

Convolutional Neural Networks for Automated Diagnosis of Diabetic Retinopathy in Fundus Images

S. Rama Krishna,¹ Naresh Cherukuri,² Y. Dileep Kumar,³ R. Jayakarthish,⁴ B. Nagarajan,⁵
Allam Balaram,⁶ and G. Divya Jyothi⁷

¹Department of Computer Science & Engineering, GSOT, GITAM (Deemed-to-be-University), Hyderabad, India

²Department of Computer Science and Engineering, Prasad V. Potluri Siddhartha Institute of Technology (Autonomous), Vijayawada, India

³Department of Electronics & Communication Engineering, School of Engineering, Mohan Babu University, Tirupati, Andhra Pradesh, India

⁴Department of Computer Science, Saveetha College of Liberal Arts and Sciences SIMATS, India

⁵Department of Computer Science and Engineering, KPR Institute of Engineering and Technology, Coimbatore, India

⁶Department of Computer Science and Engineering, MLR Institute of Technology, Hyderabad, India

⁷Department of Computer Science and Engineering, MLR Institute of Technology, Dundigal, Medchal, Telangana, India

(Received 14 April 2023; Revised 31 May 2023; Accepted 12 June 2023; Published online 24 August 2023)

Abstract: Diabetic retinopathy (DR), a long-term complication of diabetes, is notoriously hard to detect in its early stages due to the fact that it only shows a subset of symptoms. Standard diagnostic procedures for DR now include optical coherence tomography and digital fundus imaging. If digital fundus images alone could provide a reliable diagnosis, then eliminating the costly optical coherence tomography would be beneficial for all parties involved. Optometrists and their patients will find this useful. Using deep convolutional neural networks (DCNNs), we provide a novel approach to this problem. Our approach deviates from standard DCNN methods by exchanging typical max-pooling layers with fractional max-pooling ones. In order to collect more subtle information for categorization, two such DCNNs, each with a different number of layers, are trained. To establish these limits, we use DCNNs and features extracted from picture metadata to train a support vector machine classifier. In our experiments, we used information from Kaggle's open DR detection database. We fed our model 34,124 training images, 1,000 validation examples, and 53,572 test images to train and test it. Each of the five classes in the proposed DR classifier corresponds to one of the steps in the DR process and is given a numeric value between 0 and 4. Experimental results show a higher identification rate (86.17%) than those found in the existing literature, indicating the suggested strategy may be effective. We have jointly developed an algorithm for machine learning and accompanying software, and we've named it deep retina. Images of the fundus acquired by the typical person using a portable ophthalmoscope may be instantly analyzed using our technology. This technology might be used for self-diagnosis, at-home care, and telemedicine.

Keywords: automated diagnosis of diabetic retinopathy; DCNNs; diabetic retinopathy; self-diagnosis

I. INTRODUCTION

The global cost of diabetes and its complications in 2017 was estimated to be \$850.0 billion. The leading cause of permanent blindness in people is diabetic retinopathy (DR), one of diabetes' most common and serious complications. Diabetics also tend to suffer from this condition at a high rate. In 2017, the International Diabetes Foundation estimated that 451 million individuals throughout the globe were diabetic. This is more than a third of the global population (IDF). There is a high prevalence of vision impairment and blindness in this region. Worldwide, 693 million people will have diabetes by 2045, according to projections. As an added complication, the signs and symptoms of diabetes may be rather subtle, which is why almost half of people with the illness are uninformed of their condition for a considerable period of time [1].

Extremely high blood sugar levels, on the other hand, have been linked to consequences including cardiovascular disease and vision loss through damaging blood vessels and neurons. It may be feasible to arrest the worsening of depression if it is diagnosed and treated at an early stage [1].

According to [1], a thorough retinal examination is required for diagnosing the presence and severity of DR. New aberrant blood vessels form in the retina, a process known as neovascularization, in proliferative diabetic retinopathy (PDR) or neovascular diabetic retinopathy. Damage to these new blood vessels may cause scar tissue and raise the risk of retinal detachment. PDR is a severe type of DR that needs immediate medical attention to avoid irreversible vision loss. According to Hindawi Mobile Information Systems, the main features of PDR include neovascularization and associated consequences, such as retinal detachment and the first symptoms of vitreous hemorrhage, whereas non-proliferative diabetic retinopathy (NPDR) is characterized by exudation and

Corresponding author: S. Rama Krishna (e-mail: rsankara@gitam.edu).

Table I. The several types of diabetic retinopathy categorization

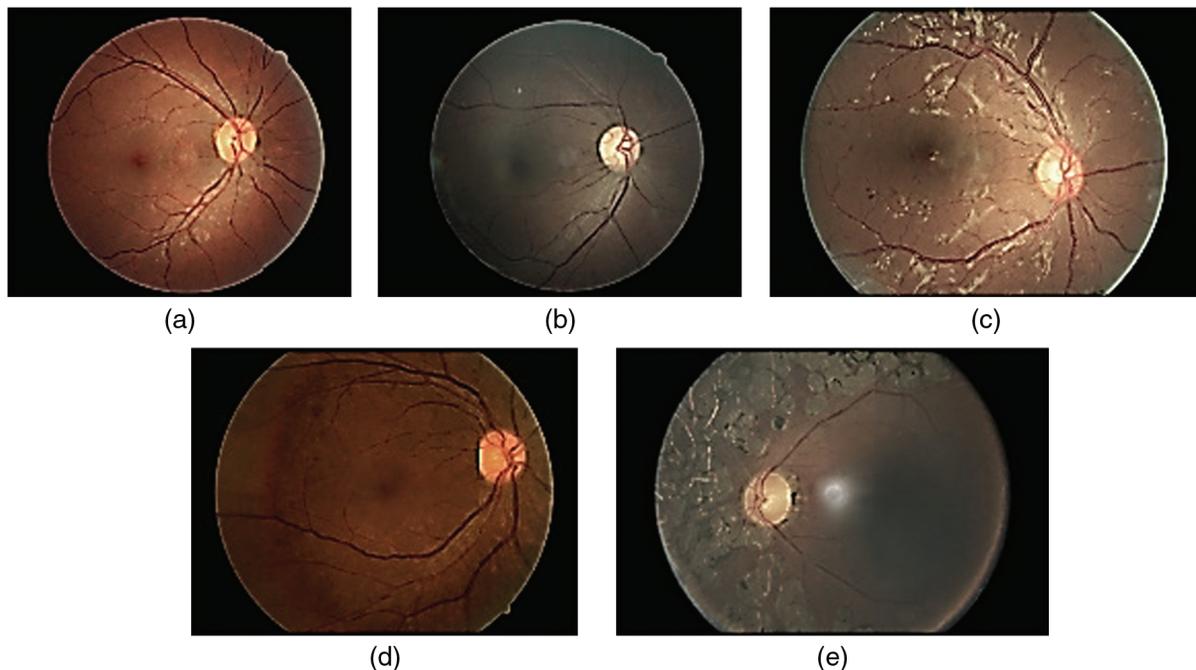
Category	Level	Manifestation
–	1	Neither major ruptures nor minor bleeds
Retinopathy caused by diabetes that does not progress to cancer	2	Conditions such as micro-aneurysms, bleeding, secretions, and venous beading
–	3	Blood leaking into the retina in all four corners
Retinopathy caused by diabetes that has progressed to a proliferative stage	4	Angiogenesis in the optic nerve or retina

ischemia of variable degrees but without retinal detachment or retinal hemorrhage [2]. Examples of microvascular problems related to NPDR include micro-aneurysms, dot-and-blot hemorrhages in the retina, lipid exudates, venous beading change, and intraretinal microvascular abnormalities. Lesions in the NPDR are classified into one of three categories based on their frequency and severity: Micro-aneurysms or a few tiny retinal hemorrhages are the hallmarks of mild NPDR. The presence of more serious micro-aneurysms, bleeding, or soft exudate, but not yet progressing to the point of severe NPDR, defines moderate NPDR. Retinal hemorrhage may affect any or all four quadrants of the eye in patients with severe NPDR, and venous beading can affect at least two of the four. The many DR symptoms are listed in Table I [1]. Those with advanced cases of diabetes and DR are more likely to have micro-aneurysms, hemorrhage, and soft exudates. The greater likelihood of certain events justifies this conclusion. Retinal micro-aneurysms, if left untreated, may cause major consequences due to blood component leakage into surrounding tissue. Yellow deposits of lipids and proteins called soft exudates may occur as a result. Retinal blood vessels are delicate art and easily ruptured, which

may lead to bleeding and subsequent blood accumulation in the retina's layers. These signs and symptoms are clinically significant markers of vascular damage and the severity of DR.

For a long time, manual grading was the only type of DR screening used by ophthalmologists. Automated detection of DR has the potential to be a useful and effective screening approach, especially in light of the rising number of diabetes patients and the state of current technology. Automated identification of DR using convolutional neural networks (CNNs) is an exciting new screening tool. In order to achieve accurate and efficient DR detection from fundus images, recent research has focused on the novel and promising nature of CNN-based techniques. The idea of CNNs serves as the foundation for these techniques. These cutting-edge health informatics solutions efficiently and effectively identify patients via the use of deep learning and image analysis. Therefore, they could help doctors and nurses with early diagnosis and treatment. If doctors and nurses can spend less time on screen thanks to CNNs, more people might benefit. The detecting method has been computerized, allowing for this. The ability of CNNs to sift through massive amounts of data in search of subtle symptoms of illness makes a compelling argument for sophisticated health informatics. This opens the door to earlier diagnosis and treatment, which may avoid permanent vision loss [3]. Current automated retinal image analysis (ARIA) techniques like iGradingM, Retmarker, and Eye Art are primarily intended to detect retinal damage and locate referable DR. Regrettably, ARIAs lack the cognitive abilities necessary to distinguish various DR intensities. For this reason, detecting the little variations across DR levels remains a substantial problem for the medical image processing approach [4]. Figure 1 displays sample fundus images for each kind of lesion.

Retinal cross-sectional pictures with great accuracy may be obtained using optical coherence tomography (OCT). Since it necessitates cutting-edge technology and specialized equipment, it comes with a heftier price tag compared to other imaging modalities. The high price tag of OCT may limit its broad usage and

**Fig. 1.** Illustrations of various diseases on the fundus, including some examples.

accessibility; however, fundus pictures combined with CNN may automatically diagnose DR. Although OCT is more costly than other diagnostic tools, it gives more precise data for spotting DR.

According to [2], it is crucial that medical pictures be processed with absolute accuracy, but it is also important that medical examination equipment be adaptable and easy to transport. Digital photos of the fundus need the patient's cooperation and sitting in front of the fundus camera with the lights turned down or lowered as much as possible. Infrared fundus imaging allows for pinpointing the exact area of interest by having the patient gaze straight into a camera that is lighted from above. Posterior pole detection triggers automated focusing and photo taking by the camera. The use of a flash and an RGB image sensor is still necessary for photography in the visible spectrum. Most digital fundus imagers used in hospitals and clinics are too bulky and costly to be practical. This limits their usefulness as a screening tool, which is why they are not more extensively used [1].

Using CNNs for automated detection of DR in fundus pictures presents a number of important obstacles. A lack of high-quality, diversified datasets makes training models even more challenging. Retinal pictures, due to their complexity and variety (including artifacts and light oscillations), need proper preprocessing. The lack of annotations and the subjective character of diagnosis provide challenges, despite the fact that specialized expertise and experience with the topic are necessary for effective interpretation of fundus pictures. Some people question the transferability of CNN models to other data sets and populations. It will be difficult to develop trustworthy and efficient CNN-based systems for automated DR diagnosis until these problems are resolved. The architecture of cloud computing makes it possible for the system under consideration to be as mobile as needed without sacrificing processing performance. Due to the system's migration to the cloud, where computer resources may have their capacity dynamically enhanced in response to an increase in the amount of work, this is now achievable. Our deep learning algorithm often returns a result to the diagnostic query in about 10 seconds. On theory, this algorithm may be run in the cloud [1].

In addition, the cloud-based architecture facilitates the collection of massive volumes of data. There is a direct correlation between the availability of end devices, such as portable fundus cameras, and the volume of fundus images saved in the cloud. We have been able to save all of the fundus imaging data we have acquired throughout the years, so we can utilize that data effectively. Retraining machine learning models, investigating undiscovered characteristics, and engaging in cross-domain data mining

are just a few of the potential solutions to issues in ophthalmology. Incorporating AI, development of mobile, and big data analytics, this research presents a revolutionary system architecture for DR screening. Our method's technical details are summarized below. Figure 2 illustrates the proposed system in practice. This innovation will drastically improve access to telemedicine in remote areas that have hitherto been neglected [5].

II. LITERATURE REVIEW

Li *et al.* [6] argue that DR may be identified by inspecting the blood vessels of the retina for any signs of leakage (DR). By mapping out a patient's veins and analyzing the wall thickness of their veins, it is possible to identify whether or not the patient has DR. However, due to the fact that fundus pictures often display a great deal of additional information in addition to the vessels themselves, vessel monitoring may be rather challenging. When using a CNN for automated detection of DR in fundus pictures, the quality of the setup is critical for the initial vascular segmentation step. The accuracy of data obtained is dependent on the network architecture used. To avoid overfitting and underfitting, it is possible to adjust GBM's hyper-parameters this research used XGBoost GBM software. XGBoost outscored SVM and Random Forest in our tests; therefore, we implemented it. MXNet (short for "Multi-Expanding Network") was used to build CNN in R. Enjoy the trained neural networks. Several different methods have been proposed for segmenting vessels, including vascular tracking, matching filtering, morphological processing, deformation models, and machine learning.

Saranya and Prabakaran [7] argue that the tracking of blood vessels may be done in a number of different ways [8,9]. To trace the path of the vascular system from beginning to end requires beginning at one point, traveling around in a circle, and continuing this pattern until there are no more blood vessels to follow. The process in question is referred to as vascular tracking. The quality of the first configuration directly affects how well the first vessel can be segmented. At the moment, determining the baseline may either be done manually or automatically, and you have the option to choose either one.

Pratt *et al.* [10] argue that the first kind of adaptive vascular imaging was the use of X-ray angiograms to reconstruct the vasculature. The authors start with an "extrapolation-update" approach, which calculates local vessel trajectories given an initial location and direction within the vessel. This is done after the first set of instructions, and direction has been given. After a vessel

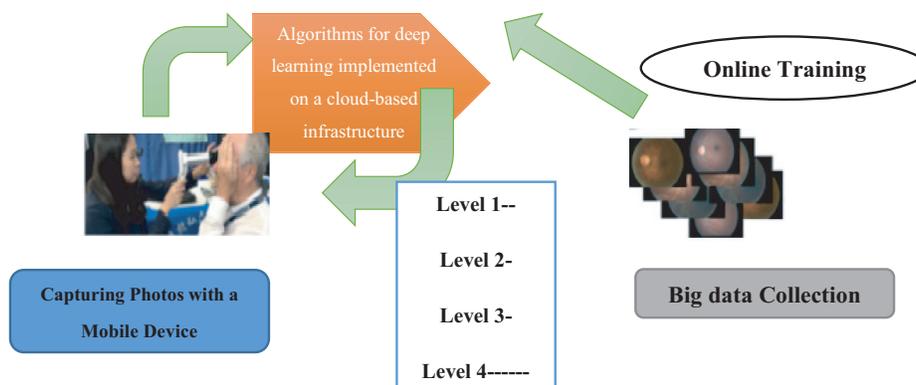


Fig. 2. Basic form of proposed arrangement.

fragment has been traced, it is no longer visible in the picture. This process is performed as many as necessary until the vascular tree is gone [10]. The user is left to decide where the vessels should begin in three dimensions, since the algorithm seems rigid in this respect. These two issues are flaws in the strategy. As a central part of its tracking process, this technique gathers vascular local minimum points, which are often located in the vessel's middle. Post-processing method has morphological consequences for tracking lines of varying thicknesses. This approach relies on the Bayesian methodology and iterative tracking to get the retinal vascular tree. Vascular tracking's potential to provide light on regional factors like artery breadth and flow direction is a major plus. Blood vessel monitoring technologies may lose accuracy due to arterial branching and crossing, making it more difficult to reliably identify and track blood vessels in the retinal vasculature. Because of these anatomical intricacies, it is sometimes difficult to identify specific arteries and their paths. Computational computing technologies, such as CNNs and image processing algorithms, are employed by intelligent health informatics to solve these issues. These techniques aim to alleviate the challenges posed by vascular branching and crossing in order to improve the diagnostic accuracy of automated systems. DR, for example, may now be diagnosed with greater accuracy because of health informatics' improved monitoring and identification of blood vessels. However, arterial branching and crossing may reduce the reliability of vascular tracking identification. One could expect this to happen often.

Wan *et al.* [11] argue that there are a variety of matched-filtering approaches. Vessel detection is especially dependent on the quality of the filter used in matched-filtering methods due to the large number of matched filters typically used throughout the extraction process. Since the grayscale distribution of fundus vessels follows a Gaussian distribution, using the maximum response of filtered photos provides a straightforward method for locating vessel sites. Monitoring vessels often makes use of a multi-scale Gaussian filter strategy because of the large size and width variations that may occur.

Alghamdi [12] argues that certain vascular parameters are considered with the usage of Gaussian filters for vessel tracking. Blood vessels are distinguished by a number of visual cues, including their darker hue than the surrounding tissue, their variable width (between 2 and 10 pixels), and their radial expansion from the optic disc. To identify ships heading in any of the twelve possible directions, we use two-dimensional Gaussian filters. However, a lot of processing power is needed for this technology, and tracking errors might occur when dark lesions have similarities to blood vessels. Segmenting blood vessels in retinal pictures is made easier with the use of a new technique that involves constantly checking to see whether the present position is a vascular point. The method takes into account both macro- and micro-features of the vessel.

Recent years have seen major advancements in the use of CNNs for the automated diagnosis of DR in fundus pictures. In recent years, there has been a proliferation of papers covering this ground. For instance, [13] designed a multi-scale CNN architecture that used both local and global data in order to appropriately classify DR. To enhance feature representation and diagnostic precision, [14] presented a hybrid CNN model that incorporates attention processes. The goal of developing this model was to improve feature representation. To further enhance the accuracy of DR detection and reduce the effect of picture distortions, [15] used an adversarial training technique using a deep residual network. In

order to automate the detection of DR in fundus pictures, researchers have developed CNN-based algorithms, and recent publications highlight this development.

Acharya *et al.* [16] argue that in light of these developments, several investigations toward better filters have been carried out. For a more comprehensive approach, a technique that takes into account the detection of numerous thresholds, must be considered. First, a local vessel cross-section analysis is performed, and then, local bilateral thresholding is used to perform the matching filtering procedure. In an attempt to enhance the extraction of minuscule vessels, a coordinated effort to mass-produce a matching filter in a range of sizes is of utmost importance.

Pao *et al.* [17] argue that methods for Handling Morphological Information through analysis and processing of the underlying structural properties of a binary image, as in morphological processing, makes object segmentation and identification much easier. Through analysis and processing of the underlying structural properties of a binary picture, morphological processing makes object segmentation and identification much easier. The image may serve as a data source to make this a reality. Therefore, the linear and circular components of blood arteries may be selected, allowing for the elimination of superfluous information and the concentration on the core structure. Morphological processing may also fill in any gaps in the picture and smooth out its contour, both of which contribute to a reduction in noise. However, this method overprioritizes structural elements and ignores vessel-specific details.

Kishor *et al.* [18] argue that based on characteristics of vasculature, a morphologically based mathematical method was created to distinguish healthy tissue from potentially dangerous patterns. For the purpose of pinpointing the vascular ridge and refining the boundaries, fundus pictures were analyzed using the curve let transform and morphological reconstruction of multi-scriptural properties. Vessels were segmented and localized using the Curvelet transform, morphological reconstruction of multi-scriptural components, and strongly connected component analysis.

III. METHODOLOGY

A. DATA ENRICHMENT

The vast majority of DR-related annotated picture sets were very small. We make use of data provided by the Kaggle community in our daily operations. It is well known that there are drawbacks to working with very tiny picture datasets and that these datasets need to be artificially enlarged by data augmentation through label-preserving manipulation. Reason being, expanding available data is crucial [19]. Thus, the algorithm's overall performance may increase, and the possibility of overfitting to the picture data may decrease. A portable fundus camera is an essential, versatile, and transportable piece of medical examination equipment primarily used for the automated diagnosis of DR in fundus photos. These cameras are designed to take detailed pictures of the retina, which may then be analyzed using CNNs for automated diagnosis. Screening for and diagnosing DR is made easier with the use of portable fundus cameras. These cameras are convenient for usage in a variety of clinical and field settings due to their small size and mobility. These devices are particularly well-suited for usage in settings with restricted access to specialized imaging equipment, such as mobile clinics and remote places, due to their mobility and flexibility. In order to conduct this study, the tagged dataset will be physically modified. As part of this process, we will enlarge, flip, and invert the dataset. There is a detailed explanation of the

Table II. Parameters for enhancing data

Transformation type	Description
Rotation	0°–360°
Shearing	Randomly with angle between –15 and 15
Flipping	0 or 1
Translation	Randomly with shift between –10 and 10 pixels
Rescaling	Randomly with scale factor between 1/1.6 and 1.6

changes in Table II, and some updated examples of frames may be seen in Fig. 3. This study employs five different forms of transformations, including translation, rotation, flipping, shearing, and resizing. Table II presents the categorized data on the parameters [19].

When employing CNNs for automated identification of DR in fundus pictures, it is necessary to often retrain the underlying machine learning models. Constant refinement and revision are required to discover new features and improve model performance. The models' longevity is ensured by the retraining process, which allows them to adapt to new data and market changes. Many different approaches have been taken to the problem of vessel segmentation, and they all have their merits. Vascular tracking refers to the process of following the course of a vessel in an image. Possible solutions include the use of active contours or area expansion techniques. Enhancing vessel structures using matching filters highlights features specific to vessel architecture. Vessel-like structures may be retrieved based on their form and connectivity using mathematical approaches in morphological processing.

These methods supplement the CNN-based method by adding further preprocessing stages and segmentation possibilities. When combined, these techniques have the potential to enhance vascular segmentation's accuracy and resilience, allowing for a more precise diagnosis of DR. Retrained machine learning models and the use of alternative vascular segmentation techniques have contributed to the development of automated retinopathy diagnosis. This enhancement enhances both the efficiency and accuracy of the system as a whole.

B. CLASSIFICATION OF IMAGES USING CNNs

CNNs are a kind of feed-forward ANNs that have many similarities with biological neural networks. Among the several deep learning architectures, the CNN is particularly prevalent. Because of the tiling pattern used in their representation, individual neurons may

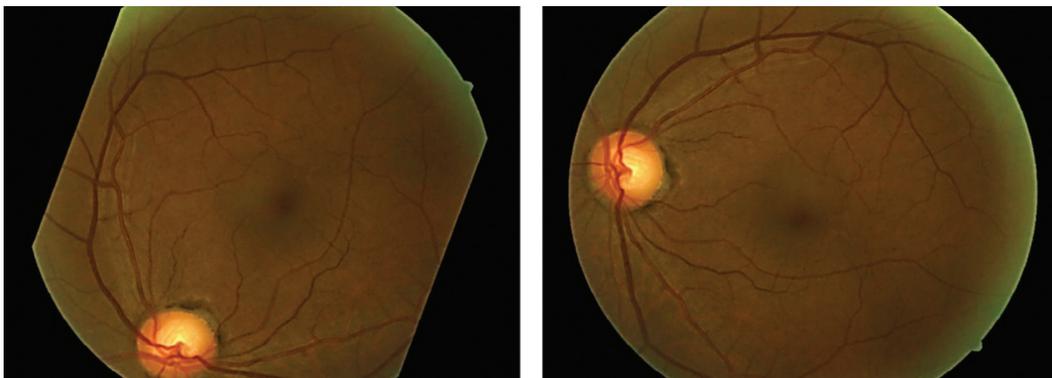
respond to overlapping visual fields. Inspired by biological neural networks, CNNs are an important class of applications that may be thought of as learnable representations [19]. A cloud based architecture enables the storage of huge amount of data (fundus images) over the cloud thus reducing the storage overhead and enable automatic detection of DR in fundus images. This architecture enable researches to efficiently improve the space complexity and reduce the need of high computational resources, and hence processing of images are efficiently managed. The more nodes there are, the more data can be gathered and used to train CNN models. One example of such a consumable is the portable fundus camera. The adoption of a scalable and easily accessible cloud-based infrastructure may increase the precision of DR diagnostics. Many additional formulas have been proposed during the last few years. Nonetheless, the basics remain the same. In order to construct CNNs, several convolutional phases are interleaved with pooling phases. Pooling layers have been found to increase performance by decreasing computation time and boosting spatial and configuration invariance between convolutional layers. To construct the last stages so close to the outputs, we shall use a sequence of interconnected, one-dimensional layers. For the sake of detail, let's say that a feed-forward neural network is conceptualized as a function F that translates data points from an input vector x into the form below [20].

$$f(x) = f_L(\dots f_2(f_1(x_1, w_1), w_2) \dots, w_L) \quad (1)$$

Each function f_i has a set of adjustable parameters w_i and an associated input value x_i (where x_i is the input data). Indicative of a neural network's depth is its value of L . A suitable mapping function, represented by f , is often constructed manually (in terms of its type and sequence), although its parameters may be learned discriminatively from instances. An MNC array provides a more accurate mathematical representation of each element in a CNN. Since our problem can be reduced to a binary classification problem, the loss function for the CNN may be expressed as follows [20]:

$$L_{(w)} = 1/n \times \left(\sum_{i=1}^n l(z_i, f(x_i; w)) \right) \quad (2)$$

According to Gao *et al.* [20], sample I would be designated as Z_i if n was the entire number of samples. Learning may be reframed as instructing a neural network to make decisions that will result in the lowest possible value of a loss function L . As can be seen in Fig. 4, a CNN network is made up of many layers of relatively small neurons. By tiling the data from one set such that it overlaps

**Fig. 3.** Transformed frames.

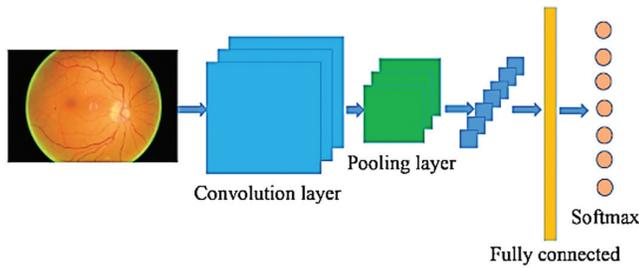


Fig. 4. An excellent example of a CNN design.

with data from another set, a more accurate representation of the original picture is obtained. Each successive layer in a convolutional network is fed a rectangle subset of the neurons in the layer below it. As an added degree of complexity, each convolutional layer might contain many grids, each with its own filter [20]. A pooling layer is added after each convolutional layer and is supplied with subsamples from the previous layer. Finding an average, a maximum, etc., are only two of many possible strategies for this sort of pooling. To characterize the whole input image, a fully connected layer is constructed from the outputs of the previous layer, and this compact feature is then used. Optimization procedures, such as backpropagation and stochastic gradient descent, are used by the network to obtain its peak performance. It is crucial to account for any variations in forward and inverse propagations due to layer type [20].

As part of our research, we have investigated and analyzed a number of distinct CNN architectures, and we have made some of these models accessible to you [2]. The size of the convolution kernel may be anywhere from 1 to 5, and the level of the neural network being evaluated can be anywhere from 9 to 18. We reduced the size of the image such that it has 224 by 224 by 3 pixels in order for it to be processed correctly by the CNN's input size limitations. The completed network architecture for the inquiry is shown in Table III. A pair of probabilities will be returned by the network in response to each input, with the overall probability equaling 1. It's just a straightforward issue of dividing everything up into two groups at this point. In this particular experiment, the neural network is trained using 800 labeled photographs; however, only 200 of those photos are utilized to evaluate its overall performance [2].

Table III. The experiment took advantage of the CNN

Output Shape	Description
$212 \times 212 \times 2$	input
$210 \times 210 \times 30$	30 filter
$208 \times 208 \times 29$	29 filter
$103 \times 103 \times 27$	2×2 max-pooling
$100 \times 100 \times 55$	55 filter
$100 \times 100 \times 55$	55 filter
$45 \times 45 \times 55$	2×2 max-pooling
$45 \times 45 \times 111$	111 filter
$35 \times 35 \times 111$	111 filter
$39 \times 39 \times 115$	2×2 max-pooling
$20 \times 20 \times 233$	233 filter
$10 \times 10 \times 233$	233 filter
$8 \times 8 \times 233$	2×2 max-pooling

Automating the diagnosis of DR in fundus photos requires a number of steps, the first of which is image categorization using CNNs. CNNs are built to recognize complicated patterns and features in pictures, which is crucial for accurately differentiating between the various phases of the illness. Many layers of convolutional and pooling approaches are used by CNNs to assist them comprehend the pictures they are presented with. These representations are then classified using fully linked layers. During the training phase, the network's parameters are adjusted with the use of labeled data. In this way, CNN may learn how to properly categorize retinal pictures. CNN-based picture categorization has outperformed more conventional machine learning methods, producing extremely accurate diagnostic findings for DR.

IV. RESULTS AND DISCUSSIONS

Machines trained using CNNs and gradient boosting methods were used to categorize data to see how well the suggested method worked. The photographs are also labeled by a human expert as ground-truth, allowing for a comparison between the findings achieved by automated classification algorithms and the performance of human judgment. In this way, the accuracy of automated classification systems may be evaluated. Specific feature extraction techniques have been used for the four tasks of detecting blood vessels, distinguishing hard exudates and red lesions, and spotting micro-aneurysms [2]. Over \$760 billion will be spent worldwide this year on diabetes care and related issues. The fast increase in diabetes prevalence may be attributed to many factors, including changes in lifestyle, an older population, urbanization, better diagnostic tools, and more public awareness. Inactivity, poor nutrition, and excess body fat may all play a role in the rising incidence of type 2 diabetes. We can help the healthcare system save money and keep up with increased demand by automating the diagnosis of DR using CNNs. Early diagnosis and treatment have the potential to improve patient outcomes. Classifiers using the aforementioned extracted features and the GBM classification approach, and classifiers based on CNNs both required to be trained for the classification job (with or without data augmentation). Specifically, the GBM's hyper-parameters are initialized with the values of 2 for the number of classes and 6 for the maximum depth. The GBM software utilized in this study was known as eXtreme Gradient Boosting, or XGBoost for short. We found that XGBoost outperformed the competition in our testing; therefore, we've adopted it (i.e., Support Vector Machine, Random Forest). To construct CNN, we used an R package known as MXNet (short for "Multi-Expanding Network"). Here, for your viewing enjoyment, are the trained neural networks [2].

Table IV presents the findings that were obtained from the investigations that were carried out in order to validate the classifications. The data shown in the table demonstrate that the

Table IV. Evaluate the effectiveness of various methods

Method	Accuracy (%)
Solid discharges + GBM	79.4
Papules (Red) + GBM	78.7
Micro-aneurysms + GBM	76.2
Indicators of blood vessels + GBM	69.1
Typical CNN without extra data	81.5
Enhanced CNN with Extra Data	84.5

CNN-based technique achieves better results than the other alternatives, adding support to the premise that was presented before. The table also indicates that the improved performance of the CNN outperforms the original performance of the CNN even without the inclusion of any extra data. Because the data augmentation may help the CNN deal with minute rotations or translations while it is collecting data, the results obtained using the CNN with data augmentation are probably superior to those obtained using the CNN without data augmentation. This is due to the fact that the data augmentation was used [21].

A. DISCUSSION

The fundus image data were first collected via participation in a Kaggle competition titled “Identify indicators of DR in eye photos.” It’s possible that something in the neighborhood of 90 thousand pictures are stored here [21]. In order to ensure that our model is accurate, we use one thousand examples from the dataset that it was trained on in the beginning. Details pertaining to the dataset itself, along with our two unique network topologies. Our approach makes use of two DCNNs each having a fractional max-pooling layer. The two DCNNs will output a one-by-five vector indicating the lesion prediction probabilities for each fundus picture that is fed into them (category). A dimension 24 attribute is shaped by the probability distribution in conjunction with other variables. A breakdown of all 24 features may be found here [21]:

- One can observe that the combined standard deviation of the original picture and the cropped version with a centering ratio of 50% is when we compare these two images to one another.
- A single fundus image might reveal up to twelve distinct characteristics. Then, we take a snapshot of the fundus of the second eye of the same patient and add it to the study, which results in the addition of 12 more variables. Because of this, the total length of the feature vector is 24, and the S algorithm accepts as input vectors feature vectors that have 24 dimensions [21].

Only by doing a comprehensive examination of the retina can DR be discovered and its severity evaluated. One such tool is fundus photography, which provides a detailed visual assessment of the retina’s anatomy and function. OCT creates cross-sectional pictures, expanding the scope of the examination in novel ways. Images may be used to evaluate the retinal thickness and identify fluid buildup. Fluorescein angiography may detect aberrant blood arteries or leakage, in addition to assessing blood flow. These studies allow for a correct diagnosis of DR and monitoring of the condition, allowing for early intervention and therapy.

In order to train a multiclass support vector machine, the 24-dimensional vector is employed, and the TLBO method is used to optimize the SVM’s parameters. In contrast, the reference system makes use of an ensemble classification method that has some similarities to the one under evaluation (RF). On fine-tuning the SVM’s parameter values, we used TLBO for the validation set’s data. The range of values that may be entered for various parameters can be reused as used by the researchers in [22]. Five hundred kids took part in each iteration. The best we’ve done on a five-class DR task is 86.17%, and on a binary class classification task, we achieved a 91.05% success rate. If you want to make a simple yes/no diagnosis, this test is 0.893 sensitive and 0.9089 specific. Damage to the retinal blood vessels, known as DR, is more common in those with advanced cases of diabetes. DR increases the likelihood of vision loss or total blindness. This condition has

Table V. Particulars about each data set

Data set	Images	DR Lesions				
		0 (%)	1 (%)	2 (%)	3 (%)	4 (%)
Train	35134	74	8	17	5	5
Justification	1101	69	9	17	6	5
Test	54542	69	11	17	3	3

the potential to cause irreversible vision loss due to fluid leaks, aberrant blood vessel development, and scar tissue formation. Early detection of DR using CNNs might lead to more effective therapy and a lower risk of irreversible visual loss. Our binary classification strategy uses a *t*-test in addition to counting precision to provide the highest possible degree of accuracy. In the humanities and social sciences, the *t*-test is more often known as the Student’s *t*-test. This is a statistical test, more specifically, of the assumption that the statistic follows a normal distribution [22].

The *t*-test is often used to compare the degree of dissimilarity between two data sets in order to conclude whether or not the differences are statistically significant. Null hypothesis compatibility is ensured by the results of a paired samples *t*-test performed with a binary class classification and a random judgment: 1, 0, and a confidence interval of 1. This is true even when using a 5% threshold of significance [22]. Using the index produced by the hypothesis test, it is feasible to ascertain whether or not two data sets are from the same distribution. If all the data come from the same place, the outcome of the hypothesis test will be extremely close to 0. But if the data sources are really distinct, the resulting number will be closer to 1, indicating a discernible difference. The *p* value represents the likelihood of accepting the hypothesis that there is a difference between the two data sets. A larger likelihood that the data are deceptive is indicated by a smaller *p* value. We also developed an app for mobile devices. Remote monitoring, in-depth analysis, and telemedicine are all made possible by the programmer. We employ a custom-built machine learning technique to validate a user-selected fundus picture once it has been submitted to our server. The chance of each lesion will be shown in less than 10 seconds. The accuracy of this estimate is proportional to the bandwidth available. The first checkup may be performed either locally at the district office or remotely via mobile device. This might be especially helpful in rural areas that have been left out of hospital expansion plans [22].

Table V compares the accuracy attained by different classifiers and methods for fine-tuning the parameters for each dataset. The SVM outperforms the RF by a wide margin when using the default values for both the validation and test sets (and optimization is not done). Parameter optimization utilizing the default parameter searching strategy that is offered in the software package yields extremely high accuracy in the five-fold cross-validation experiment but substantially lower validation and test accuracies. This is true despite the fact that we use the method. This evidence leads us to the conclusion that overfitting happens at all stages of the SVM optimization process, particularly during parameter optimization [22].

B. EVALUATION AND CONSEQUENCES

This research set out to see whether fundus pictures might be automatically analyzed using CNNs to identify DR. Compared to more conventional approaches to feature extraction, the new CNN

architecture designed for this challenge produced much better results. The CNN model successfully classified retinal pictures according to the presence or absence of DR. The performance was improved by using several data augmentation procedures, and generalization was enhanced as a consequence. The CNN model showed early promise, with accuracy on par with or better than that reported by human graders. More clinical studies are suggested in the paper to see whether the proposed CNN-based technique might be integrated into a diagnostic tool. Larger and more varied datasets, patient demographics, picture quality, and illness severity will all be considered in these assessments. The results show the potential of CNNs for automated diagnosis, which is especially important given the limitations of healthcare financing. Due to its superior capacity to detect minute differences and patterns, CNNs provide a cutting-edge and reliable way of detecting DR. Before the CNN model can be employed in real-world applications, further study and testing are required. Important for the accuracy and speed of medical diagnosis is the study's empirical demonstration that CNNs may effectively automate the identification of DR in fundus pictures.

The outcomes of the studies may be used to divide individuals into two categories: those who are healthy and those who are ill. Figure 5 presents the findings in a visual format, whereas Fig. 6

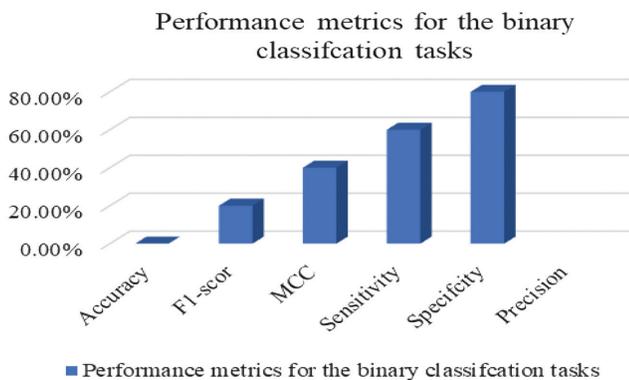


Fig. 5. Measures of success for the use in binary classification.

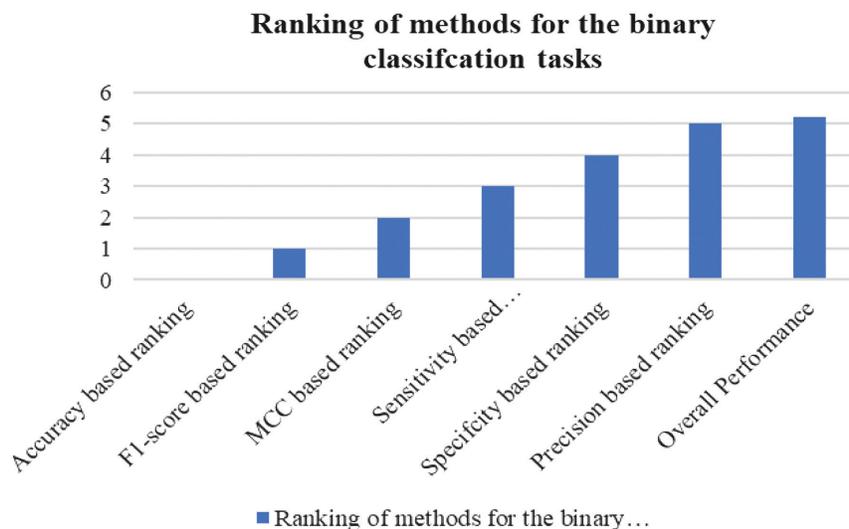


Fig. 6. Graphical representation of performance measures used to rank strategies for the binary classification tasks.

shows the same data in a different format. Both Figs. 5 and 6 include the same information. Several other measurements, including Matthews, F1, and Accuracy, have been used.

This research looked at the feasibility of autonomously diagnosing DR in fundus pictures using CNNs. It was determined that the new CNN architecture was superior to the older feature extraction techniques. Images of DR on the retina were correctly detected by the CNN model. Using data augmentation techniques boosted the model's capacity to generalize. Initial findings using the CNN model were promising, with accuracy on par with or greater than that reported by human graders. The feasibility of incorporating a CNN-based technique into a diagnostic tool requires more clinical research. These kinds of studies need larger and more diverse information, in addition to patient characteristics, picture quality, and illness severity. The results show that CNNs can automate DR diagnosis, providing a solution to the problems caused by the scarcity of available healthcare resources. CNNs are a very effective and sophisticated diagnostic technique. Improving and verifying the CNN model for use in real applications require constant research and testing. This research suggests that CNNs might help speed up the diagnosis of DR, which would have positive implications for healthcare delivery.

Due to the limited size of our available training and validation datasets, it is likely that we are unable to tune DL architectures to their full potential. More training examples are required for modern DL architectures to prevent overfitting. Our study suffers in large part from not being verified on a multicenter validation set, which would be necessary for true clinical use. Finally, if our pilot study using 3D CNN structures with data augmentation procedures is successful, it may be advantageous for eye care practitioners to deploy DL approaches for clinical use.

V. CONCLUSION

The use of DCNNs for automated DR detection has the potential to greatly advance the field. The present physician shortage might be alleviated by a technology that can autonomously diagnose a huge number of retinal pictures. According to the results, CNNs perform better than conventional feature extraction techniques when it

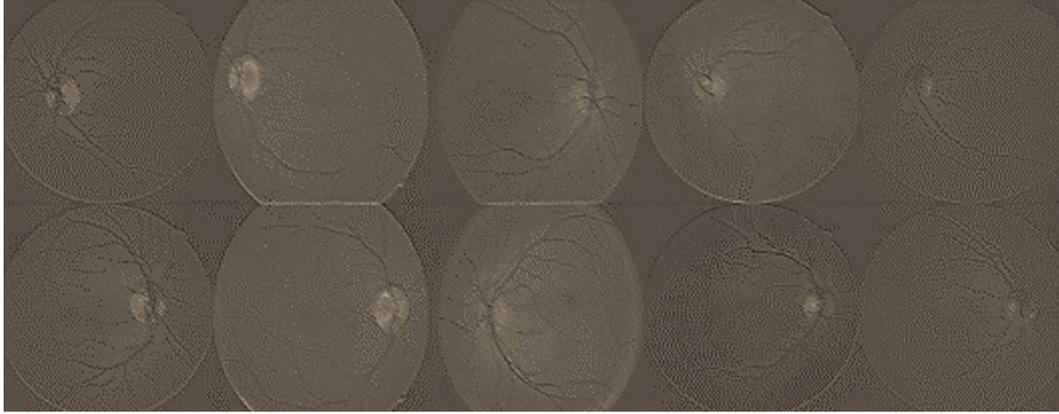


Fig. 7. Illustration of the learned neural networks graphically.

comes to categorizing optical pictures. DR image diagnosis using a custom-built neural network architecture yields impressive results. The proposed technique might benefit from the addition of a data augmentation mechanism. Comparing these preliminary findings to those of human grading suggests great potential for real-world use. The results raise the possibility that diagnostic tools based on CNNs might improve the speed and precision with which DR is diagnosed, reducing strain on already overcrowded healthcare systems. The study's authors draw the conclusion that DCNNs may be used to automate the diagnosis of DR, which has important benefits in contexts where there is a shortage of trained medical practitioners. In particular, the application of CNNs to the task of classifying optical images has shown to be a very effective approach of feature extraction. The accuracy and overall performance of the system have been greatly improved because of the development of a unique neural network architecture tailored to the categorization of DR. The suggested approach already produces impressive results, but it might perform much better with the help of data augmentation tools. These promising first findings highlight the potential use of the CNN-based technique for use in DR diagnostics. If this proves effective, it has the potential to significantly alter the industry by facilitating better and faster diagnosis in areas with restricted access to healthcare.

VI. FUTURE SCOPE

Arden Dertat asserts that CNNs are superior to traditional methods. Presently, "CNNs" are the most often used model in deep learning. Their efforts have made deep learning more accessible for general use. CNNs have steadily replaced previous models and are now being used in many other domains, such as robotics and natural language processing, because of their superior accuracy. CNNs are used to detect and prevent fraud in the mobile communication sector of the telecoms industry. The goal here is, of course, to reduce instances of fraud. It has been shown that DCNNs are superior to more conventional machine learning methods for detecting fraudulent activity in consumer data. Future plans for CNNs include ensuring their sustained preeminence, penetrating new industries, and enhancing their ability to identify and prevent fraud. These developments demonstrate the efficacy and adaptability of CNNs across fields and their capacity to outperform conventional methods.

REFERENCES

- [1] K. Xu, D. Feng, and H. Mi, "Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image," *Molecules*, vol. 22, p. 2054, 2017. DOI: [10.3390/molecules22122054](https://doi.org/10.3390/molecules22122054)
- [2] A. Samanta, A. Saha, S. C. Satapathy, S. L. Fernandes, and Y.-D. Zhang, "Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset," *Pattern Recognit. Lett.*, vol. 135, pp. 293–298, 2020. DOI: [10.1016/j.patrec.2020.04.026](https://doi.org/10.1016/j.patrec.2020.04.026)
- [3] F. Iqbal, A. Abbasi, A. R. Javed, A. Almadhor, Z. Jalil, S. Anwar, and I. Rida, "Data augmentation-based novel deep learning method for Deepfaked images detection," in *ACM Trans. Multimed. Comput., Commun., Appl.*, Association for Computing Machinery (ACM), 2023. DOI: [10.1145/3592615](https://doi.org/10.1145/3592615)
- [4] A. Kukkar, D. Gupta, S. M. Beram, M. Soni, N. K. Singh, A. Sharma, R. Neware, M. Shabaz, and A. Rizwan, "Optimizing deep learning model parameters using socially implemented IoMT systems for diabetic retinopathy classification problem," in *IEEE Transactions on Computational Social Systems: Institute of Electrical and Electronics Engineers (IEEE)*, 2023, pp. 1–12. DOI: [10.1109/tcss.2022.3213369](https://doi.org/10.1109/tcss.2022.3213369)
- [5] Y.-P. Liu, Z. Li, C. Xu, J. Li, and R. Liang, "Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network," *Artif. Intell. Med.*, vol. 99, p. 101694, 2019. DOI: [10.1016/j.artmed.2019.07.002](https://doi.org/10.1016/j.artmed.2019.07.002)
- [6] Y.-H. Li, N.-N. Yeh, S.-J. Chen, and Y.-C. Chung, "Computer-assisted diagnosis for diabetic retinopathy based on fundus images using deep convolutional neural network," *Mobile Inf. Syst.*, vol. 2019, p. e6142839, 2019. DOI: [10.1155/2019/6142839](https://doi.org/10.1155/2019/6142839)
- [7] P. Saranya and S. Prabakaran, "Automatic detection of non-proliferative diabetic retinopathy in retinal fundus images using convolution neural network," *J. Ambient Intell. Human Comput.*, 2020. DOI: [10.1007/s12652-020-02518-6](https://doi.org/10.1007/s12652-020-02518-6)
- [8] Y. Fang, S. H. Ko, and G. S. Jo, "Learning a Unified Tracking-and-Detection Framework with Distractor-aware Constraint for Visual Object Tracking," in *Proc. of the 3rd Asian Conf. on Artificial Intelligence Tech.*, July 05–07, 2019.
- [9] Y. Fang, S. H. Ko, and G. S. Jo, "Robust visual tracking based on global-and-local search with confidence reliability estimation," *Neurocomputing*, vol. 367, pp. 273–286, 2019.
- [10] H. Pratt, F. Coenen, D.M. Broadbent, S.P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Proc.*

- Comput. Sci., 20th Conf. Med. Image Understand. Anal. (MIUA 2016)*, vol. 90, pp. 200–205, 2016. DOI: [10.1016/j.procs.2016.07.014](https://doi.org/10.1016/j.procs.2016.07.014)
- [11] S. Wan, Y. Liang, and Y. Zhang, “Deep convolutional neural networks for diabetic retinopathy detection by image classification,” *Comput. Electr. Eng.*, vol. 72, pp. 274–282, 2018. DOI: [10.1016/j.compeleceng.2018.07.042](https://doi.org/10.1016/j.compeleceng.2018.07.042)
- [12] H. S. Alghamdi, “Towards explainable deep neural networks for the automatic detection of diabetic retinopathy,” *Appl. Sci.*, vol. 12, p. 9435, 2022. DOI: [10.3390/app12199435](https://doi.org/10.3390/app12199435)
- [13] X. Li, H. Chen, X. Qi, Q. Dou, C. W. Fu, and P. A. Heng, “Multi-scale convolutional neural networks for automated diagnosis of diabetic retinopathy,” *Pattern Recognit.*, vol. 113, p. 107863, 2021.
- [14] Y. Xu, M. Li, X. Ding, H. Li, and L. Zhang, “Hybrid attention-based convolutional neural network for automated diagnosis of diabetic retinopathy,” *Neurocomputing*, vol. 479, pp. 190–201, 2022.
- [15] Z. Zhang, Y. Zhang, Z. Wang, M. Xu, and M. Tan, “Deep residual network with adversarial training for robust diabetic retinopathy detection,” *IEEE Trans. Med. Imaging*, vol. 42, no. 1, pp. 227–238, 2023.
- [16] V. Acharya, V. Ravi, T. D. Pham, and C. Chakraborty, “Peripheral blood smear analysis using automated computer-aided diagnosis system to identify acute myeloid leukemia,” in *IEEE Transactions on Engineering Management: Institute of Electrical and Electronics Engineers (IEEE)*, 2022, pp. 1–14. DOI: [10.1109/tem.2021.3103549](https://doi.org/10.1109/tem.2021.3103549)
- [17] S.-I. Pao, H.-Z. Lin, K.-H. Chien, M.-C. Tai, J.-T. Chen, and G.-M. Lin, “Detection of diabetic retinopathy using bichannel convolutional neural network,” *J. Ophthalmol.*, vol. 2020, p. e9139713, 2020. DOI: [10.1155/2020/9139713](https://doi.org/10.1155/2020/9139713)
- [18] A. Kishor and C. Chakraborty, “Early and accurate prediction of diabetics based on FCBF feature selection and SMOTE,” in *International Journal of System Assurance Engineering and Management: Springer Science and Business Media LLC*, 2021. DOI: [10.1007/s13198-021-01174-z](https://doi.org/10.1007/s13198-021-01174-z)
- [19] C. Raja and L. Balaji, “An automatic detection of blood vessel in retinal images using convolution neural network for diabetic retinopathy detection,” *Pattern Recognit. Image Anal.*, vol. 29, pp. 533–545, 2019. DOI: [10.1134/S1054661819030180](https://doi.org/10.1134/S1054661819030180)
- [20] Gao, Z., Li, J., Guo, J., Chen, Y., Yi, Z., and Zhong, J., “Diagnosis of diabetic retinopathy using deep neural networks,” *IEEE Access*, vol. 7, pp. 3360–3370, 2019. DOI: [10.1109/ACCESS.2018.2888639](https://doi.org/10.1109/ACCESS.2018.2888639)
- [21] K. Xu, L. Zhu, R. Wang, C. Liu, and Y. Zhao, “SU-F-J-04: automated detection of diabetic retinopathy using deep convolutional neural networks,” *Med. Phys.*, vol. 43, pp. 3406–3406, 2016. DOI: [10.1118/1.4955912](https://doi.org/10.1118/1.4955912)
- [22] P. Prentašić and S. Lončarić, “Detection of exudates in fundus photographs using convolutional neural networks,” in *2015 9th Int. Symp. Image Signal Process. Anal. (ISPA)*. Presented at the 2015 9th International Symposium on Image and Signal Processing and Analysis (ISPA), 2015, pp. 188–192. DOI: [10.1109/ISPA.2015.7306056](https://doi.org/10.1109/ISPA.2015.7306056)