Enhancing Early Detection of Diabetic Retinopathy Through the Integration of Deep Learning Models and Explainable Artificial Intelligence

Prof. Supriya. S. Walunj Author[1], Prof. Nita S. Salunke Author[2], Prof. Sandip F. Jadhav Author[3]
[1,2,3] *Member, S. P. I. T. Polytechnic, Kurund,*

Abstract: **People are susceptible to a wide range of illnesses, some of which have no clear cures. Diabetic retinopathy (DR) is a condition that can damage one or both of the eyes in humans. It can cause vision problems and finally lead to permanent blindness. It is among those many intricacies. Therefore, early detection of DR can significantly reduce the risk of vision impairment by appropriate treatment and necessary precautions. The primary aim of this study is to leverage cutting-edge models trained on diverse image datasets and propose a CNN model that demonstrates comparable performance. Specifically, we employ transfer learning models such as DenseNet121, Xception, Resnet50, VGG16, VGG19, and InceptionV3, and machine learning models such as SVM, and neural network models like (RNN) for binary and multi-class classification. It has been shown that the proposed approach of multi-label classification with softmax functions and categorical cross-entropy works more effectively, yielding perfect accuracy, precision, and recall values. However, our proposed CNN model shows superior performance, achieving an accuracy of 95.27% on this dataset, surpassing the state-of-the-art Xception model. Moreover, for single- label (binary classifications), our proposed model achieved perfect accuracy as well. Through exploration of these advances, our objective is to provide a comprehensive overview of the leading methods for the early detection of DR. The aim is to discuss the challenges associated with these methods and highlight potential enhancements. In essence, this paper provides a high-level perspective on the integration of deep learning techniques and machine learning models, coupled with explainable artificial intelligence (XAI) and gradient-weighted class activation mapping (Grad-CAM). We present insights into their respective accuracy and the challenges they face. We anticipate that these insights will prove valuable to researchers and practitioners in the field Our goal is that this thorough analysis and comparison of models will guide and motivate further research initiatives, which will ultimately improve illness detection in medical imaging and benefit medical practitioners.**

Index Terms**: CNN, Xception, inception, Grad-CAM. Diabetic retinopathy, transfer learning.**

I. INTRODUCTION

Diabetic retinopathy (DR) is a common condition for which there is unfortunately no known cure, but some methods have been developed to reduce its symptoms. It is quite concerning that DR is so common, and the scientific community is increasingly coming to the conclusion that people with diabetes are at a significant and increasing risk of getting DR. As stated in [1], Of the 37 billion incidents of blindness worldwide, 4.8% are caused by this illness. A complicated collection of specific traits and locations in the retinal pictures are used to diagnose DR. The World Health Organization (WHO) estimates that 422 million people worldwide have diabetes, and that the prevalence of the disease has risen from 4.7% to 8.5% globally, as published in [2]. It is projected that by 2030, there will be 191.0 million people with diabetes worldwide, as described in [3]. Currently, DR and diabetic macular edema (DME) have respective prevalence rates of 35% and 7%, according to [4]. Diagnosing diabetic retinopathy in its early stages is exceptionally challenging because microaneurysms, capillary outflow, retinal hemorrhages, and blood vessel breaches are typically concealed at that point. To reduce the burden on ophthalmologists, researchers have developed a computerized technique that can detect and classify anomalies in the retina based on their severity. Nevertheless, the intended outcomes cannot be achieved by just identifying microaneurysms. Thus, KNN, SVMs, and ensemble-based techniques were used to offer more means of differentiating aneurysms and rating DR. Retinal pictures are the main source of information used by computer-aided diagnosis to identify and classify diseases [5]. Nonetheless, the screening services for diabetic retinopathy—whether in developed or developing nations— are uneven and confront persistent obstacles, like the absence of precise recommendations regarding the best screening techniques (e.g., fundus photography versus clinical examination) and the growing cost of setting up and maintaining extensive DR screening programs. As a result, DR is becoming a more significant public health issue, particularly in many middle-class to low-income nations, where access to modern eye care treatments, such as laser and intravitreal therapies, and qualified eye care practitioners can be limited. Moreover, it is essential to acknowledge that the manual procedure employed for conducting eye examinations is not only time-consuming but also a meticulous and labor- intensive endeavor, as highlighted in [6]. It is worth noting that diabetes and its associated complications impose a substantial economic burden on individuals and their families, as well as on national economies, mainly due to the mounting medical expenses and the resultant loss of wages. Consequently, the prospects of vision loss can be significantly reduced, possibly by as much as 90%, through the early identification of diabetes and its ensuing consequences. The chief structural modifications induced by DR manifest within the intricate layers of retinal tissues. The retina, an exceedingly thin layer comprised of nerve cells that possess light sensitivity, spans the posterior wall of the eye and is situated in close proximity to the optic nerve, as expounded upon in [7]. It is here that incoming light is absorbed by optic nerve cells, subsequently undergoing transformation into neural impulses. These impulses are then conveyed by the optic nerves to the brain, and our capacity to perceive and discern visual stimuli hinges upon the intricate processing of these signals by the brain. It is imperative to recognize that impairment of vision can arise from retinal damage triggered by a multitude of medical ailments.

A. RESEARCH OBJECTIVE

DR is a very serious illness that causes its poor victims to have permanently reduced eyesight. Concern over the rising incidence of diabetes, which is fueled by things like longer life expectancies and opulent lives, is growing. People with diabetes and DR need to receive prompt and appropriate care in order to ensure treatment efficacy and cost-effectiveness. However, because symptoms do not appear until the disease has progressed to a later stage, starting treatment is a complex challenge. Therefore, our primary goal is to predict DR early by using a variety of approaches, such as state-of-the-art deep learning and machine learning models. These models accurately determine the stage of the disease based on a retinal image as input. It is important to remember that there are inherent complications when using deep learning in the medical field. In order to make artificial intelligence easier to interpret in the field of health, it is imperative that these models and their results be presented in a way that medical experts can understand. Additionally, we use gradient-weighted class activation mapping (Grad-CAM), an elucidative visualization technique, to create heat maps, guaranteeing that the reasoning behind the model's decision-making can be clearly conveyed We have thoughtfully selected a range of models and subjected them to rigorous comparison, assessing their performance in terms of accuracy, precision, and recall scores. The successful early prediction of DR by a functioning model paves the way for the prompt referral of patients to ophthalmologists for further evaluation and tailored care. The implementation of this kind algorithm could greatly reduce the incidence of DR- related vision loss worldwide.

B. RESEARCH PROBLEM

Together with the proposed model, we use a variety of transfer learning, machine learning, and deep learning techniques in this work, and we perform a comparative analysis. The dataset has certain constraints, notably an uneven distribution of classes, which could potentially affect the outcomes for both established and the proposed models. As a solution, we implemented a binary classification for the proposed model, which yielded outstanding results. Additionally, the scarcity of datasets specifically designed for research required our utilization of the Kaggle competition dataset.

C. CONTRIBUTION AND IMPACT

The contributions of this study can be summarized in the following manner:

- For both DR and No_DR, we provide a brand-new tailored CNN model that achieves an astounding 95.27% classification accuracy. Developments in deep learning models have had a major impact on several fields, but especially on medical image processing, where they have improved jobs like DR diagnosis.

- Our research confirms the effectiveness of our two-pronged strategy, which combines our recently suggested model with a detailed assessment of several transfer learning approaches, including both machine learning and deep learning models. The best model for diagnosing DR may be found thanks to this thorough process, which is backed by a detailed study that assesses the model's performance in a number of different areas using a variety of metrics.
- To improve the quality of retinal fundus images by removing noise and artifacts, the study used a variety of image pre-processing techniques, such as Zoom, Resample, Constant Filling, Random Rotation, and Horizontal and Vertical Flip.
- In addition, an augmentation technique was used to grow the image collection, which led to a significant rise to 28,059 images in the enhanced dataset.
- In addition, various optimizers such as Adam, Nadam, RmsProps, and stochastic gradient descent (SGD) were employed to determine the most suitable optimizer for the model, thus optimizing its performance.
- The utilization of the XAI technique Grad-CAM provided valuable visual insights into the predictions generated by the CNN model, facilitating the identification and highlighting of specific areas within an input image that significantly influence the ultimate classification or decision made by the network.
- This study encompassed the application of both single-label and multi-label classification techniques, utilizing established models such as SVM, RNN, K-nearest neighbors (KNN), DenseNet-121, Xception, ResNet-50, VGG-16, VGG-19, Inception, and our innovative model. It is worth mentioning that the proposed model has shown remarkable accuracy in multi-label classification scenarios. Additionally, it performed exceptionally well in single-label classification tasks with minimal ambiguity. To deepen our understanding of model behavior, we conducted experiments involving both pre-trained and randomly initialized weights in both of these methodologies.
- Using two different datasets, Messidor2 and IDRiD, we further investigated the capabilities of our suggested CNN model, obtaining remarkable accuracies of 93.86% and 91.18%, respectively. These outcomes highlight the suggested model's efficacy and efficiency in detecting diabetic retinopathy.

II RELATED WORKS

DR is a microvascular consequence of diabetes that, if left untreated, can result in blindness. To avoid vision loss, early detection and treatment of DR are critical. The gold standard for DR screening is fundus photography, however, it is time- consuming and requires skilled individuals to interpret the results. In recent years, there has been a sharp increase in interest in the creation of automated DR diagnosis tools. These methods use artificial intelligence (AI) to analyze fundus photos and identify DR symptoms. DR diagnosis systems with AI capabilities have shown remarkable accuracy. On the other hand, the majority of AI-based DR diagnosis algorithms were developed using enormous photo datasets from specialized clinics or research. These datasets may not accurately represent the population and the precision of AI-driven systems that detect retinopathy (DR), in real-world scenarios may vary. Multiple studies have evaluated the accuracy of AI-based DR diagnosis systems in real-world environments revealing that their performance can differ. For example, the study [8] found that an AI-based DR diagnosis system can achieve an accuracy of 90.6% in a real-world setting. However, in [9], research found that the accuracy of an AI-based DR diagnosis system dropped to 85%. Various factors can influence the accuracy of these systems, such as image quality, concurrent eye diseases, and the stage of DR itself. Generally speaking, AI-driven DR diagnosis systems tend to excel at identifying cases, then the early stages of DR. Naive Bayes (NB) is a straightforward machine-learning technique for classification tasks. It operates by estimating the chance of a data point belonging to each class based on the class's prior probabilities and the likelihood of the data point assigned to each class.

With an accuracy of roughly 85%, NB has been employed for DR diagnosis. It is a simple algorithm to implement that is also computationally efficient. However, NB is not very sensitive to changes in the data and can be overfitted.

A dataset of 5,000 fundus images was analyzed in the study; 2,500 of these images were healthy, and the remaining 2,500 showed DR. The accuracy of the NB classifier in the test set was 75.13%. SVM is a popular model that may be applied to both regression and classification. SVM separates the data points into two groups by locating a hyperplane in the data. SVM has been used to detect DR with a reported accuracy of 90%. SVM is more resistant to data changes than NB and thus less prone to overfitting. However, training SVM can be computationally costly, particularly for large datasets. For example, in [10] suggested an SVM-based model for the identification of DR. A separate test set yielded an accuracy of 94.5% for the model, which was trained on a dataset of 10,000 fundus images.

Random Forest (RF) combines forecasts from multiple decision trees to produce a more accurate prediction. RF can be used to find complex correlations in the data and is more resistant to overfitting than individual decision trees (DTs). It has been claimed that RF has a 91% accuracy rate in detecting DR. RF is computationally efficient and reasonably easy to implement. The study [11] demonstrates that RF is useful for DR diagnosis. 89.5% accuracy was attained by the model on a dataset including 200 rental photos. Different phases of dementia can be recognized and categorized using an AI-based model known as the bi-directional long short-term memory (BiLSTM) model. Before using the LSTM model, the researchers used a pre-processing method called multiscale retinex with chromaticity preservation (MSRCP) to improve the contrast and quality of the fundus images. This was done to increase the accuracy of the results and make the images easier for the LSTM model to interpret. The study [12] used the LSTM network to diagnose diseases with a 92% accuracy rate. In a similar vein, the study [13] presented a cutting-edge LSTM model to illustrate the effective use of KOA severity categorization and early detection based on the BiLSTM network.

The algorithm achieved an accuracy of 78.57% in cross-validation and 84.09% in testing. One kind of neural network that does well on image identification tasks is CNN. CNN has been used to diagnose DR with an accuracy rate of above 93%. However, CNNs can be challenging and computationally costly. CNN may frequently be difficult to comprehend. The study [14] used a radial basis function neural network to detect DR with an accuracy of 89.4%. Using a retrained AlexNet CNN, the study [15] obtained 92.5% accuracy for DR diagnosis in another investigation.

In the study [16], they employed a CNN model known as the DenseNet-169 classifier, which achieved an impressive accuracy rate of 90%. The authors utilized two datasets namely the diabetic retinopathy detection 2015 and the APTOS 2019 Blindness Detection dataset obtained from Kaggle. These datasets contained images characterized by black backgrounds, black corners, and varying sizes, and all of these challenges were effectively addressed during the data pre-processing phase. Furthermore, to address the issue of imbalanced data, the authors implemented data augmentation techniques. Subsequently, using the prepared data, their proposed model showed an accuracy of 95% along with a validation accuracy of 90%. It should be noted that the reported accuracy surpasses that of other models used for comparison, including SVM which has an accuracy of 85.6%, regression that has an accuracy of 78%, Decision Tree (DT) with an accuracy of 85.1%, and KNN with an accuracy of 55.17%.

Interestingly, the VGG model was the most accurate, with an accuracy rate of 67% and an AUC (Area Under the Curve) of 0.67, making it the most successful model. The researchers also experimented with several image pre-processing methods and included various variables, such attention maps, to better optimize the neural networks. Similarly, in the research reported in [18], a machine learning approach was employed to anticipate whether an image contains any signs of DR, using data sourced from the UCI dataset. In this investigation, the sigmoid activation function was chosen for its soft-switching capabilities and straightforward derivative. The study reported promising results in the field of DR diagnosis, indicating the potential effectiveness of the machine learning approach. Likewise, a double contribution was provided in the study [19]. First, in contrast to conventional feature extraction-based techniques, they demonstrated exceptional performance by introducing a specific neural network architecture created for the goal of DR image recognition. Furthermore, their suggested algorithm's total efficiency was increased by the application of data augmentation. By correctly detecting micro aneurysms and carefully labeling retinal fundus images, they created a deep CNN in the study [20] that can be used for the early identification of DR. These images are then classified into five different groups. The authors used an augmentation technique in combination with a CNN network based on a 4 × 4 kernel to optimize their implementation.

Their dataset comprised 88,702 images from the Kaggle EyePACS dataset, ensuring an equal distribution of images across all disease stages. Particularly, their model excelled in the low-high DR classification, achieving a sensitivity of 98% and a specificity of 94% for detection in the very early- stage. Additionally, they attained a quadratic weighted kappa score of 0.851 on the Kaggle dataset's severity scoring test set and an AUC (Area Under the Curve) score of 0.844. These impressive results validated the utility of their proposed model for clinical applications.

The research presented in [21] delves into the examination of two categories of DR: non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). The study utilized image datasets sourced from diabetic retinopathy databases to train neural networks

The author [22] highlighted diagnosing diabetic retinopathy (DR) from fundus images is a specialty of deep learning. The goal of this work was to develop a highly precise and effective system for DR stage classification. A unique CNN model named RetNet-10 was created by pre-processing images, combining three datasets, and using data augmentation. RetNet-10 outperformed six state-of- the-art models and achieved an astounding 98.65% testing accuracy in the DR stage classification. RetNet-10 also shows resilience in a variety of k-fold validation tests.

III METHODOLOGY

In this experimental investigation, we improved the accuracy and performance of six pre-trained models from the Keras library by using transfer learning techniques. The models that were chosen were InceptionV3, VGG16, VGG19, ResNet50, Xception, and DenseNet-121. Fig 1 provides an overview of our proposed work with Explainable AI (XAI). We tested a collection of images from the Kaggle Aptos dataset on each of these models. The images, which represented five classes of diabetic retinopathy (DR) — 0 (No DR), 1 (Mild DR), 2 (Moderate DR), 3 (Severe DR), and 4 (Proliferative DR) — were processed by the pre-trained models to classify them into their respective categories. Following the experimentation, we performed a comprehensive comparison of the models to determine which one exhibited the highest levels of accuracy, precision, and recall based on the results obtained. We used Tensorflow ML framework and the version is 2.14.0. Additionally, we incorporated Scikit-learn and Keras into our methodology for training and generating valuable results.

Experimental Setup Specifications: The study employed an AMD Ryzen 7 5800H processor with 8 cores, 16 threads, and a 3.2 GHz clock speed. The system is equipped with 16 GB of RAM and a 512GB m.2 Gen3 SSD for all experiments. Additionally, it features an Nvidia GeForce RTX 3060 GPU with 6 GB of video RAM (VRAM). The setup utilizes the Nvidia Cuda Toolkit version 12.2, TensorFlow 2.14.0, and Jupyter Notebook version 7.0.4 as the integrated development environment (IDE).

Data Processing: This study used the 'APTOS 2019 Blindness Detection' dataset, sourced from Kaggle, which encompasses 3,662 high-resolution color fundus retinal images with associated tags. The dataset is publicly available in the link https://www.kaggle.com/competitions/aptos2019-blindness-detection/data. Since these images were collected with multiple cameras over an extended period of time from different clinics, there is a great deal of variation and inconsistency in the aspect ratio, image quality, and other aspects of the images, as highlighted in a study in [24].

Within this dataset, each of the five classes corresponds to a distinct stage of the disease, categorized on a severity scale ranging from 0 to 4, as elucidated in [16]. Specifically, among the 3,662 ocular images, there are 1,805 images representing the absence of dry eye disease (DR), 999 images with moderate manifestations, 370 images indicating mild severity, 295 images reflecting proliferative conditions, and 193 images depicting severe stages of the disease. The description of different stages of diabetic retinopathy is given below.

- No diabetic retinopathy: This state denotes a completely healthy eye free of retinal damage, as well as the absence of any type of diabetic retinopathy.
- Mild diabetic retinopathy- At this early stage of diabetic retinopathy, there is a noticeable expansion of the blood vessels within the eye's retina, a phenomenon referred to as microaneurysms. In addition, a small amount of fluid can seep into the retina, resulting in enlargement of the macula, a critical region of the retina located at its center.
- Moderate diabetic retinopathy - In the second stage of diabetic retinopathy, inflammation affects the blood vessels, impeding the normal blood supply to the retina. This blockage leads to the entrapment of essential nutrients and other supplies, causing the macula to become obscured by a buildup of blood and other fluids.Severe diabetic retinopathy- This stage is characterized by more intense vascular occlusion, affecting a larger portion of the retinal blood vessels

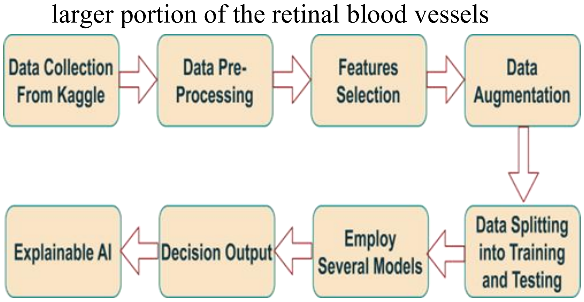


FIGURE 1. Top-level overview of our proposed work with XAI

This extensive occlusion results in reduced blood flow to various parts of the retina, signaling the body to initiate the development of new blood vessels.

- Severe diabetic retinopathy: This stage affects a greater percentage of the retinal blood vessels and is distinguished by more severe vascular occlusion. The body starts to create new blood vessels as a result of the diminished blood supply to different areas of the retina caused by this widespread blockage.
- Proliferative diabetic retinopathy- Representing the most severe stage of diabetic retinopathy, this condition involves excessive growth of blood vessels within the retina. Weaker blood vessels at this stage are prone to fluid leakage, leading to various vision-related complications, including blurriness, narrowing of the field of vision, and, in severe cases, potential blindness.

Image Transformation: Edge detection, semantic seg- mentation, and grayscale image colorization are among the numerous applications based on image transformation tasks, as discussed in [25]. These tasks involve processing images where the input and output are visual data. Pre- processing is necessary to prepare image data for model input, as highlighted in the work in [26]. The pre-processing of images not only streamlines the time required for model training but also expedites the model's inference process.

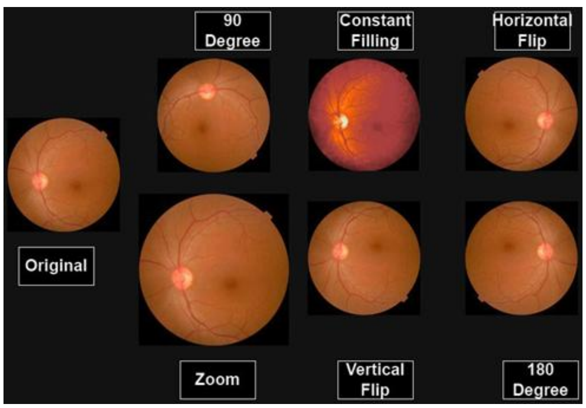
Resizing input images, for example, may not impact the model's performance, however it can significantly reduce the time needed for training, particularly when dealing with excessively large images. To select the most appropriate pre-processing techniques for enhancing model performance, a comprehensive understanding of the problem, the collected data, and the production environment is indispensable. In our study, we used a range of image manipulation methods to achieve the goal.

i) Image flipping: Image flipping involves the rotation of a picture along either a horizontal or vertical axis. This technique introduces randomness in the flipping process, allowing the model to learn to recognize and correctly label flipped images. As a result, when the model encounters a horizontally or vertically flipped image, even if it pertains to a challenging case such as a '3 - Severe DR' classification, it can identify and label the image accurately, ensuring there are no irregularities or inconsistencies.

- Horizontal flipping: Horizontal flipping is reversing an image's full row and column pixels horizontally (left to right) where the flipping happens on the vertical axis.
- Vertical flipping: Vertical flipping involves reversing the order of pixels in each row and column, creating an upside-down version of the original image as the flipping action occurs along the horizontal axis.
- Random Rotation: The random rotation augmentation method introduces variability by randomly rotating images in a clockwise

direction. This rotation can range from 0 to 360 degrees, ensuring a diverse set of image orientations for training and evaluation.

ii) Resample=Image.LANCZOS: Employing the Lanczos interpolation method yields a more refined and sharper image as opposed to a blurry one When it comes to preserving intricate details, reducing aliasing effects, and minimizing artifacts, Lanczos interpolation is con- sidered the most optimal choice This method involves filtering the input samples from the original image through the Lanczos kernel, effectively reconstructing the image. However, it is important to note that along the sharper edges of the resulting image, there may be a presence of dark or bright halos, contributing to the perceived image sharpness.



iii) Image Zoom: This technique involves magnifying an image to make its features more visible and clear. By zooming in, unnecessary regions can be removed, allowing the model to focus on the pertinent areas of interest. Focusing on particular areas of the image becomes crucial when it comes to accurately classifying the extent of eye injury. Since each pixel's value affects model training during prediction, more accurate classifications result from reducing the impact of pixels outside the image boundary.

iv)Constant Filling: To ensure a clean and uncluttered image classification process, it is imperative to eliminate the unnecessary black regions surrounding the retinal images. This adjustment is crucial because dark areas in the images could potentially impact the model's predictions, and it is essential to exclude them from consideration.

Dataset Augmentation: In the Aptos datasets, we have addressed the issue of imbalanced labels by implementing data augmentation. The distribution of images across dif- ferent labels is uneven, with 'No DR' having the highest number of images and 'Moderate' being the second highest. Conversely, the 'Severe' label has a notably low count of 193 images. To mitigate this imbalance, we applied various data augmentation techniques, encompassing six types, including Random Rotation (0 degrees to 360 degrees), Horizontal Flip, Vertical Flip, Resample, Constant filling, and Zoom. Fig. 2 shows all types of augmentation processes of the dataset. Following augmentation, we also increased the image for No DR 3 times, Mild 15 times, Moderate 6 times, Severe 30 times and PDR 16 times so that after augmentation, each of the label images will be almost the same length, and finally, the total number of images in our dataset increased to 28,059.

VI CONCLUSION

The goal of this study is to develop a model for the early diagnosis of diabetic retinopathy that can classify images based on the disease's development. In this study, we have used high-resolution color fundus retinal pictures. Since a wide range of fundus images were present, there were some imbalances and challenges present as well. However, a variety of image modification and data augmentation techniques have been used in this study:

Image flipping (Horizontal, Vertical and Random Rotation), Image resampling, Image zooming and Constant Filling. Additionally, the number of images has been increased so that upon augmentation, every label image would have almost the same length. As a result, the dataset had 28,059 images overall and went through several cycles to fine-tune the pre-processing phases to produce the greatest possible image quality. The dataset was augmented and several pre-processing techniques have also been applied. In this work, 1928 pictures were set aside for validation and a total of 28059 photos were used for training. Transfer learning models such as DenseNet121, Xception, Resnet50, VGG16, VGG19, and InceptionV3, machine learning models such as SVM, and neural network models like RNN have been employed for binary and multi-class classification followed by a comprehensive comparison among them. Two different image sizes have been used for this study, for the 32 × 32 image size, the average time per epoch was roughly 20 seconds, and for the 16 × 16 image size, the average time per epoch was only 10 seconds.

Our future plan is to develop a well-balanced dataset, particularly focusing on regions with a high prevalence of diabetes like in Bangladesh. Moreover, we aim to create a platform allowing users to upload images for real-time output, facilitating immediate decision-making for this disease.

REFERENCE

[1] Kazi Ahnaf Alavee1, Mehedi Hasan1, Abu Hasnayen Zillanee1, Moin Mostakim1, Jia Uddin2, Eduardo Silva Alvarado3,4,5, Isabel De La Torre Diez6, Imran Ashraf7, (Member, Ieee), And Md Abdus Samad7, (Member, Ieee) "Enhancing Early Detection Of Diabetic Retinopathy Through The Integration Of Deep Learning Models And Explainable Artificial Intelligence" 27 May 2024, date of current version 3 June 2024.

[2] M. Rahman, A. Kamrul-Hasan, M. Islam, A. Hasan, F. Chowdhury,O. Miah, M. Islam, S. Wadud, and A. Akhanda, ''Frequency and risk factors of diabetic retinopathy among patients with type 2 diabetes mellitus: A single-center study from Bangladesh,'' *Mymensingh Med. J.*, vol. 29, no. 4, pp. 807–814, 2020.

[3] G. Roglic, *Global Report on Diabetes*. Geneva, Switzerland: World Health Organization, 2016.

[4] D. S. W. Ting, G. C. M. Cheung, and T. Y. Wong, ''Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges: A review,'' *Clin. Experim. Ophthalmol.*, vol. 44, no. 4, pp. 260–277, May 2016.

[5] D. S. Dhoot, K. Baker, N. Saroj, R. Vitti, A. J. Berliner, C. Metzig, D. Thompson, and R. P. Singh, ''Baseline factors affecting changes in diabetic retinopathy severity scale score after intravitreal aflibercept or laser for diabetic macular edema,'' *Ophthalmology*, vol. 125, no. 1, pp. 51–56, Jan. 2018.

[6] F. Nawaz, M. Ramzan, K. Mehmood, H. U. Khan, S. H. Khan, and M. R. Bhutta, ''Early detection of diabetic retinopathy using machine intelligence through deep transfer and representational learning,'' *Comput., Mater. Continua*, vol. 66, no. 2, pp. 1631–1645, 2021.

[7] H. H. Vo and A. Verma, ''New deep neural nets for fine-grained diabetic retinopathy recognition on hybrid color space,'' in *Proc. IEEE Int. Symp. Multimedia (ISM)*, Dec. 2016, pp. 209–215.

[8] H. Zhang and K. Ogasawara, ''Grad-CAM-based explainable artificial intelligence related to medical text processing,'' *Bioengineering*, vol. 10, no. 9, p. 1070, Sep. 2023.

[9] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado, L. Peng, and D. R. Webster, ''Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning,'' *Nature Biomed. Eng.*, vol. 2, no. 3, pp. 158–164, Feb. 2018.

[10] K. Gunasekaran, R. Pitchai, G. K. Chaitanya, D. Selvaraj, S. A. Sheryl, H. S. Almoallim, S. A. Alharbi, S. S. Raghavan, and B. G. Tesemma, ''A deep learning framework for earlier prediction of diabetic retinopathy from fundus photographs,'' *BioMed Res. Int.*, vol. 2022, pp. 1–15, Jun. 2022.

[11] R. P. R. Priya and P. Aruna, ''SVM and neural network-based diagnosis of diabetic retinopathy,'' *Int. J. Comput. Appl.*, vol. 41, no. 1, pp. 6–12, Mar. 2012.

[12] A. M. Nasir, M. A. Abir, Z. T. Sadaf, S. C. Shrestha, C. Shahnaz, and A. Subhana, ''Automatic diabetic retinopathy detection using random forest classifier,'' in *Proc. IEEE Region 10 Symp. (TENSYMP)*, Jul. 2022, pp. 1–6.

[13] S. Xie, Z. Yu, and Z. Lv, ''Multi-disease prediction based on deep learning: A survey,'' *Comput. Model. Eng. Sci.*, vol. 128, no. 2, pp. 489–522, 2021.

[14] A. Qadir, R. Mahum, and S. Aladhadh, ''A robust approach for detection and classification of KOA based on BILSTM network,'' *Comput. Syst. Sci. Eng.*, vol. 47, no. 2, pp. 1365–1384, 2023.

[15] V. V. Kamble and R. D. Kokate, ''Automated diabetic retinopathy detection using radial basis function,'' *Proc. Comput. Sci.*, vol. 167, pp. 799–808, Jan. 2020.

[16] I. Arias-Serrano, P. A. Velásquez-López, L. N. Avila-Briones, F. C. Laurido-Mora, F. Villalba-Meneses, A. Tirado-Espin, J. Cruz-Varela, and D. Almeida-Galárraga, ''Artificial intelligence-based glaucoma and diabetic retinopathy detection using MATLAB—Retrained AlexNet convolutional neural network,'' *F1000Research*, vol. 12, p. 14, Apr. 2024.

[17] G. Mushtaq and F. Siddiqui, ‘‘Detection of diabetic retinopathy using deep learning methodology,’’ *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 1070, no. 1, 2021, Art. no. 012049.
[18] R. Cai. (2020). *Study of Convolutional Neural Networks for Early Detection of Diabetic Retinopathy*. [Online]. Available: https://ysjournal. com/medicine/study-of-convolutional-neural-networks-for-early-detection-of-diabetic-retinopathy/
[19] P. Dwivedi, ‘‘Understanding and coding a ResNet in Keras,’’ 2019. Accessed: May 22, 2024. [Online]. Available: https://towardsdatascience. com/understanding-and-coding-a-resnet-in-keras-446d7ff84d33
[20] K. Xu, D. Feng, and H. Mi, ‘‘Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image,’’ *Molecules*, vol. 22, no. 12, p. 2054, Nov. 2017.
[21] S. M. S. Islam, M. M. Hasan, and S. Abdullah, ‘‘Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images,’’ 2018, *arXiv:1812.10595*.
[22] V. K. Gurani, A. Ranjan, and C. L. Chowdhary, ‘‘Diabetic retinopathy detection using neural network,’’ *Int. J. Innov. Technol. Exploring Eng.*, vol. 8, no. 10, p. 5, 2019.
[23] M. A. K. Raiaan, K. Fatema, I. U. Khan, S. Azam, M. R. U. Rashid, M. S. H. Mukta, M. Jonkman, and F. De Boer, ‘‘A lightweight robust deep learning model gained high accuracy in classifying a wide range of diabetic retinopathy images,’’ *IEEE Access*, vol. 11, pp. 42361–42388, 2023.
[24] I. U. Khan, M. A. K. Raiaan, K. Fatema, S. Azam, R. U. Rashid, S. H. Mukta, M. Jonkman, and F. De Boer, ‘‘A computer-aided diagnostic system to identify diabetic retinopathy, utilizing a modified compact convolutional transformer and low-resolution images to reduce computation time,’’ *Biomedicines*, vol. 11, no. 6, p. 1566, May 2023. [Online]. Available: https://www.mdpi.com/2227-9059/11/6/1566
[25] S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. A. Khan, and W. Jadoon, ‘‘A deep learning ensemble approach for diabetic retinopathy detection,’’ *IEEE Access*, vol. 7, pp. 150530–150539, 2019.
[26] M. Shaban, Z. Ogur, A. Mahmoud, A. Switala, A. Shalaby, H. A. Khalifeh, M. Ghazal, L. Fraiwan, G. Giridharan, H. Sandhu, and A. S. El-Baz, ‘‘A convolutional neural network for the screening and staging of diabetic retinopathy,’’ *PLoS ONE*, vol. 15, no. 6, Jun. 2020, Art. no. e0233514.
[27] C. Affonso, A. L. D. Rossi, F. H. A. Vieira, and A. C. P. de Leon Ferreira, ‘‘Deep learning for biological image classification,’’ *Expert Syst. Appl.*, vol. 85, pp. 114–122, Nov. 2017.
[28] T. Minewiskan. (2022). *Training and Testing Data Sets*. Accessed: Oct. 19, 2023. [Online]. Available: https://learn.microsoft.com/En-Us/ Analysis-Services/Data-Mining/Training-and-Testing-Data-Sets?View= asallproducts-Allversions
[29] A. Sarkar, ‘‘Creating DenseNet 121 with TensorFlow,’’ 2020. Accessed: May 22, 2024. [Online]. Available: https://towardsdatascience.com/ creating-densenet-121-with-tensorflow-edbc08a956d8
[30] B. Pansare, N. Deorukhakar, T. Hajare, P. Nalawade, and P. Nawghare, ‘‘Deep learning for diabetic retinopathy,’’ *Int. J. Recent Adv. Multidisci- plinary Topics*, vol. 2, no. 6, pp. 27–31, Jun. 2021. [Online]. Available: https://journals.ijramt.com/index.php/ijramt/article /view/813
[31] 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 *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2019, pp. 1–6.
[32] S. Majumder and N. Kehtarnavaz, ‘‘Multitasking deep learning model for detection of five stages of diabetic retinopathy,’’ *IEEE Access*, vol. 9, pp. 123220–123230, 2021.
[33] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Deep residual learning for image recognition,’’ in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[34] J. Shaikh, ‘‘Deep learning in the trenches: Understanding inception network from scratch,’’ 2023. Accessed: May 22, 2024. [Online]. Available: https://www.analyticsvidhya.com/blog/2018/10/ understanding-inception-network-from-scratch/
[35] R. Thakur, ‘‘Step by step VGG16

implementation in Keras for beginners,'' 2019. Accessed: May 22, 2024. [Online]. Available: https://towardsdatascience.com/step-by-step-vgg16-implementation-in- keras-for-beginners-a833c686ae6c

[36] J. Brownlee, ''How do convolutional layers work in deep learning neural networks,'' 2020. Accessed: May 22, 2024. [Online]. Avail- able: https://machinelearningmastery.com/convolutional-layers-for-deep- learning-neural-networks/