Visnyk NTUU KPI Seriia – Radiotekhnika Radioaparatobuduvannia, 2023, Iss. 94, pp. 49–57

UDC 004.8.67

Investigation of Fundus Images for Detection of Diabetic Retinopathy Stage Using Deep Learning

Basarab M. R., Ivanko K. O.

Electronic Engineering Department, Faculty of Electronics,
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

E-mail: *mbasarab-ee21@lll.kpi.ua, ivanko-ee@lll.kpi.ua*

The study is dedicated to the investigation of diabetic retinopathy images by digital processing methods and further pathological outcome levels classification. The application of image processing methods to the problem of diabetic retinopathy (DR) analysis is considered in the paper. In order to investigate the possibilities of machine learning for the problem of classification of retinal images, the dataset of retinal images, which represent 5 classes: absence of DR, moderate, mild, proliferate stages, and severe DR, was used in this work. The aim of this study is to identify and compare the different image processing methods used for diabetic retinopathy detection, as well as to choose the classification method that provides the highest accuracy in the identification of the human retina condition. The convolutional neural networks with tuned parameters such as EfficientNet and ResNet were applied to determine the best classification models for computerized disease screening. The accuracy and losses of the different models were determined and compared. Based on this, a combination of image preprocessing steps and neural network models, which provide the highest accuracy of diabetic retinopathy condition recognition, reaching 91.4% for the task of recognition of 5 classes (absence of DR and 4 stages of DR) is proposed. Intermediate stages in the development of diabetic retinopathy are the most difficult to distinguish: the best model showed 85.2% of correctly defined cases of moderate stage of diabetic retinopathy and 83% of correctly defined cases of mild stage. Overall, this article highlights the significance of artificial intelligence (AI) and deep learning in the detection and classification of diabetic retinopathy. It underscores the need for improved screening methods, especially in underserved areas, and emphasizes the potential of these technologies in preserving vision, reducing healthcare professionals' workload, and promoting widespread adoption in clinical practice. The article also acknowledges the challenges associated with image variability and the potential impact on AI model performance, calling for further research and improvement in image quality and consistency.

Keywords: diabetic retinopathy; blindness; machine learning; neural network; diabetes; digital image processing; image recognition

DOI: 10.20535/RADAP.2023.94.49-57

Introduction

Diabetic retinopathy (DR) is a complication of diabetes, caused by high blood sugar levels damaging the back of the retina and it is proven to be the leading cause of blindness in the working-age population of the developed world. It is estimated to affect over 93 million people [1]. The World Health Organization on the other hand estimates that 347 million people have diabetes worldwide. Hence, around 30% of people with diabetes may end up having diabetic retinopathy as the result of long-standing diabetes.

The symptoms are diverse and extend beyond the following mentioned: a gradual decline in vision, challenges with seeing in low light conditions, sudden loss of vision, eye discomfort, redness, the presence of shapes drifting in the visual field (referred to as floaters), and blurred or irregular vision. Detecting diabetic retinopathy (DR) in a timely manner is crucial for mitigating or preventing progression to vision impairment. Nevertheless, timely detection is challenging due to the disease often exhibiting minimal symptoms until it reaches a point where effective treatment is no longer feasible [3].

Presently, the identification of diabetic retinopathy (DR) is a time-intensive and manual procedure, necessitating a trained clinician to assess digital color fundus photographs of the retina. The delayed submission and subsequent review by doctors, often occurring a day or two later, result in delayed outcomes, contributing to lost follow-up, communication errors, and postponed treatment. Clinicians rely on the identification of lesions associated with vascular abnormalities caused by DR to diagnose the condition. Although effective, this approach places significant demands on resources. The expertise and equipment essential for this method are frequently lacking in regions with high local diabetes rates, where efficient DR detection is crucial. With the growing number of individuals affected by diabetes, the existing infrastructure for preventing DR-induced blindness is expected to become increasingly inadequate [3]. If undiagnosed or/and left untreated diabetic retinopathy may result in full or partial blindness. Under the usual circumstances, these complications occur after several years.

To function properly, the human retina needs a constant supply of blood, transferred to it through the blood vessel network across the human body. Those at the very end of the vessel chain, connected to the retina get damaged with consistently high levels of sugar in three main stages:

1. Background retinopathy – tiny bulges develop in the blood vessels, which may bleed slightly but do not usually affect a person's vision.

2. Pre - proliferative retinopathy – more severe and widespread changes affect the blood vessels, including more significant bleeding into the eye.

3. Proliferative retinopathy – the stage that involves the development of scar tissue and delicate, easily-ruptured new blood vessels on the retina, potentially leading to a loss of vision. When the blood vessels in the retina become obstructed, the retina is deprived of essential oxygen and nutrients. Responding to this restricted blood supply, the retina initiates the growth of abnormal blood vessels in inappropriate areas. However, these newly formed vessels are irregular and fail to provide adequate blood flow. They are highly fragile and, during their expansion, may release blood into the vitreous. This bleeding can manifest as visual distortions known as floaters, and in severe cases, it can partially or completely impair vision.

The importance of the use of computer vision algorithms for diabetic retinopathy detection is caused by two main factors: complexity and cost. The researchers have developed a wide range of methods to use neural networks and computer vision techniques to either replace human-involved diagnostics or ease it in a major way due to the high cost of skilled medical professionals and the growing rate of uncertainties in stage classification [4].

The main problem with the algorithms used to perform these actions before is the low accuracy of early stages classification. It is above all important to detect the mild and moderate stages as they are the best time point to cure disease or put it under control.

Therefore it is essential to examine people with diabetic retinopathy for this disease, as well as to detect the high risk of developing diabetic complications. There is a lack of uniform strategies for screening, diagnosing, or predicting the risk of DR.

An effective detection and prediction algorithm for diabetic retinopathy is needed due to the high complexity of the algorithms used today and the seriousness of the consequences such as diabetic macular edema - extra fluid in the eye macula and neovascular glaucoma, both resulting in a vision loss over time. Literature sources demonstrate a wide range of deep learning algorithms employed for the detection of DR. The authors of the paper [1] propose a deep learning-based approach to automated detection of diabetic retinopathy using fundus images. The authors used a convolutional neural network (CNN) to classify retinal images as normal or diabetic retinopathy, achieving an accuracy of 91%.

Another example is the article by Rajalakshmi et al. [2], where the model's performance was reported to be 95.8% sensitivity, which demonstrates its potential as a screening tool for diabetic retinopathy with a limitation that leads to bias and reduced generalizability of the results.

The results of these studies show promising results in DR recognition; however, they differ depending on the used images, the applied methods, and the datasets used for training. Unfortunately, there are numbers of inconsistencies and difficulties, such as the variability in image quality, particularly in smartphone-based fundus photography [5].

Some images have low resolution or artifacts that can hinder the artificial intelligence (AI) model from accurately identifying diabetic retinopathy. This can impact the sensitivity and specificity of the AI model, potentially leading to false-positive or false-negative results.

Improving the detection of diabetic retinopathy through the application of AI and deep learning techniques is crucial for several reasons. Firstly, early detection and treatment of diabetic retinopathy can help prevent or delay blindness in individuals with diabetes. Therefore, timely diagnosis and management of the condition are essential for preserving vision and preventing complications.

Secondly, the use of AI and deep learning techniques can improve the efficiency and accuracy of screening for diabetic retinopathy. This can reduce the workload of healthcare professionals and increase the availability of screening services, particularly in underserved areas. The use of AI-based detection models has the potential to streamline the screening process, making it faster and more accurate, and providing healthcare professionals with more reliable results. Thirdly, improving the performance and reliability of AI-based detection models can increase the confidence of healthcare professionals in using these tools for screening and diagnosis. It can also help to build trust in machine learning methods and promote its widespread adoption in clinical practice.

One of the primary reasons for employing neural networks in the diabetic retinopathy context is their remarkable aptitude for image analysis. Diabetic retinopathy detection heavily relies on the analysis of retinal images, and neural networks excel in identifying relevant patterns, lesions, and abnormalities within these images. With their ability to automatically learn and extract intricate features from visual data, neural networks can comprehensively analyze retinal images, capturing both local and global features associated with different stages of retinopathy.

1 Materials and methods

For assessing the efficiency of the deep learning methods for DR stage recognition in fundus images, APTOS Blindness Detection dataset was used in this research [12]. The challenge, the Aravind Eye Hospital in India faced, was to inherit the ability to detect and prevent diabetic retinopathy among those living in rural areas with low or absent access to the conduction of medical screening. Currently, the technicians from the above-mentioned hospital do travel to these remote regions to provide diagnoses. Implementation of state-of-the-art technologies is today's goal – to gain the ability to automatically screen images for disease and provide informational support on the condition severity levels of the pupil.

The APTOS dataset consists of 3563 large-scale retina images, illustrating various diabetic retinopathy conditions (Fig. 1), taken using fundus photography – the ocular photo record of the pupils' retina. Images from the dataset containing noise and artifacts can be out of focus, underexposed, or overexposed. Since the dataset was gathered from different clinics over an extended period of time, the variation in the taken retina pictures is clearly visible (Fig. 2).

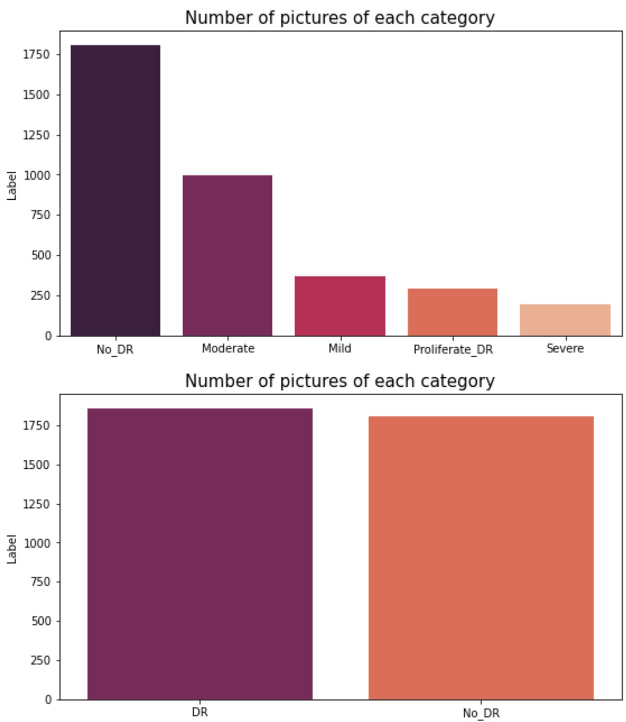


Fig. 1. The number of the fundus images corresponding to DR different stages in the APTOS dataset

Retinal images in the investigated dataset are graded from zero to four (0 - 4): 0 – no diabetic retinopathy; 1 – mild stage; 2 – moderate stage; 3 – severe stage; 4 – proliferative diabetic retinopathy stage.

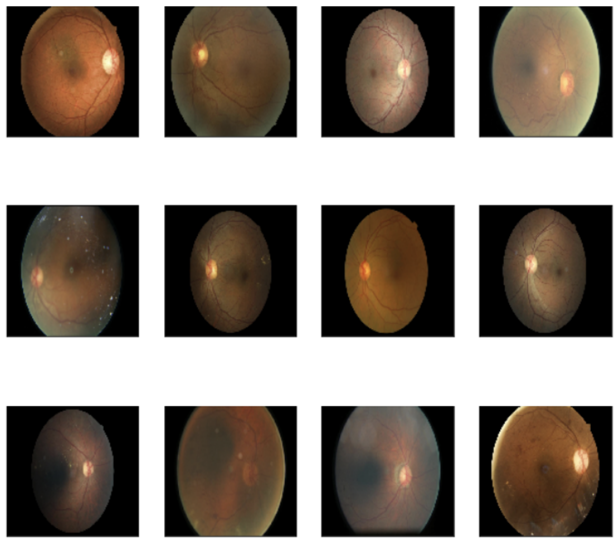


Fig. 2. Diabetic retinopathy conditions in the APTOS dataset before preprocessing

2 Preprocessing of the Fundus Images for Detection of Diabetic Retinopathy Signs

The detection of diabetic retinopathy (DR) is a challenging task that requires accurate and efficient image analysis techniques. Deep learning-based approaches have emerged as a promising solution for DR detection due to their ability to automatically learn features from raw image data. However, before applying deep learning models, the image data needs to be preprocessed to standardize the size, and scale pixel values. This preprocessing step is crucial to improve the model's performance and generalization ability.

The classification task can be split into two main sub-tasks: first - DR or no - DR classification and multi-class classification with recognition of DR stages respectively.

To prepare the dataset for deep learning-based classification of DR, on the first stage we performed the following techniques:

- *Image resizing*, which standardizes the size of the images to 224×224 pixels. This size is commonly used as input for deep learning models. The initial size of the images varied, hence the dataset was standardized to 224 by 224 pixels. With the decrease in the size of the picture, the processing speed increased. The current data was utilized for the comparison of preprocessing efficiency, taking advantage of the numerous pretrained machine learning and deep learning models.

- *Normalization*, which scales the pixel values of the images to a range between 0 and 1. Normalization is essential to prevent the model from being affected by the different intensity levels of the images.

- *Data augmentation*, which artificially increased the size of the dataset by applying random transformations to the images. This technique improves the model's performance by introducing variations in the images, making it more robust to different conditions, reaching 4000 images. Prior to augmentation, there were 3050 pictures that were classified according to their severity of DR as follows: 1500 pictures had no DR, 320 had mild DR, 820 had moderate DR, 180 had severe DR, and 220 had proliferative DR. Following augmentation, the total number of pictures increased to 4000, with the following distribution: 1500 pictures with no DR, and 500 pictures each with mild, moderate, severe, and proliferative DR.

Preprocessing of retinal images is crucial to effectively detect microaneurysms, which are small areas of blood leakage, scar tissue formation, blockage of blood vessels in the macula of the eye, and retinal detachment. Therefore, in order to categorize diabetes types accurately, it is essential to perform thorough preprocessing on the retinal images (Fig. 3).

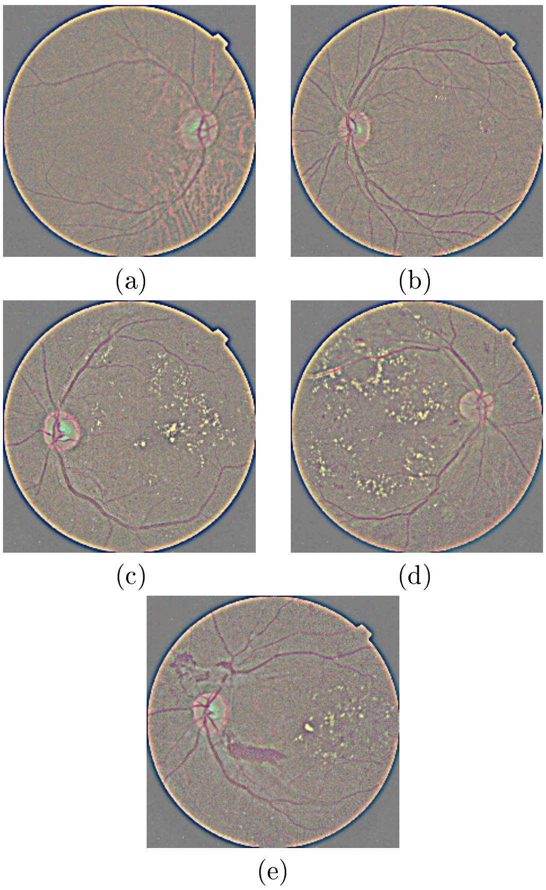


Fig. 3. Diabetic retinopathy conditions in the APTOS dataset after preprocessing: a) no diabetic retinopathy, b) mild stage, c) moderate stage, d) severe stage, e) proliferative diabetic retinopathy stage

Hence the task of preprocessing narrows itself to two main points: improving the lighting conditions of the pupil image and as a result extracting a bigger amount of information from the medical image and second – eliminating the areas that are not informative and have no value for further exploration.

First of all, we have used the Gaussian blur, as like in any other signal, images also can contain different types of noise, especially because of the source (camera sensor). Image smoothing techniques, like Gaussian blur, help in reducing noise. Gaussian filters have the properties of having no overshoot to a step function input while minimizing the rise and fall time. In terms of image processing, any sharp edges in images are smoothed while minimizing too much blurring [1].

Mathematically, using Gaussian blur is identical to the image convolution with a Gaussian function. It is also known as the Weierstrass transform: a smoothed version of $f(x)$ obtained by averaging the values of f, weighted with a Gaussian centered at x [7].

Gaussian blur is a low-pass filter. Since the Fourier transform of a Gaussian is another Gaussian, applying a Gaussian blur has the effect of reducing the image's high-frequency components [8]. The formula of a Gaussian function for two dimensions is:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}},$$

where x represents the distance from the origin along the horizontal axis, y represents the distance from the origin along the vertical axis, and σ denotes the standard deviation of the Gaussian distribution. It is crucial to emphasize that the origin for these axes is positioned at the center, specifically (0, 0). When implemented in two dimensions, this equation generates a surface with concentric circles as contours, featuring a Gaussian distribution emanating from the central point.

The results obtained from this equation are subsequently employed to construct the convolutional matrix, which is then implemented on the original image. The new value of each pixel is determined by a weighted average of its surrounding pixels. The original pixel carries the greatest weight, attributed to its highest Gaussian value, while neighboring pixels receive progressively smaller weights as their distance from the original pixel increases. This process yields a blur that effectively retains boundaries and edges, surpassing the performance of other more uniformly applied blurring filters.

The next step was to enhance the light conditions with Ben Grahams' method [6]. The key idea of its application to the DR detection task is to exaggerate the contrast ratio to make the tiny vessels and abnormalities visible.

As mentioned before, enhancements allow us to get more information from the different images, but they don't deal with unrelated and useless info, hence we had to use a crop function. The one that ended up being used is a circular crop function, results of the work of it are shown in Fig. 4.

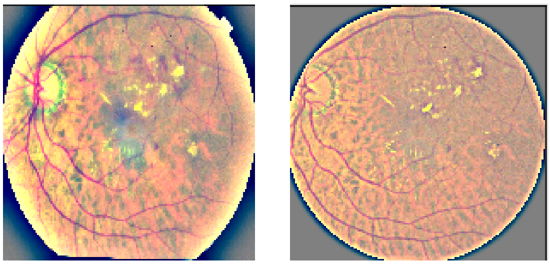


Fig. 4. The application of a circular cropping function to a fundus image

The circular crop function, denoted as $C(x,y)$, can be mathematically expressed as follows for an input image $I(x,y)$, where (x,y) represent the spatial coordinates of the image:

$$C(x,y) = 1, \quad if \quad (x-a)^2+(y-b)^2 < r^2;$$
$$C(x,y) = 0, \quad otherwise,$$

where a and b are the coordinates of the center of the circle or ellipse, and r is the radius of the circle or the semi-axes of the ellipse.

The above equation represents a binary mask where the pixels inside the circle are set to 1 and those outside the circle are set to 0. Applying this binary mask to the input image results in the cropped image that contains only the circular or elliptical region of interest.

3 Application of Deep Learning to DR Stage Classification

Convolutional Neural Networks (CNNs) have transformed the field of computer vision and are now making a significant impact in medical image classification. These deep learning models are specifically designed to process visual data, and their ability to automatically learn and extract features from images is proving invaluable in the healthcare domain.

CNNs mimic the organization of the human visual cortex and consist of interconnected layers, including convolutional, pooling, and fully connected layers. The convolutional layers apply filters to the input image, detecting local patterns and features such as edges or textures. By learning these filters during training, the network becomes adept at recognizing relevant features specific to medical images.

Following the convolutional layers, pooling layers downsample the feature maps, reducing their size while retaining essential information. This not only reduces computational complexity but also introduces translational invariance, making the network robust to slight spatial variations.

The extracted features then pass through fully connected layers, where high-level reasoning and decision-making occur. The final layer employs softmax activation to generate a probability distribution over different disease classes in medical image classification tasks.

In the medical field, CNNs have shown remarkable performance by automatically learning intricate features from medical images. By training on large datasets of labeled medical images, they excel at distinguishing between various diseases, identifying anomalies, and aiding healthcare professionals in making accurate diagnoses. The potential applications of CNNs in medical image classification are vast, including diagnosis of diabetic retinopathy. Their ability to learn and extract features from the fundus images has the potential for healthcare by enabling accurate and efficient diagnoses of DR and its stages. By training on extensive datasets of labeled retinal images, CNNs acquire the ability to discern specific features indicative of retinopathy stages. This process enables them to differentiate between normal retinal images and those displaying signs of diabetic retinopathy.

To determine the best classification models for computerized diabetic retinopathy screening, we implemented such supervised machine learning methods as: Support Vector Machines, Random Forests, Gradient Boosting, K-Nearest Neighbors and Convolutional Neural Networks. It should be noted that Gradient Boosting and K-Nearest Neighbors methods demonstrated insufficient performance for the task of DR stage recognition (near 50% of correctly defined cases).

EfficientNet introduces a methodology for model scaling known as compound scaling. Departing from arbitrary increases in width, depth, or resolution, this technique uniformly scales each dimension with predefined coefficients. EfficientNet is a family of models developed by Google that use a novel compound scaling method to achieve state-of-the-art performance with fewer parameters than other models. EfficientNet builds upon a foundational network established through neural architecture search, employing the AutoML MNAS framework. The primary objective is to fine-tune the network to achieve optimal accuracy while simultaneously considering computational complexity. To this end, penalties are imposed on excessively computationally demanding models as well as those with slow inference times. The architectural design of EfficientNet incorporates a mobile inverted bottleneck convolution, reminiscent of MobileNet V2 [13]. However, EfficientNet distinguishes itself through its larger scale, achieved by augmenting the floating-point operations per second (FLOPS). This enhanced baseline model serves as the foundation for scaling up and generating the family of EfficientNets.

The ResNet is a widely used model that uses residual connections to improve training and address the vanishing gradient problem.

In this study, two deep learning models EfficientNet and ResNet were applied to the task of DR detection with the APTOS dataset.

The EfficientNet model used in the study consisted of seven blocks, with each block using a combination of convolutional, pooling, and activation layers.

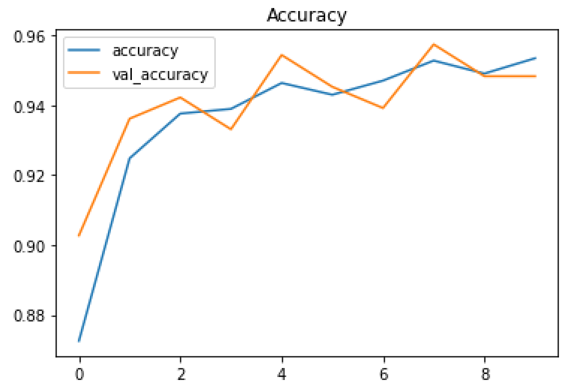
Convolutional layers are fundamental components in CNNs that enable the extraction of local features from input data. They consist of filters, also known as kernels, which are small matrices applied to the input data using a sliding window technique. Each filter performs a convolution operation by taking the dot product between its weights and a corresponding region of the input. This process generates feature maps that capture different aspects of the input, such as edges, textures, or patterns. Convolutional layers are responsible for learning hierarchical representations of the input data, enabling the network to recognize complex patterns and structures. Pooling layers are employed in CNNs to downsample the feature maps generated by convolutional layers. They reduce the spatial dimensions of the feature maps, thereby decreasing the computational complexity of subsequent layers. The model also used a global average pooling layer followed by a dense output layer with two neurons for the binary classification task.

Activation layers introduce non-linearities into the network, allowing CNNs to model complex relationships between input and output. They apply a non-linear activation function to the output of a layer, element-wise. One commonly used activation function is the Rectified Linear Unit (ReLU), which sets negative values to zero and preserves positive values. ReLU activations help introduce non-linearities into the network, enhancing its ability to learn and model complex patterns.

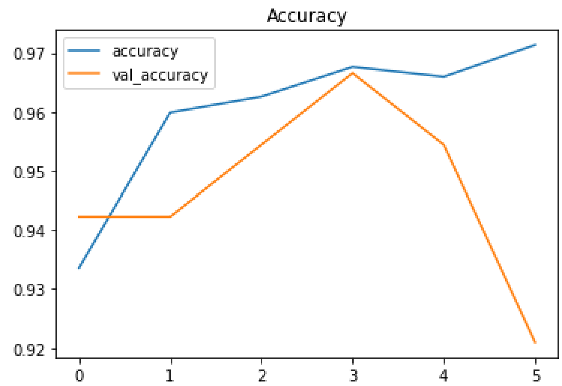
Each applied architecture requires specific parameters for optimal performance. The EfficientNet architecture consists of two convolutional layers followed by a max pooling layer and two fully connected layers. A suggested configuration includes 32 filters for the first layer, 64 filters for the second layer, 3×3 kernel size, and a dropout rate of 0.5.

The VGG-16 architecture includes 13 convolutional layers, 5 max pooling layers, and 3 fully connected layers. The configuration used in this work includes 64 filters for the first convolutional layer, double the number of filters after each max pooling layer, and a dropout rate of 0.5. The accuracy comparison of the described architectures is shown in Figure 5.

The Inception-v3 architecture has a stem module followed by 11 inception modules and a global average pooling layer. In this study, a suggested configuration includes 32 filters for the stem module, varying number of filters for each branch, and a dropout rate of 0.4.



(a)



(b)

Fig. 5. Accuracy of the EfficientNet (a) and ResNet (b) models applied to the task of DR detection with the APTOS dataset

The VGG19 architecture consists of 19 layers, including convolutional and fully connected layers. We suggest using transfer learning to leverage the pre-trained weights of VGG19 on the ImageNet dataset. With this approach, the initial layers of the VGG19 model are frozen, and the last few layers are replaced with custom fully connected layers for the specific classification task. The suggested configuration includes using an Adam optimizer with a learning rate of 0.0001 and a batch size of 32. The Adam optimizer maintains adaptive learning rates for each parameter in a neural network. It computes individual adaptive learning rates using the first and second moments of the gradients. The first moment is the mean (average) of the gradients, and the second moment is the uncentered variance of the gradients.

The algorithm updates the parameters iteratively by calculating the exponentially decaying average of past gradients (first moment) and the exponentially decaying average of past squared gradients (second moment). These averages are used to adjust the learning rate for each parameter, allowing for faster convergence and better handling of sparse gradients.

Finally, the ResNet-50 architecture has a convolutional layer followed by a stack of residual blocks and a global average pooling layer. In this study, a recommended configuration includes 64 filters for the convolutional layer, a 7×7 kernel size, and a dropout rate of 0.5. These four architectures offer unique features that could impact their predictive ability.

Data augmentation is the way the EffiecientNet model was upgraded. Data augmentation is a technique where synthetic data is generated from existing data, allowing the model to be trained on a more diverse set of examples. In the case of DR detection, data augmentation techniques such as elastic distortion, rotation, and scaling were used to generate additional training data, leading to improved generalization and robustness of the model.

In CNN, a feature map refers to the output of a convolutional layer. Each feature map corresponds to a particular learned filter or kernel that convolves across the input image, capturing different local patterns or features. Figure 6 demonstrates the behavior of the neural network decision-making process through the feature map. Lighter areas reflect the more informative zones that provide the biggest amount of information to the algorithm.

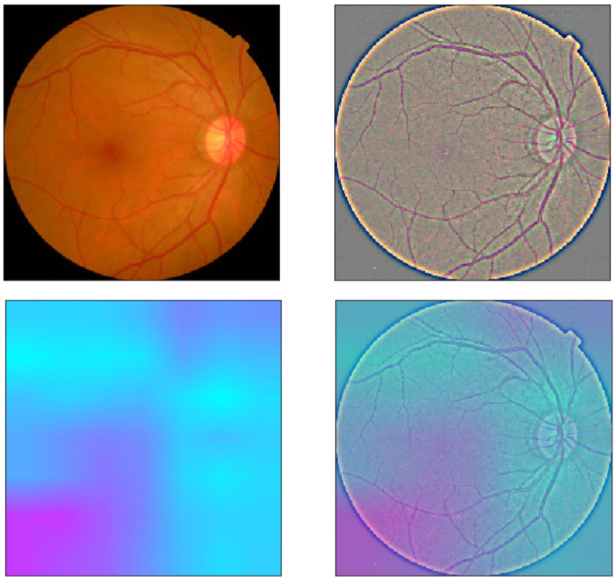


Fig. 6. Heatmap of the neural network decision-making process. From left to right: original picture, processed picture, heat map mask, heat map mask overlaid with processed picture

When applied to the task of distinguishing retinopathy classes, a CNN learns to extract relevant features from retinal images through multiple convolutional layers. These layers apply various filters to the input image, creating a set of feature maps that highlight different characteristics. As the network goes deeper, the feature maps capture increasingly complex and abstract features. For example, early feature maps might identify basic edges or textures, while deeper ones could recognize more specific retinal structures like blood vessels, exudates, or microaneurysms.

By training the CNN on a labeled dataset of retinal images, the network learns to associate specific features in the feature maps with different retinopathy classes. Through the learning process, the network identifies patterns that are indicative of healthy retinas or various stages of diabetic retinopathy.

4 Results and Discussion

The aim of this study was to identify and compare different preprocessing methods and choose the classification model that provides the highest accuracy in identifying the condition of the human retina.

The images employed in conjunction with neural networks were RGB images, as grayscale images entail the omission of significant informational facets, frequently encompassing details pertinent to blood vessels, minuscule hemorrhages, and analogous characteristics. These details, at times, play a pivotal role in informing the decision-making process of neural networks.

For the classification task, neural networks with tuned parameters, namely EfficientNet and ResNet, were utilized to determine the best models for computerized disease screening. The accuracy and losses of the different models were determined and compared.

The results of the study demonstrated the effectiveness of the preprocessing techniques in improving the performance of the deep learning models for DR classification. By standardizing the image size, scaling pixel values, and introducing data augmentation, the models were able to achieve higher accuracy in classifying the diabetic retinopathy condition.

The combination of preprocessing steps proposed in this study resulted in the higher accuracy of diabetic retinopathy condition recognition (Table 1). Provided results show the delineation of the classification accuracy of nine machine learning algorithms as per their performance metrics on the APTOS dataset. Notably, the EfficientNet neural network attains the highest accuracy, registering 92.4% overall accuracy across the five distinct classes, with particularly notable achievements surpassing 70% accuracy in the Mild and Moderate cases. The model scaling within EfficientNet's design has facilitated the fine-tuning of its structure, resulting in performance surpassing that of other prevalent transfer learning networks.

Table 1 Overall accuracy and percentage of correctly predicted cases for each class in parenthesis (Absence of DR, moderate stage, mild stage, proliferate stage, and severe DR)

#	Name	Before preprocessing	After preprocessing
1	ResNet-50	85.3% (93%, 81.6%, 79%, 82%, 91%)	87% (94%, 80%, 84%, 83.9%, 93.1%)
2	VGG19	88% (94%, 78%, 85,3%, 85%, 93.7%)	89.1% (94,6%, 79%, 84.5%, 85.3%, 93.6%)
3	**EffiecientNet**	**89.3% (97.1%, 81%, 86.4%, 87.2%, 94.8%)**	**91.4% (98%, 83%, 85.2%, 95.2%, 96%)**
4	ElasticNet	79.6% (86%, 73.6%, 74.2%, 78.3%, 86.1%)	80.4% (87.5%, 74%, 74%, 79%, 87.6%)
5	Inception-v3	74.8% (86%, 65%, 68.3%, 73%, 82%)	76% (86.3%, 65%, 68.5%, 73.7%, 82%)
6	VGG16	79.5% (85%, 72%, 73%, 78.6%, 89.1%)	86.8% (94.5%, 80%, 83%, 83.9%, 92.8%))
7	XGBoost	74.1% (85.5%, 63%, 68%, 73%, 81%)	75% (86.6%, 65%, 67%, 73%, 82.4%)
8	SVM	65.3% (76.2%, 54%, 57%, 65%, 76%)	67.8% (79%, 55.8%, 59.5%, 66%, 79%)
9	Random Forest	66.7% (76.3%, 52%, 60%, 67%, 78%)	69.4% (80%, 52.4%, 60%, 73%, 82%)

Conclusion

In conclusion, the study highlights the importance of preprocessing methods in improving the performance of deep learning models for diabetic retinopathy detection and classification. The findings support the potential of machine learning techniques to automate and enhance the screening process for DR, enabling early detection and timely intervention to prevent vision impairment.

The best results were obtained for the VGG19 and EfficientNet (overall accuracy with five classes reached 89.1% and 91.4%, respectively). The results indicate that most models experience an improvement in test accuracy after applying preprocessing techniques. The increase in accuracy suggests that preprocessing helps to enhance the models' ability to identify and classify Mild and Average cases of diabetic retinopathy (with five classes in total). Notably, models like VGG19 and EfficientNet show relatively higher accuracies both before and after preprocessing, indicating their robustness in capturing relevant features for classification.

These findings underscore the significance of preprocessing steps in the pipeline of diabetic retinopathy classification models. Further research and development in this field hold great potential to make a significant impact on healthcare, particularly in regions with limited resources and access to specialized eye care professionals. By leveraging machine learning techniques and appropriate preprocessing methods, it becomes possible to provide efficient and accurate screening for diabetic retinopathy, enabling timely intervention and reducing the risk of vision impairment.

References

[1] Li, F., Wang, Y., Xu, T., et al. (2022). Deep learning-based automated detection for diabetic retinopathy and diabetic macular edema in retinal fundus photographs. *Eye*, 36(6), 1433–1441. doi:10.1038/s41433-021-01552-8.

[2] Rajalakshmi, R., Subashini, R., Anjana, R. M., Mohan, V. (2018). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye*, 32, 1138-1144. doi: 10.1038/s41433-018-0064-9.

[3] Lam C., Yi D., Guo M., Lindsey T. (2018). Automated Detection of Diabetic Retinopathy using Deep Learning. *AMIA Jt Summits Transl Sci Proc.*, 2017:147-155.

[4] Doshi D., Shenoy A., Sidhpura D. and Gharpure P. (2016). Diabetic retinopathy detection using deep convolutional neural networks. *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, pp. 261-266. doi: 10.1109/CAST.2016.7914977.

[5] Gulshan, V., Peng L., Coram M. et al. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, 13; 316(22):2402-2410. doi:10.1001/jama.2016.17216.

[6] Asia, A.-O.; Zhu, C.-Z.; Althubiti, S. A.; et al. (2022). Detection of Diabetic Retinopathy in Retinal Fundus Images Using CNN Classification Models. *Electronics*, 11(17), 2740. doi:10.3390/electronics11172740.

[7] Uppamma P., Bhattacharya S. (2023). Deep Learning and Medical Image Processing Techniques for Diabetic Retinopathy: A Survey of Applications, Challenges, and Future Trends. *Journal of Healthcare Engineering*, Volume 2023, Article ID 2728719. doi: 10.1155/2023/2728719.

[8] Mohanty, C.; Mahapatra, S.; Acharya, B.; et al. (2023). Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy. *Sensors (Basel)*, 23(12), 5726. doi:10.3390/s23125726.

[9] Sharma T., Shah M. (2021). A comprehensive review of machine learning techniques on diabetes detection. *Visual Computing for Industry, Biomedicine, and Art*, 4(1):30. doi: 10.1186/s42492-021-00097-7.

[10] Shah P., Mishra D. K., Shanmugam M. P., Doshi B., Jayaraj H., Ramanjulu R. (2020). Validation of Deep Convolutional Neural Network-based algorithm for detection of diabetic retinopathy – Artificial intelligence versus clinician for screening. *Indian J Ophthalmol*, 68(2):398-405. doi: 10.4103/ijo.IJO_966_19.

[11] Nadeem, M. W.; Goh, H. G.; Hussain, M.; Liew, S.-Y.; Andonovic, I.; Khan, M. A. (2022). Deep Learning for Diabetic Retinopathy Analysis: A Review, Research Challenges, and Future Directions. *Sensors*, 22(18), 6780. doi: 10.3390/s22186780.

[12] *APTOS Symposium dataset*.

[13] Hollemans M. (2018). *MobileNet V2*, architecture.

Дослідження зображень очного дна для виявлення стадії діабетичної ретинопатії за допомогою глибокого навчання

Басараб М. Р., Іванько К. О.

Робота присвячена дослідженню зображень діабетичної ретинопатії за допомогою методів цифрової обробки та подальшої класифікації рівнів патологічних змін. У статті розглянуто застосування методів обробки зображень до проблеми аналізу діабетичної ретинопатії (ДР). Для вивчення можливостей машинного навчання для класифікації зображень сітківки ока у цій роботі було використано набір даних, що відображають 5 класів: відсутність ДР, помірну, легку, проліферативну стадії та важку ДР.

Метою цього дослідження є ідентифікація та порівняння різних методів обробки зображень, які застосовано для виявлення діабетичної ретинопатії, а також вибір методу класифікації, який забезпечує найвищу точність визначення стану людської сітківки у випадку ДР. Для визначення найкращих моделей класифікації діабетичної ретинопатії були застосовані нейронні мережі з налаштованими параметрами, такі як EfficientNet, ResNet та інші. Було визначено точність моделей і на основі цього запропоновано кроки попередньої обробки та комбінацію параметрів нейронної мережі, яка забезпечує найвищу точність визначення стану діабетичної ретинопатії, досягаючи 91,4% для завдання визначення 5 класів (відсутність ДР та 4 стадії ДР). Проміжні стадії розвитку діабетичної ретинопатії найважче відрізнити: найкраща модель показала 85,2% правильно визначених випадків помірної стадії діабетичної ретинопатії і 83% правильно визначених випадків легкої стадії.

Загалом, ця стаття підкреслює значущість штучного інтелекту та глибокого навчання у виявленні та класифікації діабетичної ретинопатії. Вона наголошує на необхідності поліпшення методів скринінгу, особливо у недостатньо обслуговуваних районах, та підкреслює потенціал цих технологій у збереженні зору, зменшенні робочого навантаження медичних працівників та сприянні широкому впровадженню у клінічну практику. У статті також визнаються проблеми, пов'язані з варіабельністю зображень та потенційним впливом на якість роботи моделей, що вимагає додаткових досліджень та покращення якості зображень.

Ключові слова: діабетична ретинопатія; сліпота; машинне навчання; нейронна мережа; діабет; цифрова обробка зображень; розпізнавання зображень