MICROANEURYSM DETECTION USING SPARSE PRINCIPAL COMPONENT ANALYSIS & FIREFLY BASED DEEP LEARNING MODEL

[1]Jarapala Parvathi, [2]Rajesh Perugu

[1]parvathi.cse@gmail.com, [2]rajeshp526@gmail.com

[1]Lecturer, Department of CSE, Singareni Collieries Polytechnic, Mancherial, Telangana, India.

[2] Assistant Professor Department of CSE, Balaji Institute of Technology & Science,Warangal, Telangana, India

ABSTRACT: Since microaneurysms (MAs) can be seen as the earliest lesions in diabetic retinopathy (DR), its detection plays a critical role in diabetic retinopathy diagnosis. In recent years, many machine learning methods have been developed for MA detection. Generally, MA candidates are firstly indentified and then a set of features for these candidates are extracted, finally machine learning methods are applied for candidate classification. Various machine learning and deep learning approaches have been implemented on diabetic retinopathy dataset but majority of them have neglected the aspect of data pre-processing and dimensionality reduction, leading to biased results. The dataset used in the present study is a diabetes retinopathy dataset collected from the UCI machine learning repository. At its inceptions, the raw dataset is normalized using the Standard scalar technique and then Principal Component Analysis (PCA) is used to extract the most significant features in the dataset. Further, Firefly algorithm is implemented for dimensionality reduction. This reduced dataset is fed into a Deep Neural Network Model for classification. The results generated from the model is evaluated against the prevalent machine learning models and the results justify the superiority of the proposed model in terms of Accuracy, Precision, Recall, Sensitivity and Specificity.

KEY WORDS: Microaneurysm detection, sparse PCA, Deep Neural Network Model and Firefly algorithm.

I. INTRODUCTION

Diabetic Retinopathy (DR) is a progressive disease with almost no early symptoms of vision impairment, which is the leading cause of blindness prior to the age of 50. The first detectable sign of DR is the presence of microaneurysms (MAs), which results from leakage of tiny blood vessels in the retina and manifest themselves as small red circular spots on the surface of retinas. Early detection of MAs is critical for diagnosis and treatment of DR, which has led to a great deal of research towards automatic detection of MAs.

Deep neural network is based on the concept of machine learning and artificial neural networks. DNN has successfully contributed towards analysis and decision making in the fields of computer vision, speech recognition, drug design, medical image processing and many others. The implementation of such advanced machine learning approaches as DNN had significantly contributed towards pathological screening and disease predictions thereby reducing the burden of human interpretations. With such glorifying results of application of DNN and machine learning approaches in various other fields of healthcare, application of the same in the detection of diabetic retinopathy was a natural point of interest with an objective to reduce occurrence of this disease [1,2]. Thus motivation of the present study was:

1. Early detection of the diabetic retinopathy disease giving opportunity for medical practitioners to treat and cure the same at an early stage with higher accuracy.
2. Focus on the most significant factors of the disease eliminating the irrelevant ones ensuring more accurate classification.

A Deep neural network model is used in the present study in convergence with Principal Component Analysis (PCA) and firefly algorithm for the classification of diabetic retinopathy set. The dataset is collected from the publicly available UCI machine learning repository. The data being collected from the public domain includes attributes which are irrelevant and inclusion of the same would only increase burden on the ML model. Hence Principal component analysis (PCA) algorithm is implemented for feature extraction from DR image dataset.

II. LITERATURE REVIEW

Several methods have been proposed for MA detection. Quellec explored the use of template matching in the wavelet domain. This method was further substantiated following the University of Iowa's release of the Retinopathy Online Challenge database and subsequent competition for MA detection, in which the competition winner extended the wavelet domain template matching method. Ram in [4] created a clutter rejection strategy, in which successive stages of the algorithm eliminated more and more clutter, while passing most target MAs. One recent work a comprehensive grading system is proposed for DR based on classification of 16 features that captured shape, color and intensity information and the features were extracted from candidate regions. Many existing MA detection methods rely on hand-crafted features, which are often based on low-level information. Lowlevel information is easily susceptible to signal drift artifacts and thus prevent reliable generalization among different research sites. A recent method [5] leveraged the use of deep learning for MA detection using a Stacked Sparse Autoencoder (SSAE). Deep learning approaches often learn high-level and robust attributes directly from the raw signal input, and have been successfully applied to various classification and recognition tasks. In [5], small image patches were generated from the original fundus images and used by the SSAE to learn highlevel features from pixel intensities. These patches were then classified as either MA or non-MA using the high-level features learned by SSAE.

Apart from the above mentioned MA detection approaches, a number of template matching based algorithms have been proposed for MA detection. In [3], a local template matching in the wavelet domain has been used for detecting MAs. The problem of illumination variations or high-frequency noise can be avoided effectively in this approach. In addition, Zhang in [8] employed Multi-scale Gaussian Correlation Coefficients (MSCF) method to detect MAs. In this work, MA candidates can be detected by computing the maximum correlation coefficient with five different Gaussian kernels for each pixel. And then 31 features were extracted for each candidate. Finally, true MAs can be identified by specifying the thresholds for each feature directly. However, setting threshold for all the features depends on prior knowledge of experts [8]. Also, simplify taking all features into account is not appropriate. It is inevitably to introduce some irrelevant or redundant features, which not only deteriorates classification performance, but also is time-consuming. So how to choose the useful subset of features for candidate classification should be considered.

III.PROPOSED MODEL

The output layer used sigmoid activation function to classify the Diabetic Retinopathy dataset, since it is a binary classifier. For back propagation the Root mean square propagation (RMS prop) error was used. The dataset was split into 8:2 ratios to train and test respectively. Instead of training entire 80 percent of data and then testing the model on remaining 20 percent of data at one go, for every epoch a batch of 64 records were fed to the model, out of which 80 percent of the records were used to train the model and remaining 20 percent of those records were used to test the model. The proposed model is summarized as follows:

1. Data Transformation: Normalization of the input dataset is done using Standard Scaler.

2. Dimensionality Reduction: Input the transformed dataset to the PCA for dimensionality reduction. To further refine feature engineering use firefly optimization algorithm.

3. Classification: Feed the extracted features to the DNN for classifying the Diabetic Retinopathy dataset.

4. Evaluation: Evaluate the performance of the model using several measures like Accuracy, Precision, Recall, Sensitivity and Specificity.

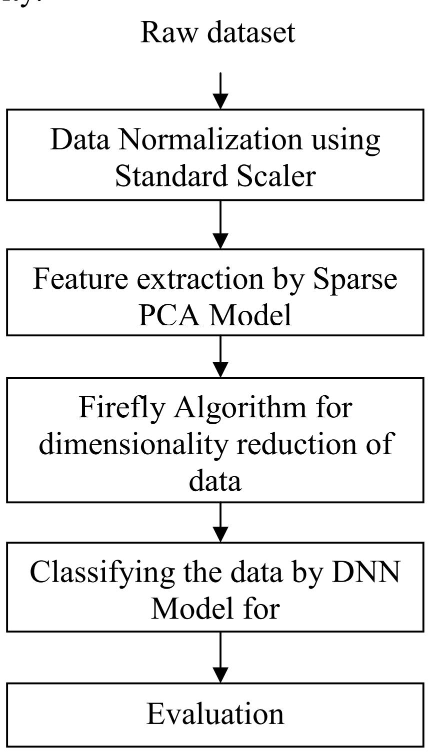

Fig. 1: Proposed Model for Diabetic Retinopathy

A. Principal Component Analysis

The concept of PCA is based on the objective of reduction in dimensionality of a data set consisting of multiple variables which are correlated with each other while retaining maximum variability in the data set from the UCI machine learning repository. The algorithm transforms the variables in the data set to a new set of orthogonal principal components ordered in a manner such that the retention of variation in the original variables decreases while traversing down the order. Hence the first principal component retains maximum variation present in the original components. The principal components are the eigenvectors in co-variation matrix which are orthogonal. The dataset to be used in PCA needs to be a scaled one and the method summarizes the data generating results which are also sensitive to relative scaling. The principal component is defined as a "linear combination of optimally weighted observed variables". The output generated from PCA are such principal components whose numbers are either lesser or equal to the original variables.

The steps involved in implementing PCA on a two dimensional data set starts with Normalization of the data. This is done by subtracting the respective means from each of the respective columns in the data set computing a data set with mean of zero. The second step involves calculation of the co-variation matrix. Then the Eigen values and Eigen vectors are calculated for the covariance matrix. The Eigen values are then ordered in a descending order to provide the order of significance for the components and the dimensionality is reduced by choosing first set of Eigen values and ignoring the rest. A matrix of vectors is formed to create a feature vector. In the final step the principal components are formed by considering the transpose of the feature vector and computing the left multiplication with the transpose of scaled version of the data set.

B. Firefly Algorithm

Firefly algorithm is a "nature—inspired" algorithm based on the behaviour of flies. Nature inspired algorithms are extensively used in several stages of machine learning process. The fireflies have natural lights emitting from their body that help them to attract or find other flying mates. It also helps them to catch their prey and protect themselves from predators. The algorithm is based on three primary assumptions:

1. The artificial fireflies are unisex and their attraction is not dependent on gender.
2. The attractiveness of a firefly is proportional to the brightness of the lights emitted and hence it decreases as they move away from each other due to absorption of the light by air. Since all fireflies emit light, the one emitting the brightest one attracts most of its neighbors. On the contrary, if there is a situation of no such bright firefly, all the fireflies move around in a haphazard fashion in any direction.
3. The brightness of the flashing light being the criteria for attraction is the objective function to be optimized in the algorithm.

The basic schema followed in this algorithm 1 is:

Algorithm 1: Pseudo Code of Firefly Algorithm

1. Objective function makespan, z = (P1, P2...Pn)
2. Generate machines (i = 1, 2...m)
3. Generate job sequences (j = 1, 2...n)
4. Evaluate makespan (P1, P2, P3. . . .Pn) for all population
5. **while** Gen. < MaxGen. **do**
6. **for** each job sequence j = (1, 2, 3. . . , n) **do**
7. **for** For each lot sizes Xij **do**
8. Move firefly in d-dimensional space
9. Determine the attractiveness based on distance Rij
10. Evaluate makespan
11. Assess light intensity
12. Select job sequences and lot sizes for Gen. + 1E

The proposed methodology is illustrated in Figure (1). The dataset used in this work has 19 contributing attributes. The values of these attributes are of different range. This variation in the range of the values of the attributes may lead to varied weights of some instances which may results in biased prediction results. In order to avoid such heterogeneity, as part of pre-processing, a Standard Scaler method is used in the proposed work. Standard scaler method normalizes the data converting it to a common range to eliminate bias in the prediction results. The Principal Component Analysis algorithm is then applied on this normalized data. The main reason for using PCA is to eliminate the insignificant attributes from consideration for training the DNN. To further strengthen the feature engineering process, one of the popular nature inspired algorithms, Firefly Optimization Algorithm, is used in this work.

The main strength of Firefly algorithm is that it tunes the parameters in such a way that this algorithm chooses the optimal parameters, whose convergence rate would be very fast avoiding local minima. This property of Firefly algorithm makes it an ideal choice for feature engineering to choose optimal parameters which influence the classification in a positive way thereby reducing training time. The dimensionally reduced dataset is then fed to DNN for classification of diabetic retinopathy datasets.

IV. RESULTS

For evaluating the proposed model, a sequential model was used to build the DNN-PCA model. Figure (2) illustrates the performance evaluation of the ML models based on the measures accuracy, precision, recall, sensitivity and specificity. It is evident from these figures that PCA-Firefly based ML models outperform than ML with

PCA model. Considering inclusion and non inclusion of dimensionality reduction and feature engineering concepts with ML algorithms, it is observed that the proposed model- DNN-PCA-Firefly performs better than the other hybrid ML algorithms considered.

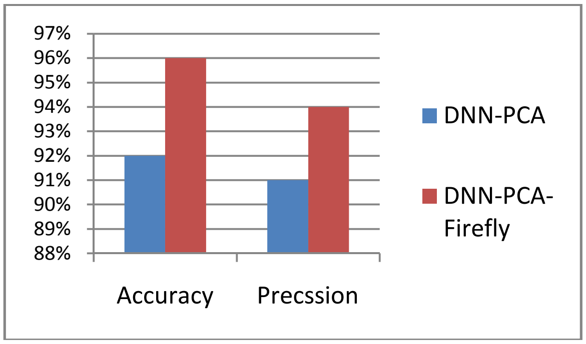

Fig. 2: PERFORMANCE EVALUATION COMPARISION OF DNN BASED MODELS

V. CONCLUSIONS

In this paper, an unsupervised classification method based on sparse PCA was developed to detect MA. Sparse PCA is applied to find the latent structure of MA data, once a model has been developed that reflect the MAs, any departure from standard MAs are detected by monitoring the statistics. Further, Firefly algorithm was used for the purpose of dimensionality reduction. This reduced dataset was fed into the deep neural network (DNN) which generated classifications results with enhanced accuracy. The results of the model were also evaluated with the predominant machine learning approaches wherein the results defended the superiority of the model in terms of the Accuracy, Precision. As part of the future study, the proposed model could be utilized for data sets in other domains. The performance of the proposed model therefore motivates to conduct similar studies in various other domains having high dimensional data. This approach can also be used for eliminating noisy data in Magneto Encephalo Graphy (MEG) data analysis contributing towards better prediction in healthcare.

VI. REFERENCES

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

[2] Venkatraman, S.; Alazab, M. Use of data visualisation for zero-day Malware detection. Secur. Commun. Netw. 2018: 2018. [Cross Ref].

[3] G. Quellec, M. Lamard, P. M. Josselin, G. Cazuguel, B. Cochener, C. Roux, "Optimal wavelet transform for the detection of microaneurysms in retina photographs," IEEE Transactions on Medical Imaging, vol. 27, no. 9, pp. 1230-41, 2008.2169-3536 (c) 2016 IEEE.

[4] K. M. Adal, D. Sidibé, S. Ali, E. Chaum, T. P. Karnowski, F. Mériaudeau, "Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning," Comput. Methods Programs Biomed, vol. 114, no. 1, pp. 1–10, 2014.

[5] B. Antal, A. Hajdu, "An ensemble-based system for microaneurysm detection and diabetic retinopathy grading," IEEE transactions on bio-medical engineering, vol. 59, no. 6, pp. 1720-6, 2012.

[6] B. Antal, A. Hajdu, "Improving microaneurysm detection using an optimally selected subset of candidate extractors and preprocessing methods," Pattern Recognition, vol. 45, no. 1, pp. 264-270, 2012.

[7] D. B. Mukamel, G. H. Bresnick, J. C. Dickinson, D. R. Cole, " A screening approach to the surveillance of patients with diabetes for the presence of vision-threatening retinopathy," Ophthalmology, vol. 107, no. 1, pp. 19-24, 2000., vol. 3, pp. 169–208, 2010.

[7] B. Zhang, X. Wu, J. You, Q. Li, F. Karray, "Detection of microaneurysms using

multi-scale correlation coefficients," Pattern Recognition, vol. 43, no. 6, pp. 2237-2248, 2010.
[9] M. Abràmoff, M. Garvin, M. Sonka, "Retinal imaging and image analysis," IEEE Reviews in Biomedical Engineering.
[10] C. E. Baudoin, B. J. Lay, J. C. Klein, "Automatic detection of microaneurysms in diabetic fluorescein angiography," Revue D Épidémiologie Et De Santé Publique, vol. 32, no. 3-4, pp. 254-61, 1984.

Jarapala Parvathi is currently working as a Lecturer in the Department of Computer Science & Engineering, Singareni Collieries Polytechnic, Mancherial, Telangana, India. She is having 10 years of experience in teaching and her area of interest includes Network Security, IOT, Cloud Computing and Machine Learning etc.

Rajesh Perugu is currently working as an Assistant Professor in the Department of Computer Science & Engineering, BITS, Warangal, India. He is having 10 years of experience in teaching and his area of interest includes IOT, Artificial intelligence, Data Mining, Cloud Computing and Machine Learning etc.