Relevance Mapping based CNN model with OSR-FCA Technique for Multi-label DR Classification

S. Hemamalini [1], V. D. Ambeth Kumar [2], R. Venkatesan[3], S. Malathi[*,1]

[1]Panimalar Engineering College, Anna University, Chennai 600123,
Tamil Nadu, India
[2] Mizoram University, Aizawl, Mizoram 796004, India
[3]Computer Science and Engineering, Karunya University, Coimbatore 641114, India

Emails: hemamalini.phd2020@gmail.com; ambeth@mzu.edu.in; rlvenkei_2000@karunya.edu; malathi.raghuram@gmail.com

Abstract

In computer vision, multi-label classification (MLC) is especially important for medical picture analysis. We use MLC to classify diverse stages of diabetic retinopathy (DR) using colour fundus pictures of varying brightness and contrast. As a result, ophthalmologists can now identify the early warning symptoms of DR and the varying stages of DR, allowing them to begin therapy sooner and prevent further difficulties. Using the outlier-based shallow regularization fuzzy clustering approach (OSR-FCA), for classification we present a deep learning method in this paper's picture segmentation task. The fundamental feature of the proposed system is the ability to identify and analyse different degenerative changes in the retina that occur alongside the progression of DR without requiring the patient to undergo costly diagnostic procedures like dye injections. Photographs are first resized, converted to grayscale, cleaned of noise, and the contrast increased by the use of histogram equalization adopting the CLAHE method. The clipping limit of CLAHE is optimized by the help of the rat optimization algorithm, which is applied throughout the histogram process. In addition, a Gaussian metric regularization to the objective function in OSR-FCA is a great way to enhance clustering approaches that use fuzzy membership with sparseness which is based on neutrosophic set. This research proposes a new approach called "Relevance Mapping on Multi-Class Label" (RMMCL) for locating and viewing regions of interest (ROI) inside a segmented picture. These representations give better explanations for the predictions of the DL model founded on a convolutional neural network-(CNN). The validation of two ML datasets showed the projected model outperformed the existing models by achieving an average correctness of 97.27 percent over five stages of the IDRID dataset.

Keywords: Multi-label classification; Relevance Mapping; Outlier-based skimpy regularization fuzzy clustering technique; Regions of interest; Convolutional neural network.

1. Introduction

Upthrow blood glucose levels are the hallmark of diabetes, it's a chronic condition. Over time, this causes significant harm to the human body's vascular system, visual system, and nervous system [1]. Nowadays in diabetics, diabetic retinopathy (DR) becomes prevalent. Step by step, it can result in sudden and irreversible blindness [2]. About 642 million adults around the world will have diabetes by 2040, according to the study, and DR will impact one out of every three persons with diabetes [3]. By 2030, where the sum of people with DR is predictable to reach 191 million, according to another study [4]. Diabetic retinopathy (DR) is a major reason for intense blindness. People with diabetes have a 25-fold increased risk of rapid blind compared to those who don't have the disease [5]. Retinal vasculopathy is the most prevalent serious consequence of diabetes, and which

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

increases slowly but it's a steady degeneration of retinal blood vessels. DR is a disease that aggravate invisibly, where it is noted only in severe stage. When blood vessels swelling, leak, or develop abnormally, it can cause severe DR [6].

The most prominent features and symptoms of the DR microaneurysms (MA), hemorrhages (HM), exudates (EX), venous loops (VL), venous reduplication (VR), and neovascularization (NV). In the retina, DR stages are classified according to the presence of one, two, or all of these characteristics [7]. Patients in the early stages of DR typically exhibit no noticeable indications; nevertheless, they may have experienced distortion, blurred vision, and a liberal loss of visual severity in the preceding grades. Therefore, DR severity can be broken down into two subtypes: NPDR (non-proliferative DR) and PDDR (PDR). NPDR, on the other hand, has three distinct subgrades [8]. The presence of a small MA indicates mild severity, while the presence of HM and/or EX indicates moderate severity. Small, aberrant, and feeble BV, known as NV, it emerges when the severity of NPDR aggravate, by reflecting the progress of retinal ischemia. The PDR is being set up in this severe grade [9]. Blood leaks into the retinal surface as a consequence of the retina's response of generating new vessels that are weak, brittle, and leaky. In this phase, aberrant new blood vessels form a branching pattern. DR was the advanced stage and which can cause permanent vision loss, if untreated [10]. In Figure 1, we have a labelled example of an aberrant fundus picture.

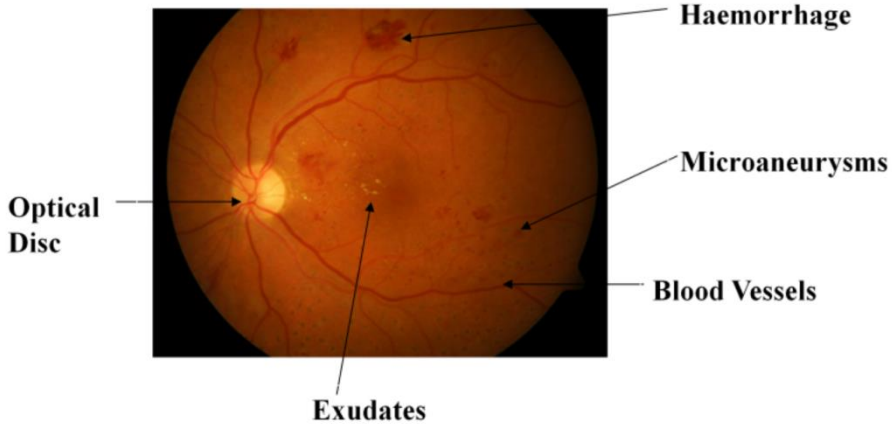


Figure 1: Labeled Fundus image

In recent years, automated DR diagnostic systems have gained widespread recognition as a valuable resource for diagnosing and prognosing diseases around the world. Automatic pathology detection plays a crucial function in the medical industry by aiding illness interpretation through the automatic analysis of disease severity, allowing patients to receive prompt treatment and follow-ups. Early detection through routine screening is crucial in reducing DR consequences [11]. The disease can be treated in its early stages by proper treatment.

Several different approaches have been taken by researchers in order to detect DR in fundus images for the purpose of early abnormality detection. However, the participation of non-affected areas with the afflicted area, which in turn offers a weak collection of characteristics, is the reason for degraded system. Streamlined set of key-points needed to show detection systems promptness. For DR detection and classification, numerous computer vision techniques were used [12]. When it comes to various computer vision and image analysis applications, the most popular DL architecture is the Convolutional Neural Network (CNN). It has shown greater performance associated to other standard machine approaches [13, 15].

Even though several potential deep learning (DL) models for modality categorization have been discussed recently, no previous works have been made to display, comprehend, or interpret their DL models' internal representations or prediction outcomes. Due to a lack of understanding Inaccuracies are occurred in answering applications of the model's behaviour. The significance of analysing and amplify the learnt behaviour of DL models towards medical organisation is a major motivation of this study. For better understanding DL models, step by step progress has been achieved in various vision tasks in recent years. By explaining the predictions of the CNN- the difficult task of classifying medical imaging modalities, we offer a unique approach of locating and visualising an area of interest (ROI) within a DR segmented picture that is regarded to be the most significant. Here, we introduce the novel idea of RMMCL. It quantifies the significant activation at a given (x,y) coordinate is in the feature maps generated by the final layer that led to the successful classification of the input image. The ROI is located and highlighted by projecting the result "significance map," in which each represents the standing back onto the input image. It's important to note that while determining how significant each feature

Doi: https://doi.org/10.54216/FPA.110207

Received: December 27, 2022 Accepted: April 25, 2023

map piece is, our RMMCL takes into account both positive and negative charities to the model prediction. There is an inherent bias against certain social classes in the resulting ROI.

The outstanding sections of the paper are as follows: The relevant analysis of the current models for DR categorization is obtainable in Section 2. The projected model explanation is provided in Section 3. Section 4 and 5 present the validation analysis and its scientific contributions, respectively.

2. Related Work

Huang et al. [16] propose a RTB that incorporates attention apparatuses at two main levels: a self-attention transformer that makes use of global dependencies among lesion features, and that enables interactions among lesion and vessel features by participating valuable vascular information to reduce ambiguity in lesion detection. Both of these transformers rely solely on the user's focus and concentration. Here, additionally propose a global that can preserve specificity in deep networks, making it a promising tool for collecting the earliest, subtlest lesion patterns. This network is able to perform concurrent segmentation of the four distinct types of lesions because it incorporates the aforementioned blocks of dual-branches. It has been demonstrated through exhaustive research using the IDRiD and DDR which performed best when compared to other models.

In order to achieve accurate multi-lesion and automatically segment a large number of diabetic retinopathy-related lesions, Guo et al. [17] propose using a cascade attentive RefineNet (CARNet). It is highly skilled in using the full granularity as well as indelicate universal data is included in the fundus image. CARNet is made up of three distinct parts: the attention refinement decoder, the local image encoder, and the global image encoder. We utilise the whole image and the patch image as input, and we send them to ResNet101, correspondingly, in order to down sample and extract lesion features. This allows us to performance in both. A dual attention approach is utilised by the high-level refinement decoder in order to combine the findings of the low-level attention refinement module with the characteristics derived from the two encoders that operate at the same level. Because of this, the model is able to acuminate on the area affected by the lesion and produce accurate predictions. In segmentation capabilities of the proposed CARNet, we resorted to the usage of the IDRiD, E-ophtha, and DDR data sets. The proposed framework has been shown, through exhaustive testing and ablation experiments on a variety of data sets, for more accurate and more resilient than the methods that are currently considered as the state-of-the-art. In order to perform precise multi-lesion segmentation, it not only gets rid of the interference caused by identical tissues and noises, but it also keeps aspects of small lesions intact.Without putting a strain on the GPU's memory consumption, all of this is accomplished

Wang et al. [18] recommend the use of a deep CNN in order to differentiate automatically between the four diverse types of fundus lesions that are related with DR. In order to make use of information on several scales, we suggest a collaborative architecture that is composed of two branches. An attention apparatus is planned to combine feature maps from all the different decoding stages in order to achieve an effective and comprehensive merging of useful branches. A new supplemental classification task that utilises an innovative supervision strategy which has been incorporated in order to significantly improve the accuracy of lesion segmentation and lessens likelihood of the model is overfitted. After conducting thorough experiments on three fundus datasets that are made available to the public, by discovering the proposed method produces an average AUC of 0.677, 0.629, and 0.581 correspondingly. The results validate that the proposed strategy is superior to other competing strategies as well as path way are considered to be state-of-the-art.

Upadhyay et al. [19] suggest a DR- extraction by using colour fundus pictures as their source material. During the pre-processing phase, it is recommended by eliminating retinal blood vessels in order to advance the visual look of red anomalies and uniformly lesion false positives. In order to improve feature learning, it is recommended by integrating both deep features and handcrafted features. The five lesion-specific features that propose here are merged with the deep features based on the U-net. Additionally, propose an innovative method known as "characteristic patch-based training" for selectively target the red lesions. This method has been demonstrated to boost the effectiveness of training. In this paper, it is undertaken as an ablation research in order to demonstrate the importance of the handcrafted features and the presentation of the baseline model. The precision-recall curve for the suggested approach is around seven percent and ten percent better than the other state-of-the-art algorithms respectively, when functional to the IDRiD and DDR datasets.

Aziz et al. [20] developed an autonomous deep learning-based bleeding detection system that can be utilised as a second interpreter by ophthalmologists for reducing the amount of time and effort required for regular routine screening procedures. During the pre-processing stage, it is able to increase the image quality and prediction about the locations of the haemorrhages. Gradient information in modified gamma alteration where used to achieve adaptive lighting of fundus images. This correction takes into consideration for different degrees of brightness present in the photographs. A Gaussian match filter are incorporated into the programme so that it could make predictions on the positions of potential candidates. Separation of the necessary objects was

Doi: https://doi.org/10.54216/FPA.110207

Received: December 27, 2022 Accepted: April 25, 2023

accomplished through the use of approximated positions and geographical variety. The well-known deep models and the novel haemorrhage network, proposed for haemorrhage categorization, which are compared and contrasted in this section. The sensitivity of the model, its specificity, its precision, and its accuracy were all examined using two different datasets. The proposed network fact is the least amount of extent by the four networks. It is conducted by simulated tests in order to evaluate the bleeding network's performance, by our evaluation minimizing the time taken to train the network and then increase in accurate classifications. The results of the classification stage were encouraging, demonstrating good accuracy while significantly reducing the amount of training time required. The findings demonstrated that increasing the number of layers in a deep network doesn't shows necessary result in improving the performance but it lengthens the duration of the training process. When developing a deep model, it is essential to make use of the appropriate architecture and to fine-tune the parameters in highest quality of results.

Nahiduzzaman et al. [21] present an original automated method for the detection of adaptive histogram equalisation so that the lesions would stand out more (CLAHE). Here employed an PCNN and an ELM, one for the purpose of feature extraction and the other for DR classification, respectively. Because it has fewer parameters and layers than a CNN does, the PCNN design can extract features in a shorter amount of time than the CNN takes. Both the Kaggle DR 2015 competition dataset (which had 3,662 FIs) were used to evaluate the method, and the results were promising. The accuracy of the project method was measured at 91.78 and 97.27 percent when it was applied to the two different data sets. A corollary result of the study in proposed framework was robust over a variety of dataset sizes and shapes, including both symmetrical and asymmetrical ones. This is the one of the study's findings.

Gu et al. [22] presented the intelligent DR classification model of fundus photos as their contribution. This method able to recognise all five stages of DR, beginning with non-DR and progressing through DR. This plan is comprised of two primary components. The grading prediction block (GPB) is in-charge of classifying the five DR phases, whereas the feature extraction block (FEB) is responsible for removing the features from the fundus images. It is easier for the FEB transformer to identify retinal haemorrhage and exudate as well as to respond to these conditions as a result of its improved capacity to concentrate on minute details. The residual attention of the GPB works exceptionally well for mapping the regions that are occupied by a wide variety of things. Extensive research performed on DDR datasets demonstrates that it is superior method, and when measured against the gold standard, it obtains on par with those obtained by employing the gold standard.

3. Proposed System

The two datasets utilised in this study are labels detailed in the paper's results and discussion sections. Next, we'll go through pre-processing, which is used to get rid of clutter like undesired details in the background or foreground..

3.1. Pre-processing

Photos should be pre-processed to get rid of noise, boost image qualities, and guarantee image consistency. Low contrast and non-illumination in retinal colour fundus photographs can be attributed to anatomical issues such as the eye fundus' lenses on cameras, pupil size variations, sensor array geometry, and movement during image collection. This highlights the significance of proper pre-processing of retinal fundus pictures. Disease identification rates are raised with the use of pre-processing techniques like images and visual assessment. In the first stage, colour fundus images from two different pixels. For quicker processing, images are rehabilitated to grayscale. In Figure 2, we see the original image resized and converted to grayscale.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

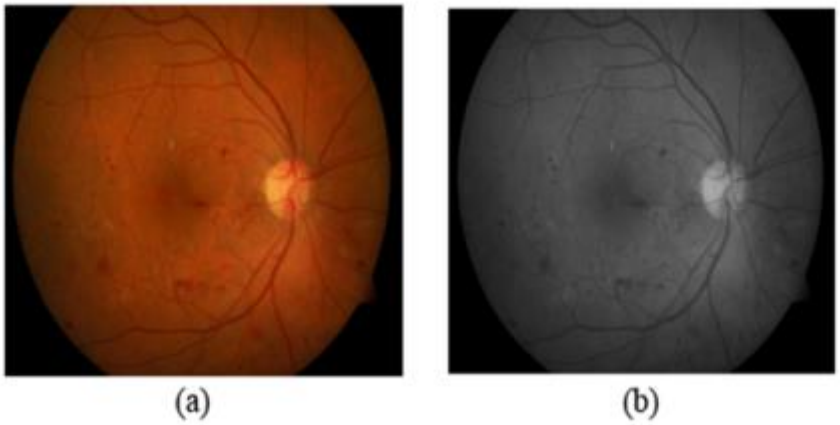


Figure 2: a) Resized RGB- copy b) Gray copy

The fundus appears in these blood vessels and the scenery. It's easy to see the choroid plexuses against the blood red background. The grayscale image will be used to pinpoint the location of the optic disc. Blue channel identification is unreliable due to noise and a lack of retinal morphological info.

Histogram equalisation is practical after grayscale change to enhance contrast and homogeneity. By applying image processing techniques, a grayscale image can be converted into a histogram-equalized one. Here, we'll be focused on bringing out the finer features of the image. The microaneurysm, optic disc, and hard exudates all are clearly visible in histogram equalised photos. CLAHE is an AHE algorithm, namely the Equalization variety. The issue of AHE's over amplification can be mitigated by the usage of clip limit and tile count settings. With the help of CLAHE, the picture is broken up into MN little tiles that can be displayed locally. Each tile has its own histogram that is calculated separately. The average number of pixels (N s) in each segment must be determined before the histogram can be calculated equation (1).

$$N_s = (N_X \times N_Y)/N_G \quad (1)$$

Where N_G is the average sum of pixels, which is found by multiplying N_X by 2 and N_y by 2, respectively. The sum of pixels in the x direction is marked by N_X, while the corresponding quantity in the y direction is written as N_y. In CLAHE, the ideal solution for the clip limit is identified by the RAT algorithm, which keeps a healthy balance bounded by exploration and exploitation [23], and this clip limit is employed to improve the quality of the image.

3.1.1. Mathematical model and optimization algorithm

The rat's fighting and pursuing behaviour are covered here. Next, we give a rundown of the Rat Swarm Optimization (RSO) algorithm.

3.1.1.1 Hurtling the prey

Rats are typically animals that exhibit agonistic social behaviour when hunting in groups. It is presumed the best agent which has a fair understanding in where the quarry is hiding. When compared to the current best search agent, the standings of the other search agents are subject to change. The equations for this mechanism are as follows:

$$\vec{T} = U \,.\vec{T}_i(x) + K \,.(\vec{T}_r(x) - \vec{T}_i(x)) \quad (2)$$

This is where the placements of rats are defined, and where $\vec{T}_r(x)$ is the best-optimal clipping limit and $\vec{T}_i(x)$ defines the position of the pixels.

U and K parameters, on the other hand, are calculated as shadows:

$$U = L - x \times \left(\frac{L}{Max_{iteration}}\right) \quad (3)$$

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

Where, $x = 0,1,..,Max_{iteration}$

$$K = 2 . rand\ () \quad (4)$$

As a result, L and K are both random statistics which fall inside the range of [1, 5]. As iterations go on, the parameters U and K are more important than ever in helping to maximise exploration and exploitation.

3.1.1.2 Fighting with prey

The subsequent equation has been projected in order to quantitatively describe the fighting process among rats and their prey.

$$\vec{T}_i(x+1) = |\vec{T}_r(x) - \vec{T}| \quad (5)$$

Specifically, the rat's current position is modified and denoted by the formula $\vec{T}_i(x+1)$. Whenever the optimal answer is recorded, all other search agents are reordered to reflect their new standing relation to the record. The rat (G, H) can shift its position (G*, H*) to keep tabs on its prey (G, H). Altering the parameters, as shown in Equations (3) and (4), will result in a different number of locations surrounding the current position (4). However, this concept can be expanded to n-dimensional space. The parameters U and K, once set, ensure both exploration and exploitation. By employing a proposed RSO approach, the best potential outcome is preserved while the fewest number of operators are used.

3.1.2 Background Removal and Foreground Identification

Background removal and foreground identification are the primary procedures in pre-processing the collected fundus pictures. Since the Optical Disc (OD) has an intensity value close to the yellow lesion, it is an unfavourable physiological structure for DR diagnosis. Morphological closing process is used to pinpoint the exact location of the OD boundary, and the major circular area is designated as the OD region [24, 25].

When comparing the actual and estimated fundus picture backgrounds, blood vessels are one of the major background objects that can be picked out using morphological reconstruction followed by a global thresholding technique. Furthermore, linked component analysis is utilized to eliminate some of the outlying regions which is not corresponded to vascular regions. The vessel and non-vessel pixels are separated using a threshold of eight connected pixels [26]. Background elimination utilizing morphological reconstruction (IMrph) includes OD localization and blood vessel removal using equation (6).

$$I_{BC} = I - I_{Mrph} \quad (6)$$

where I is the unique fundus image, I_{Mrph} is morphologically rebuilt image and I_{BC} is the background detached fundus image.

Figure 3 shows the background region as well as the foreground portions that can be distinguished in one of the photos.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

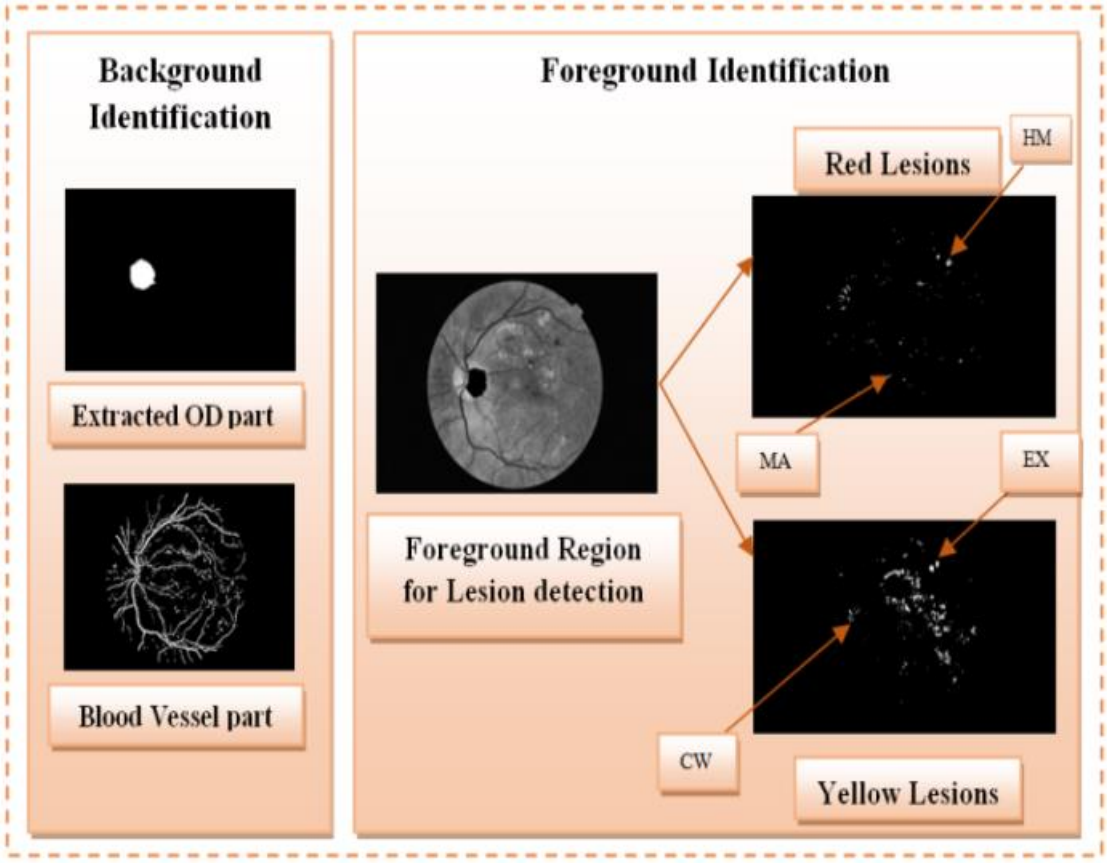


Figure 3: Representation of background and foreground identification

The foreground region is obtained first, by obtaining the original fundus image and after background removal, non-diseased fundus image does not possess any foreground lesion portion. For diseased fundus images, intensity-based foreground region discrimination is carried out. This step separates lesions into two categories: red lesions, which include MAs and HMs, and yellow lesions, which include EXs and CWs.

3.2. Segmentation

As a result, within the scope of this work, we present an updated version of the OSR-FCA which is based on neutrosophic set. With the OSR-FCA that has been suggested, it is possible to overcome quickly over-segmentation and outlier sensitivity, which results in enhanced segmentation outcomes...

3.2.1. Brief Description of OSR-FCA

It is currently available FCM algorithms such as DSFCM [27], MalFCM [28], and MEFC [29], it's not possible to construct fuzzy memberships with a small number of members. Hence fuzzy memberships require a large number of members. By introducing u^2ij as an additional penalty term, new regularisation approaches are presented in order to cope with this issue and bring it under regulator. To describe OSR-FCA is one of our primary objectives.

$$\tilde{J} = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\Phi(x_j|v_i,\sum_i) + \gamma\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^2, \qquad (7)$$

where $\Phi(x_j|v_i,\sum_i)$the distance amongx_{ji} andv_{ij} to control the sparsity of membership. In order to represent the distance between the two this is done. Changing the value enables one to adjust the degree to which the target function is robust in the face of outliers and noise.

Therefore, the first term of J in equation (7) which is little, whereas the second term will be extremely significant. As a direct consequence of this, the OSR-FCA requires a greater number of iterations to arrive at the optimal calculation than k-means but fewer than FCM. The goal function J represents an acceptable middle ground between k-means and FCM. The significance of the association is significantly lower than that of the FCM. In contrast to the k-means algorithm, the

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

membership values of fuzzy sets are not always either 0 or 1. In addition to the previous point $\Phi(x_j|v_i,\sum_i)$:

$$\boldsymbol{\Phi(x_j|v_i,\sum_i) = \ln\left(-\rho(x_j|v_i,\sum_i)\right)}$$
(8)

where $\rho(x_j|v_i\sum_i)$ is the Gaussian thickness separate as:

$$\boldsymbol{\rho(x_j|v_i\sum_i) = \frac{\exp\left(-\frac{1}{2}(x_j-v_i)^T\sum_i^{-1}(x_j-v_i)\right)}{(2\pi)^{(D/2)}|\sum_i|^{(1/2)}}} \quad (9)$$

I for the i[th] class in D is the data. I describe the dispersion within each class. When we change equation (8) to (9), we obtain the following:

$$\Phi(x_j|v_i,\sum_i) = \frac{1}{2}((x_j-v_i)^T\sum_i^{-1}(x_j-v_j) + \ln|\sum_i| + Dln(2\pi))$$
(10)

It's possible that the variable ln|i| will have a huge negative worth when applied to the data that is densely distributed; this will result in the distance metric in (10) being different from the Mahalanobis distance. The restriction that x_j/v_i i must have non-negative standards that is not satisfied because of the effect of $\Phi(x_j|v_i,\sum_i)$ This is because of the consequence of ln $|\sum_i|$

As the value gets lower it causes significant problems in distance dimension and misclassification. These issues become more severe as the value gets lower. This problem also has a solution...

$$\Phi'(x_j|v_i,\sum_i) = \begin{cases}\Phi - \min(\Phi) & \min(\Phi) < 0\\ \Phi & \text{otherwise}\end{cases} \quad (11)$$

Where, $\Phi'(x_j|v_i,\sum_i)$ contains the positive constraint of distance. After relieving $\Phi'(x_j|v_i,\sum_i)$into equation (7), the ultimate function that is impartial is separate as

$$\tilde{J}' = \sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}\Phi(x_j|v_i,\sum_i) + \gamma\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^2 \quad (12)$$

For each example$\boldsymbol{x_j}$, the $\tilde{J}'$ difficulties under constraints can be decoupled into c-level sub-problems$0 \le u_{ij} \le 1$ and $\sum_{i=1}^{c}u_{ij}^m = 1$. Then we get

$$\tilde{J}'_j = min\sum_{i=1}^{c}(u_{ij}\Phi'(x_j|v_i,\sum_i) + \gamma u_{ij}^2) \quad (13)$$

Through simplification, $\tilde{J}'_j$can be rewritten as:

$$\tilde{J}'_j = \min||u_{ij} - h_{ij}||^2 \quad (14)$$

Where $h_{ij} = -\Phi'(x_j|v_i,\sum_i)/2\gamma$. To solve (14), We optimize the method described in [30] to reach various degrees of sparsity.

By addressing each of these sub-problems, we may determine where their cluster inside the larger problem of finding a solution to the $\frac{\partial\tilde{J}'_i}{\partial v_i} = 0$,

$$\frac{\partial\tilde{J}'_i}{\partial v_i} = \sum_{j=1}^{n}u_{ij}\left(\frac{\partial\left[(x_j-v_i)^T\sum_i^{-1}(x_j-v_i) + \ln|\sum_i| + Dln(2\pi)\right]}{\partial v_i}\right)$$

$$= \sum_{j=1}^{n}u_{ij}(x_j - v_i)$$

$$= 0,$$

We get

$$v_i = \frac{\sum_{j=1}^{n}u_{ij}x_j}{\sum_{j=1}^{n}u_{ij}} \quad (15)$$

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

Furthermore, the solving $\frac{\partial \tilde{J}_i'}{\partial \sum_i} = 0$,

$$\frac{\partial \tilde{J}_i'}{\partial \sum_i} = \sum_{j=1}^{n} u_{ij} \left(\frac{\partial \left[(x_j - v_i)^T \sum_i^{-1} (x_j - v_i) + \ln |\sum_i| + Dln(2\pi) \right]}{\partial \sum_i} \right)$$

$$= \sum_{j=1}^{n} u_{ij} ((x_i - v_i)^T \sum\nolimits_i^{-2} (x_j - v_i) + \sum_i^{-1})$$

$$= 0,$$

The solution yields

$$\sum_i = \frac{\sum_{j=1}^{n} u_{ij} (x_j - v_i)^T (x_j - v_i)}{\sum_{j=1}^{n} u_{ij}} \quad (16)$$

Using the neutrosophic set (Fuzzy Set), FCM method to kick things off reduces the need for many iterations. Here is a quick rundown of our planned OSR-FCA procedure:

Set the number of clusters c, regularization parameter γ, convergence threshold η, and maximum iteration number T.
Initialize the membership $U^{(0)}$, the clustering centers $V^{(0)}$, and the covariance matrix $\Sigma^{(0)}$ using the FCM algorithm.
Set the loop counter $t = 1$.
Update $U^{(t)}, V^{(t)}, \Sigma^{(t)}$ using (14), (15), and (16), respectively.
Update the objective function $\tilde{J}'^{(t)}$ using (12).
If $\max |\tilde{J}'^{(t)} - \tilde{J}'^{(t-1)}| \leq \eta$ or $t \geq T$, stop;
otherwise, update $t = t + 1$, and go to step 4.

Input images, ground truth images, and segmented yield images are all displayed in Figure 4.

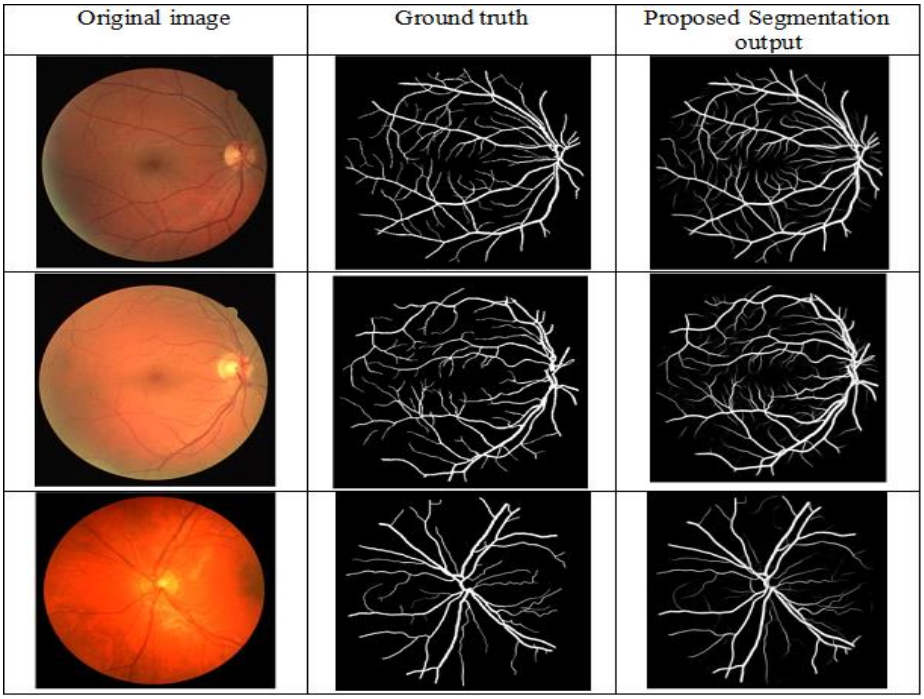


Figure 4: Sample Output image for Segmentation Procedure

3.3. Classification using CNN Configuration

The pretrained VGG16 network served as the basis to build our "multi-modality" CNN copy. To begin, on top of the reduced model we cut off the final convolutional layer of the VGG16. The

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

hierarchical feature representations in model image were learned by this model during training using modest weight updates, allowing for complete feature extraction and classification.

The ideal values for the hyper parameters [31] were determined using a hybrid particle swarm optimization with Cat swarm method (HCPSO), as will be discussed below (3.3.2). We conducted our search between the intervals [1e5, 1e1], [0.8, 0.99], and [1e10, 1e1] for the regularization, respectively. Along with the "Adam" optimizer, we employed a mini-batch size of 10, which yielded 3500 iterations per epoch. In order to create our "multi-modality" CNN model, we relied on NVIDIA's CUDA toolkit as well as the popular open-source library Tensorflow. Training was carried out at NLM's NVIDIA-DGX1 facility, which is equipped with Tesla V-100 GPUs.

3.3.1. Relevance Map for Multi-class Label (RMMCL)

To begin, we provide a quick overview of the currently obtainable techniques for locating and highlighting relevant ROI a specific sort within an input image designed to visually convey the CNN forecast. We then discuss, by using a battery of experiments, compare it to the aforementioned approaches.

3.3.1.1. Class Activation Map (CAM)

By feeding global average final output layer, a specific CNN architecture produced CAM [32]. In this context, we'll refer to w_K^C as an output node consistent to the class c, and $F^k \in R^{u \times v}$as u-by-v k-th feature map created by the final convolution layer. The output layer prediction score S_c for class c can be written as a weighted of the previous layers' prediction scores.

$$\boldsymbol{S_C = \sum_k w_k^c F^k = \sum_k w_k^c \sum_{x,y} f_k(x,y) = \sum_{x,y} \sum_k w_k^c f_k(x,y)} \quad (17)$$

where $f_k(x, y)$ signifies the activation at the spatial component (x, y) in the k-th feature map.

CAM for class c, $M_c \in R^{u \times v}$ activation at spatial component (x,y) was distinct as a weighted sum.

$$\boldsymbol{M_c(x,y) = \sum_k w_k^c f_k(x,y)} \quad (18)$$

CAM reflected the significance of the activation for the categorization of an input picture to linear sum of all components in M c (x,y). Finally, we found the region of interest (ROI) inside the image that was the most relevant image.

3.3.1.2. Gradient-Weighted (Grad-CAM)

Grad-CAM [33] used the incline information of a target class smooth back into the last convolutional layer to produce visualizations from any CNN-based DL model. The current CNN design is not modified in any way, unlike the aforementioned CAM Grad-class CAM's c maps were linked to the negative weights belonged to other classes in the image, therefore the ReLU function was employed to get rid of the negative weights' possible effect on the class of interest...

$$\boldsymbol{Grad_{M_c}(x,y) = ReLU(\sum_k \alpha_k^c f_k(x,y))} \quad (19)$$

Here, a prediction score, and α_k^c is the weight produced by calculating the gradient of $S_c$with respect to the k^{th} feature map.:

$$\boldsymbol{\alpha_k^c = \sum_{x,y} \frac{\partial S_c}{\partial f_k(x,y)}} \quad (20)$$

For CNN architectures are used with CAM, this $\alpha_k^c$was identical to ww_k^c, as shown by Equations (17) and (20). Since Gard-CAM and CAM are otherwise identical under these conditions, the only discernible difference lies in the latter's use of ReLU to filter out the spatial elements carrying a negative weight.

3.3.1.3. Relevance Map

The term "relevance" was first coined to quantify the significance of a given hidden node in a network's ability to accurately generate the desired output. The incremental MSE without the m^{th} hidden node was used to determine the hidden node's significance.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

$$\boldsymbol{R_m = \sum_p \sum_c \left(t_c^p - o_c^p(m)\right)^2 - \sum_p \sum_c \left(t_c^p - o_c^p\right)^2} \quad (21)$$

To clarify, o_c^p and o_c^p indicate in response to an input p, whereas t_c^p represents the corresponding target value. Because the MSE rose significantly after a hidden node with a high relevance value was removed, it is possible that this node played an important role in the overall classification procedure. Numerous machine learning scenarios, including network pruning, accelerated learning, etc., have successfully applied in this approach.

We developed a unique accurate After score, S_c for each node c in the layer using equation (17), we deleted the spatial component, $(l, m)$from the feature maps in the last convolutional layer

$$\boldsymbol{S_c(l,m) = \sum_{x,y \neq l,m} \sum_k w_k^c f_k(x,y)} \quad (22)$$

Then, we defined the projected RMMCL, $R(l,m) \in R^{u \times v}$, as the linear sum of MSE among S_c and $S_c(l,m)$derived from all nodes in the CNN-based DL model's output layer..

$$\boldsymbol{R(l,m) = \sum_{c=1}^{N} \{S_c - S_c(l.m)\}^2} \quad (23)$$

To aid comprehension, the VGG16-based DL model's RMMCL scored calculation process is simplified in Figure 5 to the case of two-class problem. Simply said, deep learning is a method of learning by discrimination. For this reason, the gap between the prediction scores is maximized by having a crucial feature map component from the final convolution layer contribute positively and negatively to the nodes. Because our RMMCL relied on the incremental MSE which measured across all output nodes, we anticipated a discriminatory return on investment (ROI). The aforementioned CAM and Grad-CAM algorithms, on merely used the prediction score at the node's output to determine which class an input image.

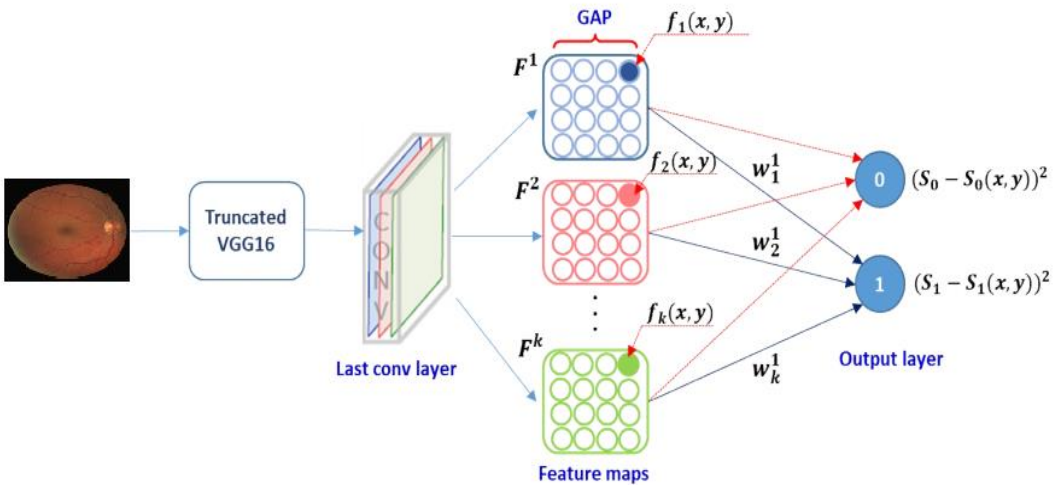


Figure 5: Using a deep learning (DL) typical based on a (CNN), we may compute a relevance mapping that is class-specific

We also established a "class-level" ROI to reflect a region that is consistently emphasised as crucial to the accurate prediction of photos all fitting to the same class. The RMMCL score of all images in a given class can be averaged to produce it.

$$\boldsymbol{avg_{R(l,m)} = \sum_{p \in c} R^p(l,m)} \quad (24)$$

Here, $R^p(l,m) \in R^{u \times v}$ identifies the RMMCL of the p^{th} image of the 'c' class. This "class-level" ROI allows us in visually explain our "multi-modality" CNN model by revealing the "inter-class" alteration and "intra-class" resemblance of ROIs across images. Its HCPSO algorithm that is used to fine-tune the suggested model's parameters.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

3.3.2. The PSO Algorithm

Russell Eberhart and James Kennedy presented the first algorithm, Particle Swarm Optimization (PSO), in 1995. The PSO algorithm mimics the collective intelligence of a flock of birds or school of fish. Most optimization problems, AI computation, and design/scheduling problems all have found uses for the PSO algorithm. Each of the N particles in the PSO algorithm stands in for a potential answer. By using a random number generator, the PSO algorithm gives each particle a completely arbitrary starting position and velocity inside the search space. Particles adjust their speed and location with each iteration using their best position $P_i$of all particles as a whole, P_g. First, the Equation (25), then both based on the Invalid source supplied (26).

$$\boldsymbol{V_i(t+1) = \omega V_i(t) + c_1 r_2 \left(P_g(t) - X_i(t)\right) + c_2 r_2 (P_i(t) - X_i(t))} \quad (25)$$

where t and t + 1 are interval [0,2], r_1 r_2 are uniformly distributed random variables in the interval [0,1], ω is an inertia weight used to strike a balance between the global and local search, and it is an interval.

$$\mathbf{X_i(t+1) = X_i(t) + V_i(t+1)} \quad (26)$$

The inertia weight ω is linearly reduced

$$\boldsymbol{\omega = \omega_{max} - t * \frac{\omega_{max} - \omega_{min}}{Numiter}} \quad (27)$$

where ω_{max}, ω_{min}, $NumIter$ is the sum of iterations, Unweight is the extreme weight, and NumMinWeight is the minimum weight.

3.3.2.1 The CSO Algorithm

A cat's hunting abilities is necessary for survival, but a house cat's insatiable curiosity and fascination with moving objects make him or her a wonderful companion. Although they spend a lot of time in sleeping, cats have keen perception when they are awake. They watch out for any danger or food and never let their guard down. Because of these two unique actions, the CSO algorithm can be run in either a searching or drawing mode. The system uses a mixing ratio to determine whether the cats should be seeking or tracking (MR).

It's possible to simulate a cat's naptime by entering a state of seeking mode. The searching mode is comprised of four parameters: the seeking memory.

How much data can be stored and retrieved in a given amount of time is quantified by a cat's searching memory capacity (SMP). With the help of a Roulette Wheel, each cat will pick a new spot to store its memories at random. Range-based stability control is a method for managing changes for values along with a given dimension provided hence they remain within a predetermined tolerance range (SRD). To determine how many parameters, need tweaking, the CDC technique is applied. The SPC is a Boolean variable. SPC (1) holds if and only if the cat has the best health of any person (0). For a value of SPC = 1, the cat will produce j new applicant positions, where j = SMP; for a value of SPC = 0, the cat will produce 1 new candidate position, where j = SMP.

In this hypothetical scenario, each cat would use the following, where it is shifted its position to account for the acceleration. According to the best feasible locations, the CSO algorithm modifies the cats' speeds. It is possible to summarize the CSO algorithm's functioning into the following eight steps:

1. Create the starting location and speed of N cats in the search space.
2. The second step is to compute the best fitness value and save it in $xbest$.
3. Third, split the cats up into two groups: one actively looking, and two actively tracing.
4. If the cat is in the seeking state, step 4 is to produce potential replacement positions using equation (28), and then select one using the random number generator.

$$\boldsymbol{x'_{j,d} = x_{j,d} \pm SRD * r * x_{j,d}} \quad (28)$$

where r is a accidental sum and $x_{j,d}$, $x'_{j,d}$ are current and novel values of measurement d correspondingly.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

5. Update the cat's position using equation (27), and its velocity using equation (29), if it's in tracking mode (30)

$$\boldsymbol{v_{k,d}(t+1) = v_{k,d}(t) + c_1 r_1 (x_{best,d}(t) - x_{k,d}(t))} \quad (29)$$

$$\boldsymbol{x_{k,d}(t+1) = x_{k,d}(t) + v_{k,d}(t+1)} \quad (30)$$

$$d = 1,2,\dots,D$$

where c_1 is the hastening constant, r_1 is an arbitrary number, t is time, $v_{k,d}(t), v_{k,d}(t+1)$are the current and new speeds, and $x_{k,d}(t), x_{k,d}(t+1)$ are the current and novel position.

6. Combine the cats from searching and drawing mode.

7. Evaluate strength value and update the best area.

8. The end point must be examined. The algorithm terminates if the condition is met; else, it returns to Step 3.

3.3.2.2. The HCPSO Procedure

In this study, we present a new hybrid approach. With the help of the popular CSO and PSO, we develop a robust metaheuristic search technique. In the HCPSO method, we use the full CSO arrangement procedure with a few minor modifications. Similar to the PSO algorithm, it keeps track of both global and regional optimums. While in this tracing mode, the PSO movement formula is applied. Also, the value of the defined dimension in the seeking mode is modified based on the best position, and the most qualified applicant for the open position is ultimately chosen. This hybridization was created to offer a more convergent approach without dramatically increasing execution time. Thus many steps of HCSPO algorithm are described here..

To do this, first construct the initial positions (X) and velocities (V) of N different people in the search space ([0,1]). (V).

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1D} \\ x_{21} & x_{22} & \cdots & x_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{ND} \end{bmatrix}, x_{kd} \in [0,1] \quad (31)$$

$$\boldsymbol{V} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1D} \\ v_{21} & v_{22} & \cdots & v_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ v_{N1} & v_{N2} & \cdots & v_{ND} \end{bmatrix} \quad (32)$$

where D is the sum of items types.

First, use equation (1) to (34).

$$\boldsymbol{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1D} \\ y_{21} & y_{22} & \cdots & y_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ y_{N1} & y_{N2} & \cdots & y_{ND} \end{bmatrix} \quad (33)$$

$$y_{kd} = round(x_{kd} * (b_d - a_d)) \quad (34)$$

Verify all of the limitations. Be certain that every one of the solutions lies in an infeasible region, defined as a set of solutions that satisfies the MBKP-MC requirements. Figure out which solutions are the most profitable overall, or have the highest fitness value. Be sure to write down C_k for the best possible position for each person and C_g for the best possible global answer. seekers and trackers. If people are actively looking for something. Make duplicates using each node's optimal position C_kEquation (35), then C_g Equation (36).

$$x'_{j,d} = C_{k,d}, \quad d = 1,2,\dots D \quad (35)$$

$$x'_{j,d} = C_{g,d} \pm SRD * r * C_{g,d} \quad (36)$$

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

If people are actively tracking something, then 1. Iterate the velocity and location using the PSO formula in equation (37) and (38)

$$\boldsymbol{V_i(t+1) = \omega V_i(t) + c_1 r_1 \left(C_g(t) - X_i(t)\right) + c_1 r_1 (C_i(t) - X_i(t))} \quad (37)$$

$$\boldsymbol{X_i(t+1) = X_i(t) + V_i(t+1)} \quad (38)$$

Combining the cats in tracing mode into one group, making sure no positions are outside the range of [0,1], and repeat. The solution needs to be interpreted using equation if it is larger space (39).

$$\boldsymbol{x_k} = \begin{cases} \frac{x_k - \min(x_k)}{\max(x_k) - \min(x_k)}, & \boldsymbol{if}\ \min(x_k) < 0 \\ \frac{x_k}{\max(x_k)}, & \boldsymbol{if}\ \max(x_k) > 1 \end{cases} \quad (39)$$

Replace Y with the term corresponding to the new position X in the MBKP-MC solution. First, the fitness value is determined once all constraints have been patterned. Both $C_k$andC_g, which represent the best local and global positions, should be updated. The end point must be examined. If this condition is met, the algorithm will exit with the result C_g. If the criteria are not met, proceed to step 6.

4. Results and Discussion

The proposed technique was evaluated using 1,200 colour fundus images from the industry-standard MESSIDOR dataset [34]. Matlab 2018a, together with a handful of additional packages, is used to put the suggested model into action. Tabular 1 displays the dataset's classification of these photos into four groups. Stage 1 includes only a few micro aneurysms seen in some of the photos. Stage 2 includes photos that have micro aneurysms and haemorrhages, while stage 3 includes images with the both category. The 10-fold cross validation procedure is used for testing...

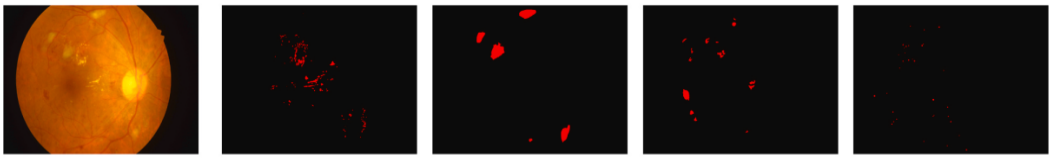
Table 1: Dataset portrayal

Levels of DR	Label	Details
Normal	Healthy	Zero abnormality
Mild NPDR	Stage 1	Occurrence of micro aneurysms
Moderate NPDR	Stage 2	Few micro aneurysms
Severe NPDR		Microvascular abnormalities inside the retina, including venous beading (IRMA)
PDR	Stage 3	Pre-retinal hemorrhage

4.1. The IDRID Dataset

In their announcement of the challenge, the creators of IDRiD claim that it is the only dataset that contain examples of both typical structures recognised at the pixel level [35]. This collection contains images of diabetic retinopathy and diabetic macular edoema, together with data on the severity of the corresponding diseases. This makes an ideal research into and testing of image processing algorithms for diabetic retinopathy screening. In specifically, the challenge's lesion segmentation job seeks to segment retinal lesions associated with diabetic retinopathy. These lesions can take from the form of microaneurysms, haemorrhages, hard exudates, or soft exudates. Accurately locating and dividing up the optic disc is also part of the mission. A retinal doctor at an Eye Clinic in Nanded founded the dataset consists of 516 photos taken from a pool of thousands of examinations (from which 81 were designated for the lesion segmentation sub-challenge). Professionals checked and confirmed that all photos are of high enough quality and clinical importance, that no duplicates exist, and that the disease stratification is reasonably representative of DR and diabetic macular edoema (DME). The complete set of 516 photos displaying a wide range of pathological DR and DME disorders were evaluated and scored by medical professionals. Figure 6 is an example from this data set's photographs.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023



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

Figure 6: (a) The following is a sample IDRiD image with ground truth masks for (b) firm exudates, (c) lenient exudates, (d) haemorrhages, and (e) microaneurysms.

All photographs were taken of 50 degrees and are focused on or very close to the macula. The photos are 4288x2848 pixels in size and saved as jpg files. Every picture is roughly 800 KB in size. There are 81 colour fundus images with DR symptoms in this dataset for the lesion segmentation mini-challenge. Accurate pixel-level annotation of DR anomalies such as micro aneurysms (MA), and haemorrhages (HE) is supplied as a binary mask to appraise the effectiveness of different lesion segmentation methods. Binary masks based on lesions are included in addition to colour fundus photos (.jpg files) (.tif files). This table lists the total sum of photos that have binary masks available for each lesion type (SE). For each of the 81 included photos, we also include a binary mask of the optic disc region, which helps to hide any anomalies.

4.2. Performance Measures

This method's efficacy was evaluated using a battery of four criteria: sensitivity, specificity, accuracy, and precision factor. Eqs. (40)-(43) provide the functions used to derive the value:

$$\boldsymbol{Accuracy\ (ACC)} = \frac{\boldsymbol{TN+TP}}{\boldsymbol{TP+TN+FN+FP}} \times \mathbf{100} \tag{40}$$

$$\boldsymbol{F-measure\ (F-M)} = \frac{\boldsymbol{2TP}}{\boldsymbol{(2TP+FP+FN)}} \times \mathbf{100} \tag{41}$$

$$\boldsymbol{Precision\ (PR)} = \frac{\boldsymbol{TP}}{\boldsymbol{(FP+TP)}} \times \mathbf{100} \tag{42}$$

$$\boldsymbol{Recall\ (RC)} = \frac{\boldsymbol{TP}}{\boldsymbol{(FN+TP)}} \times \mathbf{100} \tag{43}$$

The different stage classification and binary classification for IDRID dataset are shown in Table 2 and 3.

Table 2: Results of five stages (IDRID dataset) of DR classification

Category	Recall	F1-score	Accuracy	Precision
Moderate DR	0.97	0.96	--	0.96
Severe DR	0.89	0.92	--	0.94
No DR	0.99	0.99	--	0.98
Mild DR	0.96	0.96	--	0.97
Proliferative DR	0.97	0.97	--	0.97
Average	0.95	0.96	97.27	0.97

For normal, the proposed model has 98% to 99% of precision, recall and F1-score, where the proposed model has 94% of precision, 89% of recall and 92% of F1-score for severe DR. Finally, for average, the proposed model has 97% of precision, 95% of recall, 96% of F1-score and 97.27% of accuracy. Figure 7 provides the graphical analysis for various stages.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

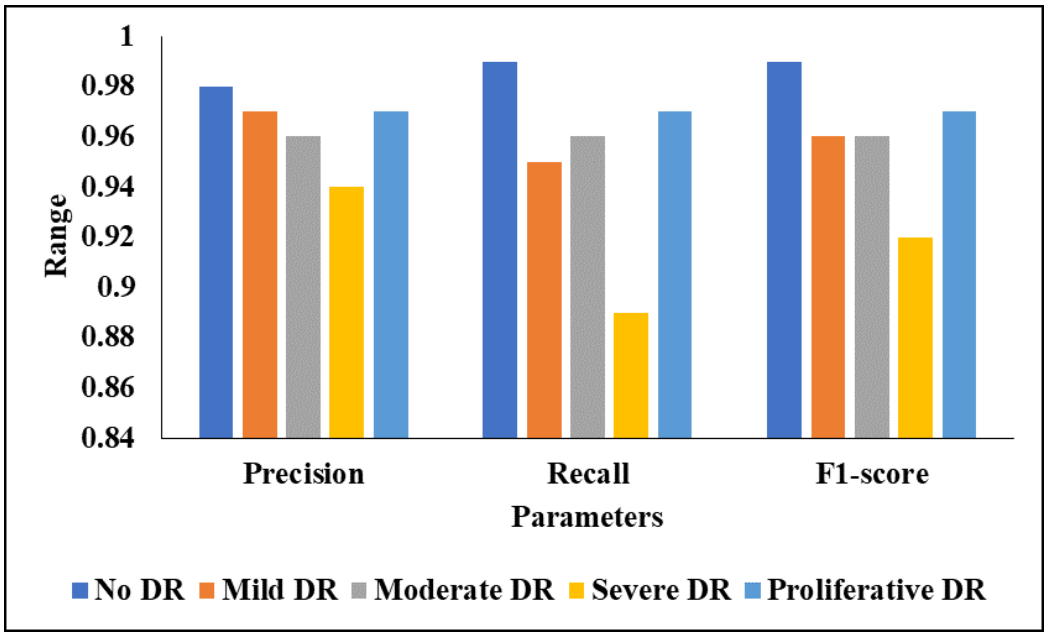


Figure 7: Various Stage of DR Analysis on IDRID dataset

Table 3: Results for binary (IDRID dataset) classification.

Type	Recall	F1-score	Accuracy	Precision
No DR	0.99	0.99	--	0.98
DR	0.98	0.98	--	0.99
Average	0.99	0.99	98.90%	0.99

For No DR, the projected model has 98% of precision, recall and 99% of F1-score and the same model achieved 99% of precision, 99% of precision, 99% of F1-score and 98.90% of accuracy for average analysis. Table 4 and 5 provides the experimental analysis of multi-stage and binary stage DR classification.

Table 4: Results of five stages-(MESSIDOR dataset) of DR classification.

Category	Recall	F1-score	Accuracy	Precision
Severe DR	1.00	1.00	--	1.00
Proliferative DR	0.86	0.92	--	1.00
No DR	0.99	0.97	--	0.96
Mild DR	0.94	0.94	--	0.94
Moderate DR	0.91	0.94	--	0.97
Average	0.94	0.96	96.26%	0.98

The proposed model has nearly 94% to 97% of precision for no DR, mild DR and moderate DR, 100% of precision for severe DR. The proposed model has 96.26% of accuracy, 96% of F1-score, 94% of recall and 98% of precision for average analysis. The model attained 91% to 99% of recall for first three stage and achieved 94% to 97% of F1-score for the no DR, mild DR and moderate DR. Figure 8 presents the graphical analysis.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

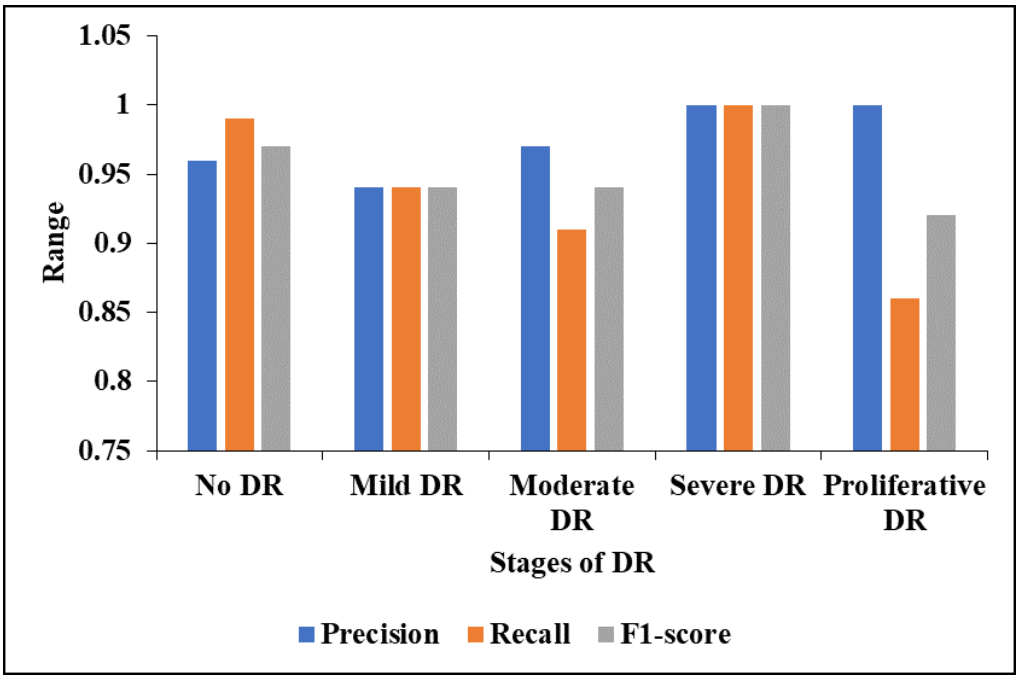


Figure 8: Various Stage of DR Analysis on MESSIDOR dataset

Table 5: Results for binary (MESSIDOR dataset) classification.

Category	Recall	F1-score	Accuracy	Precision
No DR	1.00	1.00	--	0.99
DR	0.99	1.00	--	1.00
Average	1.00	1.00	99.73%	1.00

For average analysis, the proposed model has 100% of precision, 100% of recall and 99.73% of accuracy. For DR analysis, the proposed model has 99% of recall and 100% of precision and recall.

4.3. Comparative Investigation of Projected Model with existing Procedures

The existing techniques such as RTB [16], CARNet [17], U-Net [19] and FEB-GPB [22] are used IDRID dataset for DR classification, therefore, these techniques are implemented and tested with MESSIDOR dataset, then results are averaged in Table 6 and 7.

Table 6: Results of five phases (IDRID dataset) of DR classification.

Model	Accuracy	Precision	Recall	F-Score
RTB [16]	0.6770	0.69	0.90	0.81
CARNet [17]	0.8371	0.88	0.74	0.92
U-Net [19]	0.8985	0.91	0.83	0.95
FEB-GPB [22]	0.9421	0.92	0.87	0.92
RMMCL-HCPSO	0.9727	0.97	0.95	0.96

When the techniques are tested with accuracy, the RTB achieved only 67.70%, CARNet achieved 83.71%, U-Net has 89%, FEB-GPB has 94.21% and proposed model has 97.27%. In the analysis of F-score, the existing models achieved nearly 92% to 95%, RTB has 81% and proposed model has 96%. When comparing with all techniques, CARNet has less recall (74%), RTB has 90%, U-Net and FEB-GPB has 83% to 87% of recall and proposed model has 95% of recall. The reason for better performance is that the relevance mapping is used and ROI is considered for five different stages. In addition, the hyper-

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

parameters of the RMMCL are carried out using HCPSO and reaches better classification accuracy. Figure 9 shows the graphical analysis.

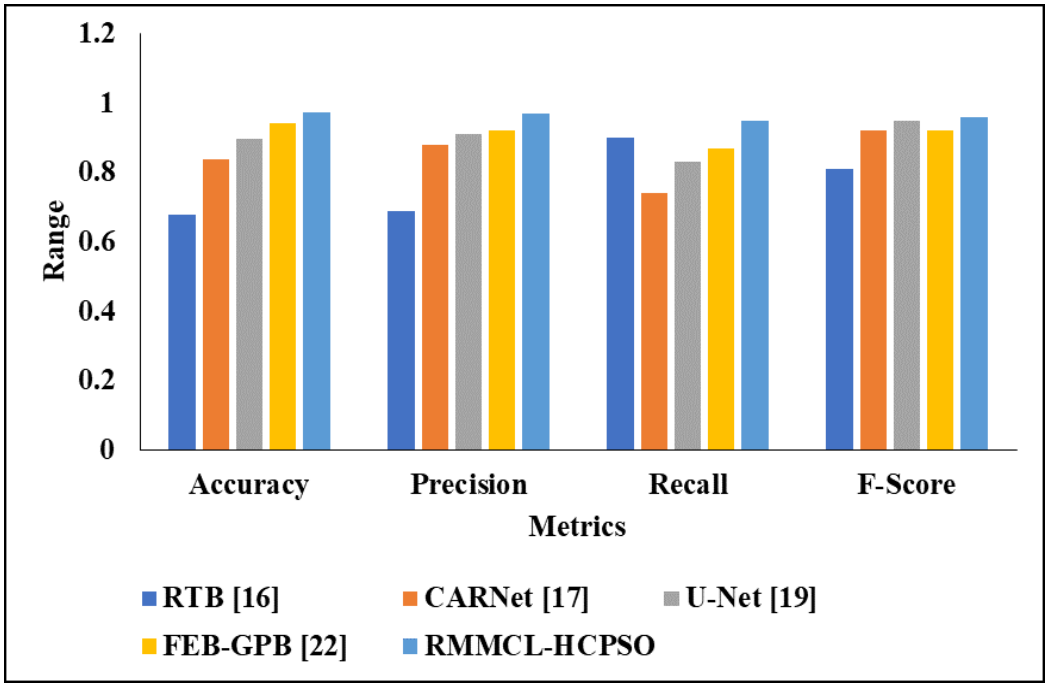


Figure 9: Proportional Analysis of Projected on IDRID dataset

Table 7: Consequences of five phases (MESSIDOR dataset) of DR classification.

Model	Accuracy	Precision	Recall	F-Score
RTB [16]	0.8271	0.88	0.86	0.92
CARNet [17]	0.8476	0.89	0.88	0.90
U-Net [19]	0.8915	0.91	0.90	0.93
FEB-GPB [22]	0.9576	0.95	0.92	0.94
RMMCL-HCPSO	0.9626	0.98	0.94	0.96

In the second dataset called MESSIDOR, the projected model achieved 96.26% of accuracy, 98% of precision, 94% of recall and 96% of F-score. But, the existing techniques such as RTB, CARNet, U-Net and FEB-GPB has nearly 84% to 95% of accuracy, 89% to 95% of precision, 88% to 92% of recall and 90% to 94% of F-score. Figure 10 provides graphical representation.

Doi: https://doi.org/10.54216/FPA.110207
Received: December 27, 2022 Accepted: April 25, 2023

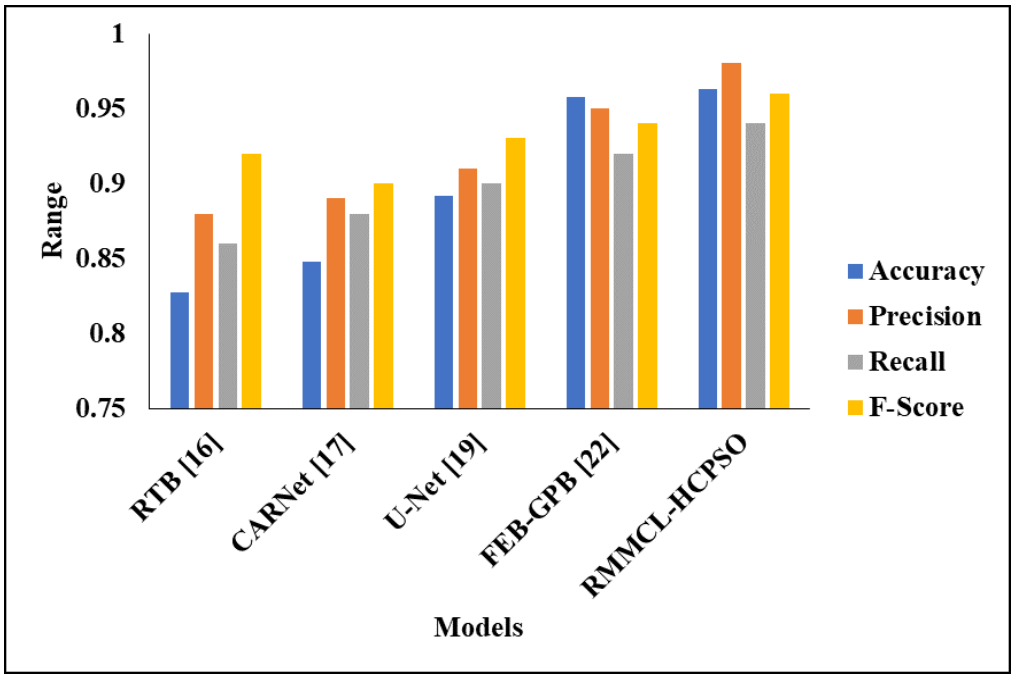


Figure 10: Analysis of Proposed Model on MESSIDOR dataset

6. Conclusion

In this work, we use a multi-step procedure that begins with data cleaning and ends with segmentation and classification. In the first stage of the research, the clipping limit of CLAHE is set by the rat optimization method, and any unwanted noise, size variants, or colour variants are eliminated and incorporated into the final product. This study proposes a new sparse fuzzy c-means method for the segmentation process. It has been observed that deep learning approaches, and convolutional neural networks in particular, outperform more traditional machine learning methods when it comes to retrieving and processing information related to the imaging modalities with minimal human interaction. While many studies have used DL to classify modalities, none of these investigations have attempted to understand the DL models' internal representations or prediction outcomes. Specifically, we introduce a new approach called RMMCL for detecting and highlighting regions in an image that are most discriminative in discriminating DR picture modalities, with the goal of providing a visual explanation for the behaviour taught by CNN-based DL models... While prior models achieved an accuracy of about 89% to 92% on both datasets across several stages of classification, the experimental findings show that the suggested model achieved an accuracy of 97.27% on the IDRID dataset and 96.26% on the MESSIDOR dataset. Better results can be expected in the future from DR segmentation and feature extraction if a new DL model is used.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Wan, C., Chen, Y., Li, H., Zheng, B., Chen, N., Yang, W., Wang, C. and Li, Y., 2021. EAD-net: a novel lesion segmentation method in diabetic retinopathy using neural networks. Disease Markers, 2021.

[2] Xu, Y., Zhou, Z., Li, X., Zhang, N., Zhang, M. and Wei, P., 2021. Ffu-net: Feature fusion u-net for lesion segmentation of diabetic retinopathy. BioMed Research International, 2021.

[3] Garifullin, A., Lensu, L. and Uusitalo, H., 2021. Deep Bayesian baseline for segmenting diabetic retinopathy lesions: Advances and challenges. Computers in Biology and Medicine, 136, p.104725.

[4] Abdelmaksoud, E., El-Sappagh, S., Barakat, S., Abuhmed, T. and Elmogy, M., 2021. Automatic diabetic retinopathy grading system based on detecting multiple retinal lesions. IEEE Access, 9, pp.15939-15960.

[5] Dai, L., Wu, L., Li, H., Cai, C., Wu, Q., Kong, H., Liu, R., Wang, X., Hou, X., Liu, Y. and Long, X., 2021. A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nature communications, 12(1), p.3242.

[6] Zhou, Y., Wang, B., Huang, L., Cui, S. and Shao, L., 2020. A benchmark for studying diabetic retinopathy: segmentation, grading, and transferability. IEEE Transactions on Medical Imaging, 40(3), pp.818-828.

[7] Van Craenendonck, T., Elen, B., Gerrits, N. and De Boever, P., 2020. Systematic comparison of heatmapping techniques in deep learning in the context of diabetic retinopathy lesion detection. Translational vision science & technology, 9(2), pp.64-64.

[8] Sambyal, N., Saini, P., Syal, R. and Gupta, V., 2020. Modified U-Net architecture for semantic segmentation of diabetic retinopathy images. Biocybernetics and Biomedical Engineering, 40(3), pp.1094-1109.

[9] Erciyas, A. and Barışçı, N., 2021. An effective method for detecting and classifying diabetic retinopathy lesions based on deep learning. Computational and Mathematical Methods in Medicine, 2021, pp.1-13.

[10] Porwal, P., Pachade, S., Kokare, M., Deshmukh, G., Son, J., Bae, W., Liu, L., Wang, J., Liu, X., Gao, L. and Wu, T., 2020. Idrid: Diabetic retinopathy–segmentation and grading challenge. Medical image analysis, 59, p.101561.

[11] Lakshminarayanan, V., Kheradfallah, H., Sarkar, A. and Jothi Balaji, J., 2021. Automated detection and diagnosis of diabetic retinopathy: A comprehensive survey. Journal of Imaging, 7(9), p.165.

[12] Colomer, A., Igual, J. and Naranjo, V., 2020. Detection of early signs of diabetic retinopathy based on textural and morphological information in fundus images. Sensors, 20(4), p.1005.

[13] Qiao, L., Zhu, Y. and Zhou, H., 2020. Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. IEEE Access, 8, pp.104292-104302.

[14] Mateen, M., Wen, J., Nasrullah, N., Sun, S. and Hayat, S., 2020. Exudate detection for diabetic retinopathy using pretrained convolutional neural networks. Complexity, 2020, pp.1-11.

[15] Zago, G.T., Andreão, R.V., Dorizzi, B. and Salles, E.O.T., 2020. Diabetic retinopathy detection using red lesion localization and convolutional neural networks. Computers in biology and medicine, 116, p.103537.

[16] Huang, S., Li, J., Xiao, Y., Shen, N. and Xu, T., 2022. RTNet: relation transformer network for diabetic retinopathy multi-lesion segmentation. IEEE Transactions on Medical Imaging, 41(6), pp.1596-1607.

[17] Guo, Y. and Peng, Y., 2022. CARNet: Cascade attentive RefineNet for multi-lesion segmentation of diabetic retinopathy images. Complex & Intelligent Systems, 8(2), pp.1681-1701.

[18] Wang, X., Fang, Y., Yang, S., Zhu, D., Wang, M., Zhang, J., Zhang, J., Cheng, J., Tong, K.Y. and Han, X., 2023. CLC-Net: Contextual and Local Collaborative Network for Lesion Segmentation in Diabetic Retinopathy Images. Neurocomputing.

[19] Upadhyay, K., Agrawal, M. and Vashist, P., 2023. Characteristic patch-based deep and handcrafted feature learning for red lesion segmentation in fundus images. Biomedical Signal Processing and Control, 79, p.104123.

[20] Aziz, T., Charoenlarpnopparut, C. and Mahapakulchai, S., 2023. Deep learning-based hemorrhage detection for diabetic retinopathy screening. Scientific Reports, 13(1), p.1479.

[21] Nahiduzzaman, M., Islam, M.R., Goni, M.O.F., Anower, M.S., Ahsan, M., Haider, J. and Kowalski, M., 2023. Diabetic Retinopathy Identification Using Parallel Convolutional Neural Network Based Feature Extractor and ELM Classifier. Expert Systems with Applications, p.119557.

[22] Gu, Z., Li, Y., Wang, Z., Kan, J., Shu, J. and Wang, Q., 2023. Classification of Diabetic Retinopathy Severity in Fundus Images Using the Vision Transformer and Residual Attention. Computational Intelligence and Neuroscience, 2023.

[23] Dhiman, G., Garg, M., Nagar, A., Kumar, V. and Dehghani, M., 2021. A novel algorithm for global optimization: rat swarm optimizer. Journal of Ambient Intelligence and Humanized Computing, 12(8), pp.8457-8482.

[24]. Bhardwaj C, Jain S, Sood M: Automated Optical disc segmentation and blood vessel extraction for fundus images using ophthalmic image processing. In International Conference on Advanced Informatics for Computing Research. Springer, Singapore, pp. 182–194, 2018.

Doi: https://doi.org/10.54216/FPA.110207

Received: December 27, 2022 Accepted: April 25, 2023

[25]. Niemeijer M, Staal J, Ginneken BV, Loog M, Abramoff MD: Comparative study of retinal vessel segmentation methods on a new publicly available database. In Medical Imaging 2004: Image Processing. International Society for Optics and Photonics. 5370, pp. 648–657, 2004.

[26]. Rahim SS, Palade V, Shuttleworth J, Jayne C: Automatic screening and classification of diabetic retinopathy and maculopathy using fuzzy image processing. Brain Inf. 3(4), 249–267, 2016

[27] Y. Zhang, X. Bai, R. Fan, and Z. Wang, ''Deviation-sparse fuzzy C-means with neighbor information constraint,'' IEEE Trans. Fuzzy Syst., vol. 27, no. 1, pp. 185–199, Jan. 2019

[28] L. Guo, L. Chen, X. Lu, and C. L. P. Chen, ''Membership affinity lasso for fuzzy clustering,'' IEEE Trans. Fuzzy Syst., vol. 28, no. 2, pp. 294–307, Feb. 2020

[29] S. Miyamoto and M. Mukaidono, ''Fuzzy C-means as a regularization and maximum entropy approach,'' in Proc. 7th Int. Fuzzy Syst. Assoc. World Congr., Prague, Czech Republic, 1997, pp. 86–92.

[30] J. Huang, F. Nie, and H. Huang, ''A new simplex sparse learning model to measure data similarity for clustering,'' in Proc. 24th Int. Joint Conf. Artif. Intell., Buenos Aires, Argentina, 2015, pp. 3569–3575.

[31] Santoso, K.A., Kurniawan, M.B., Kamsyakawuni, A. and Riski, A., 2022, February. Hybrid Cat-Particle Swarm Optimization Algorithm on Bounded Knapsack Problem with Multiple Constraints. In International Conference on Mathematics, Geometry, Statistics, and Computation (IC-MaGeStiC 2021) (pp. 244-248). Atlantis Press.

[32] Zhou, B.; Khosla, A.; Lapedriza, A.; Oliva, A.; Torralba, A. Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 2921–2929.

[33] Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the International conference, Computer Vision Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 618–626.

[34] Messidor Dataset, [online]. Available: http://www.adcis.net/en/third-party/messidor/

[35] Furtado, P., Baptista, C. and Paiva, I., 2020, February. Segmentation of Diabetic Retinopathy Lesions by Deep Learning: Achievements and Limitations. In Bioimaging (pp. 95-101).

Doi: https://doi.org/10.54216/FPA.110207

Received: December 27, 2022 Accepted: April 25, 2023