INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2026, VOL. 72, NO. 1, PP. 1–7
Manuscript received December 15, 2025; revised February, 2026. doi: 10.24425/ijet.2026.157868

Deep learning approach for retinopathy identification in combined clinical datasets

Michał Zmonarski, Ewa Skubalska-Rafajlowicz, Aleksandra Zgryźniak, and Sławomir Zmonarski

***Abstract*—This work presents a system for automatic detection of various stages of diabetic retinopathy (DR) based on fundus images of patients. The system was built based on a relatively new and little-used image database: "Dataset of fundus images for the study of diabetic retinopathy" version v3 CastilloBenitez21. The primary dataset was expanded using clinical fundus photographs acquired from the Department of Nephrology at Wroclaw Medical University. The diagnostic system was developed based on various variants of convolutional neural networks (CNNs) that were pre-trained on ImageNet data. The CNN classifier, based on VGG16 with transfer learning, proved to be effective and gave a global accuracy of 83.15%. The evaluation of discrimination between the non-DR and the DR state resulted in an accuracy of 89.7%, with a sensitivity of 94.9%, a specificity of 88.3%, and a Matthews Correlation Coefficient of 0.7665.**

***Keywords*—Diabetic retinopathy detection; deep learning; CNN-VGG16; sequential transfer learning**

I. Introduction

DIABETIC retinopathy (DR) [4], [34] represents a specific complication of diabetes, frequently remaining asymptomatic during its initial phases. It can lead to vision impairment or even blindness. Research [11], [10] suggests a correlation between the presence of retinopathy and the increased susceptibility to nephropathy and cardiovascular disorders among diabetic patients. Consequently, proactive detection and management of DR can positively influence the broader clinical prognosis in diabetes care.

The implementation of large-scale screening and early diagnosis is feasible through the computational analysis of fundus photography, utilizing a synergy of digital signal processing and artificial intelligence algorithms. The necessity for such automated systems is reflected in the extensive body of computer science literature dedicated to the classification of retinal images, addressing both disease identification and severity grading. With clinical data indicating that nearly one-third of the diabetic population is affected by DR [11], automated screening offers a pathway to improved diagnostic throughput, cost reduction, and reduced reliance on manual expert evaluation.

Digital image databases offer insight into typical changes in diabetic retinopathy and normal retinal structures, visualized at the pixel level. Representative sets of retinal images of various retinas are essential for the development and testing of automated screening algorithms for DR symptoms at different levels. In the area of automatic DR classification algorithms, the International Clinical Diabetic Retinopathy (ICDR) Severity Scale is used [34]. Standard fundus images are divided into five classes according to the severity of the DR.

TABLE I
GRADING SCALE OF DIABETIC RETINOPATHY SEVERITY USED IN THE STUDY

Class Label	Clinical Description
Class 0	Normal (no signs of DR)
Class 1	Mild Non-Proliferative DR (NPDR)
Class 2	Moderate NPDR
Class 3	Severe NPDR
Class 4	Proliferative Diabetic Retinopathy (PDR)

In [4], Class 5 (Clinically significant macular oedema) is added. These symptoms are observed as large red disks. In the case of the EPDRS (Early Treatment of Diabetic Retinopathy Study) classification [34], the division of DR symptoms is even more detailed. Occasionally, other detailed classification systems occur. For example, the taxonomy adopted in [18] distinguishes six classification levels, defining Class 4 as very severe NPDR and assigning PDR samples to Class 5. The fundus image data set for the study of diabetic retinopathy (DFISDR) [3] consists of 757 color images (DFISDRv2). In addition, an extended database (DFISDRv3) containing 1,437 images is available. The classification of fundus images has been done in 7 categories: No DR signs, Non-Proliferative diabetes Retinopathy (NPDR mild, moderate, severe, very severe), Proliferative diabetes Retinopathy (PDR) and Advanced PDR.

Unlike more serious lesions in the retinal vessels, class 1 is the most difficult to identify. The development of microaneurysms (MA) in the retina is the first sign of DR. MA usually manifests itself as punctate dark red lesions, with a diameter marginally larger than that of the finest retinal capillaries [29], [9].

M. Zmonarski and E. Skubalska-Rafajłowicz are with Faculty of Information and Communication Technology, Wroclaw University of Science and Technology, Wroclaw, Poland (e-mail: michal.zmonarski, ewa.skubalska-rafajlowicz@pwr.edu.pl).

S. Zmonarski is with Dept. of Nephrology and Transplantation Medicine, Wrocław Medical University, Wrocław, Poland (e-mail: slawomir.zmonarski@umw.edu.pl).

A.Zgryźniak is with Clinic of Ophthalmology, University Teaching Hospital, Wrocław, Poland (e-mail: aleksandra.zgrzyniak@gmail.com).

This paper continues the problem presented in [36]. The purpose of this paper is to investigate the usefulness of creating an image database consisting of fundus images obtained from patients in a nephrology clinic at Wrocław Medical University (WMU). Unfortunately, that data set was relatively small, consisting of only 71 images. Therefore, the study utilized the DFISDRv2 dataset, a recently introduced and less common collection, distinguished by its rigorous expert annotation. Subsequently, a straightforward convolutional neural network architecture was constructed to evaluate classification performance within a controlled environment. The average classification accuracy obtained was 84.7%. The classifier built on data from a source other than WMU did not correctly classify WMU images. After combining both sets, the new classifier was trained on 80% of the images and tested on the remaining 20%; the classification accuracy was similar to the previous results. It should be emphasized that the images from the WMU and the DFISDRv2 database were divided in the same proportions, and all 13 test images from Wrocław were correctly classified. In the binary classification task (Healthy vs. DR), the average accuracy reached 96%. This result suggests that the algorithm proposed in [36] is effective, but not fully robust. We believe that the primary limitation is the relatively small training data set, which reduces confidence, even for this basic identification task.

In this article, we investigate the properties of a classifier based on the combination of two datasets from different sources, with the external set chosen as DFISDRv3 [3], which is twice as large as DFISDRv2 used in [36] and [3].

The paper is structured as follows. The next section presents a review of the current literature on automatic systems for identifying and classifying DR based on fundus images and publicly available retinal datasets. Information about the data sets used in our experiments, the DR detection method, performance metrics, and experiment results is explained in Section 3. Finally, we discuss our approach and results and conclude the article.

II. Problem statement

The purpose of this paper was to investigate the extent to which it is possible to create our own system for the automatic classification of patients with nephrology at risk for the potential or observed development of DR. At the same time, due to further research concepts, it is essential that the image database used by the system in the future come from a population selected for kidney diseases.

Currently, due to the available data from WMU, the main significant problem is discrimination between asymptomatic images (no DR) and those indicating the presence of changes associated with the existence of DR. From this point of view, the DFISDRv3 database is balanced, as it contains 711 sample images in class 0, which is about half of all data in this database containing a total of 1437 images.

The choice of the DFISDRv3 [3] database was driven not only by its relatively small size, but also by the fact that it was unlikely to be used as data for an automatic classification system. This allows both sets of data to be considered new and unencumbered by any additional prior knowledge beyond their source.

Limiting the size of the database, which we intend to treat as a supporting dataset, should allow the construction of a practical classification algorithm with a relatively small number of parameters. This algorithm was the CNN VGG16 network mentioned earlier [36] and was presented at the DepCoS-RELCOMEX 2025 conference.

III. Related works

Computational experiments in the field of automatic DR detection are based on many publicly available image databases, such as Kaggle base DRD [8], APTOS19 [35], MESSIDOR [7], EyePACS [6], IDRiD [26], Dataset for Diabetic Retinopathy (DDR) [14], among many others. A specific database is the DeepDRiD database [16], which helps to evaluate and improve the quality of fundus images intended for the construction of DR classification systems.

Automated systems for detecting and grading Diabetic Retinopathy typically utilizes the following computational strategies:

- image processing algorithms related to improving the quality of photographs, variation in color attenuation, intensity conversion, denoising, and contrast enhancement
- advanced algorithms for detecting objects, image segmentation, or background analysis,
- classification methods based on machine learning algorithms that analyze the selected features of images subject to classification,
- deep learning methods such as various variants of convolutional networks (CNN),
- hybrid methods, visual transformers.

A. Advances methods of image processing and analysis

Reference [20] describes Computer-Aided Diagnosis (CAD) frameworks designed specifically for the binary classification of DR pathology. Computational experiments were conducted on the Kaggle 2015 dataset. The main component of the algorithm was a combination of image processing with EGMM-based Retinal Blood Vessel Detection and Segmentation, Blob analysis and Connected Components Analysis, Retinal Blood Vessel ROI Identification, Alaxnet-based feature extraction, Feature Selection, dimensionality reduction, and finally SVM-based DR classification.

The crucial problem in quickly starting appropriate therapy is the detection of the early stages of DR symptoms, such as microaneurysm (MA) lesions (mild NPDR) [9]. Some articles deal with the detection of MA [18], [22]. The paper [29] applies traditional image processing methods related to the occurrence of visual symptoms on a multi-scale in the context of MA.

B. Machine learning methods

Beyond a general literature survey, reference [12] investigates specific image enhancement protocols and feature extraction techniques used in DR staging. The study evaluates

a broad spectrum of computational models, ranging from classical classifiers (SVM, Random Forest, Naive Bayes, k-NN, k-Means) to neural network architectures, including Auto-encoders.

C. Deep learning methods

Reference [1] benchmarked 26 distinct pre-trained CNN architectures using transfer learning. The study utilized the DRGF dataset, comprising 3,662 Gaussian-filtered images derived from the APTOS19 archive and resized to 224×224 pixels. Despite the variety of models tested, the classification accuracy for DR detection remained below the 80% threshold.

The paper [25] presents a multiclass classification of fundus images to detect eye diseases (including DR identification) based on a transfer learning approach and various CNN network architectures. The authors showed that eliminating bad quality images using an integrated quality evaluation system improves classification performance.

The study in [33] evaluates both binary detection and multi-class DR classification. The authors proposed an ensemble framework integrating established architectures (InceptionV3, Inception-ResNetV2, Xception) across various input resolutions with data augmentation. Although the reported accuracy was exceptional, it relied on a massive, non-public dataset.

In contrast, reference [32] focuses on hyperparameter optimization. The authors demonstrated that the adjustment of global parameters could elevate the classification accuracy from a baseline of 0.48–0.83 to a peak of 0.96. These experiments utilized the 5-class Kaggle 2015 dataset [8] and standard backbones (AlexNet, VGGNet, GoogLeNet, ResNet) facilitated by transfer learning.

In [19] a multitask deep learning framework is introduced for DR staging. The architecture comprises two distinct components: a primary classification network and a stand-alone regression model. Their outputs are fused and subsequently processed by a Multilayer Perceptron (MLP) to generate the final prediction.

D. Hybrid methods

Hybrid methods combine deep learning methods (DL) with classical methods of digital image processing and analysis. On the other hand, hybrid models can use CNNs in combination with simpler machine learning algorithms such as multilayer perceptron (MLP), support vector machine (SVM), random forest (RF), transformers and vision transformers (ViT), or boosting methods. Adaboost, Gradient boost, and Extreme gradient boost (XGBOOST) are based on combining built sequentially simple predictors [1]. Naturally, almost all methods, even those that we do not call hybrid, use some kind of pre-processing of input images.

Reference [23] leverages Vision Transformers (ViT) to capture long-range spatial dependencies within image data. The methodology was validated on the FGADR dataset, a recently introduced large-scale repository featuring fine-grained annotations, which served as the benchmark for both DR identification and grading tasks. This data set has 1,842 images with pixel-level DR-related lesion annotations, and 1,000 images with image-level labels graded by board-certified ophthalmologists. Because the FGADR base is an unbalanced data set, the authors of the paper combined several techniques to deal with this issue. It achieved satisfactory results compared to the baseline ViT models, namely $F_1 = 0.825$ and accuracy $= 0.825$. The paper [30] uses for a similar purpose a dynamic recurrent CNN (R-CNN) solution. In both works, small image patches are recursively used to determine long-range dependencies in images. In this work, the authors achieved very high classification accuracy using the Kaggle DRD base. The Kaggle DR dataset (DRD) consists of 88,702 color fundus images, including 35,126 samples for training. After initial preprocessing, all images are processed using the Gray Level Co-occurrence Matrix (GLCM), the method proposed by Haralick et.al. in 1973.

The method proposed in [24] is based on the independent extraction of local information from multiple rectangular image regions. The machine learning system to detect diabetic retinopathy from fundus images uses a multiple learning paradigm with an integrated attention mechanism. Reference [23] leverages Vision Transformers (ViT) to capture long-range spatial dependencies. The method was validated on the FGADR dataset—a recently introduced large-scale repository with fine-grained annotations. Crucially, the architecture employs an attention mechanism that explicitly localizes pathological regions and DR-induced lesions to facilitate classification.

Only selected examples of the methods have been presented in more detail. The topic of DR diagnostics is very popular, as evidenced by the large number of review articles published over the last few years. See, for example, [1], [2], [12], [14], [18], [21], [19], [22], [27], [28], among many others.

IV. Data characterization

The primary training material consists of 1437 color retinal images from the database DFISDRv3 [3]. These clinical samples were acquired at the Department of Ophthalmology, Hospital of Clínicas (FCM-UNA), Paraguay. We selected this recently published repository to assess the feasibility of automatic DR classification as a training set, we chose ”Dataset from fundus images for the study of DR” (DFISDRv3) described in [3], which is a relatively new set, and we even found a preliminary examination of its usefulness in the detection and classification of DR interesting and useful. The data set containing 1437 color fundus images was acquired in the Department of Ophthalmology of the Hospital of Clínicas, Facultad de Ciencias Médicas (FCM), Universidad Nacional de Asunción (UNA), Paraguay. The fundus images are initially divided by the authors of the database into seven classes: Normal (no signs of DR)- 711 images, mild (early) NPDR - 6 images, moderate NPDR - 110 images, severe NPDR - 210 images, very severe NPDR - 139 images, PDR - 116 images, and advanced PDR -145 images. For the purpose of this study, the original 7-class grading system was assigned to 5 classes to balance the data set. Specifically, ’Advanced PDR’ was excluded due to the limited number of samples, and ’PDR’ was merged with ’Very Severe NPDR’. The ground

truth in DFISDRv3 base was established by annotation by two independent ophthalmology specialists. The repository contains JPEG files at a resolution of 2124×2056 pixels. During data set construction, samples that exhibited blur, artifacts, or insufficient quality were rigorously excluded. The primary objective of this collection is to facilitate research into the detection of both Non-Proliferative (NPDR) and Proliferative (PDR) stages.

Figure 1 presents sample photographs from the DFISDRv3 base. The image in the upper left corner is an asymptomatic example, while the other three illustrate different types of DR symptoms.

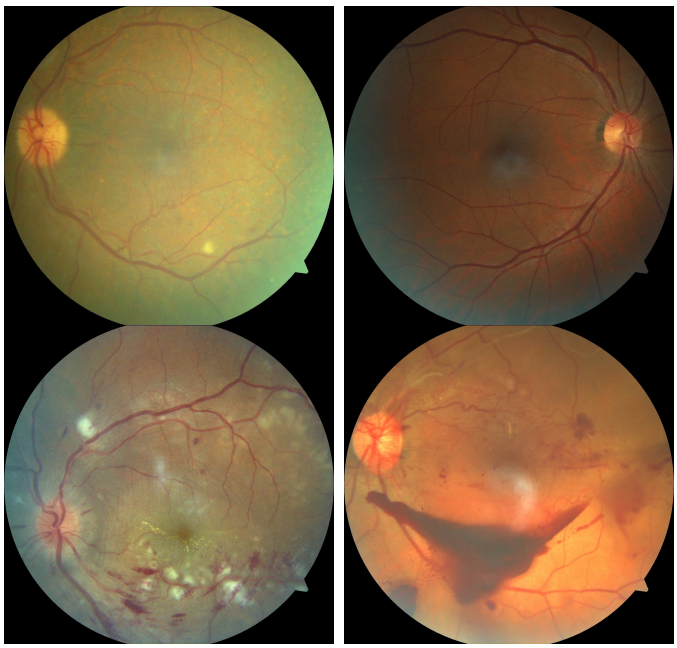

Fig. 1. Example of 4 original, unprocessed images from the DFISDRv3 dataset

The Wrocław Medical University image set WMU consists of 71 images. All images belong to the same class: Normal (no signs of DR). Figure 2 presents four photographs from the WMU base.

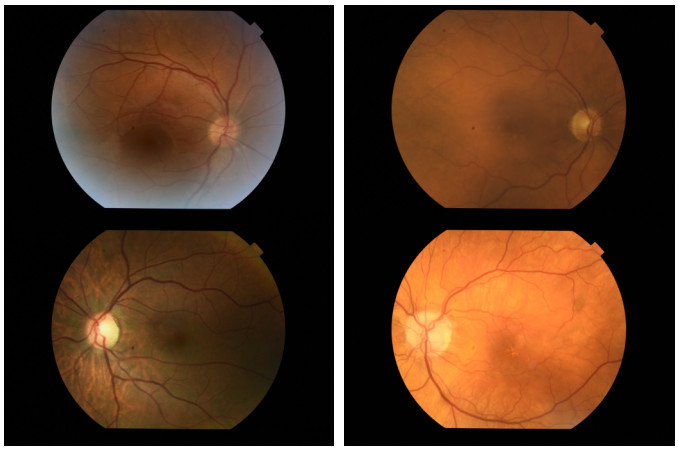

Fig. 2. Example of 4 original, unprocessed images from the WMU dataset

A. Data preprocessing

Recognizing the critical role of input fidelity, the study implemented an automated preprocessing pipeline to ensure accurate data processing. This stage involved histogram equalization to enhance local contrast, followed by a resizing operation to standardize the spatial dimensions of the DFISDRv3 dataset against the WMU samples. Due to the different profiles of the photos from both databases, a special mask was used to eliminate these differences. A specific preprocessing technique was applied in which the green channel of the fundus image was extracted, histogram-equalized to enhance contrast, and then replicated to form a 3-channel input tensor compatible with the VGG16 backbone.

B. Metrics for evaluation of the binary classification performance

To quantify the performance of the proposed solution, we adopted the standard set of evaluation metrics widely used in medical imaging tasks. These include:

a) Recall:

$$R = TP/(TP + FN)$$

b) Precision:

$$P = \frac{TP}{TP + FP}$$

c) Specificity:

$$S = \frac{TN}{TN + FP}$$

d) F1 score:

$$F1 = \frac{2TP}{2TP + FP + FN}$$

e) Accuracy:

$$ACC = \frac{TP + TN}{P + N}$$

f) bACC: bACC denotes balanced accuracy and is used when dealing with imbalanced data

$$bACC = \frac{R + S}{2}$$

g) MCC: MCC denotes the Matthews Correlation Coefficient [5]. MCC provides a comprehensive evaluation, even when classes are imbalanced. MCC is usually more informative than accuracy, F1 score, and balanced accuracy.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

where P represents the number of positive examples and N represents the number of negative examples, respectively. Positive examples can be classified correctly or erroneously. TP equals the number of positive examples correctly indicated by the classifier. TN (truly negative) refers to the number of truly detected negative examples. The false positive example number, FP, represents the number of negative instances classified as positive and FN represents the number of positive objects classified erroneously as negative.

C. Algorithms

The classification system is built on the VGG16 convolutional neural network architecture, using a transfer learning approach with weights pre-trained on the ImageNet dataset.

The VGG16 architecture was selected for its structural simplicity and proven robustness in medical imaging tasks. In contrast to more complex models, VGG16 relies on a uniform arrangement of small convolutional filters (3×3), which facilitates stable feature extraction and reduces the risk of overfitting in limited datasets. To adapt the model to detect diabetic retinopathy, the original fully connected layers were replaced with a custom classifier block. This block consists of two dense layers with 4,096 neurons, each followed by a ReLU activation function and a Dropout layer (rate = 0.5) to mitigate overfitting. The final output layer contains 5 neurons, equivalent to the severity classes labeled (0–4) defined in the study. The model accepts input images resized to a resolution of 176×176 pixels. The network is optimized using the CrossEntropyLoss function with label smoothing (0.1) to improve generalization.

The training process employs a two-stage fine-tuning strategy. Initially, the convolutional base is frozen and only custom classifier layers are trained to map high-level features to the 5 severity classes of DR. This initial phase prevents the large gradients typical of randomly initialized layers from distorting the pre-trained feature extractors (catastrophic forgetting). Upon reaching the $20th$ epoch (set specifically for this experimental setup), the entire network is unfrozen. At this critical stage, the learning rate is reduced by a factor of ten. Simultaneously, the AdamW optimizer and scheduler are re-initialized to update the weights of the full model. This approach ensures stable adaptation of feature extraction layers. As evidenced in Table IV (specifically experiment exp 1), this configuration yielded the highest validation accuracy compared to other experimental setups, confirming the benefits of the controlled fine-tuning procedure.

D. Optimizers

As an optimization method, we used the AdamW method, which works faster than its classic version, Adam [15]

AdamW [17] is a fast variant of the Adam optimizer. The AdamW optimizer distinguishes itself by decoupling weight decay from gradient adaptation. Unlike the standard L_2 regularization, which modifies the objective function, AdamW applies the decay term directly to the parameters during the update step. Separating the regularization from the optimization step in AdamW improves its generalization performance.

E. Hyper-parameter selection

The computations were performed on an Nvidia RTX 3000 Ada Generation Laptop GPU using mixed precision (FP16) to optimize performance. In every training cycle, the model utilized pixel normalization for input data, along with Dropout and L2 Regularization (Weight decay) to prevent overfitting. In addition, we have employed an adaptive learning rate to ensure stable convergence.

TABLE II
HYPER-PARAMETER SELECTION

	Exp 1	***Exp 2***	***Exp 3***	***Exp 4***	***Seq 5***
Epoch no.	50	100	100	50	50
Batch size	16	64	64	16	16
LR value	3.1e-07	1.95e-08	1.95e-08	7.81e-08	1.56e-07
Train ACC	99.36%	99.36%	99.63%	99.45%	99.54%
Test ACC	83.15%	82.78%	81.68%	80.59%	76.92%
Time	00:11:43	00:11:52	00:11:46	00:11:33	00:11:43

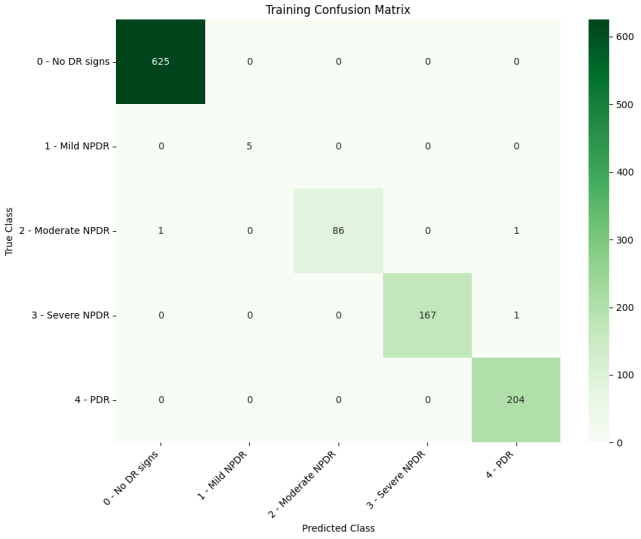

Fig. 3. Confusion matrix for training data

Based on the comparative analysis of the experimental results presented in Table I, the configuration labeled 'Seq 1' demonstrated the highest classification performance in the test set (83.15% accuracy). Consequently, the final model was established using a batch size of 16 and trained for 50 epochs. The optimization process was driven by the AdamW algorithm with an initial learning rate of 10^{-7}, which was dynamically adjusted by the adaptive scheduler to ensure optimal convergence. A dropout rate of 0.5 was maintained to effectively regularize the network.

V. EXPERIMENTS

Both sets, DFISDRv3 and WMU, were combined with all WMU images included in class 0 (no DR). The DFISDRv3 database contains 711 sample images in class 0 and the WMU consists of 71 images, all of class 0. Thus, a total of 782 images with no signs of DR are present in the combined database. Figs. 3 and 4 show the results of training the classifier and the results of testing in the form of confusion matrices for training and testing, respectively. Table III presents the results of discriminating between the class without DR (healthy eyes) and images representing different classes of DR. These results, especially the MCC measure, show that the classifier used has good discriminatory properties. In the examined database, the number of images without DR (class 0) and images with

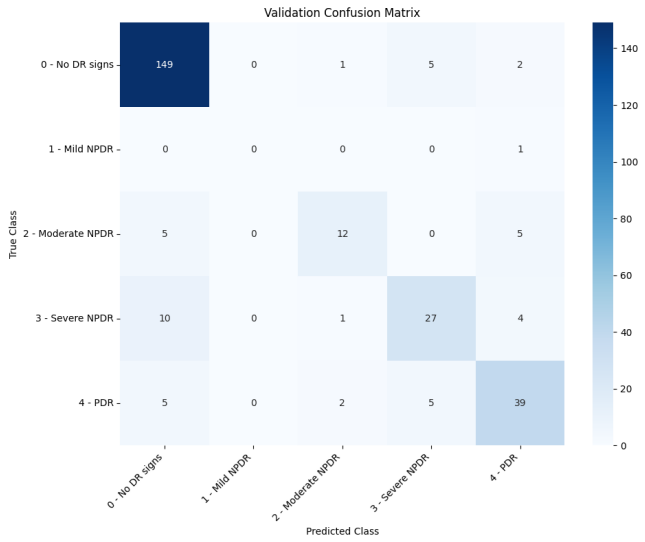

Fig. 4. Confusion matrix for testing data

TABLE III
RESULTS OF TEST CLASSIFICATION - NO DR CLASS IDENTIFICATION

ACC	Recall	Prec	Spec	F1	bACC	MCC
0.897	0.949	0.882	0.883	0.911	0.916	0.7665

DR symptoms was nearly balanced. We have used a multi-class classifier to assess the DR detection level, allowing us to also evaluate the level of identification of the remaining image classes extracted from the database. The classification of the remaining classes was more difficult due to the smaller number of representatives. We used a multi-class classifier to assess the DR detection level, allowing us to also evaluate the level of identification of the remaining image classes extracted from the database. A very low number of images from the mild NPDR class caused us to decide not to include it in the evaluation of the classification results. The identification results for the remaining classes are presented in Table IV in a similar way to the previous presentation for class 0.

TABLE IV
DISCRIMINATION RESULTS OF SYMPTOMATIC CLASSES

class	ACC	Recall	Prec	Spec	F1	bACC	MCC
m NPDR	0.949	0.545	0.462	0.984	0.632	0.765	0.614
s NPDR	0.908	0.643	0.730	0.957	0.684	0.800	0.632
PDR	0.912	0.765	0.765	0.946	0.765	0.855	0.711

VI. CONCLUSIONS

The main contributions of our article are as follows:

- The study utilized the DFISDRv3 dataset (1,437 images), supplemented by a unique collection of 71 images from Wrocław Medical University (WMU). Experiments confirmed that combining public datasets with local clinical data is a viable approach to train automatic screening systems.
- The CNN classifier, based on VGG16 with transfer learning, proved to be effective for this specific combination of data. The optimal configuration (Sequence 1) yielded a test ACC of 83.15%. This suggests that a standard architecture with limited parameters is sufficient for this classification task.
- A important contribution is the validation of a specific two-stage fine-tuning strategy (involving controlled unfreezing and optimizer re-initialization). This procedural adjustment effectively mitigated catastrophic forgetting and ensured stable convergence, proving critical for adapting the pre-trained backbone to the limited medical dataset.
- Detailed analysis showed strong discrimination between Class 0 (No DR) and Class 4 (PDR/Advanced). This suggests that the system is particularly robust in distinguishing healthy eyes from severe pathology, which is a primary requirement for medical triage.
- The evaluation of the binary classification task (Healthy vs. DR) resulted in a test accuracy of 89.7%, with a sensitivity of 94.9% and a specificity of 88.3%. The Matthews Correlation Coefficient value significantly above 0.7 indicates that the classifier makes a very good, though not perfect, prediction.
- A remaining challenge is the accurate classification of early-stage disease, specifically Mild NPDR (Class 1). We will attempt to incorporate more examples from this class to establish a comprehensive initial representation. Future work will focus on enhancing the detection of these subtle features and expanding the model's performance across the full seven-level grading scale by acquiring suitable images from WMU.

REFERENCES

[1] I.Y. Abushawish, S. Modak, E. Abdel-Raheem, A. Soliman, S.A. Mahmoud, J.A. Hussain, "Deep Learning in Automatic Diabetic Retinopathy Detection and Grading Systems: A Comprehensive Survey and Comparison of Methods", IEEE Access vol.12, pp.84785–84802, 2024.
[2] M.Z. Atwany , A.H. Sahyoun , M. Yaqub, "Deep Learning Techniques for Diabetic Retinopathy Classification: A Survey", IEEE Access vol.10, pp.28642–28655, 2022.
[3] V.E. Castillo Benítez, I. Castro Matto, J.C. Mello Román, J. Vázquez Noguera, M. García-Torres, J. Ayala, D.P. Pinto-Roa, P.E. Gardel-Sotomayor, J. Facon, S.A. Grillo, "Dataset from fundus images for the study of diabetic retinopathy", Data in Brief vol.36, 107068, 2021.
[4] N. Cheung, P. Mitchell, T. Yin Wong, "Diabetic retinopathy", The Lancet, Vol. 376, Iss. 9735, 124 – 136, 2010.
[5] D. Chicco, M. J. Warrens, G. Jurman, "The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment," in IEEE Access, vol. 9, pp. 78368–78381, 2021.
[6] J. Cuadros, G. Bresnick, "EyePACS: An adaptable telemedicine system for diabetic retinopathy screening", J. Diabetes Sci. Technol., vol.3, no.3, 509–516, 2009.
[7] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, J.C. Klein, "Feedback on a publicly distributed image database: The MESSIDOR database", Image Anal. Stereology, vol.33 , no.3, 231, 2014.
[8] Diabetic Retinopathy Detection, 2015, https://www.kaggle.com/c/diabetic-retinopathy-detection/data
[9] D. J. Derwin, S. T. Selvi, O.J. Singh, B. Priestly Shanda, "A novel automated system of discriminating Microaneurysms infundus images", Biomedical Signal Processing and Control, vol.58, 101839, 2020. https://doi.org//10.1016/j.bspc.2019.101839

[10] M. Gupta , I. R. Rao , S. P. Nagaraju , S. V. Bhandary , J. Gupta , and G.T.C. Babu, "Diabetic Retinopathy Is a Predictor of Progression of Diabetic Kidney Disease: A Systematic Review and Meta-Analysis", International Journal of Nephrology, Vol. 2022, Article ID 3922398, 11 pages, 2022. https://doi.org/10.1155/2022/3922398

[11] J. Hong, A. Surapaneni, N. Daya, E. Selvin, J. Coresh, M.E. Grams, and S. H. Ballew, "Retinopathy and Risk of Kidney Disease in Persons With Diabetes", Kidney Medicine Vol.3, no. 5, 808–815, 2021. https://doi.org/10.1016/j.xkme.2021.04.018

[12] A. Ikram, A. Imran, J. LI, A. Alzubadi, S. Fahim, A. Yasin, H. Fathi, "A Systematic Review on Fundus Image-Based Diabetic Retinopathy Detection and Grading: Current Status and Future Directions", IEEE Access vol.12, pp.96273–96303, 2024. https://0.1109/ACCESS.2024.3427394

[13] M. Karthik, D. Sohier, "APTOS 2019 Blindness Detection" https://kaggle.com/competitions/aptos2019-blindness-detection, 2019.

[14] T. Li , Y. Gao , K. Wang , S. Guo , H. Liu , H. Kang, "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening", Information Sciences, vol.501, 511–522, 2019.

[15] L. Liu, H. Jiang, P. He, W. Chen, X. Liu, J. Gao, J. Han, "On the Variance of the Adaptive Learning Rate and Beyond", Proceedings of the Eighth International Conference on Learning Representations, 2020.

[16] R. Liu et al., "DeepDRiD: Diabetic Retinopathy—Grading and Image Quality Estimation Challenge", Patterns 3, 100512, 2022. https://doi.org/10.1016/j.patter.2022.100512

[17] I. Loshchilov, F. Hutter, "Decoupled Weight Decay Regularization", ICLR, 2019. https://doi.org/10.48550/arXiv.1711.05101

[18] V. Mayya, S. Kamath, U. Kulkarni, "Automated microaneurysms detection for early diagnosis of diabetic retinopathy: A Comprehensive review", Computer Methods and Programs in Biomedicine Update vol.1, 100013, 2021.

[19] S. Majumder, N. Kehtarnavaz, "Multitasking Deep Learning Model for Detection of Five Stages of Diabetic Retinopathy", IEEE Access vol.9, 123220–123230, 2021. https://doi.org/10.1109/ACCESS.2021.3109240

[20] R.F. Mansour, "Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy", Biomed. Eng. Lett. vol.8, no.1, pp.41–57, 2018. https://doi.org/10.1007/s13534-017-0047-y

[21] A. M. Mutawa, S. Alnajdi, S. Sruthi, "Transfer Learning for Diabetic Retinopathy Detection: A Study of Dataset Combination and Model Performance", Applied Sciences (MDPI) vol.13, 5685, 2023. https://doi.org/10.3390/app13095685

[22] D. Nagpal, S.N. Panda, M. Malarvel, P.A. Pattanaik, K. Zubair Khan, "A review of diabetic retinopathy: Datasets, approaches, evaluation metrics and future trends", Journal of King Saud University – Computer and Information Sciences vol34, pp.7138–7152, 2022.

[23] W. Nazih, A.O. Aseeri, O.Y. Atallah, S. EL-Sappagh, "Vision Transformer Model for Predicting the Severity of Diabetic Retinopathy in Fundus Photography-Based Retina Images", IEEE Access, Vol.11, pp. 117546–117561, 2023. https://doi.org/10.1109/ACCESS.2023.3326528

[24] A. Papadopoulos, F. Topouzis, A. Delopoulos, "An interpretable multiple-instance approach for the detection of referable diabetic retinopathy in fundus images", Scientific Reports (Nature) vol.11, 14326, 2021. https://doi.org/10.1038/s41598-021-93632-8

[25] K. Pin, J. Ho Chang and Y. Nam, "Comparative Study of Transfer Learning Models for Retinal Disease Diagnosis from Fundus Images", Computers, Materials & Continua, vol. 70, no.3, pp.5821–5833, 2022. https://doi.org/10.32604/cmc.2022.021943

[26] P. Porwal, S. Pachade , R. Kamble , M. Kokare, G. Deshmukh, V. Sahasrabuddhe V., F. Meriaudeau, "Indian Diabetic Retinopathy Image Dataset (IDRiD)", IEEDataPort, 2018. doi:10.21227/H25W98 https://doi.org/10.3390/data3030025

[27] G. Rajarajeshwarti, G. Chemmalar Selvi, "Application of Artificial Intelligence for Classification, Segmentation, Early Detection, Early Diagnosis, and Grading of Diabetic Retinopathy From Fundus Retinal Images", A Comprehensive Review. IEEE Access vol.12, pp.172499–172536, 2024. https://doi.org/10.1109/ACCESS.2024.3494840

[28] A. Sebastian, O. Elharrouss, S. Al-Maadeed, N. A. Almaadeed, "A Survey on Deep-Learning-Based Diabetic Retinopathy Classification", Diagnostics, vol.13, 345, 2023. https://doi.org/10.3390/diagnostics13030345

[29] I. Soares, M. Castelo-Branco, A. Pinheiro, "Microaneurysms detection in retinal images using a multi-scale approach", Biomedical Signal Processing and Control, vol.79, pp.104184–104195, 2023.

[30] V. Srisravya, P. Naga Srinivasu, J. Shafi, W. Hołubowski, A. Zielonka, "Advanced diabetic retinopathy detection with the R–CNN: a unified visual healthy solution", Int. J. Appl. Math. Comput. Sci., Vol. 34, No. 4, 661–678, 2024. https://doi.org/10.61822/amcs-2024-0044

[31] M.B. Sudhan, M. Sinthuja, S. Pravinth Raja, J. Amutharaj, G. Charlyn Pushpa Latha, S. Sheeba Rachel, T. Anitha, T. Rajendran , Y. Asrat Waji "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model", Journal of Healthcare Engineering, vol 2022, Article ID 1601354, 2022. https://doi.org/10.1155/2022/1601354

[32] Wan S., Liang Y., Zhang Y.: Deep convolutional neural networks for diabetic retinopathy detection by image classification", Computers and Electrical Engineering, vol.72, pp. 274–282, 2018.

[33] W. Zhang , J. Zhong , S. Yang , Z. Gao , J. Hua, Y. Chen, Y. Zhang, "Automated identification and grading system of diabetic retinopathy using deep neural networks", Knowledge-Based Systems vol. 175, pp.12–25, 2019.

[34] Z. Yang , T-E. Tan, Y. Shao , T.Y. Wong, X. Li, "Classification of diabetic retinopathy: Past, present and future", Frontiers in Endocrinlogy, 2022. DOI:10.3389/fendo.2022.1079217

[35] M. Karthik, D. Sohier, "APTOS 2019 Blindness Detection" https://kaggle.com/competitions/aptos2019-blindness-detection, 2019.

[36] M. Zmonarski, E. Skubalska-Rafajlowicz, A. Zgryźniak, S. Zmonarski, "Creating an Automatic Classifier for Detecting Diabetic Retinopathy Based on the Combining a New 7-Class Dataset Fundus Images and a Collection of Images from Another Source", W. Zamojski et al. (Eds.): DepCoS-RELCOMEX 2025, LNNS 1427, pp. 246–255, 2025.