Deep Convolutional Neural Network and Emotional Learning Based Breast Cancer Detection using Digital Mammography

Sravani Vulisetti, M Tech student, CSE:BVCE, Email:sravanivulisetti7@gmail.com
Gunamani Jena, Prof CSE:BVCE, drgjena@gmail.com
Chandra Mouli VSA, Prof. CSE:BVCE, E-mail: mouliac@yahoo.co.in
V Dhanaraj, Prof. ECE:BVCE,drvalluri@hotmail.com
Shubhashish Jena CSE,OUTR, sjena1998@gmail.com

ABSTRACT

Breast cancer is one of the deadly diseases among women. However, the chances of death are highly reduced if it gets diagnosed and treated at its early stage. Mammography is one of the reliable methods used by the radiologist to detect breast cancer at its initial stage. Therefore, an automatic and secure breast cancer detection system that accurately detects abnormalities not only increases the radiologist's diagnostic confidence but also provides more objective evidence. In this work, an automatic Diverse Features based Breast Cancer Detection (DFeBCD) system is proposed to classify a mammogram as normal or abnormal. Four sets of distinct feature types are used. Among them, features based on taxonomic indexes, statistical measures and local binary patterns are static. The proposed DFeBCD dynamically extracts the fourth set of features from mammogram images using a highwaynetwork based deep convolution neural network (CNN). Two classifiers, Support Vector Machine (SVM) and Emotional Learning inspired Ensemble Classifier (ELiEC), are trained on these distinct features using a standard IRMA mammogram dataset. The reliability of the system performance is ensured by applying 5-folds crossvalidation. Through experiments, we have observed that the performance of the DFeBCD system on dynamically generated features through highway network-based CNN is better than that of all the three individual sets of ad-hoc features. Furthermore, the hybridization of all four types of features improves the system's performance by nearly 2–3%. The performance of both the classifiers is comparable using the individual sets of ad-hoc features. However, the ELiEC classifier's performance is better than SVM using both hybrid and dynamic features. We adding new algorithm called ELM and then training with Brisk features and this combination is giving accuracy closer to 100%.

Keywords:CNN, ELiECSVM, ELM, Brisk features, LBA (local binary patterns)
DOINumber:10.14704/nq.2022.20.10.NQ55729 **NeuroQuantology 2022;20(10):7392-7405**

I. INTRODUCTION:

Breast cancer has become one of the deadliest diseases among women around the world. This starts due to the uncontrolled cell division that results in the formation of a tumor in the breast [1]. The symptoms of breast cancer include abnormalities like breast aches, change in breast skin color, change in dimension and shape, and formation of a breast mass. Normally, imaging techniques like ultrasound, X-rays and magnetic imaging are used to

analyze breast cancer. However, for early breast cancer detection, one of the most effective techniques is mammography [2] which uses low-dose of X-rays for image formation. The abnormalities like calcification and masses and other subtle signs like architectural distortion and bilateral asymmetry could be detected using mammography. A mass, density, nodule or distortion in the mammogram represent potential abnormalities. However, not all abnormalities are cancerous. For example, a smooth and well-defined bordered mass (lump) is normally benign. On the other hand, an irregular bordered mass (lump) with a starburst appearance (speculated) might be cancerous, and to verify it, a biopsy is needed. Micro-calcifications are basically a small set of calcium clusters that might be of benign, indeterminate or suspicious nature. Mostly these micro-calcifications clusters are of benign nature. However, in some of the cases, these micro-calcifications might appear in a certain form of clusters and patterns. In that case, they indicate precancerous cells or may represent an initial stage of breast cancer, which could be verified by biopsy. To diagnose breast cancer, radiologists analyze these mammogram images. However, the opinion of the radiologists regarding the presence of breast cancer may not be consistent due to the differences in their previous knowledge and experiences. Therefore, a machine learning based breast cancer detection system may be used to increase the radiologist confidence and can also be used as a second opinion for the detection of breast cancer [3]. In the past two decades, machine learning algorithms have gained popularity by solving problems of complex nature, like clustering, prediction and classification [3–5].

Machine learning (ML) based techniques have shown exemplary performance for different image recognition problems [6]. Based on complexity, experimental data and biological description, different mathematical models [7] exist for the control [8] and prediction [9]of the glucose-insulin system. Previously, many ML-based computer-aided diagnostic systems have been successfully developed for the diabetic retinopathy problem [10,11]. In this regard, SVM is often applied to classification problems because of its strong discrimination ability [12]. To detect the breast cancer at its early stage, researches are continuously developing interesting machine learning-based systems. The major difference in these works is the usage of different feature extraction methodologies, mammogram datasets and the machine learning model. The problem of detecting breast cancer can be treated as a hierarchical classification problem [13,14]. First, the mammograms are classified into normal or abnormal. If it is detected as abnormal, it can be further classified as benign and malicious. Most of the previous work commonly used models based on statistics, texture, or signal processing for feature extraction. The majority of the reported works has used Support Vector Machine (SVM) as a classifier and is evaluated by taking varying subsets of either the Digital Database for FIGUREing Mammography (DDSM) or the database of Mammographic Image Analysis Society (MIAS). Harefa et al. showed that using grey level co-occurrence matrices (GLCM), SVM outperformed k-Nearest Neighbour (k-NN) in detecting breast cancer abnormalities with an accuracy of 93.88% on the MIAS database [15]. De Oliveira et al. used the taxonomic distinctness index and taxonomic diversity index and achieved a maximum accuracy of 98.88% with the SVM classifier [16]. Gorgela ¨ et al. used local seed growing technique with spherical wavelet transformation in combination with SVM classifier for the classification of mass/non-mass and reported an accuracy of 94% [17]. Gabor filters of various scales and directions are employed by Hussain and achieved 0.98 area under ROC [18]. Berbar et al. exploited hybrid features based on local binary patterns and statistical measures and thus reported accuracies of 98.63% and 97.25% with SVM and k-NN, respectively [19]. Nithya et al. reported an accuracy of 98% using the statistical, grey level, and horizontal based features [20]. The above discussion shows that the texture-based features seem more feasible to achieve good breast cancer detection results with SVM.

1.1 Mammography

Mammography is specialized medical imaging that uses a low-dose x-ray system to see inside the breasts. A mammography exam, called a mammogram, aids in the early detection and diagnosis of breast diseases in women.An x-ray exam helps doctors diagnose and treat medical conditions. It exposes you to a small dose of ionizing radiation to produce pictures of the inside of the body. X-rays are the oldest and most often used form of medical imaging.Three recent advances in mammography include digital mammography, computer-aided detection and breast tomosynthesis.*Digital mammography*, also called full-field digital mammography (FFDM), is a mammography system in which the x-ray film is replaced by electronics that convert x-rays into mammographic pictures of the breast. These systems are similar to those found in digital cameras and their efficiency enables better pictures with a lower radiation dose. These images of the breast are transferred to a computer for review by the radiologist and for long term storage. The patient's experience during a digital mammogram is similar to having a conventional film mammogram.*Computer-aided detection (CAD)* systems search digitized mammographic images for abnormal areas of density, mass, or calcification that may indicate the presence of cancer. The CAD system highlights these areas on the images, alerting the radiologist to carefully assess this area.

1.2 Breast cancer

Breast cancer is the most diagnosed cancer among women worldwide, accounting for 1 in 4 cancer cases. It is the most frequent cancer amongst both sexes and is the leading cause of death from cancer in women. The estimated 2.3 million new cases indicate that one in every 8 cancers diagnosed in 2020 is breast cancer. In 2020, there were an estimated 684,996 deaths from breast cancer, with a disproportionate number of these deaths occurring in low-resource settings.[1]

Breast cancer cells usually form a tumour that can often be seen on an x-ray or felt as a lump. If spread outside the breast through blood vessels and lymph vessels, it becomes advanced breast cancer. When breast cancer spreads to other parts of the body (such as the liver, lungs, bones or brain), it is said to have metastasised, and is referred to as metastatic breast cancer.Survival rates for breast cancer are very high when the cancer is detected early and where treatment is available. Unfortunately, 50 to 80% of breast cancer cases are diagnosed at an advanced stage[2] in many low- and middle-income countries, when the cancer is more difficult to treat, is more expensive to do so, and is usually incurable. To tackle the growing breast cancer burden, it is critical that improvements are made in access to early detection, timely access to treatment and care, palliative and survivorship care, and comprehensive data collection through robust cancer registries.

1.3 Motivation:

Breast cancer is one of the deadly diseases among women. However, the chances of death are highly reduced if it gets diagnosed and treated at its early stage. Mammography is one of the reliable methods used by the radiologist to detect breast cancer at its initial stage.

1.4 Problem Statement:

According to a study for Breast Cancer Care, we have discovered that 42% of National Health Service (NHS) trusts say that they do not have the staff to assign individuals with limited breast cancer specialist nurse. It is the most important reason that can cause low survival rate of breast cancer all around the world. Due to lack of breast cancer specialist nurse or doctor, it will cause late diagnosis of breast cancer, lack of compliance to optimal detection or treatment, and inequity of access to optimal treatment. Therefore breast cancer detection is developed to perform effectiveness in both abnormalities and classification breast detection. This is to assist and diagnose the breast cancer.

1.5 Objective:

The major goal of this article is to address breast cancer disease lots of peoples are losing their lives and this death rate can be reduced by diagnosing this disease in time.

1.6 Scope of work:

Due to breast cancer disease lots of peoples are losing their lives and this death rate can be reduced by diagnosing this disease in time. In propose paper author is applying 4 features extraction technique on IRMA mammography dataset to efficiently predict such cancer.

II. Literature Survey

1. COVID-19 Detection in Chest X-Ray Images Using a New Channel Boosted CNN

COVID-19 is a highly contagious respiratory infection that has affected a large population across the world and continues with its devastating consequences. It is imperative to detect COVID-19 at the earliest to limit the span of infection. In this work, a new classification technique CB-STM-RENet based on deep Convolutional Neural Network (CNN) and Channel Boosting is proposed for the FIGUREing of COVID-19 in chest X-Rays. In this connection, to learn the COVID-19 specific radiographic patterns, a new convolution block based on split-transform-merge (STM) is developed. This new block systematically incorporates region and edge-based operations at each branch to capture the diverse set of features at various levels, especially those related to region homogeneity, textural variations, and boundaries of the infected region. The learning and discrimination capability of the proposed CNN architecture is enhanced by exploiting the Channel Boosting idea that concatenates the auxiliary channels along with the original channels. The auxiliary channels are generated from the pre-trained CNNs using Transfer Learning. The effectiveness of the proposed technique CB-STM-RENet is evaluated on three different datasets of chest X-Rays namely CoV-Healthy-6k, CoV-NonCoV-10k, and CoV-NonCoV-15k. The performance comparison of the proposed CB-STM-RENet with the existing techniques exhibits high performance both in discriminating COVID-19 chest infections from Healthy, as well as, other types of chest infections. CB-STM-RENet provides the highest performance on all these three datasets; especially on the stringent CoV-NonCoV-15k dataset. The good detection rate (97%), and high precision (93%) of the proposed technique suggest that it can be adapted for the diagnosis of COVID-19 infected patients.

2. Coronavirus Disease Analysis Using Chest X-Ray Images and a Novel Deep Convolutional Neural Network

Purpose: The novel coronavirus (COVID-19) is quickly spreading throughout the world, but facilities in the hospitals are limited. Therefore, diagnostic tests are required to timely identify COVID-19 infected patients, and thus reduce the spread of COVID-19. Methods: The proposed method exploits the learning capability of the convolutional neural network (CNN) to classify COVID-19 infected versus healthy patients. The classification is accomplished using a new CNN architecture suitable for pneumonia-based analysis of

COVID-19 chest X-ray images. The proposed COVID-19 RENet is an encoder-based CNN architecture that is well suited for feature extraction and image analysis. It is observed that the systematic dimensionality reduction through several layers combined with the synchronization of max-pooling (edge-based information extraction) and average pooling (Region-based information extraction) is well suited for image analysis. Finally, the deep features are extracted from CNN architecture and fed into the SVM classifier to improve the classification performance. The proposed technique is evaluated and compared with existing existing techniques using 5-fold cross-validation on the COVID-19 X-ray dataset. Implementation of the model is available at https://github.com/m-mohsin-zafar/shk_covid_pytorch. Results: The proposed technique shows good performance and in most of the cases, outperforms the current techniques using metrics such as the accuracy, F-score, and ROC curve. The proposed approach (concatenated deep features of both the COVID-RENet and COV-VGGNet model) achieved the highest classification performance on COVID-19 X-ray images. Objective evaluation of proposed approach achieved an accuracy of 98.3%, AUC: 0.98, F-score: 0.98, Recall: 0.97, and Precision: 0.99, respectively. Conclusions: The performance analysis of the proposed deep classification scheme and its comparison with the existing CNN models suggests that the concatenated deep feature-based model outperforms the existing models and is expected to help medical practitioners for the diagnosis prediction of COVID-19 infected patients. Moreover, it has the potential to adopt in the future for the analysis of different types of chest X-ray abnormalities.

3. Classification and Region Analysis of COVID-19 Infection Using Lung CT Images and Deep Convolutional Neural Networks

COVID-19 is a global health problem. Consequently, early detection and analysis of the infection patterns are crucial for controlling infection spread as well as devising a treatment plan. This work proposes a two-stage deep Convolutional Neural Networks (CNNs) based framework for delineation of COVID-19 infected regions in Lung CT images. In the first stage, initially, COVID-19 specific CT image features are enhanced using a two-level discrete wavelet transformation. These enhanced CT images are then classified using the proposed custom-made deep CoV-CTNet. In the second stage, the CT images classified as infectious images are provided to the segmentation models for the identification and analysis of COVID-19 infectious regions. In this regard, we propose a novel semantic segmentation model CoV-RASeg, which systematically uses average and max pooling operations in the encoder and decoder blocks. This systematic utilization of max and average pooling operations helps the proposed CoV-RASeg in simultaneously learning both the boundaries and region homogeneity. Moreover, the idea of attention is incorporated to deal with mildly infected regions. The proposed two-stage framework is evaluated on a standard Lung CT image dataset, and its performance is compared with the existing deep CNN models. The performance of the proposed CoV-CTNet is evaluated using Mathew Correlation Coefficient (MCC) measure (0.98) and that of proposed CoV-RASeg using Dice Similarity (DS) score (0.95). The promising results on an unseen test set suggest that the proposed framework has the potential to help the radiologists in the identification and analysis of COVID-19 infected regions.

4. Using Genetic Algorithm for Identification of Diabetic Retinal Exudates in Digital Color Images

Blood vessels in ophthalmoscope images play an important role in diagnosis of some serious pathology on retinal images. Hence, accurate extraction of vessels is becoming a main topic of this research area. In this paper, a new hybrid approach called the (Genetic algorithm and vertex chain code) for blood vessel detection. And this method uses geometrical parameters of retinal vascular tree for diagnosing of hypertension and identified retinal exudates automatically from color retinal images. The skeletons of the segmented trees are produced by thinning. Three types of landmarks in the skeleton must be detected: terminal points, bifurcation and crossing points, these points are labeled and stored as a chain code. Results of the proposed system can achieve a diagnostic accuracy with 96.0% sensitivity and 98.4% specificity for the identification of images containing any evidence of retinopathy.

5. Evolutionary computing enriched computer-aided diagnosis system for diabetic retinopathy

Complications caused due to diabetes mellitus result in significant microvasculature that eventually causes diabetic retinopathy (DR) that keeps on increasing with time, and eventually causes complete vision loss. Identifying subtle variations in morphological changes in retinal blood vessels, optic disk, exudates, microaneurysms, hemorrhage, etc., is complicated and requires a robust computer-aided diagnosis (CAD) system so as to enable earlier and efficient DR diagnosis practices. In the majority of the existing CAD systems, functional enhancements have been realized time and again to ensure accurate and efficient diagnosis of DR. In this survey paper, a number of existing literature presenting DR CAD systems are discussed and analyzed. Both traditional and varoius evolutionary approaches, including genetic algorithm, particle swarm optimization, ant colony optimization, bee colony optimization, etc., based DR CAD have also been studied and their respective efficiencies have been discussed. Our survey revealed that evolutionary computing methods can play a vital role for optimizing DR-CAD functional components, such as proprocessing by enhancing filters coefficient, segmentation by enriching clustering, feature extraction, feature selection, and dimensional reduction, as well as classification.

6. An emotional learning-inspired ensemble classifier (ELiEC)

In this paper, we suggest an inspired architecture by brain emotional processing for classification applications. The architecture is a type of ensemble classifier and is referred to as 'emotional learning-inspired ensemble classifier' (ELiEC). In this paper, we suggest the weighted k-nearest neighbor classifier as the basic classifier of ELiEC. We evaluate the ELiEC's performance by classifying some benchmark datasets.

7. Towards a standard reference database for computer-aided mammography

Because of the lack of mammography databases with a large amount of codified images and identified characteristics like pathology, type of breast tissue, and abnormality, there is a problem for the development of robust systems for computer-aided diagnosis. Integrated to the Image Retrieval in Medical Applications (IRMA) project, we present an available mammography database developed from the union of: The Mammographic Image Analysis Society Digital Mammogram Database (MIAS), The Digital Database for FIGUREing Mammography (DDSM), the Lawrence Livermore National Laboratory (LLNL), and routine images from the Rheinisch-WestfälischeTechnischeHochschule (RWTH) Aachen. Using the IRMA code, standardized coding of tissue type, tumor staging, and lesion description was developed according to the American College of Radiology (ACR) tissue codes and the ACR breast imaging reporting and data system (BI-RADS). The import was done automatically using scripts for image download, file format conversion, file name, web page and information file browsing. Disregarding the resolution, this resulted in a total of 10,509 reference images, and 6,767 images are associated with an IRMA contour information feature file. In accordance to the respective license agreements, the database will be made freely available for research purposes, and may be used for image based evaluation campaigns such as the Cross Language Evaluation Forum (CLEF). We have also shown that it can be extended easily with further cases imported from a picture archiving and communication system (PACS).

8. Two-phase deep convolutional neural network for reducing class skewness in histopathological images based breast cancer detection

Different types of breast cancer are affecting lives of women across the world. Common types include Ductal carcinoma in situ (DCIS), Invasive ductal carcinoma (IDC), Tubular carcinoma, Medullary carcinoma, and Invasive lobular carcinoma (ILC). While detecting cancer, one important factor is mitotic count - showing how rapidly the cells are dividing. But the class imbalance problem, due to the small number of mitotic nuclei in comparison to the overwhelming number of non-mitotic nuclei, affects the performance of classification models. This work presents a two-phase model to mitigate the class biasness issue while classifying mitotic and non-mitotic nuclei in breast cancer histopathology images through a deep convolutional neural network (CNN). First, nuclei are segmented out using blue ratio and global binary thresholding. In Phase-1 a CNN is then trained on the segmented out 80×80 pixel patches based on a standard dataset. Hard non-mitotic examples are identified and augmented; mitotic examples are oversampled by rotation and flipping; whereas non-mitotic examples are undersampled by blue ratio histogram based k-means clustering. Based on this information from Phase-1, the dataset is modified for Phase-2 in order to reduce the effects of class imbalance. The proposed CNN architecture and data balancing technique yielded an F-measure of 0.79, and outperformed all the methods relying on specific handcrafted features, as well as those using a combination of handcrafted and CNN-generated features.

2.2 Research Contribution:

In this paper, CNN algorithm, SVM, ELIEC (Emotional Learning inspired Ensemble Classifier), Extreme Learning Machine (ELM) and then training with Brisk features and this combination is giving accuracy closer to 100% for based breast cancer detection.

III. Proposed System

In this proposed work we applied 4 features extraction technique on IRMA mammography dataset to efficiently predict such cancer. Used Statistical Measures, Taxonomy, LBA (local binary patterns) and Dynamic Features Extraction using CNN algorithm and in all features extraction CNN features are giving better result. Author has evaluated performance of all 4 features extraction algorithm by using SVM and ELIEC (Emotional Learning inspired Ensemble Classifier). ELIEC is the super form of WEIGHTED KNN Algorithm, adding new algorithm called ELM and then training with Brisk features and this combination is giving accuracy closer to 100%

3.1 Algorithms/ techniques:

In this study, we useCNN algorithm, SVM, ELIEC (Emotional Learning inspired Ensemble Classifier), Extreme Learning Machine (ELM)

3.1.1 CNN ALGORITHM

Concept of Neural networks:

In a regular Neural Network there are three types of layersInput Layers: It's the layer in which we give input to our model. The number of neurons in this layer is equal to the total number of features in our data (number of pixels in the case of an image).

Hidden Layer:

The input from the Input layer is then feed into the hidden layer. There can be many hidden layers depending upon our model and data size. Each hidden layer can have different numbers of neurons which are generally greater than the number of features.

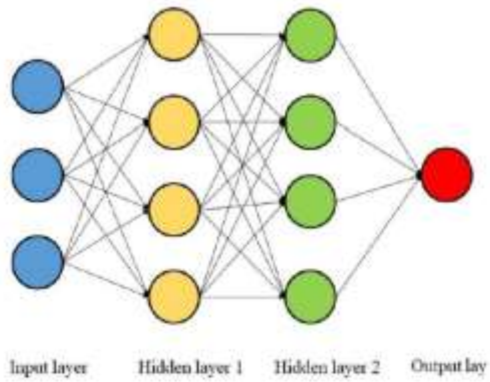

Fig.1. CNN

Output Layer:

The output from the hidden layer is then fed into a logistic function like sigmoid or soft max which converts the output of each class into the probability score of each class.

3.1.2 SVM

Support Vector Machine (SVM) is a supervised machine learning algorithm. This model has better prediction performance in short and medium term compared to long term. Every algorithm has its way of learning patterns and then predicting. The predictability of financial trend with SVM model by evaluating the weekly trend of NIKKEI 225 index. SVM is a boundary that best separates two classes with employing a line or hyperplane. The decision boundary is defined in Equation. SVMs convert non-separable classes to separable ones by kernel functions such as linear, non-linear, sigmoid, radial basis function (RBF) and polynomial.

3.1.3 ELIEC (Emotional Learning inspired Ensemble Classifier):

The ELiEC model is a general purpose classification method that aims to reduce the misclassification rate and time complexity in classification applications. The ELiEC model can be categorized as a type of ensemble classifiers.

3.1.4 Extreme Learning Machine (ELM)

Extreme learning machines are feed-forward neural networks having a single layer or multiple layers of hidden nodes for classification, regression, clustering, sparse approximation, compression, and feature learning, where the hidden node parameters do not need to be modified. These hidden nodes might be assigned at random and never updated, or they can be inherited from their predecessors and never modified. In most cases, the weights of hidden nodes are usually learned in a single step which essentially results in a fast learning scheme.

These models, according to their inventors, are capable of producing good generalization performance and learning thousands of times quicker than backpropagation networks. These models can also outperform support vector machines in classification and regression applications, according to the research.

3.3 Architecture/Framework:

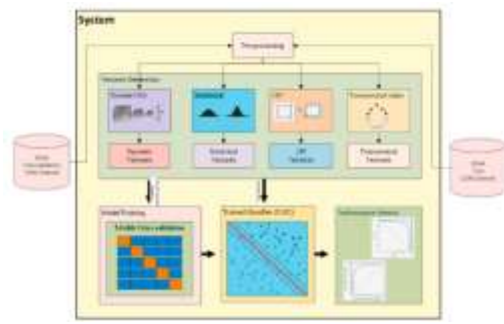

Fig.2. Framework

3.4 Algorithm and Process Design:

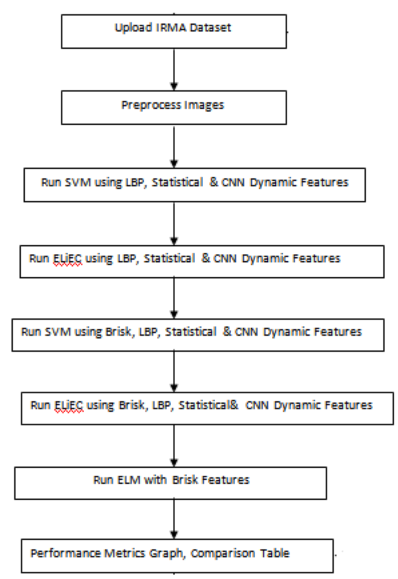

Fig.3. Process Design

Upload IRMA Dataset: using this module we will upload dataset images to application

1) Preprocess Images: using this module we will read all images and then extract Statistical features, LBP and Dynamic CNN features
2) **Run SVM using LBP, Statistical & CNN Dynamic Features:** using this module we will train SVM with 3 different features such as Statistical features, LBP and dynamic CNN and then apply this trained SVM model on test data to calculate ROC-AUC, Precision and Accuracy
3) **Run ELiEC using LBP, Statistical & CNN Dynamic Features:** using this module we will train ELIEC with all 3 features and then apply this trained model on test data to calculate performance metrics like accuracy etc.
4) **Performance Metrics:** using this module we will plot accuracy for both algorithms
   Comparison Table: using this module we will show comparison values for both algorithms on all 3 features

4Implementation and Outcome

4.1 Data collection

The most popular dataset used by the researchers is the https://www.kaggle.com IRMA dataset is one of the renowned datasets that have been used by many researchers for the development of breast cancer detection system. This dataset consists of 2220 image patches of size 128x128 from mammogram images of the DDSM dataset. Among these images, 1030 are from normal cases, while 1190 are from abnormal cases.

We are using IRMA Dataset images to train both algorithms

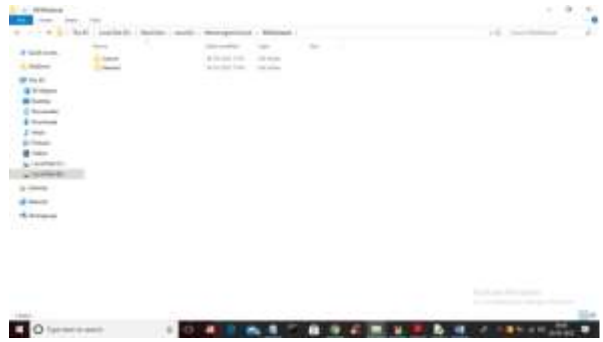

Fig.4. NORMAL and CANCER class

In above FIGURE we have two classes such as NORMAL and CANCER (abnormal) and go inside any folder to view images

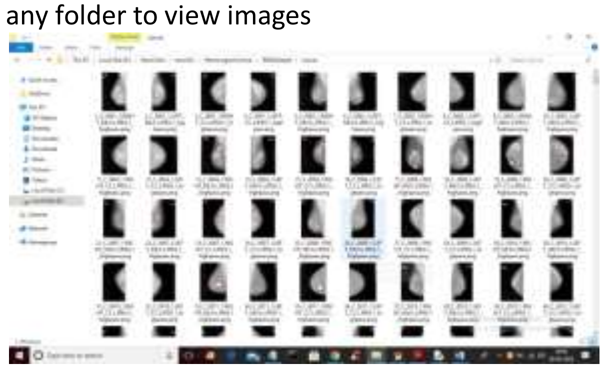

Fig.5. view images

4.2 Evaluation Metrics:

F1-Score, Accuracy and Receiver Operating Characteristics-Area Under the Curve (ROC-AUC) metrics are employed to evaluate the performance of our models. For Computing F1-score and Accuracy, Precision and Recall must be evaluated by

FPR=False Positive Rate
TPR=True Positive Rate
Accuracy
Precision
Recall
F1-score

For this, the calculation of values is measured based on:

- True positive (TP) = No. of events, correctly determined.
- False negative (FN) = No. of events, inaccurately anticipated and not required.
- False-positive (FP) = No. of events, incorrectly predicted.
- True negative (TN) = No. of events, correctly anticipated and not required.

eISSN1303-5150

Accuracy: It is the most important performance measure and it is easily done by a ratio of correctly predicted observations to the total observations.

Accuracy=(TN+TP)/(TP+FP+TN+FN)

4.3 Outcome:

To run project double click on 'run.bat' file to get below FIGURE

Fig.6. Upload IRMA Dataset

In above FIGURE click on 'Upload IRMA Dataset' button to upload dataset and to get below FIGURE

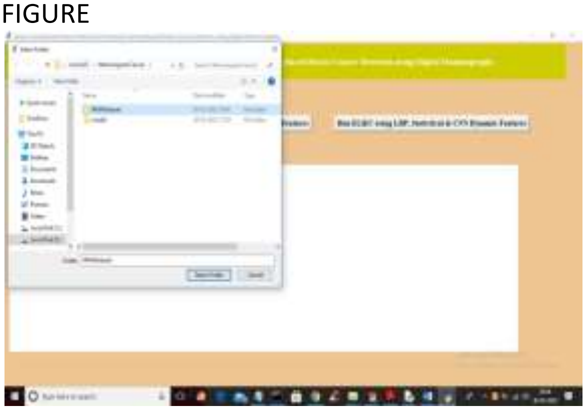

Fig.7uploading 'IRMADataset'

In above FIGURE selecting and uploading 'IRMADataset' folder button and then click on 'Select Folder' button to load dataset and to get below FIGURE

Fig.8. dataset loaded

In above FIGURE dataset loaded and now click on 'Preprocess Images' button to load all images and then extract statistical, LBP and CNN features and get below output

Fig.9. images loaded

In above FIGURE images loaded and features extracted properly and for sample I am showing one sample image and in dataset we can see total 2220 images are found and its contains normal and cancer classes images. 7401

Now click on 'Run SVM using LBP, Statistical & CNN Dynamic Features' button to train SVM with all 3 features and get below output

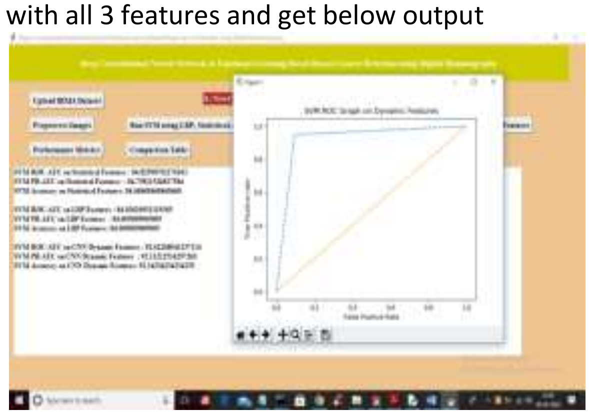

Fig.10. trained SVM

In above FIGURE we trained SVM with statistical features and got accuracy as 84.68 and with LBP we got 84.009 and with CNN features we got SVM accuracy as 92% so CNN features are giving better performance with SVM and now click on 'Run ELiEC using LBP, Statistical & CNN Dynamic Features' button to train all 3 features with ELIEC algorithm and calculate accuracy

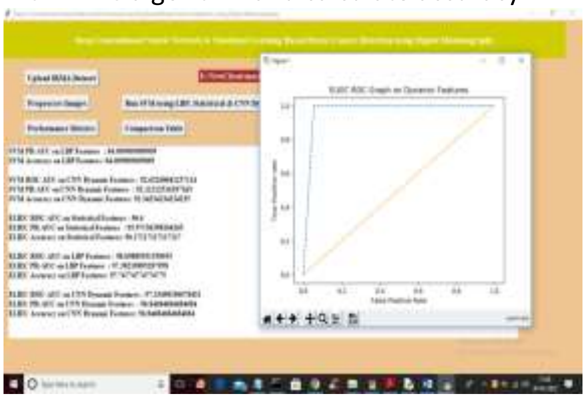

Fig.11.FIGURE with ELIEC

In above FIGURE with ELIEC and statistical we got 96% accuracy and with LBP we got 97 and with CNN we got 96.84 so propose ELIEC is giving better performance compare to SVM and in above ROC graph x-axis represents false positive rate and y-axis represents true positive rate and blue line is the correct prediction which reached closer to 100% and now click on 'Performance Metrics' button to get below graph

Fig.12. algorithm names vs performance

In above graph x-axis represents algorithm names and y-axis represents performance values and each different colour bar represents different metric and in graph we can see propose ELIEC got better performance. Now click on 'Comparison Table' button to get below output

Fig.13.LBP

In above comparison table we can see for same LBP, statistical and CNN propose ELIEC got better performance compare to SVM

Extension outcomes:

Deep Convolutional Neural Network & Emotional Learning Based Breast Cancer Detection using Digital Mammography

In this project as extension we have added 2 algorithms such as

1) BRISK Features Extraction: using Brisk algorithm we are extracting BRISK descriptor from images and then training with SVM and ELIEC
2) Extreme Learning Machine (ELM): adding new algorithm called ELM and then training with Brisk features and this combination is giving accuracy closer to 100%

In below FIGURE you can read red colour comments to know about brisk features extraction

Fig.14. know about BRISK features

In above FIGURE see red line comments to know about BRISK features extraction and this features will get train with all algorithms.
To run project double click on 'run.bat' file to get below FIGURE

Fig.15. upload dataset

In above FIGURE upload dataset

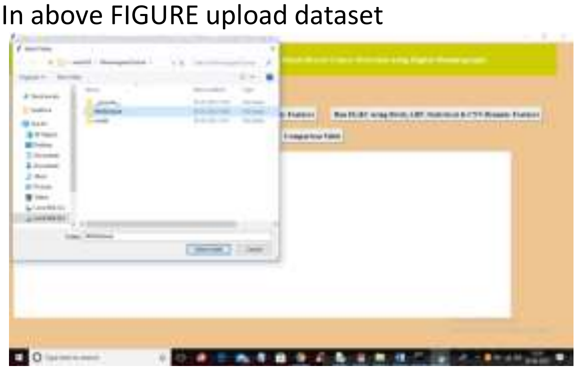

Fig.16. preprocess

After dataset upload click on preprocess button to get below output

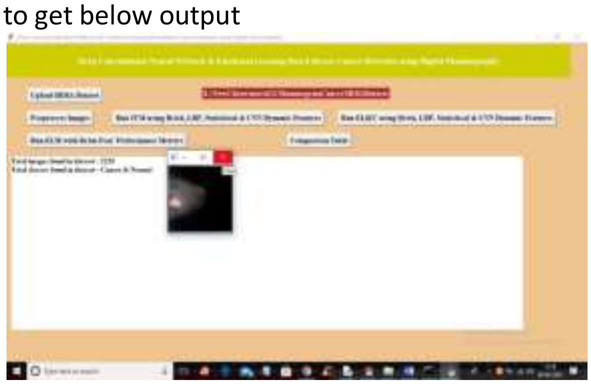

Fig.17. features extracted

In above FIGURE image processing completed and all features extracted and now click on 'Run SVM using Brisk, LBP, Statistical & CNN Dynamic Features' button to train SVM with 4 different features such as Brisk, LBP, CNN and Statistical and get below accuracy

Fig.18. SVM got 97% accuracy

In above FIGURE with BRISK features SVM got 97% accuracy which is higher than any other features. Now click on 'Run ELiEC using Brisk, LBP, Statistical& CNN Dynamic Features' button to train ELIEC with all features and get below output

Fig.19. ELIEC BRISK99.26% accuracy

In above FIGURE with ELIEC BRISK features we got 99.26% accuracy and now click on 'Run ELM with Brisk Features' to get below output

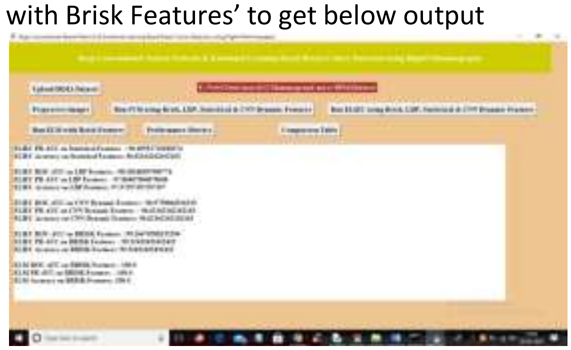

Fig.20. ELM and Brisk got 100% accuracy

In above FIGURE with ELM and Brisk features we got 100% accuracy and now click on 'Comparison Table' to get below output

Fig.21. Brisk features has got high accuracy

In above table we can see all algorithms with Brisk features has got high accuracy

Conclusion

An automatic and secure breast cancer detection system that accurately detects abnormalities not only increases the radiologist's diagnostic confidence but also provides more objective evidence. In this work, an automatic Diverse Features based Breast Cancer Detection (DFeBCD) system is proposed to classify a mammogram as normal or abnormal. The proposed system uses a deep Highway-Network for extracting dynamic features. Using one of the renowned datasets, IRMA, first, the images are pre-processed and then split into three sets as TrainCNN, TrainValidation, and Test. TrainCNN images (838 images), along with their corresponding labels, are used to train the highwaynetwork based CNN. Then, four sets of features, i.e., Statistical, LBP, Taxonomical, and Dynamic features, are extracted for TrainValidation and Test datasets using the procedure. The four different sets of features are then hybridized to generate a final information-rich feature space, and an emotional-Intelligence based classifier is used to exploit this hybrid feature space. The experimental results show that the performance of the dynamic features generated by the proposed HighwayNetwork based CNN is better than all of the three individual sets of ad-hoc features by both SVM and ELiEC classifier. Furthermore, hybridization of all four sets of features yields better performance using both SVM and ELiEC classifiers. Lastly, using a hybrid feature space, the ELiEC classifier gives better performance than SVM. And we adding new algorithm called ELM and then training with Brisk features and this combination is giving accuracy closer to 100%

References:

[1] N. Wahab and A. Khan, "Multifaceted fused-CNN based scoring of breast cancer whole-slide histopathology images," Appl. Soft Comput. J., vol. 97, 2020.

[2] A. Rodríguez-Ruiz, et al., Detection of breast cancer with mammography: effect of an artificial intelligence support system, Radiology 290 (3) (2019) 1–10.

[3] S.H. Khan, A. Sohail, A. Khan, COVID-19 Detection in Chest X-Ray Images Using a New Channel Boosted CNN, 2020, pp. 1–26.

[4] S. Hussain, A. Khan, Coronavirus Disease Analysis Using Chest X-Ray Images and a Novel Deep Convolutional Neural Network, 2020, pp. 1–31, https://doi.org/ 10.13140/RG.2.2.35868.64646. April.

[5] S.H. Khan, A. Sohail, A. Khan, Y.S. Lee, Classification and Region Analysis of COVID-19 Infection Using Lung CT Images and Deep Convolutional Neural Networks, 2020. Mcc.

[6] R.F. Mansour, Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy, Biomed. Eng. Lett. 8 (1) (2018) 41–57.

[7] M.W. Khan, M. Abid, A.Q. Khan, Fractional order Bergman's minimal model-A better representation of blood glucose-insulin system, in: 2019 International Conference on Applied and Engineering Mathematics, ICAEM 2019 - Proceedings, 2019, pp. 68–73.

[8] M.W. Khan, M. Abid, A.Q. Khan, G. Mustafa, Controller design for a fractionalorder nonlinear glucose-insulin system using feedback linearization, Trans. Inst. Meas. Contr. 42 (13) (2020) 2372–2381.

[9] M.W. Khan, M. Abid, A.Q. Khan, G. Mustafa, M. Ali, A. Khan, Sliding mode control for a fractional-order non-linear glucose-insulin system, IET Syst. Biol. 14 (5) (2020) 223–229.

[10] R.F. Mansour, Using genetic algorithm for identification of diabetic retinal exudates in

digital color images, J. Intell. Learn Syst. Appl. 4 (3) (2012) 188–198.

[11] R.F. Mansour, Evolutionary computing enriched computer-aided diagnosis system for diabetic retinopathy: a survey, IEEE Rev. Biomed. Eng. 10 (2017) 334–349.

[12] R.F. Mansour, E.M. Abdelrahim, A.S. Al-Johani, Identification of diabetic retinal exudates in digital color images using support vector machine, J. Intell. Learn Syst. Appl. 5 (3) (2013) 135–142.

[13] R. Cerri, R.C. Barros, A.C.P.L.F. De Carvalho, Hierarchical multi-label classification using local neural networks, J. Comput. Syst. Sci. 80 (1) (2014) 39–56.

[14] P. Shi, J. Zhong, A. Rampun, H. Wang, A hierarchical pipeline for breast boundary segmentation and calcification detection in mammograms, Comput. Biol. Med. 96 (2018) 178–188.

[15] J. Harefa, M. Pratiwi, Comparison classifier: support vector machine (SVM) and Knearest neighbor (K-NN) in digital mammogram images, Juisi 2 (2) (2016) 35–4.

[16] Narmadha, R., Latchoumi, T. P., Jayanthiladevi, A., Yookesh, T. L., & Mary, S. P. (2022). A Fuzzy-Based Framework for an Agriculture Recommender System Using Membership Function. InApplied Soft Computing:Techniques and Applications(pp. 207-223). CRC Press.