

Retinal lesions classification for diabetic retinopathy using custom ResNet-based classifier

Silpa Ajith Kumar, James Satheesh Kumar

Department of Electronics and Instrumentation Engineering, Dayananda Sagar College of Engineering (DSCE) Bengaluru, India and Visvesveraya Technological University, Belagavi, India

Article Info

Article history:

Received Jul 5, 2023

Revised Nov 2, 2023

Accepted Nov 8, 2023

Keywords:

Batch normalization

Convolutional neural network

Custom ResNet-based classifier

Deep learning

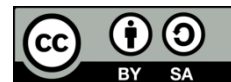
Diabetic retinopathy

Retinal lesions

ABSTRACT

Failure to diagnose and treat retinal illnesses on time might lead to irreversible blindness. The focus is on three common retinal lesions associated with diabetic retinopathy (DR): microaneurysms (MAs), haemorrhages, and exudates. The proposed solution leverages deep learning, employing a customized residual network (ResNet) based classifier trained on real-time retinal images meticulously annotated and graded by ophthalmologists. Annotation noise was a significant obstacle addressed by downsampling and augmenting the data. Compared to cutting-edge techniques, this one performs better with test-set accuracy of 93.34% across all classes. This approach holds great promise for enhancing early detection and treatment of DR by automating the recognition of these vital retinal abnormalities. The ability to automatically classify these symptoms can aid clinicians in making more precise diagnosis and starting treatments sooner. This research shows that deep learning-based approaches are highly effective, especially when combined with a customised ResNet-based classifier and thorough pre-processing steps. We observed that this method provides the ability to better the lives of patients and lower the rate of permanent blindness resulting from retinal disorders.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Silpa Ajith Kumar

Department of Electronics and Instrumentation Engineering, Dayananda Sagar College of Engineering (DSCE) Bengaluru, India and Visvesveraya Technological University

Bengaluru, Karnataka, India

Email: silpaajithkumar@gmail.com

1. INTRODUCTION

Retinal problems are crucial indicators of general eye health. When light enters the eye, it is transformed into electrical impulses that travel along the optical nerve to the brain. The retina can exhibit diseases and disorders that can disrupt the normal vision and perception of an individual. These diseases can sometimes cause permanent blindness if unchecked. Therefore, it is crucial to get frequent retinal examinations, as early identification can prevent degenerative illnesses and prevent a person from going blind [1].

Retinal exudates can be either hard or soft. Exudates are yellow specks composed of lipid remnants that form as a result of serious leaking from injured capillaries. Increased vascular permeability in the retina, indicated by the presence of hard exudates, can be the result of several potentially devastating or even fatal diseases. Yellow-white cotton-wool-like structures represent soft exudates [2]. Microaneurysms (MAs) are usually the earliest visible manifestation of diabetic retinopathy (DR) [3]. They appear as scattered red dots on the retina [4]. Retinal haemorrhage is bleeding from the blood vessels in the retina. This problem can be caused by several conditions, including diabetes, hypertension, anaemia, and leukaemia. Physical trauma or exposure to sudden changes in pressure can also cause retinal haemorrhage.

Artificial intelligence's (AI) recent rise to prominence has had profound effects on medical care. The AI-based solutions in healthcare aim to assist specialists and improve diagnosis. In radiology, AI is frequently utilised as a second reader, and in certain situations, as the primary reader. It aims to counter human fatigue and bias that can often occur. The field of ophthalmology has a promising potential for AI-based solutions that can assist in the early diagnosis and prevention of vision-threatening diseases [5], [6]. Convolutional neural networks (CNNs) have paved way for training image-based classifiers. Furthermore, complex CNN architectures can detect patterns which are identifiable by humans. Deep CNNs impact on medical imaging applications [7], [8].

The deep learning model created by Bajwa *et al.* [9] was trained on a proprietary dataset which identified test representations as DR-positive and negative. The model's accuracy, sensitivity, and specificity were all independently verified at 93.40%, 97.42%, and 89.45%, respectively. Due to its reliance on modified CNNs, which necessitate a high number of training images, and a limited, testing dataset gathered with the same device, the proposed model has limitations. Ali *et al.* [10] proposed a deep learning model named incremental modular network synthesis (IMNS), that efficiently classified several kinds of medical applications, which include malaria, DR, and tuberculosis, with an accuracy of 97% for malaria, 97.9% for DR, and 88.6% for tuberculosis, despite using a relatively small dataset.

Fundus pictures gathered from diverse sources are used by Rodriguez *et al.* [11] to create a special multi-label classification approach for the diagnosis of various retinal illnesses. The base of the MuReD dataset is a collection of other existing datasets. It's a one-of-a-kind multi-label dataset for diagnosing diseases of the fundus. To guarantee the precision and variety of the diseases they address, a wide range of post-processing techniques are used once the images have been collected. The approach is shown to outperform state-of-the-art research on the identical issue by 7.9% and 8.1%, respectively, in terms of the AUC score for disease classification and sickness detection. Since morphological changes frequently take place before the physical signs of these disorders. Optical coherence tomography (OCT) can offer a method for the early identification of a variety of diseases [12]. Moreover, follow-up imaging can evaluate a disease's recurrence and the efficacy of a treatment. The results and conclusions drawn from OCT images in the field of retinal diseases, and how to employ these conclusions in clinical settings, necessitate a review of the literature in this area [13], [14].

Eladawi *et al.* [15] provide an overview of the methods currently being used to assess how well OCT pictures may be used for the early identification and diagnosis of retinal disorders. The report also lists many difficulties encountered by researchers while analysing OCT retinal images. The effectiveness of AI equipped with deep learning has been proved in several areas of medicine [16], [17]. Data can be represented as a 1D signal (e.g., an electrocardiogram), 2D or 3D diagnostic imaging (e.g., a colour fundus photograph or OCT), or a structured electronic medical record. According to Litjens *et al.* [18], its performance is optimal for well-defined clinical tasks in which the vast majority of relevant information is already included in the data. Here we'll examine how some of the most popular AI systems currently available are being put to use in the medical field, where they have been shown to at least keep up with human specialists. Ophthalmologists can benefit greatly from implementing AI into their practices. Li *et al.* [19] the related study develops a collaborative automotive network (CANet), an AI model based on deep network architecture, in which a disease-specific learning module reliably predicts a wide range of illness features. DR and diabetic macular edema (DME) will be the module's first goals. Multiple deep learning-based approaches exist for DR categorization [20], [21]. Many current approaches, however, do not deal with the difficulty of classifying real-time datasets of exudates, MAs, and haemorrhages simultaneously [22], [23].

Given the diversity of retinal diseases and the paramount importance of early detection, this study's main contributions involve the application of AI to solve the difficult problems associated with the detection and categorization of retinal lesions. The research employs CNNs and deep learning techniques to create a model capable of simultaneously distinguishing MAs, haemorrhages, and exudates-critical indications of numerous retinal disorders, including DR. The study uses CNNs and deep learning methods to develop a model that can detect MAs, haemorrhages, and exudates all at once. This is a significant step in improving diagnostic accuracy since it offers a thorough evaluation of retinal illness that goes beyond the identification of specific lesions. Residual network (ResNet)-18 represents one of the most cutting-edge architectures used in this work, which pioneers the integration of AI-driven categorization to effectively tackle retinal pathologies.

Important findings and new contributions made in this work are the effective creation of a CNN-based model, which has a remarkable accuracy of 93.34% on test data. The model's ability to simultaneously recognise and categorise several retinal lesions such as MAs, haemorrhages, and exudates is a significant advancement beyond traditional techniques that concentrate only on individual lesion recognition. This all-encompassing method aids in the accurate diagnosis of retinal illness and provides ophthalmologists with a deeper understanding of disease progression. The model's effectiveness even in the presence of concurrent diseases is indicative of its robustness and therapeutic potential. Thus, this study paves the way for future

improvements in multi-class identification of retinal diseases, which offers increased precision and effectiveness in disease diagnosis.

The paper discusses the information on the datasets employed, the data preparation, pre-processing procedures, the CNN design, and the deep learning model's training in section 2. In section 3 explains performance metrics like receiver operating characteristic (ROC) curves and the confusion matrix. In section 4, we discuss the conclusion and the potential significance of the future work of the proposed study.

2. METHOD

2.1. Datasets

Literature review indicates that present algorithms [24]–[29] for diagnosing DR rely heavily on publicly available datasets like Kaggle, Messidor, STARE, DRIVE, and ORIGA. A real-world hospital dataset containing information on 450 patients with MAs, haemorrhages, and exudates at various stages of DR are tested across several DR phases. The different stages of the fundus images like MAs, haemorrhages, and exudates were graded by two ophthalmologists before the deep learning model is trained. The Figure 1 depicts the lesions present in the different stages of DR. The Figures 1(a), 1(b), and 1(c) depicts the lesions of MAs, haemorrhage, and exudates respectively that can be seen at different stages during the detection of DR. The images are grouped into 3 disorders. Some images, however, depict more than one disorder. In such cases, haemorrhage was given the highest priority, followed by MA and exudate with the lowest. This decision was made in considering the disease's severity, and the classifier would be trained to focus on identifying haemorrhages.

We begin by outlining the challenges posed by real-time clinical datasets for recognising various lesions such as MAs, haemorrhages, and exudates. After that, we provide a clear outline of the datasets used in the proposed work and data preparation, a set of preprocessing operations performed on the collected real-time fundus images. We next look into the implementation of the CNN architecture using ResNet 18. Finally, the training environments for detecting retinal lesions are outlined so they can be detected simultaneously.

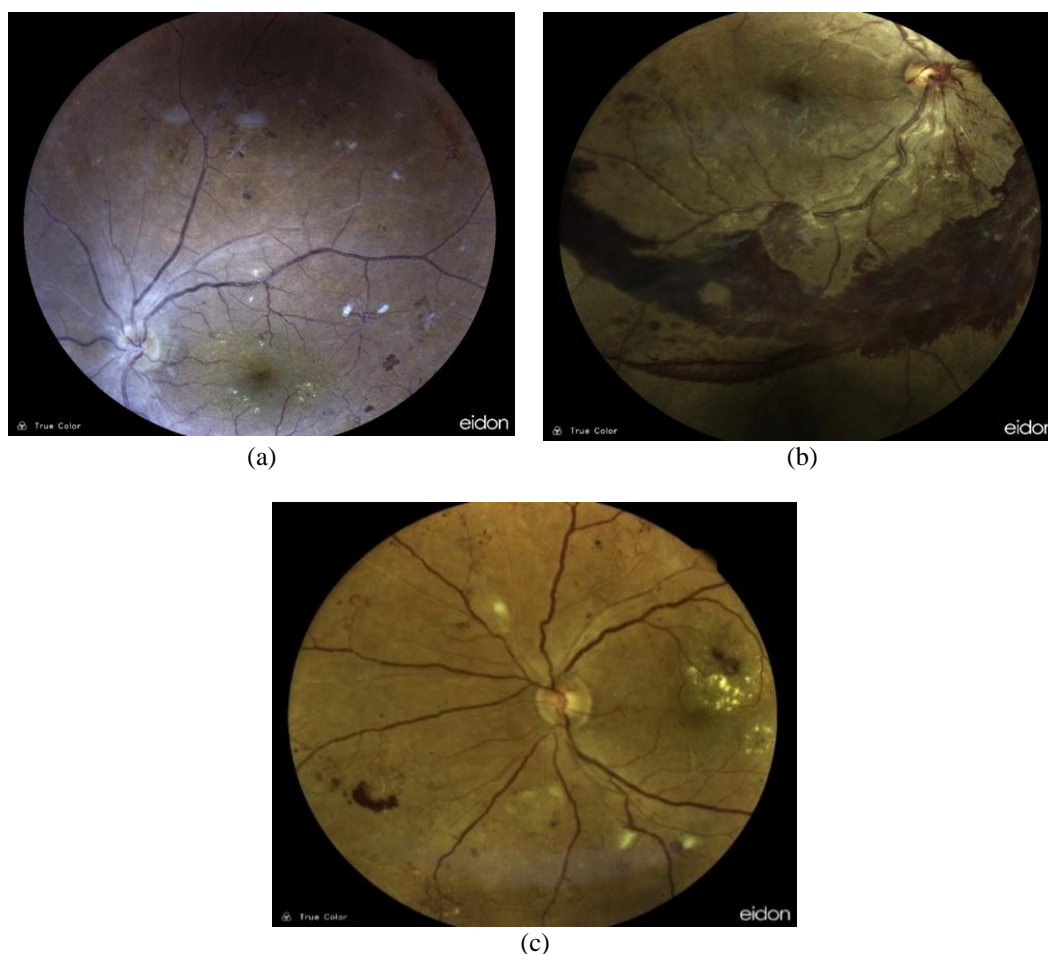


Figure 1. Lesions present in DR; (a) micro aneurysms, (b) haemorrhages, and (c) exudates

2.2. Data preparation

The three types of lesions namely exudates, haemorrhages, and MAs-have their own separate CSV files that are used to prepare the data for training the deep learning model. The path of each image stored in specific folders is retrieved and labels are assigned. The three different labels are assigned namely for hemorrhages: 1, exdates: 2 and MAs: 3. Training is done using the one hot classification technique. The input which is fed to the ResNet 18 network is based on one hot classification. One-hot encoding is a commonly used technique in classification tasks. It's used to represent categorical labels as binary vectors, where each class is represented by a unique combination of binary values. In the work that is being proposed, each label for a class is transformed into a binary vector, where each element indicates whether or not the class is present. For a three-class problem, the one-hot encoded vectors would look like: exudates: [1, 0, 0], hemorrhages: [0, 1, 0] and MAs: [0, 0, 1].

2.3. Pre-processing steps

The dataset is proprietary and was acquired for the specific purpose of this research. Images of the retinal fundus are of excellent quality, measuring 3,680 by 3,288 pixels. Training images of such a large dimension proved to be a challenge with current hardware constraints. Therefore, the images were down-sampled to 1/3rd the resolution to 1,226×1,096. The images were smoothed using a Gaussian filter with a 3×3 kernel, the PyTorch's torchvision. Transforms are applied for the preprocessing steps such as resizing, normalization, and data augmentation if needed. Figure 2 shows the results of applying the various pre-processing approaches discussed in the review papers [30]–[33] for improving the noise and resolution of fundus pictures. When employed correctly, these methods improved key performance indicators for classification models in the aforementioned assessments of the research literature. When compared to other methods that rely on publicly available datasets, the suggested method's effectiveness is enhanced by the excellent quality of the fundus images, which necessitates no additional pre-processing.

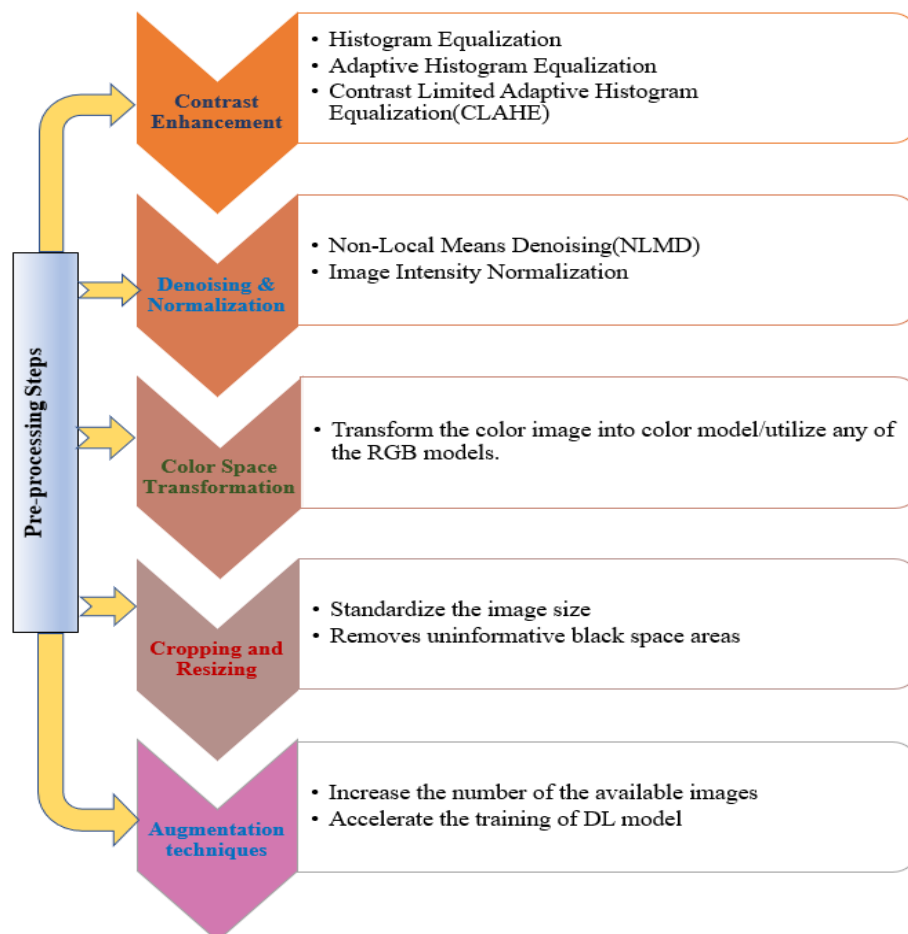


Figure 2. Steps in DR detection prior to analysis of DR as per the review papers

2.4. Proposed algorithm and architecture of the CNN model

A ResNet-18 CNN-based network is employed as the training model. ResNet is a popular deep CNN design because it outperforms its predecessors in training very deep networks. The ResNet-18 model design is shown in Figure 3, which illustrates the suggested model. The network is fed with the whole RGB images as input. The input images are reduced in resolution by a factor of three. The network begins with a convolutional layer, and its first layer has a 7×2 stride kernel. This is then followed by a rectified linear unit (ReLU) activation layer and a batch normalisation layer. The input is processed via a max pooling layer with a 3×3 kernel and a 2×2 stride before reaching the residual blocks. The convolutional and batch normalisation layers are arranged sequentially across the four residual blocks.

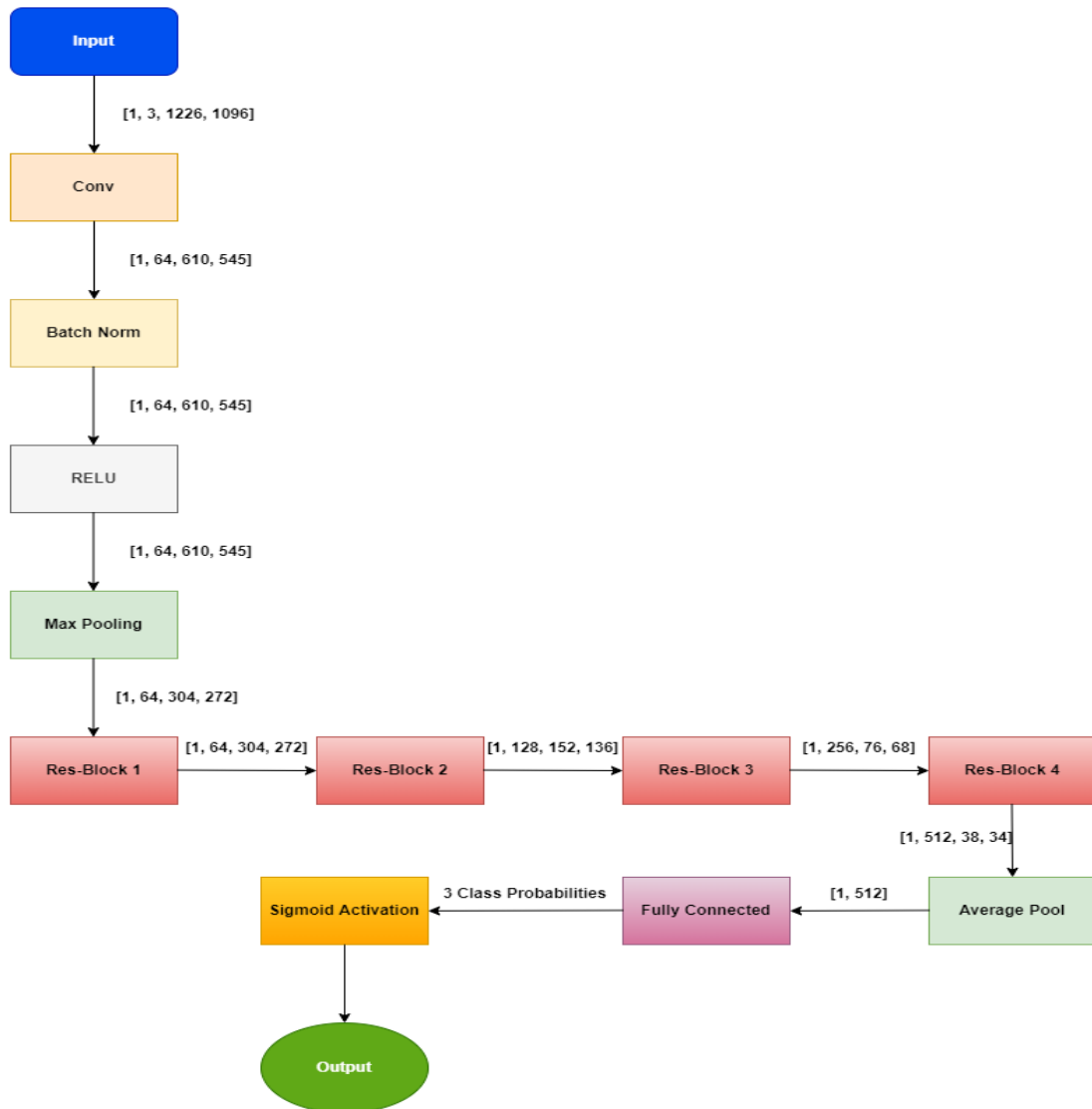


Figure 3. Architecture of the proposed network

By introducing the idea of residual blocks, ResNet helps to solve the vanishing gradient problem, allowing for the training of very deep networks. The network receives a 4-dimensional input. According to the methods described in the literature papers [34]–[36], these residual blocks are linked together using skip connections. In ResNets, the output of one layer is added to the input of the next layer via a skip connection, which prevents the degradation of learning as one progress from the first to the deeper layers. CNN models often employ activation functions like the ReLU. The Figure 4 illustrates the normal connection and the residual connection present in the ResNet architecture. Figure 4(a), explains how, in standard CNN models with no residual connections, x is fed into the weight layers and Figure 4(b) demonstrates how x is passed on

to ResNet's weight layers and residual connections. Skipping the second or third non-linear layer ReLU in a CNN-based ResNet model is used to learn the skip weights, and an extra weight matrix is employed to do so. This makes it easier to fix the problem of vanishing gradients in deep neural networks. To further process the data, the results from the previous average pooling layer are passed to a fully connected layer with three outputs (class probabilities). The results are then activated using a sigmoid function.

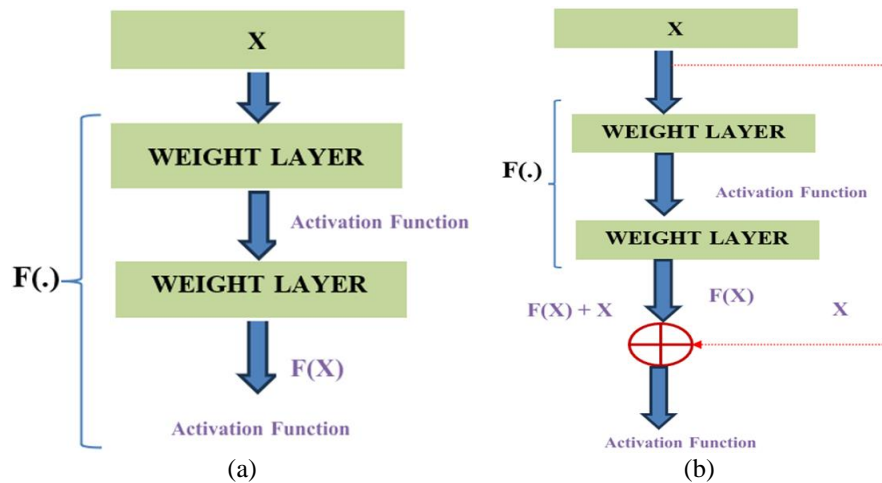


Figure 4. Analysis of the normal connection and the residual connections used in ResNet (a) CNN model with no residual connections and (b) CNN model with residual connections

2.5. Training parameters

Separate training and validation sets were created. The validation set consisted of 20% of the images used in the training set. Due to computational constraints, a batch size of 10 was utilised. The loss function was a binary cross entropy (BCE) function.

$$BCE = -\frac{1}{N} \sum_{i=0}^N y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i) \tag{1}$$

training a network entail finding the optimal way to reduce the loss function defined by (1).

$$y = \text{minimize}(BCE(\theta)) \tag{2}$$

where: θ =Training parameters

BCE loss was chosen due to its track record for performing multi-class classification. ADAM, with a learning rate of 0.001, is the optimizer of choice throughout training we used an early stopping criterion of 5 epochs to complete training on the network for 30 epochs. Early termination was determined by a combined weight of training and validation loss. A conditional learning rate decaying method was used where the learning rate was reduced every 10 epochs if there was a significant reduction in training and validation loss between the start and end of the current learning rate cycle. The best-trained model was from the 26th epoch where early stopping kicked in on the 21st epoch.

3. RESULTS AND DISCUSSION

The proposed methodology is implemented on PyTorch vision which is a popular package built on top of the PyTorch deep learning framework. The system requirements for the proposed work are a RAYZEN 4,800 processor, 8 core, 16 thread, 64 GB RAM. The RTX 3,070 Ti with 8GB VRAM is the GPU specification. In total, 450 pictures were used to train the network, comprising 120 from the exudate class, 175 from the haemorrhage class, and 155 from the MA class. Validation was performed throughout the training process on a random sample of 90 images (20%). The efficiency of the network was evaluated using a dataset of 30 images (10 images for each class).

3.1. Performance measures

On the test data, the proposed method was 93.34% accurate. The outcomes of several performance indicators are displayed in Table 1. Apart from accuracy, sensitivity, specificity and false positive rate (FPR) were used as performance metrics to determine the validity of the proposed network. The network identified MAs with a 100% sensitivity, whereas haemorrhages and exudates were identified with a sensitivity of 90%. The network misclassified a haemorrhage as MA because the haemorrhage image also contained MAs. A similar situation was seen in exudates where MAs were present. Exudate and haemorrhage classes were predicted with 100% specificity, no false positives of these classes were present. However, MA was falsely identified in exudate and haemorrhage classes with 90% specificity.

Table 1. Performance measures on test data

Performance metrics	Microaneurysms	Hemorrhages	Exudates
Sensitivity	1	0.9	0.9
Specificity	0.9	1	1
False positive rate	0.1	0	0

3.2. Confusion matrix

Prediction accuracy by class is shown in Figure 5 as the confusion matrix. Starting with exudates, 9/10 exudates were correctly predicted as exudates. No exudates were misclassified as haemorrhage, but one exudate was misclassified as MA class. Nine out of ten haemorrhages were recognised without any confusion with exudate. One haemorrhage, however, was incorrectly labelled as a MAs effect. A perfect 100% sensitivity was achieved when classifying all 10 MAs. One haemorrhage and one exudate were incorrectly identified as MA, which decreased the test's specificity.

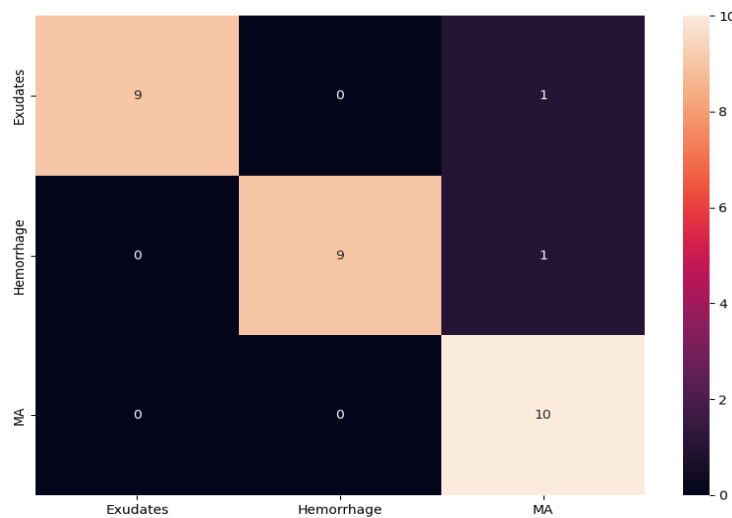


Figure 5. Confusion matrix

3.3. Validation loss and error curve

Important methods for evaluating deep learning models' efficacy include validation loss and the error curve. They help researchers and practitioners understand how well a model is learning from the data and when it might be overfitting or underfitting. It quantifies how well the model generalizes to unseen data, which is essential for assessing its ability to make accurate predictions on new, unseen examples. Figure 6 shows the progression of validation loss and error throughout 30 epochs.

Loss functions such as contrastive loss [37] and perceptual loss [38] could be explored as alternatives to the current BCE loss. The network could be trained to detect multiple disorders in one single image. This could be beneficial for this application since MAs and exudates are present in almost all classes. The network could determine the presence of multiple classes with a confidence score for each class. Different CNN-based architectures apart from ResNet could be used and compared with the current method. The advent of transformer architectures [39] and its importance not only in the natural language processing domain but also

in the computer vision space could also be utilized here. Transformer networks for classification could be tested for this application. The current classification work can be expanded to include segmentation and detection (bounding box) [40] of disorders. This could be especially beneficial for determining haemorrhages.

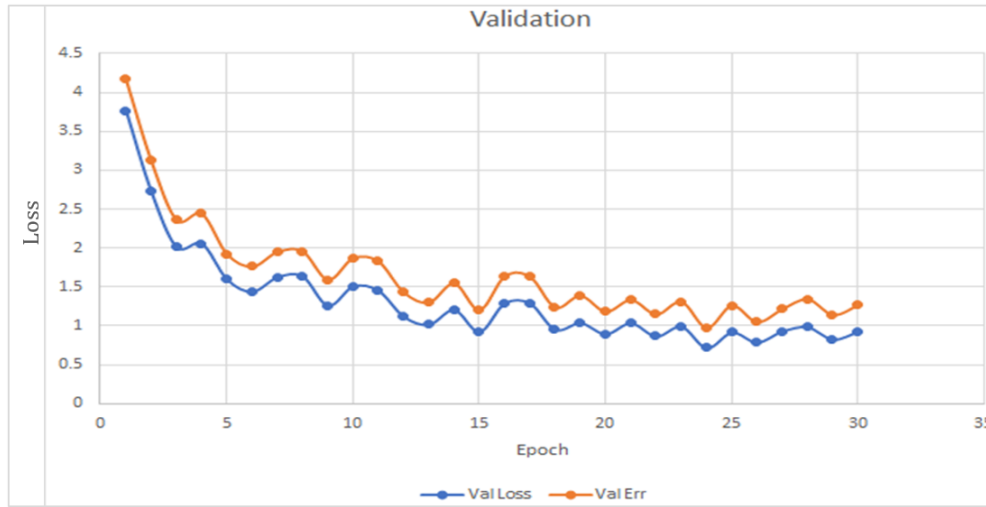


Figure 6. Validation error and loss curve

3.4. ROC curves

ROC curves offer a visual and quantitative method to assess and compare model performance. OC curves are created by plotting the true positive rate (TPR) against the FPR at various classification thresholds. The Figure 7 depicts the ROC representation of exudates Vs MAs and haemorrhage vs MA. Figure 7(a) demonstrates the ROC curve for exudates against MAs. In the predictions for the exudate class, one image was misclassified as MAs. Therefore, the classifier has 1 false positive. Figure 7(b) demonstrates the ROC for haemorrhage vs MAs; the same conclusion can be drawn as above.

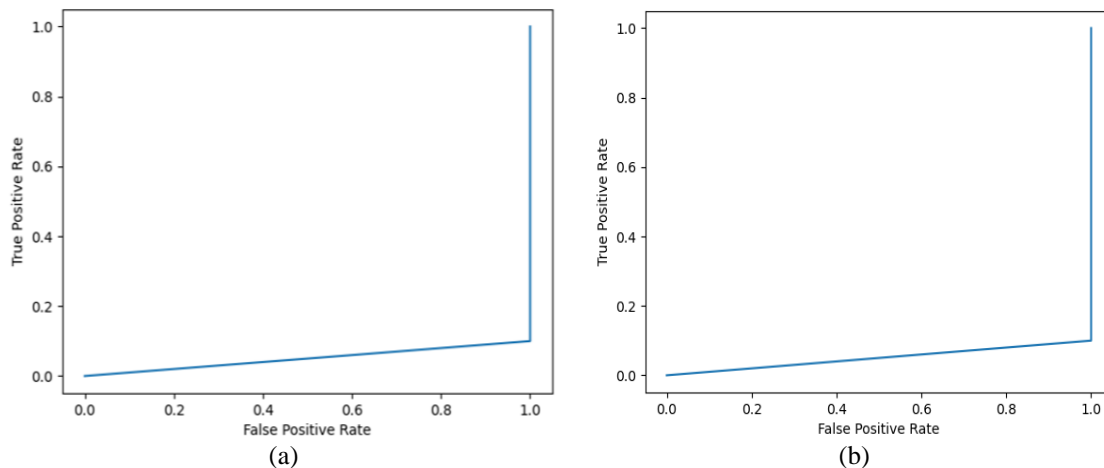


Figure 7. ROC representation (a) exudates vs MAs and (b) haemorrhage vs MAs

3.5. Comparison and interpretation

To solve the problem of detecting and classifying retinal lesions such as exudates, haemorrhages, and MAs, we employed a ResNet-18 architecture-based CNN. Our method produced encouraging outcomes, with test data accuracy coming in at 93.34%. Sensitivity and specificity measures demonstrated the network's proficiency in properly classifying the various illnesses, even when the fundus images comprised numerous

abnormalities. A noteworthy finding was that the network showed a modest bias towards identifying MAs, which are common across all classes. This shows that the network is flexible enough to accommodate different retinal illnesses, offering a foundation for eventual multi-class detection. In Table 2, we observe how the proposed approach stacks up against existing ones.

Table 2. Comparison between the proposed method and the existing techniques

Authors	Retinal lesions detected	Fundus images/datasets used	Methodology	Performance measures
Kassani <i>et al.</i> [41]	No DR, mild DR, moderate DR, severe DR, proliferative DR	APTOS 2019	Modified Xception	Accuracy: 83.09% Sensitivity: 88.24% Specificity: 87.00%
Zago <i>et al.</i> [42]	No DR and DR images	Messidor, Kaggle, IDRiD, DDR, DIARETDB0	VGG 16	Accuracy: 91.2% Sensitivity: 90.0% Specificity: 87.00%
Liu <i>et al.</i> [43]	No DR, mild DR, moderate DR, severe DR, proliferative DR	Kaggle	EfficientNetB4, EfficientNetB5, NASNetLarge, InceptionResNetV2, Xception	Accuracy: 86.34% Sensitivity: 98.77% Specificity: 74.76%
Le <i>et al.</i> [44]	No DR and DR images	Real-time clinical dataset	VGG 16	Accuracy: 87.27% Sensitivity: 83.76% Specificity: 90.82%
Proposed work	MAs, hemorrhages and exudates	Real-time clinical dataset	Custom ResNet 18	Accuracy: 93.34% Sensitivity (MAs): 100% Sensitivity (Hemorrhages and Exudates): 90% Specificity (MAs): 90% Specificity (Hemorrhages and Exudates): 100%

4. CONCLUSION

Images from exudate, haemorrhage, and MAs were included in the data used for this study. Some of the images of exudate and haemorrhage, however, also featured MAs. Some images labelled as haemorrhage also contained exudates. In view of what has been mentioned so far, haemorrhage was given the highest priority, and as a result, haemorrhage images containing exudates were identified correctly as haemorrhage. The network was tuned to achieve this behaviour. However, in the 1 misclassified haemorrhage image, many MAs tilted the network in favour of identifying and classifying the image as MA. Overall, the network is slightly biased to detect MAs as they are present in almost all images in every class. However, the network can identify other pathologies even in the presence of MAs with a high level of confidence as shown in the results. The dataset for this research was small compared to deep learning standards. Therefore, we expect the network to overfit this distribution and data characteristics. However, the proposed method's performance to solve this problem is encouraging and can be expanded upon.

The performance of the current method can be tested on an unseen dataset with different characteristics from that of the training set. This would determine the real-world performance of the proposed method. To overcome the data shortage problem, synthetic images with augmentations could be used to enlarge the dataset size. However, it would be beneficial to gather more annotated data. The outcomes of this investigation have important significance for the early detection and treatment of retinal disorders. The created model demonstrates how AI might help ophthalmologists identify key disorders, allowing for prompt therapies and averting permanent vision damage. Future research should entail the validation of the suggested strategy on bigger and more varied datasets, assuring robustness and generalizability because the study used a very small private dataset. Exploring data augmentation methods and including synthetic pictures might produce an enhanced training dataset, which would help to lessen the restrictions of data scarcity. The network's efficiency could be enhanced by exploring alternative loss functions and using transformer-based topologies. The beneficial significance of the suggested AI model might be further increased by expanding the existing classification technique to include segmentation and detection tasks since this would give detailed insights into the location and severity of retinal defects.

REFERENCES




- [1] J. Krause *et al.*, "Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy," *Ophthalmology*, vol. 125, no. 8, pp. 1264–1272, Aug. 2018, doi: 10.1016/j.ophtha.2018.01.034.
- [2] W. Kusakunniran, Q. Wu, P. Ritthipravat, and J. Zhang, "Hard exudates segmentation based on learned initial seeds and iterative graph cut," *Computer Methods and Programs in Biomedicine*, vol. 158, pp. 173–183, May 2018, doi: 10.1016/j.cmpb.2018.02.011.
- [3] S. Wang *et al.*, "Localizing microaneurysms in fundus images through singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 990–1002, May 2017, doi: 10.1109/TBME.2016.2585344.

- [4] J. H. Tan *et al.*, “Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network,” *Information Sciences*, vol. 420, pp. 66–76, Dec. 2017, doi: 10.1016/j.ins.2017.08.050.
- [5] M. M. Fraz, M. Badar, A. W. Malik, and S. A. Barman, “Computational methods for exudates detection and macular edema estimation in retinal images: a survey,” *Archives of Computational Methods in Engineering*, vol. 26, no. 4, pp. 1193–1220, Sep. 2019, doi: 10.1007/s11831-018-9281-4.
- [6] C. Lam, D. Yi, M. Guo, and T. Lindsey, “Automated detection of diabetic retinopathy using deep learning,” in *AMIA Jt Summits Transl Sci Proc*, 2018, pp. 147–155.
- [7] E. H. Ali, H. A. Jaber, and N. N. Kadhim, “New algorithm for localization of iris recognition using deep learning neural networks,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 1, pp. 110–119, Jan. 2023, doi: 10.11591/ijeecs.v29.i1.pp110-119.
- [8] A. Saini, A. R. Khaparde, S. Kumari, S. Shamsher, J. Joteeswaran, and S. Kadry, “An investigation of machine learning techniques in speech emotion recognition,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 2, pp. 875–882, Feb. 2023, doi: 10.11591/ijeecs.v29.i2.pp875-882.
- [9] A. Bajwa, N. Nosheen, K. I. Talpur, and S. Akram, “A prospective study on diabetic retinopathy detection based on modify convolutional neural network using fundus images at sindh institute of ophthalmology and visual sciences,” *Diagnostics*, vol. 13, no. 3, p. 393, Jan. 2023, doi: 10.3390/diagnostics13030393.
- [10] R. Ali, R. C. Hardie, B. N. Narayanan, and T. M. Kebede, “IMNets: deep learning using an incremental modular network synthesis approach for medical imaging applications,” *Applied Sciences (Switzerland)*, vol. 12, no. 11, p. 5500, May 2022, doi: 10.3390/app12115500.
- [11] M. A. Rodriguez, H. AlMarzouqi, and P. Liatsis, “Multi-label retinal disease classification using transformers,” *IEEE Journal of Biomedical and Health Informatics*, pp. 1–13, 2022, doi: 10.1109/JBHI.2022.3214086.
- [12] M. Majumder and B. Dam, “A clinical study on optic neuritis in a patient of paediatrics age group attending O . P . D . in a tertiary care hospital in assam,” *International Journal of Science and Reserch*, vol. 5, no. 2018, pp. 2013–2016, 2020, doi: 10.9790/0853-1706123444.
- [13] X. Liu *et al.*, “Deep learning to detect OCT-derived diabetic macular edema from color retinal photographs: a multicenter validation study,” *Ophthalmology Retina*, vol. 6, no. 5, pp. 398–410, May 2022, doi: 10.1016/j.oret.2021.12.021.
- [14] A. Patil and C. Chakravorty, “Detection of hard exudate using retinal optical coherence tomography (OCT) images,” *Global Transitions Proceedings*, vol. 2, no. 2, pp. 566–570, Nov. 2021, doi: 10.1016/j.gltp.2021.08.067.
- [15] N. Eladawi *et al.*, “Classification of retinal diseases based on OCT images,” *Frontiers in Bioscience - Landmark*, vol. 23, no. 2, pp. 247–264, 2018, doi: 10.2741/4589.
- [16] D. Wang and L. Wang, “On OCT image classification via deep learning,” *IEEE Photonics Journal*, vol. 11, no. 5, pp. 1–14, Oct. 2019, doi: 10.1109/JPHOT.2019.2934484.
- [17] T. Tsuji *et al.*, “Classification of optical coherence tomography images using a capsule network,” *BMC Ophthalmology*, vol. 20, no. 1, p. 114, Dec. 2020, doi: 10.1186/s12886-020-01382-4.
- [18] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [19] X. Li, X. Hu, L. Yu, L. Zhu, C. W. Fu, and P. A. Heng, “CANet: cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 5, pp. 1483–1493, May 2020, doi: 10.1109/TMI.2019.2951844.
- [20] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, “Diabetic retinopathy detection through deep learning techniques: a review,” *Informatics in Medicine Unlocked*, vol. 20, p. 100377, 2020, doi: 10.1016/j.imu.2020.100377.
- [21] S. A. Kumar, J. S. Kumar, P. Neraniki, K. A. Yadav, and S. S. Basha, “A review of deep learning techniques on fundus images for detecting diabetic retinopathy on public datasets,” *Journal of Innovative Image Processing*, vol. 4, no. 4, pp. 226–236, Dec. 2022, doi: 10.36548/jiip.2022.4.002.
- [22] M. Mateen, J. Wen, Nasrullah, S. Song, and Z. Huang, “Fundus image classification using VGG-19 architecture with PCA and SVD,” *Symmetry*, vol. 11, no. 1, p. 1, Dec. 2019, doi: 10.3390/sym11010001.
- [23] J. Wang, J. Luo, B. Liu, R. Feng, L. Lu, and H. Zou, “Automated diabetic retinopathy grading and lesion detection based on the modified R-FCN object-detection algorithm,” *IET Computer Vision*, vol. 14, no. 1, pp. 1–8, Feb. 2020, doi: 10.1049/iet-cvi.2018.5508.
- [24] K. Kangra and J. Singh, “Comparative analysis of predictive machine learning algorithms for diabetes mellitus,” *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 3, pp. 1728–1737, Jun. 2023, doi: 10.11591/eei.v12i3.4412.
- [25] R. C. Wihandika, P. P. Adikara, S. Adinugroho, Y. A. Sari, and F. Utamingrum, “Retinal blood vessel segmentation using multiple line operator-based methods,” *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 3, pp. 1696–1705, Jun. 2022, doi: 10.11591/eei.v11i3.3026.
- [26] Y. S. Devi and S. P. Kumar, “A deep transfer learning approach for identification of diabetic retinopathy using data augmentation,” *IAES International Journal of Artificial Intelligence*, vol. 11, no. 4, pp. 1287–1296, Dec. 2022, doi: 10.11591/ijai.v11.i4.pp1287-1296.
- [27] A. Rodtook and S. Chucherd, “Optic disc localization using graph traversal algorithm along blood vessel in polar retinal image,” *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 6, pp. 3301–3312, Dec. 2022, doi: 10.11591/eei.v11i6.3618.
- [28] M. Al-Smadi, M. Hammad, Q. B. Baker, and S. A. Al-Zboon, “A transfer learning with deep neural network approach for diabetic retinopathy classification,” *International Journal of Electrical and Computer Engineering*, vol. 11, no. 4, pp. 3492–3501, Aug. 2021, doi: 10.11591/ijece.v11i4.pp3492-3501.
- [29] C. L. Lin and K. C. Wu, “Development of revised ResNet-50 for diabetic retinopathy detection,” *BMC Bioinformatics*, vol. 24, no. 1, p. 157, Apr. 2023, doi: 10.1186/s12859-023-05293-1.
- [30] N. Tsiknakis *et al.*, “Deep learning for diabetic retinopathy detection and classification based on fundus images: a review,” *Computers in Biology and Medicine*, vol. 135, p. 104599, Aug. 2021, doi: 10.1016/j.compbiomed.2021.104599.
- [31] S. A. Kumar and J. S. Kumar, “A review on recent developments for the retinal vessel segmentation methodologies and exudate detection in fundus images using deep learning algorithms,” in *Advances in Intelligent Systems and Computing*, vol. 1108 AISC, 2020, pp. 1363–1370.
- [32] P. M. Krishnegowda and K. Ganesan, “Modelling on-demand preprocessing framework towards practical approach in clinical analysis of diabetic retinopathy,” *International Journal of Electrical and Computer Engineering*, vol. 12, no. 1, pp. 585–595, Feb. 2022, doi: 10.11591/ijece.v12i1.pp585-595.
- [33] A. Pradeep and X. F. Joseph, “Binary operation based hard exudate detection and fuzzy based classification in diabetic retinal fundus images for real time diagnosis applications,” *International Journal of Electrical and Computer Engineering*, vol. 10, no. 3, pp. 2305–2312, Jun. 2020, doi: 10.11591/ijece.v10i3.pp2305-2312.




- [34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2016, vol. 2016-Decem, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [35] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, Apr. 2017, doi: 10.1109/TPAMI.2016.2572683.
- [36] K. S. Kim and Y. S. Choi, "Hyadam: a new adam-based hybrid optimization algorithm for convolution neural networks," *Sensors*, vol. 21, no. 12, p. 4054, Jun. 2021, doi: 10.3390/s21124054.
- [37] M. Zheng *et al.*, "Weakly supervised contrastive learning," in *Proceedings of the IEEE International Conference on Computer Vision*, Oct. 2021, pp. 10022–10031, doi: 10.1109/ICCV48922.2021.00989.
- [38] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9906 LNCS, 2016, pp. 694–711.
- [39] A. Vaswani *et al.*, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 2017-December, pp. 5999–6009, 2017.
- [40] G. Plastiras, C. Kyrkou, and T. Theocharides, "Efficient convnet-based object detection for unmanned aerial vehicles by selective tile processing," in *ACM International Conference Proceeding Series*, Sep. 2018, pp. 1–6, doi: 10.1145/3243394.3243692.
- [41] S. H. Kassani, P. H. Kassani, R. Khazaeinezhad, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Diabetic retinopathy classification using a modified Xception architecture," in *2019 IEEE 19th International Symposium on Signal Processing and Information Technology, ISSPIT 2019*, Dec. 2019, pp. 1–6, doi: 10.1109/ISSPIT47144.2019.9001846.
- [42] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. Teatini Salles, "Diabetic retinopathy detection using red lesion localization and convolutional neural networks," *Computers in Biology and Medicine*, vol. 116, p. 103537, Jan. 2020, doi: 10.1016/j.compbiomed.2019.103537.
- [43] H. Liu, K. Yue, S. Cheng, C. Pan, J. Sun, and W. Li, "Hybrid model structure for diabetic retinopathy classification," *Journal of Healthcare Engineering*, vol. 2020, pp. 1–9, Oct. 2020, doi: 10.1155/2020/8840174.
- [44] D. Le *et al.*, "Transfer learning for automated octa detection of diabetic retinopathy," *Translational Vision Science and Technology*, vol. 9, no. 2, pp. 1–9, Jul. 2020, doi: 10.1167/tvst.9.2.35.

BIOGRAPHIES OF AUTHORS



Silpa Ajith Kumar    is an Assistant Professor in the Department of Electronics and Instrumentation Engineering, Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India. She holds an M.E., degree in Applied Electronics. Her research areas are machine learning, deep learning, medical image analysis, biomedical instrumentation and automation in process control. She is pursuing a Ph.D. in the area of biomedical image processing. She has authored or co-authored over 10 internationally refereed articles her research interests include computer vision, biomedical instrumentation, image processing, IoT, medical image analysis and VLSI design. She can be contacted at email: silpaajithkumar@gmail.com.



Dr. James Satheesh Kumar    received his B.E., from Madurai Kamaraj University, India, M.Tech., from NIT Trichy, India and Ph.D. degree from Karunya Institute of Technology and Sciences (Deemed University) India, in 1999, 2005, and 2014, respectively. He started his career as a teaching faculty during the year 2001. He is currently associated with Dayananda Sagar College of Engineering, Bengaluru, India as an Associate Professor. He has authored or co-authored over 25 internationally refereed articles. His current research interests include biomedical imaging/control engineering/process control. Presently he is guiding research scholars from Visvesvaraya Technological University in interdisciplinary domain. He can be contacted at email: jsatheeshngl@gmail.com.