International Conference on Recent Trends in Data Science and its Applications
DOI: rp-9788770040723.185

Diabetic Retinopathy Detection and Classification with the Aid of Deep Convolutional Neural Network (DCNN)

A.Chinnasamy
Department of Data Science and Business Systems
School of Computing, SRMIST
Kattankulathur Campus Chennai
chinnasamyamb@gmail.com

P.Selvakumari
Department of Computer Science and Engineering
Chennai Institute of Technology, Anna University
Chennai
selvakumaricp@gmail.com

***Abstract* — The prevalence of DR is steadily rising, which necessitates the automatic disease severity extraction and classification. About 2% of those with this illness are entirely blind as a consequence of the diabetic mellitus problem and 10% get vision impairment after 15 years of diabetes as a result of the DR complication. It is also a significant contributor to blindness in both middle aged and older age groups. The patient may develop to severe stages of irreversible blindness if the condition is not detected early. The growing number of diabetic patients face a major issue due to a lack of ophthalmologists. It is suggested that an automated DR screening system be created to aid the ophthalmologist in making decisions. One of the primary symptoms of the DR is hard exudates. The detection of hard exudates is crucial for screening purposes and aids in disease monitoring and diagnosis. Thus, utilising Lloyd's clustering technique, this work offered a unique techniques to segment the exudates and irregularities in DR were found.**

Keywords— Diabetes Mellitus, Cat Swarm Optimization,, Diabetic Retinopathy, Mean Squared Error.

1. INTRODUCTION

The primary consequences of diabetes mellitus are DR. For working people, it is the main source of acquired blindness. Before the age of 50, DR is the main ocular pathologic cause of blindness in adults of working age. To diagnose DR, ophthalmologists typically look for haemorrhages, exudates and micro aneurysms. Ophthalmologists are able to identify microaneurysms and haemorrhages utilising fluoresce in angiograms. NPDR and PDR are two general categories for DR. The stages of DR can be determined based on the presence of characteristics on the retina. Diabetic Macular Edoema (DME) is the main cause of people with diabetes for losing their central vision. DME is characterised by Hard Exudates (HE), macula-regional blot Diabetes affects the blood vessels in the human retina. DR is one of the main causes of vision loss and patients might avoid losing their vision if the disease is not identified earlier a dot-like appearance. It may results in detachments, optic nerve ischemia, vitreous haemorrhage or retinal vascular occlusions. In the recent years, DL techniques have been widely used for a variety of automatic classification problems. When classifying images, the typical process entails first extracting the crucial features using a set of convolutional layers and then using these features, classification is carried out. Thus, an innovative hybrid classification method is suggested in this work to identify DR in order to avoid above discussed problems.

1.1 Proposed System Model

The two main subgroups of DR disease are PDR and NPDR. Mild Non-Proliferative Retinopathy, Moderate Non-Proliferative Retinopathy, Severe Non-Proliferative Retinopathy and Proliferative Retinopathy are the four stages of the DR classification. The ML technique was used to recognise these four steps. It is crucial to classify and forecast normal and abnormal circumstances in the DRearly. Numerous different classifiers are discussed in the literature for the detection of DR, but they fall short of flawless detection. The suggested methodology introduces the DCNN based methodology to address these shortcomings. Figure 7.1 depicts the block diagram of the suggested technique.

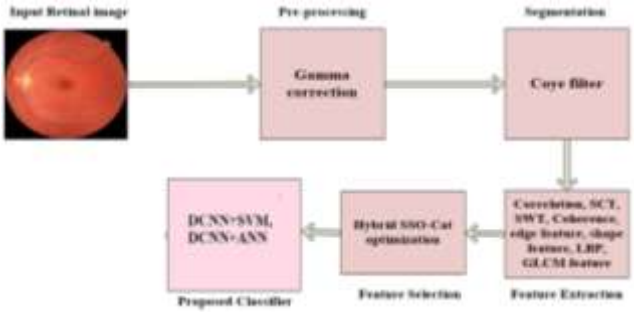

Fig. 1. Architecture of the proposed system

The open source system is used to acquire the retinal image datasets. Five distinct processes, including

- Pre-processing,
- Segmentation,
- Feature Extraction,
- Feature Selection and
- Classification

are included in the proposed system's design. Gamma correction is used at the pre-processing stage to get rid of noise in the input image. Unwanted noises are eliminated and increased contrast is accomplished with the use of pre-processing techniques, which are crucial for the prediction and classification of DR. Important features should be retrieved from the segmented image to detect DR. The features like Correlation, SWT, LBPetc., are retrieved from the segmented image. Among those, the essential features are chosen from the retrieved features with the aid of the hybrid SSO with Cat optimization technique. Finally, the DCNN classifier receives the extracted features as input to make predictions. Features are used to train the classifier during the training phase and retinal images are used during

DOI: rp-9788770040723.185

the testing phase to achieve the best results. UNet and modified Unet models are considered for classification in the suggested system.

1.2 Pre-processing Phase

The pre-processing stage is crucial for DR prediction because it removesnoise present in the input retinal image. The local brightness and contrast variation will not be homogeneous in retinal pictures. At certain points, the illuminations may be greater. Retinal imaging cannot show lesions clearly in low contrast and low brightness conditions. It is best to improve image clarity based on colour and quality. Pre-processing techniques are therefore crucial for enhancing image contrast and reducing noise. Gamma correction is applied to the input retinal image during pre-processing.

1.3 Segmentation Phase

The retinal image should typically be a light-sensitive tissue lining the inside surface of the eye. It can also be a layered structure with multiple layers of neurons connected by synapses. In the middle of the nasal side, the central retinal and vein retinal artery are present close to one another. When compared to the other portions of the retinal images, the segmented images include information. The blood vessels can serve as a marker throughout the segmentation process and helps to classify the severity of disorders. Edges from the retinal image are emphasised in the first stage. Following edge enhancement, retinal images can be converted to grayscale to remove uneven noise using filters and histogram equalisation.

1.4 Coye Filter

Coye filter is a novel method for segmenting retinal blood vessels in which cosiders RGB of the retinal image rather than only the green channel. The RGB image is transformed into a grey image and then PCA is used to find hybrid lesions. Following is the algorithm of Tyler Coye adopted for segmentation of images.

Coye Filter Algorithm

1. Consider retinal RGB image.
2. Apply CLAHE over gray image to increase contrast.
3. Remove the background by using an average filter and comparing the grayscale and average filtered images.
4. Binarize the image.
5. Divide the blood vascular network into segments

II. Feature Extraction

For the purpose of illness prediction, many features are retrieved from retinal images. With the use of various feature extraction approaches, which are described below, the characteristics are extracted:

2.1 Correlation

The similarity between the grayscale distribution of microaneurysms and the Gaussian function can be evaluated using the correlation coefficient. The correlation coefficient will be high if the two images coincide and low if they don't. The coefficient's ranges from 0 to 1. The following are the definitions of the correlation coefficient:

$$r = \frac{\sum m \sum n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\sum m \sum n (A_{mn} - \bar{A})^2 \sum m \sum n (B_{mn} - \bar{B})}}$$

Where,

A and B - mean.

Different sigma values for the Gaussian kernel are needed because microaneurysm sizes may vary.

1.2.2 Discrete Cosine Transform (DCT)

The discrete cosine transform of an image, is defined by,

$$F(u,v) = C(u)C(v) \sum_{X=0}^{(N-1)} \sum_{Y=0}^{(N-1)} f(x,y) \cos \frac{(2x+1)u\pi}{2N} cos \frac{(2Y+1)v\pi}{2N}$$

The inverse transform is defined by,

$$F(x,y) = \sum_{X=0}^{(N-1)} \sum_{Y=0}^{(N-1)} C(u)C(v)f(u,v) \cos \frac{(2x+1)u\pi}{2N} cos \frac{(2Y+1)v\pi}{2N}$$

Where,

$$C(u) = C(v) = \frac{1}{\sqrt{N}}, for\ u, v = 0$$

$$C(u) = C(v) = \frac{1}{\sqrt{N}}, for\ u, v \neq 0$$

2.2 Stationary Wavelet Transform (SWT)

The fundamental idea behind the wavelet transform is to visualise any function as a superposition of waves created by stretching and shrinking the mother function. To make the wavelet decomposition invariant, the Stationary Wavelet Transform (SWT) was developed. The SWT approach is characterised as follows.

$$v^i = \frac{\mu^i}{\sigma^i} (i = 1, \ldots . 7)$$

where

,- mean and standard deviation.

2.3 Shape Features

Shapes can be described using a few basic geometrical characteristics. Simple geometric features are employed as filters to weed out false positives or in conjunction with other shape descriptors to distinguish shapes because they can only distinguish shapes with significant variances. Center of gravity, digital bending energy, circularity ratio, etc., are some of the shape attributes. In this work,

DOI: rp-9788770040723.185

Eccentricity, solidity and circularity are taken into consideration as shape characteristics. The ratio of the distance between the main and focal axes of an ellipse can be used to define eccentricity. It ranges from 0 to 1. The formula utilized to calculate shape characteristics can be expressed as

$$Ecentricity = \frac{\sqrt{(major\ axis^2 - minor\ axis^2)}}{major\ axis}$$

$$Circularity = \frac{area\ of\ shape}{area\ of\ bounding\ circle}$$

$$Solidity = \frac{area}{convex\ area}$$

Many properties that are computed based on these aforementioned formulas are believed to be the major features for recognising DR situations. The following is an explanation of how the LBP characteristics are utilised to extract features from input retinal images:

2.4 Local Binary Pattern (LBP)

An image is converted into an array of integer labels using the LBP approach, which trains the texture image's pixel-level in detail. These labels can be seen as a histogram, which can be used to determine the texture of the image under analysis. LBP has two key characteristics that make it appealing for describing textures. It is invariant to monotonic grey level variations, such as those brought on by changes in illumination and it is computationally simple. The operation of LBP is depicted below.

$$s(x) = \begin{cases} 1, if \quad x \geq 0 \\ 0, otherwise \end{cases}$$

III. Feature Selection

Utilizing the various feature extraction approaches stated above, the features are extracted. With the help of a hybrid SSA-CSO optimization technique, necessary features are selected from among the many retrieved features. With the aid of the CSO algorithm, the SSA can be further enhanced in order to update the salps process. The required features can be quickly and efficiently chosen from the entire list of features. The following section contains a presentation of the SSA and CSO algorithms.

3.1 Salp Swarm Algorithm (SSA)

The salp swarm algorithm was created by Mirajalili to address optimization issues. In this case, SSA is used to choose the necessary features from the features collection. In essence, the salp belongs to a family called Salpidae. The SSA is based on the salps' swarming behaviour. They are able to set up cooperative chains while foraging in deep oceans. Through pursuing the food source, salps use this behaviour to get greater kinetic energy. For escaping from the local optima problem in the optimization, the salp chain can be used as a guide. The salp in the SSA algorithm is made up of two distinct classes that are dependent on both followers and leaders. At the top of the chain, the leader salp is situated, and the followers obey him. Members of the chain are also referred to as the followers. While the followers make use of their fellow followers, the leader salp aids the direction and movement of the swarm. The design of salp chain is illustrated in the Figure 7.2.

Fig. 2 Structure of Salp chain

For searching in an n-dimensional space, each salp's position vector is described. The SSO's initial population is made up of N salp and d dimensions. The below equation is the N-dimensional matrix that represents the salps' location vector:

$$Y_i = \begin{bmatrix} Y_1^1 & Y_2^1 & \dots & Y_d^1 \\ Y_1^2 & Y_2^2 & \dots & Y_d^2 \\ \vdots & \vdots & \dots & \vdots \\ Y_1^N & Y_2^N & \dots & Y_d^N \end{bmatrix}$$

All slaps are directed to the food supply. The following equation describes the leader's position:

$$Y_j^1 = \begin{cases} E_j + D_1((UB_j - LB_j)D_2 + LB_j) & D_3 \geq 0.5 \\ E_j - D_1((UB_j - LB_j)D_2 + LB_j) & D_3 \leq 0.5 \end{cases}$$

Where,

and - random vectors generating values which mentioned as the limit [0, 1],

- upper limit of j^{th} dimension,

- the lower limit of j^{th} dimension,

- Position of food source,

- Leaders position of salp and the core parameters of is described as

$$D_1 = 2e^{-\left(\frac{4t}{T_{max}}\right)^2}$$

Where,

- exploration and exploitation tendencies of SSO algorithm in balanced state,

- iteration

- maximum number of iterations.

Additionally, the position of follower salps can be expressed as.

DOI: rp-9788770040723.185

$$Y_j^i = \frac{Y_j^i + Y_j^{i-1}}{2}$$

Where,

- start of i^{th}salp at the j^{th} dimension.

3.2 Cat Swarm Optimization (CSO)

In CSO, the initial number of cats will be chosen. Every cat has a unique position made up of M dimensions, velocities for each dimension, a fitness value that shows how well the cat has adapted to the fitness function, whether the cat is searching or tracing. The best position in one of the cats would be the final solution because CSO retains the best answer till the end of iterations.

Seeking Mode

This model is intended to represent the scenario of the cat, which is lazing around, scanning its surroundings and looking for a new spot to settle down in.

$$P^i = \frac{|FS_i - FS_b|}{FS_{max} - FS_{min}}, where, 0 < i < j$$

The mutative ratio for the chosen dimensions is announced by SRD. When a dimension is chosen to change in searching mode, the difference between the old and new values won't be outside of the SRD defined range. The CDC reveals how many variables will be changed. All of these elements are significant contributors to the searching mode. The value of SMP will not be impacted by whether the SPC value is true or not.

Tracing Mode

Simulating the case of the cat when tracing some targets is the tracing mode. Once a cat enters the tracing mode, it moves in all directions at its own speeds. The following is a description of the tracing mode:

$$V_{k,d} = V_{k,d} + r_1 \times c_1 \times \left(X_{best,d} - X_{k,d}\right) \ \ where\ d = 1,2,\ldots.M$$

Process of CSO algorithm

Step 1: To update the salp position, create N cats in the process.

Step 2: Distribute the cats randomly over the M-dimensional solution space. At random, choose values for each cat's velocity that fall within the range of its maximum velocity. Select a random number of cats, place some of them in seeking mode according to MR, and others in tracing mode.

Step 3: Assess each cat's fitness value by plugging their locations into the fitness function and memorise the best one.

Step4. Reposition the cats in accordance with their flags. Apply the cat to the searching mode process if cat is in that mode. If not, apply the cat to the tracing mode process.

Step 5: Select the same number of cats this time, place them in tracing mode in accordance with MR and the remaining cats in seeking mode.

Step 6: Verify the termination condition. If it is met, end the programme. otherwise, go back and repeat steps 3 through 5.

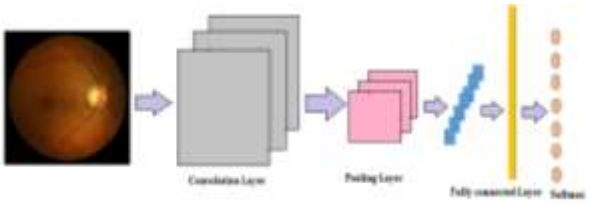

Fig. 3 Architecture of the DCNN network

UNet

The need for localization in biomedical image processing necessitates the implementation of deep learning algorithms to acquire required output. However, Standardized CNN are ineffective for biomedical image segmentation because they are designed for classification tasks, such as assigning a class label to each image rather than identifying the segmented region. Hence, a deep learning methodology based on fully CNN named U-Net model was developed for segmentation purpose.

comprises two 3 × 3 unpadded convolutions, trailed by a rectified linear unit (ReLU) and a 2 × 2 max pooling operation which can be utilized for down sampling.

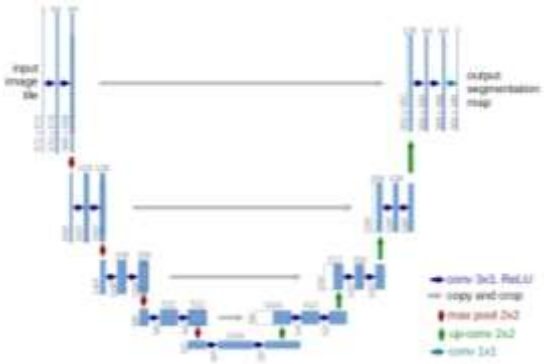

Fig. 4 U-Net architecture

In this study, during every downsampling stage, the feature channels number is increased. An unsampled feature map followed by a 2X2 convolution network reduces the feature channels to half the number. However, the loss occurring at the boundary pixels necessitates cropping technique. Thus, a 1x1 convolution is incorporated at the last layer of the topology, which converts 64 component feature vector into the anticipated number of classes. As a result, U Net networks are divided into two parts:

Modified UNet algorithm

Patch-based networking is suggested in this network architecture. The encoder path is on the left side, while the decoder path is on the right, much like in the original U-Net architecture. To create multichannel encoder feature maps, each encoder layer conducts convolution (with batch normalisation and rectified linear unit (ReLU) activation) and densely connected convolution. Thus, the down sampling operation follows this.

The decoder path up-samples the feature maps using the deconvolution layer. The skip connections combine the

DOI: rp-9788770040723.185

encoder path's feature map with the appropriate up-sampled decoder feature map. A softmax function activates the feature maps. It is possible to derive two channel probabilities: one for vessels and the other for non-vessels.

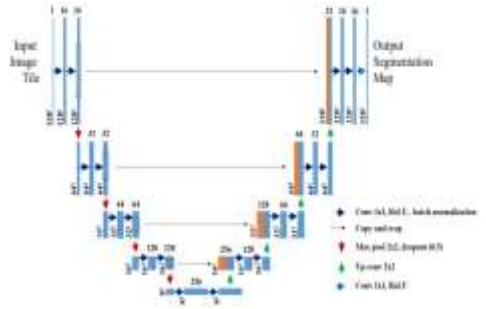

Fig. 5 Network architecture of the modified U-net

III. Results and Discussion

Table 1.1 displays the effectiveness of the proposed method's on the messidor dataset. Figure 1.8 displays the segmentation results of the two chosen images The original retinal and ground truth images are displayed in Figures 1.8(a) and (b). The segmentation results of the traditional UNettopology are shown in Figures 1.8(c1) and 1.8(c2), while the segmentation results of the modified UNet topology are shown in Figures 1.8(d1) and 1.9(d2). The segmented images demonstrate how effectively the proposed approach can identify retinal blood vessels.

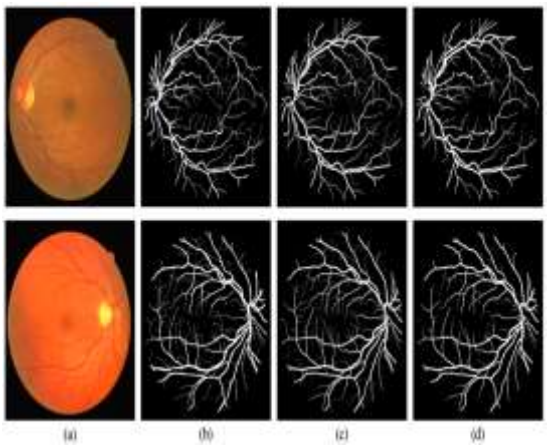

Fig. 6 Segmented analysis of proposed topologies

TABLE 1.1 Performance Measures (Hybridized CNN)

S.No	Classifier	Accuracy	Sensitivity
1.	DCNN (modified Unet)+SVM	98.12%	96.59%
2.	DCNN (modified Unet)+ANN	92.56%	88.25%
3.	DCNN (Unet) + SVM	97.45%	96.29%
4.	DCNN (Unet) + ANN	91.13%	86.51%
5.	DCNN+SVM	97.38%.	96.07%
6.	DCNN+ANN	90.19%	85.29%

The accuracy and sensitivity of the proposed topology is discussed in the table 1.1. From the table 1.1, it is observed that the accuracy of the proposed topology with DCNN and SVM classifier is about 97.38% and it outperforms than ANN. Similarly, the SVM with Modified Unet is 98.12%. The accuracy obtained with traditional Unet is 97.45%. Hence, it can be concluded that the proposed topology with modified Unet as classifier will exhibit higher accuracy. While considering the sensitivity, it's about 96.59% for modified Unet, 96.29% for traditional Unet and 96.07% for DCNN.

Thus, the results demonstrated that the proposed method extracted the blood vessels quite skillfully and precisely. The hybrid classifier comparison demonstrates that the DCNN with modified Unet and SVM outperforms all other known topologies.

IV. Conclusion

In this study, the DCNN based algorithm is used to detect and classify DR. Gamma correction is used at the pre-processing stage to get rid of noise in the input images. The Coye filter is introduced during the segmentation process. With the aid of a hybrid Salp Swarm Optimization (SSO) and Cat Swarm Optimization (CSO) method, the necessary features are chosen from the extracted features. Following extraction, the blood vessels is identified using a DCNN, which combines the U-net and modified U-net structures. The hybrid classifier comparison demonstrates that the DCNN architecture with modified Unet and SVM outperforms than all other known topologies.

References

[1] S. Long, X. Huang, Z. Chen, S. Pardhan, and D, Zheng, "Automatic Detection Of Hard Exudates In Color Retinal Images Using Dynamic Threshold And SVM Classification: Algorithm Development And Evaluation," BioMed Research International, 2019.

[2] Maksoud, Eman A Abdel, Mai Ramadan, Sherif Barakat and Mohammed Elmogy, "A Computer-Aided Diagnoses System For Detecting Multiple Ocular Diseases Using Color Retinal Fundus Images," In Machine Learning In Bio-Signal Analysis And Diagnostic Imaging, Academic Press, pp. 19-52, 2019.

[3] D. Marin, M.E. Gegundez-Arias, B. Ponte, F. Alvarez, J. Garrido, C. Ortega, and J.M. Bravo, "An exudate detection method for diagnosis risk of diabetic macular edema in retinal images using feature-based and supervised classification," Medical and Biological Engineering and Computing, vol. 56, no. 8, pp. 1379-1390, 2018.

[4] M. Mateen, J. Wen, M. Hassan, N. Nasrullah, S. Sun, and S. Hayat, "Automatic Detection of Diabetic Retinopathy: A Review on Datasets, Methods and Evaluation Metrics," IEEE Access, vol. 8, pp. 48784-48811, 2020.

[5] A.M. Mendonc¸ A. Sousa, L. Mendonc and A. Campilho, "Automatic localization of the optic disc by combining vascular and intensity information," Comput Med Imaging Graph, vol. 37, no. 5–6, pp. 409–417, 2013.

[6] Sitharthan, R., Vimal, S., Verma, A., Karthikeyan, M., Dhanabalan, S. S., Prabaharan, N., ...&Eswaran, T. (2023). Smart microgrid with the internet of things for adequate energy management and analysis.Computers and Electrical Engineering, 106, 108556.

[7] M.R.K. Mookiah, U.R. Acharya, H. Fujita, J.H. Tan, C.K. Chua, S.V. Bhandary, and L. Tong, "Application of different imaging modalities for diagnosis of diabetic macular edema: a review," Computers in Biology and Medicine, vol. 66, pp. 295-315, 2015.

[8] M.R.K. Mookiah, U.R. Acharya, R.J. Martis, C.K. Chua, C.M. Lim, E.Y.K. Ng, and A. Laude, "Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: A hybrid feature extraction approach," Knowledge-based systems, vol. 39, pp. 9-22, 2013.

[9] Moshika, A., Thirumaran, M., Natarajan, B., Andal, K., Sambasivam, G., &Manoharan, R. (2021).Vulnerability assessment in heterogeneous web environment using probabilistic arithmetic automata. IEEE Access, 9, 74659-74673.

[10] K. Nirmala, K. Saruladha, and K. Dekeba, "Investigations of CNN for Medical Image Analysis for Illness Prediction," Computational Intelligence and Neuroscience, 2022.

[11] J. Novo, M.G. Penedo, and J Santos, "Localisation of the optic disc by means of GA-optimised Topological Active Nets," Image Vis Comput., vol. 27, no. 10, 2009.