artmed.2014.09.002 PMID:25241903

Raad, A., Kalakech, A., & Ayache, M. (2012). Breast Cancer Classification using Neural Network Approach: MLP AND RBF. *The 13th international Arab conference on information technology,* 10 – 13.

Rajendra Acharya, U. (2011, May). Automated Diagnosis of Glaucoma Using Texture and Higher Order Spectra Features. *IEEE Transactions on Information Technology in Biomedicine*, *15*(3).

Rendon-Gonzalez & Ponomaryov. (2016). Automatic Lung Nodule Segmentation and Classification in CT Images Based on SVM. *International conferences IEEE.*

Roberts, T., Newell, M., Auffermann, W., & Vidakovic, B. (2017). Wavelet-based scaling indices for breast cancer diagnostics. *Statistics in Medicine*, *36*(12), 1989–2000. doi:10.1002im.7264 PMID:28226399

Roizenblatt, R., Schor, P., Dante, F., Roizenblatt, J., & Jr, R. B. (2004). Iris recognition as a biometric method after cataract surgery. *BioMedical Engineering Online, 3*(2). www.biomedical-engineering-online/com/content/3/1/2

Sachdeva, & Singh. (2015). Automatic Segmentation and Area Calculation of Optic Disc in Ophthalmic Images. *2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS).*

Sahiner, B., Heang-Ping, C., Patrick, N., Wei, D. M. A., Helie, D., Adler, D., & Goodsitt, M. M. (1996). Classification of Mass and Normal Breast Tissue: A Convolution Neural Network Classifier with Spatial Domain and Texture Images. *IEEE Transactions on Medical Imaging*, *15*(5), 598–610. doi:10.1109/42.538937 PMID:18215941

Sahlol, A. T., Kollmannsberger, P., & Ewees, A. A. (2020). Efficient classification of white blood cell leukemia with improved Swarm optimization of deep features. *Scientific Reports*, *10*(1), 1–11. doi:10.103841598-020-59215-9 PMID:32054876

Salam, Akram, Abbas, & Anwar. (2015). Optic Disc Localization using Local Vessel Based Features and Support Vector Machine. *IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE).*

Santana, M. A., Pereira, J. M. S., Silva, F. L., Lima, N. M., Sousa, F. N., Arruda, G. M. S., Lima, R. C. F., Silva, W. W. A., & Santos, W. P. (2018). Breast cancer diagnosis based on mammary thermography and extreme learning machines. *Research on Biomedical Engineering*, *34*(1), 45–53. doi:10.1590/2446-4740.05217

Santana, M., Pereira, J., Lima, N., Sousa, F., Lima, R., & Santos, W. (2017). Classificação de lesões em imagens frontais de termografia de mama a partir de sistema inteligente de suporte ao diagnóstico. In *Anais do I Simpósio de Inovação em Engenharia Biomédica*. SABIO.

Santos, W. P., Souza, R. E., & Santos Filho, P. B. (2017). Evaluation of Alzheimer's Disease by Analysis of MR Images using Multilayer Perceptrons and Kohonen SOM Classifiers as an Alternative to the ADC Maps. In *2017 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE.

Santos, W. P., Assis, F. M., Souza, R. E., Mendes, P. B., Monteiro, H. S. S., & Alves, H. D. (2009a). A Dialectical Method to Classify Alzheimer's Magnetic Resonance Images. In *Evolutionary Computation*. IntechOpen. doi:10.5772/9609

Santos, W. P., Assis, F. M., Souza, R. E., Mendes, P. B., Monteiro, H. S. S., & Alves, H. D. (2010). Fuzzy-based Dialectical Non-supervised Image Classification and Clustering. *International Journal of Hybrid Intelligent Systems*, *7*(2), 115–124. doi:10.3233/HIS-2010-0108

Santos, W. P., Assis, F. M., Souza, R. E., & Santos Filho, P. B. (2008a). *Evaluation of Alzheimer's Disease by Analysis of MR Images using Objective Dialectical Classifiers as an Alternative to ADC Maps. In 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE.

Santos, W. P., Assis, F. M., Souza, R. E., & Santos Filho, P. B. (2009). Dialectical Classification of MR Images for the Evaluation of Alzheimer's Disease. In *Recent Advances in Biomedical Engineering*. IntechOpen. doi:10.5772/7475

Santos, W. P., Assis, F., Souza, R., Santos Filho, P. B., & Neto, F. L. (2009b). Dialectical Multispectral Classification of Diffusion-weighted Magnetic Resonance Images as an Alternative to Apparent Diffusion Coefficients Maps to Perform Anatomical Analysis. *Computerized Medical Imaging and Graphics*, *33*(6), 442–460. doi:10.1016/j.compmedimag.2009.04.004 PMID:19446434

Santos, W. P., Souza, R. E., Santos Filho, P. B., Neto, F. B. L., & Assis, F. M. (2008b). A Dialectical Approach for Classification of DW-MR Alzheimer's Images. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on COmputational Intelligence)*. Hong Kong: IEEE. 10.1109/CEC.2008.4631023

Sapra, P., Singh, R., & Khurana, S. (2013). Brain Tumor Detection using Neural Network. *International Journal of Science and Modern Engineering*, *1*(9).

Sathies Kumar, T. (2017). Brain Tumor Detection Using SVM Classifier. *IEEE 3rd International Conference on Sensing, Signal Processing and Security (ICSSS).*

Sazonov, E. S., Fulk, G., Hill, J., Schutz, Y., & Browning, R. (2011). Monitoring of posture allocations and activities by a shoe-based wearable sensor. *IEEE Transactions on Biomedical Engineering*, *58*(4), 983–990. doi:10.1109/TBME.2010.2046738 PMID:20403783

Seng & Mirisaee. (2009). Evaluation of a content-based retrieval system for blood cell images with automated methods. *Journal of Medical Systems*, *35*, 571–578.

Seo, H., & Khuzani, M. B. (2020). Machine learning techniques for biomedical image segmentation: An overview of technical aspects and introduction to state-of-art applications. *American Association of Physicists in Medicine*, *45*(5), 148–167. PMID:32418337

Seyeddain, O., Kraker, H., Redlberger, A., Dexl, A. K., Grabner, G., & Emesz, M. (2014). Reliability of automatic biometric iris recognition after phacoemulsification or drug-induced pupil dilation. *European Journal of Ophthalmology*, *24*(1), 58–62. doi:10.5301/ejo.5000343 PMID:23873488

Shahzad, M., Umar, A. I., Khan, M. A., Shirazi, S. H., Khan, Z., & Yousaf, W. (2020). Robust Method for Semantic Segmentation of Whole-Slide Blood Cell Microscopic Images. *Computational and Mathematical Methods in Medicine*, *2020*, 2020. doi:10.1155/2020/4015323 PMID:32411282

Shallu, R. M., & Mehra, R. (2018). Breast cancer histology images classification:Training from scratch or transfer learning? *ICT Express*, *4*(4), 247–254. doi:10.1016/j.icte.2018.10.007

Shelke, M. P. K., Sontakke, M. S., & Gawande, A. (2012). Intrusion detection system for cloud computing. *International Journal of Scientific & Technology Research*, *1*, 67–71.

Shen, S., Sandham, W., Granat, M., & Sterr, A. (2005). MRI Fuzzy Segmentation of Brain Tissue using Neighborhood Attraction with Neural-Network Optimization. *IEEE Transactions on Information Technology in Biomedicine*, *9*(3), 459–497. doi:10.1109/TITB.2005.847500 PMID:16167700

Singh, S., Saini, S., & Singh, M. (2012). Cancer Detection using Adaptive Neural Network. *International Journal of Advancements in Research and Technology*, *1*(4).

Smart Sensors Ltd. (2013). *MIRLIN SDK*, *2*, 23.

Sorensen, Shaker, & Bruijne. (2010). Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Trans. Med. Imaging, 29*(2), 559-569.

Suckling, J., Parker, J., Dance, D., Astley, S., Hutt, I., Boggis, C., Ricketts, I., Stamatakis, E., Cerneaz, N., Kok, S., Taylor, P., Betal, D., & Savage, J. (1994). The mammographic image analysis society digital mammogram database. In *2nd International Workshop on Digital Mammography*. Excerpta Medica.

Suganya, R., & Shanthi, R. (2012). Fuzzy C- Means Algorithm. *RE:view*, *2*(11), 1–3.

Sung, M., Marci, C., & Pentland, A. (2005). Carl Marci, Alex Pentland, "Wearable feedback systems for rehabilitation. *Journal of Neuroengineering and Rehabilitation*, *2*(1), 1–12. doi:10.1186/1743-0003-2-17

Sun, S., Li, W., & Kang, Y. (2015). Lung Nodule Detection Based on GA and SVM. *8th International Conference on Bio Medical Engineering and Informatics (BMEI 2015).*

Sutra, G., Dorizzi, B., Garcia-Salitcetti, S., & Othman, N. (2013, April 23). *A biometric reference system for iris. OSIRIS version 4.1*. Available: http://svnext. it-sudparis.eu/svnview2-eph/ref syst/Iris Osiris v4.1/

Swapnil, R. T. (2016). Detection of brain tumor from MRI images by using segmentation & SVM. *World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave).*

Tang, H., Wu, E. X., Ma, Q. Y., Gallagher, D., Perera, G. M., & Zhuang, T. (2000). MRI brain image segmentation by multi-resolution edge detection and region selection. *Computerized Medical Imaging and Graphics*, *24*(6), 349–357. doi:10.1016/S0895-6111(00)00037-9 PMID:11008183

Tan, X., & Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, *19*(6), 1635–1650. doi:10.1109/TIP.2010.2042645 PMID:20172829

Thuy, Hai, & Thai. (n.d.). *Image Classification using Support Vector Machine and Artificial Neural Network.* Academic Press.

Trokielewicz, M., Czajka, A., & Maciejewicz, P. (2014). Cataract influence on iris recognition performance. Proc. SPIE 9290, Photonics Applications in Astronomy, Communications, Industry, and High-Energy Physics Experiments. doi:10.1117/12.2076040

Trujillo, C., Piedrahita-Quintero, P., & Garcia-Sucerquia, J. (2020). Digital lensless holographic microscopy: Numerical simulation and reconstruction with ImageJ. *Applied Optics*, *59*(19), 5788–5795. doi:10.1364/AO.395672 PMID:32609706

Unique Identification Authority of India. (n.d.). *AADHAAR*. Available: https://uidai.gov.in/what-is-aadhaar.html

Urban, G. (2014). Multi-modal brain tumor segmentation using deep convolutional neural networks. *MICCAI BraTS (Brain Tumor Segmentation) Challenge. Proceedings*, 31-35.

Veras, R. (2015). SURF descriptor and pattern recognition techniques in automatic identification of pathological retinas. *2015 Brazilian Conference on Intelligent Systems*.

Vijaya, Suhasini, & Priya. (n.d.). Automatic detection of lung cancer in CT images. *IJRET: International Journal of Research in Engineering and Technology.*

Vijay, J., & Subhashini, J. (2013). An Efficient Brain Tumor Detection Methodology Using K-Means Clustering Algorithm. *International conference on Communication and Signal Processing*.

Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P.-A. (2008). Extracting and Composing Robust Features with Denoising Autoencoders. In *25th International Conference on Machine Learning*. New York: ACM. 10.1145/1390156.1390294

Wadhwani & Saraswat. (2009). Classification of breast cancer using artificial neural network. *Current Research in Engineering. Science and Technology Journals*.

Wang. (2017). The multimodal brain tumor image segmentation based on convolutional neural networks. *ICCIA*.

Wang, D., Yuan, F., & Sheng, H. (2010). An Algorithm for Medical Imaging Identification based on Edge Detection and Seed Filling. In *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*. Taiyuan: IEEE.

Wang, Y., & Cao, Y. (2019). Quick leukocyte nucleus segmentation in leukocyte counting. *Computational and Mathematical Methods in Medicine*, *2019*, 2019. doi:10.1155/2019/3072498 PMID:31308855

Wells, W. M., Grimson, W. E. L., Kikinis, R., & Jolesz, F. A. (1996). Adaptive segmentation of MRI data. *IEEE Transactions on Medical Imaging*, *15*(4), 429–442. doi:10.1109/42.511747 PMID:18215925

Werlinger, R., Hawkey, K., Muldner, K., Jaferian, P., & Beznosov, K. (2008). The challenges of using an intrusion detection system: is it worth the effort? *Proceedings of the 4th symposium on Usable privacy and security*, 107-118. 10.1145/1408664.1408679

Wu, Y., Wang, N., Zhang, H., Qin, L., Yan, Z., & Wu, Y. (2010). Application of Neural Networks in the Diagnosis of Lung Cancer by Computed Tomography. *Sixth International Conference on Natural Computation*. 10.1109/ICNC.2010.5583316

Xiao, Y., Wu, J., Lin, Z., & Zhao, X. (2018). Breast Cancer Diagnosis Using an Unsupervised Feature Extraction Algorithm Based on Deep Learning. In *2018 37th Chinese Control Conference (CCC)*. IEEE. 10.23919/ChiCC.2018.8483140

Xing, T., Huang, D., Xu, L., Chung, C.-J., & Khatkar, P. (2013). Snortflow: A openflow-based intrusion prevention system in cloud environment. *Research and Educational Experiment Workshop (GREE)*, 89-92. 10.1109/GREE.2013.25

Xu, Q., & Zhang, L. (2015). The Effect of Different Hidden Unit Number of Sparse Autoencoder. In *The 27th Chinese Control and Decision Conference (2015 CCDC)*. IEEE. 10.1109/CCDC.2015.7162335

Xue, Long, & Antani, Jeronimo, & Thoma. (2008). A Web-accessible content-based cervicographic image retrieval system. *Proceedings of the Society for Photo-Instrumentation Engineers*, 6919.

Xueming, Hua, Chen, & Liangjun. (2011). PLBP: An effective local binary patterns texture descriptor with pyramid representation. *Pattern Recognition*, *44*, 2502–2515.

Xu, S., Zhang, Y., Jia, L., & Kyle, E. (2014). Soft microfluidic assemblies of sensors, circuits, and radios for the Skin. *Science*, *344*(6179), 70–74. doi:10.1126cience.1250169 PMID:24700852

Yadav, A., & Kumar, N. (2016). A Survey of Authentication Methods in Cloud Computing. *International Journal of Innovative Research in Computer and Communication Engineering*, *4*, 19529–19533.

Yao, C.-H., & Chen, S.-Y. (2003). Retrieval of translated, rotated and scaled color textures. *Pattern Recognition*, *36*(4), 913–929. doi:10.1016/S0031-3203(02)00124-3

Yarlagadda, V. K., & Ramanujam, S. (2011). Data security in cloud computing. *Journal of Computer and Mathematical Sciences*, *2*, 1–169.

Yasiran, S. S., Salleh, S., & Mahmud, R. (2016). Haralick texture and invariant moments features for breast cancer classification. In AIP Conference Proceedings. AIP Publishing. doi:10.1063/1.4954535

Ye, S., Zheng, S., & Hao, W. (2010). Medical image edge detection method based on adaptive facet model. In *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*. Taiyuan: IEEE.

Yezzi, A. J., Kichenassamy, S., Kumar, A., Olver, P., & Tannenbaum, A. (1997). A geometric snake model for segmentation of medical imagery. *IEEE Transactions on Medical Imaging*, *16*(2), 199–209. doi:10.1109/42.563665 PMID:9101329

Yi, D. (2016). *3-D convolutional neural networks for glioblastoma segmentation*. arXiv preprint arXiv:1611.04534.

Yuan, X., Zhou, H., & Shi, P. (2007). Iris recognition: A biometric method after refractive surgery. *Journal of Zhejiang University. Science A*, *8*(8), 1227–1231. doi:10.1631/jzus.2007.A1227

Yurdakul, Subathra, & Georgec. (2020). Detection of Parkinson's Disease from gait using Neighborhood Representation Local Binary Patterns. *Biomedical Signal Processing and Control, 62*.

Zakeri, F. S., Behnam, H., & Ahmadinejad, N. (2010). Classification of benign and malignant breast masses based on shape and texture features in sonography images. *Journal of Medical Systems*, *36*(3), 1621–1627. doi:10.100710916-010-9624-7 PMID:21082222

Zhang, B., Gao, Y., Zhao, S., & Liu, J. (2010). Local derivative pattern versus local binary pattern: Face recognition with higher-order local pattern descriptor. *IEEE Transactions on Image Processing*, *19*(2), 533–544.

Zhang, D., Yang, G., & Li, F. (2020). Detecting seam carved images using uniform local binary patterns. *Multimedia Tools and Applications*, *79*, 8415–8430. doi:10.100711042-018-6470

Zheng, S. (2015). Conditional random fields as recurrent neural networks. *Proceedings of the IEEE international conference on computer vision*. 10.1109/ICCV.2015.179

Zhou, H., Stone, T., Hu, H., & Harris, N. (2008). Use of multiple wearable inertial sensors in upper limb motion tracking. *Medical Engineering & Physics*, *30*(1), 123–133. doi:10.1016/j.medengphy.2006.11.010 PMID:17251049

Zikic, D. (2014). Segmentation of brain tumor tissues with convolutional neural networks. *Proceedings MICCAI-BRATS*, 36-39.

Zulpe & Chowhan. (2011). Statical Approach For MRI Brain Tumor Quantification. *International Journal of Computer Applications, 35*(7).

About the Contributors

Kalaivani Anbarasan, an Academician and Researcher, received her M.C.A degree from Presidency College, Madras University, 1998. She received her M.Phil Computer Science from Alagappa University and also received M.E Gold Medal Computer Science and Engineering degree from St. Peter's University. She has completed her prestigious Ph.D degree in Computer Science Engineering from Anna University Chennai, 2017. She has gained her 17+ years of teaching experience from removed universities and Engineering Colleges. She received South Indian "Best Non-Circuit Faculty Award" from ASDF Organization during April 2018. She also received Best Teacher award from Lions Club International during October 2018. She received Best Paper Award in International / National Conferences. She is acting as a Trustees for P.T. Lee Chengalvaraya Naicker Trust by Honourable High Court of Madras for the period of three years (2017-2020). She is appointed as a Governing Council Member and School Committee Member for P.T. Lee Chengalvaraya Naicker Trust for the period of three years (2017-2020). She has attended national workshop, FDP and Training programs and also completed NPTEL online certification course. She has published papers in high reputed journal and presented papers in IEEE and Springer Conference. She acted as a resource person and delivered lecturers to both faculties and students on research focus and upcoming technologies. She served as a technical advisory committee member for reputed international conferences. She is acting as an reviewer and editor for high reputed international journals and international conferences.

* * *

Arthi B. holds a Ph.D. degree in the field of Computer Science and Engineering from Anna University. She has 14 years of experience in teaching. Her area of interest includes Software Engineering, IOT, Cloud Computing and Green Computing. She has published several articles in various reputed journals. She has presented papers in various national and international conferences and attended many workshops, seminars and faculty development programs in order to be in track with the changing technology and teaching methodology. She is a member of various scientific and professional bodies. She has been awarded the IET Inspiring Young Teacher Award for the year 2016-2017 for the IET Chennai Local Network.

D. K. Chaturvedi is working in Dept. of Elect. Engg, Faculty of Engg, D.E.I., Dayalbagh, Agra since 1989. Presently he is Professor. He did his B.E. from Govt. Engineering College Ujjain, M.P. then he did his M.Tech. and Ph.D. from D.E.I. Dayalbagh. He is gold medalist and received Young Scientists Fellowship from DST, Government of India in 2001-2002 for post doctorial research at Univ. of Calgary, Canada. Also, he had research collaboration with different organizations at national and international level. He is the Fellow - The Institution of Engineers (India), Fellow - Aeronautical Society of India, Fellow - IETE, Sr. Member IEEE, USA and Member of many National and International professional bodies such as IET, U.K., ISTE, Delhi, ISCE, Roorkee, IIIE, Mumbai and SSI etc. The IEE, U.K. recognized his work in the area of Power System Stabilizer and awarded honorary membership to him in 2006. He did many R&D projects of MHRD, UGC, AICTE etc. and consultancy projects of DRDO. He contributed in the national mission of ICT of Govt. of India as Virtual Power Lab Developer. He has guided 10 Ph.Ds., 65 M.Tech. Dissertations and published more than 300 International and National Papers. He has chaired and Co-Chaired many International and National Conferences. He is referee of many International Journals including IEE Proceedings and IEEE Transactions. He is Head of Dept. of Footwear Technology, Convener, Faculty Training and Placement Cell, and Advisor, IEI Students' Chapter (Elect. Engg.), D.E.I. Dayalbagh, Agra.

Washington Wagner da Silva holds a PhD in Computer Science from the Federal University of Pernambuco - UFPE (2017). He holds a Master's degree in Computer Science from the Federal University of Pernambuco - UFPE (2011). He holds a degree in Systems Analysis from the Salgado de Oliveira University - UNIVERSO (2004). He has a postdoctoral degree in the Department of Biomedical Engineering of the Federal University of Pernambuco - UFPE (10/2017 until 10/2019) having as supervisor Professor. Dr. Wellington Pinheiro dos Santos. He was a Test Engineer of the CIn / Motorola project (from May 10, 2006 to October 31, 2007). He has experience in Computer Science, acting on the following subjects: Software Testing Engineering, Artificial Intelligence, Artificial Neural Networks, Hybrid Intelligent Systems, Handwriting Character Recognition, Pattern Recognition and Biomedical Engineering.

Maira de Santana has a Master's degree in Biomedical Engineering at the Federal University of Pernambuco (UFPE) and member of the Biomedical Computing Research Group. She holds a degree in Biomedical Engineering from the Federal University of Pernambuco (2017). She has fluency in Portuguese (native language) and English, as well as basic knowledge of Spanish and German. She completed an internship in Clinical Engineering at Hospital das Clínicas de Pernambuco (09/2016 - 01/2017). She was a Scholarship for Science for Borders Program of the Federal Government / CAPES in the United States for one year (edict 180: 08/2015 - 08/2016), of which nine months (08/2015 - 05/2016) were dedicated to She has a BA in Biomedical Engineering at the University of Alabama at Birmingham (UAB), AL (USA) and in the last three months (05/2016 - 08/2016) she worked as a researcher at the Carl E Ravin Advanced Imaging Laboratories (RAILabs) - Duke University, NC, USA - deepening specific knowledge of the area of image processing (sub-area of Biomedical Engineering) and acquiring experience in laboratory stage, scientific production, programming in MATLAB language and Office package.

Wellington Pinheiro dos Santos received a bachelor's degree in Electrical Electronics Engineering (2001) and MSc in Electrical Engineering (2003) from the Federal University of Pernambuco, and Ph.D. in Electrical Engineering from the Federal University of Campina Grande (2009). He is currently a Professor of the Department of Biomedical Engineering at the Federal University of Pernambuco, acting in Undergraduate and Graduate Programs in Biomedical Engineering. He is a member of the Graduate Program in Computer Engineering from the Polytechnic School of Pernambuco, University of Pernambuco, since 2009. He has experience in the area of Computer Science, acting on the following themes: digital image processing, pattern recognition, computer vision, evolutionary computation, numerical methods of optimization, computational intelligence, computer graphics, virtual reality, game design and applications of Computing and Engineering in Medicine and Biology. He is a member of the Brazilian Society of Biomedical Engineering (SBEB), the Brazilian Society of Computational Intelligence, and the International Federation of Medical and Biological Engineering (IFMBE).

Balanagireddy G. is currently working as an Assistant Professor in the department of Electronics and Communication Engineering, Rajiv Gandhi University of Technologies-AP Dr.APJ Abdul Kalam Campus Ongole, India. He did his M.Tech with Jawaharlal Nehru Technical University Hyderabad, India and pursuing Ph.D with Visveswaraiah Technological University Belgaum, India. His area of interest includes Nanoelectronics, MEMS and Ad-hoc networks. He has very good achievements in NPTEL online courses. He is an active member of ISSS, ISTE and The Indian science congress Association.

Padmapriya Govindhan received the M.E., degree in Computer Science and Engineering from K. S. Rangasamy College of Technology, Tiruchengode in 2008 and Ph.D in Information and Communication Engineering at Anna University, Chennai in 2016. She worked as Assistant Professor in Department of Computer Science and Engineering at K.S.R. College of Engineering (Autonomous), Tiruchengode from 2008, Associate Professor in Department of Computer Science and Engineering at Vidyaa Vikas College of Engineering and Technology, Tiruchengode from 2016 and currently working as Associate Professor at the Department of Computer Science and Engineering, at Saveetha School of Engineering, SIMATS, Chennai, India.. She has published more than 28 papers in refereed journals and conference proceedings. Her current research interest includes Data Mining, Text Mining, Information Retrieval, Natural Language processing Deep Learning and Neural Networks. She is a member of ISTE, IAENG, and CSTA.

Mayank Gupta is acting as System and IT Analyst in Tata Consultancy services, Noida and expert of Data sciences and Business Analytics. He has skill to visualize the situations from different perspectives and explore the real facts through critical Analysis. He has deep interest in Human health domains.

Ananthajothi K. (1983) is an Assistant Professor in the Department of Computer Science and Engineering at Misrimal Navajee Munoth Jain Engineering College, Chennai,(INDIA). He obtained his Master degree (M.E) in Computer Science and Engineering and Ph.D in Computer Science and Engineering from Anna University in the year 2010 and 2020 respectively. His research focuses are Data Mining and Big-data. He published book title of Theory of computation.

Sujatha Kesavan is presently working as Professor in EEE, Department at Dr. M.G.R Educational and Research Institute, Chennai, Tamil Nadu, India and heading the Research centre 'Center for Electronics, Automation and Industrial Research (CEAIR). She has 20 years of teaching experience in various Engineering colleges. She completed her BE in the year 1999 from Bharathiyar University, ME in 2004 and Ph. D in 2011 from Anna University. She has presented/published papers in National/International conferences/journals and also published many books with Elsevier and Springer publisher. She is also a reviewer for journals published by Springer and Elsevier publishers. Presently doing her research in the area of Image Processing for Process Control. She is awarded the Best Researcher award for the academic year 2011–2012 and 2014 by IET. Also obtained travel grant from DST in 2014 for attending the conference. She is also awarded the young researcher award at the international conference at China in year 2015. She has also published 4 patents including one international patent with the Chinese University at Huaiyin

Institute of Technology, China and also initiated international cooperation research between Huaiyin Institute of Technology, China and Dr.MGR Educational and Research Institute.

Aruna M. holds Ph.D. Degree in the Faculty of Information and Communication Engineering from Anna University, Chennai. She has 14 years of experience in teaching and currently working as Assistant Professor (SG) in the Department of Computer Science and Engineering at Saveetha School of Engineering, SIMATS, Chennai, India. Her area of interest includes Cloud Computing, IoT, Artificial Intelligence and Machine Learning, Green Computing and Software Engineering, Engineering. She has published several articles in various reputed Journals and Conferences proceedings in order to be in track with the changing technology and teaching methodology. She is a member of various Scientific and Professional bodies such as ISTE, IAENG, CSTA, SDIWC.

Malathi M. received her B.E degree in Electronics and Instrumentation engineering from Madurai Kamaraj University, in 2004, M.E degree in Applied Electronics from Sathyabama University in 2011. She is currently working as Associate Professor in the Department of Electronics and communication Engineering, Chennai Institute of Technology, Chennai, India. She has completed Ph.D from Dr.MGR. Educational Research Institute in 2018. She obtained Anna university guide ship under the area of Medical Image processing.

Deepa Narayanan is currently working as Assistant Professor in Saveetha School of Engineering in Computer Science and Engineering Department. She has published papers in several Scopus, journals and presented papers in both national and international conferences. Her areas of interest includes image processing, big data and video analytics. She has guided many students in various domains and has motivated students for paper presentation and publication.

Elakkiya R. is an Assistant Professor in the Department of Computer Science and Engineering, School of Computing, SASTRA University, Thanjavur. She received her Ph.D. (Information & Communication Engineering) and an M.E (Software Engineering) from Anna University, Chennai, India in 2018 and 2012, respectively. She received her Bachelor's degree in Computer Science & Engineering from Anna University, Chennai, India in 2010. She has more than 7 years of research experience in Machine Learning and Computer vision. She has published more than 20 research articles in leading journals, conference proceedings and book including IEEE, Elsevier and Springer. Currently, she is an editor of Information Engineering and Applied Computing journal and also, she is a Life time member of International Association of Engineers.

Rohit Rastogi received his B.E. degree in Computer Science and Engineering from C.C.S.Univ. Meerut in 2003, the M.E. degree in Computer Science from NITTTR-Chandigarh (National Institute of Technical Teachers Training and Research-affiliated to MHRD, Govt. of India), Punjab Univ. Chandigarh in 2010. Currently he is pursuing his Ph.D. In computer science from Dayalbagh Educational Institute, Agra under renowned professor of Electrical Engineering Dr. D.K. Chaturvedi in area of spiritual consciousness. Dr. Santosh Satya of IIT-Delhi and dr. Navneet Arora of IIT-Roorkee have happily consented him to co supervise. He is also working presently with Dr. Piyush Trivedi of DSVV Hardwar, India in center of Scientific spirituality. He is a Associate Professor of CSE Dept. in ABES Engineering. College, Ghaziabad (U.P.-India), affiliated to Dr. A.P. J. Abdul Kalam Technical Univ. Lucknow (earlier Uttar Pradesh Tech. University). Also, He is preparing some interesting algorithms on Swarm Intelligence approaches like PSO, ACO and BCO etc. Rohit Rastogi is involved actively with Vichaar Krnati Abhiyaan and strongly believe that transformation starts within self.

Anusuya S. is Professor & Head Department of Information Technology Saveetha School of Engineering Saveetha Institute of Medical and Technical Sciences.

Kannan S. is Asst. Professor, Dept of Computer Science and Engineering Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences.

Ganesh T. R. is currently working as a professor, Department of Electronics and Communication Engineering, Muthayammal Engineering College, Rasipuram, Namakkal District. He has obtained his A M I E degree in Electronics and Communication Engineering from The Institution of Engineers (India) in 1993. He received his ME degree in Applied Electronics from Karunya Institute of Technology, Coimbatore, in 2001. He obtained his Ph.D from Anna University, Chennai, India in 2014 in the field of Medical Image Processing. He has authored over Seventy eight research publications in international and national journals conferences and books. Published 1 Patent and filed 1 patent. His special areas of interest are Signal Processing, Wireless Networks, Image Processing, Control system and Biomedical Instrumentation. He guided three Ph.D Scholars and pursuing eight scholars from Anna University, Chennai, India. Reviewed around ten referred journals like Springer, Elseware, Biomedical research, and Inder Science publishers. He received distinguish faculty award in Venus international Foundation, Chennai, India .He is a editorial board member of internal scientific journal of Contemporary research in engineering science and Management .He is a Life time member of IETE and IE.

Devi Thiyagarajan is currently working as Assistant Professor in Saveetha School of Engineering in Computer Science and Engineering Department. She has published papers in several Scopus, Web of Science, SCI indexed journals and presented papers in both national and international conferences. Her areas of interest includes image processing, cloud computing and video analytics. She has guided many students in various domains and has motivated students for paper presentation and publication.

Sudha V. is currently pursuing her Ph.D in Bio-medical Image Processing under Anna University, Chennai registered in 2017. She has completed her P.G in Anna University of Technology, Coimbatore in 2012. Currently, She is working as Assistant Professor in the Department of Electronics & Communication Engineering at Sona College of Technology, (Autonomous) Salem. She possess 9 years of teaching experience and has made significant contributions in the area of Deep learning using Convolutional neural networks, Medical image analysis for Diabetic Retinopathy. She has made significant publications in International and National journals.

Index

P

R

S

T

W