Leveraging Convolutional Neural Networks for Automated Detection and Grading of Diabetic Retinopathy from Fundus Images

Ibnu Uzail Yamani[1] and Basari[1,2*]

[1]*Biomedical Engineering Program, Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia Kampus Baru Universitas Indonesia, Depok, Jawa Barat, 16424, Indonesia*

[2]*Research Center for Biomedical Engineering, Faculty of Engineering, Universitas Indonesia Kampus Baru Universitas Indonesia, Depok, Jawa Barat, 16424, Indonesia*

[*]*Corresponding author. Email: basari.st@ui.ac.id*

Abstract— This study addresses the critical challenge of Diabetic Retinopathy (DR) detection and severity grading, aiming to advance the field of medical image analysis. The research problem focuses on the need for an accurate and efficient model to discern DR conditions, thereby facilitating early diagnosis and intervention. Employing a Convolutional Neural Network (CNN), our methodology is developed to strike a balance between precision and computational efficiency, a pivotal aspect in the context of healthcare applications. The research leverages the APTOS 2019 dataset, a comprehensive collection of fundus photographs, to evaluate the efficacy of our proposed model. The dataset allows for a thorough investigation into the model's performance in binary-class and multi-class classifications, providing a robust foundation for analysis. The most important result of our study manifests in the achieved accuracy rates of 98.67% and 87.81% for binary-class and multi-class classifications, respectively. These outcomes underscore the model's reliability and innovation, surpassing established machine learning algorithms and affirming its potential as a valuable tool for early DR detection and severity assessment. In conclusion, the study marks a significant advancement in leveraging deep learning for ophthalmic diagnoses, particularly in the nuanced landscape of DR. The implications of our findings extend to the broader realm of AI-driven healthcare solutions, presenting opportunities for enhanced clinical practices and early intervention strategies. Future research endeavors could explore further refinements to the model, considering additional datasets and collaborating with healthcare professionals for real-world validation, ensuring the continued progress of AI applications in the medical domain.

Keywords— APTOS 2019; Convolutional Neural Network; deep learning; Diabetic Retinopathy; early diagnosis; fundus image

I. Introduction

Diabetic Retinopathy (DR) is a condition that affects the blood vessels in the retina due to poor blood sugar control [1] . It is a complication of diabetes mellitus and is the leading cause of blindness [2]. The disease is characterized by damage to the blood vessels in the retina, which can result in vision loss [3]. It is caused by chronic effects of diabetes mellitus and is considered an inflammatory, neuro-vascular complication [4]. The condition can be prevented or treated if identified in its early stages [5]. Early detection and treatment are crucial to address neurovascular damage before clinical microvascular damage occurs. DR can be diagnosed using specific algorithms and neural networks that analyze retinal images. Timely monitoring and diagnosis of DR can help prevent vision loss and improve patient outcomes.

Deep learning has shown promise in the detection and classification of DR [6]. Various deep learning models, such as DenseNet-121, VGG16, and MobileNetV2, have been used for this purpose [7]–[9]. These models have been trained and tested on publicly available datasets, such as the APTOS 2019 Blindness Detection Kaggle Dataset [10]. Data augmentation techniques, such as Enhanced Super-resolution Generative Adversarial Networks (ESRGAN) and Histogram Equalization (HIST), have been employed to enhance image quality and balance imbalanced datasets. These methods have the potential to assist ophthalmologists in the early detection and diagnosis of DR, leading to timely treatment and reduced risk of vision loss.

Deep learning research for DR detection with the APTOS 2019 dataset has identified several shortcomings that need improvement. These include imbalanced datasets, inconsistent annotations, limited sample images, inappropriate performance evaluation metrics, and the high cost of large annotated datasets [7], [11]. The imbalanced image class distribution in the APTOS 2019 dataset has been addressed through appropriate balancing techniques [12]. Furthermore, the accuracy of the models used for DR detection and classification needs improvement, with the hybrid network achieving an accuracy of 79.50% and the DenseNet 121 model achieving an accuracy of 97.30% [13]. The use of deep learning techniques, such as Convolutional Neural Networks (CNNs) and transfer learning, has shown promise in simplifying the detection step and improving the accuracy of DR diagnosis [14]. However, further research is needed to overcome these shortcomings and enhance the performance of deep learning models for DR detection with the APTOS 2019 dataset.

Received 15 November 2023, Revised 27 March 2024, Accepted 28 March 2024.

DOI: https://doi.org/10.15294/jte.v15i2.48769

To bridge these gaps, this paper proposes a novel deep learning method based on CNNs with a custom architecture, allowing for easier fine-tuning and the incorporation of data augmentation techniques. This approach aims to overcome the shortcomings identified in previous research and enhance the overall performance of deep learning models for DR detection with the APTOS 2019 dataset. The motivation behind this study lies in the potential to improve early detection and diagnosis of DR, thereby reducing the risk of vision loss and ultimately benefiting patients with diabetes.

The Methods section details our deep learning approach, the custom CNN architecture, and augmentation techniques for improved DR detection. In Results and Discussion, we evaluate the model's accuracy and clinical relevance. The Conclusion highlights our findings' impact and suggests research directions to better detect and diagnose DR, helping to prevent vision loss in diabetics.

II. Method

A. Dataset

This investigation employs the APTOS 2019 dataset, available as an integral component of the Kaggle Blindness Detection Challenge – 2019 [15]. The dataset comprises 3662 fundus eye photography images categorized into five severity levels of DR classes (0, 1, 2, 3, 4), sequentially named as Normal, Mild, Moderate, Severe, and Proliferative Diabetic Retinopathy (PDR). It is imperative to cite the aforementioned reference for further contextualization. The dataset encompasses 1805 images denoted as No DR, 370 images of Mild DR, 999 images portraying Moderate DR, 193 images representing Severe DR, and 295 images depicting Proliferative DR, as delineated in Table I.

This research capitalizes on the rich diversity encapsulated within the APTOS 2019 dataset, specifically designed for the Kaggle Blindness Detection Challenge. The dataset's stratification into distinct severity levels of DR, ranging from Normal to Proliferative DR, offers a nuanced and comprehensive framework for model training and evaluation. The tabulated distribution of images across severity classes provides a clear snapshot of the dataset's composition, laying the foundation for an in-depth exploration into the intricate patterns characterizing different DR stages. Furthermore, the robustness of the research methodology hinges on the utilization of this meticulously curated dataset, ensuring a substantive and rigorous analysis of DR detection and severity assessment.

TABLE I. SUMMARY OF APTOS 2019 DATASET

DR Classes	Number of Instance
Normal	1805
Mild	370
Moderate	999
Severe	193
PDR	295
Total	3662

B. Data Pre-Processing

The DR dataset exhibits an imbalance, with 49% allocated to the No DR class, 8% to PDR, 5% to Severe, 27% to Moderate, and 10% to Mild, posing a potential risk of overfitting. To enhance the model's performance, refinement of the DR dataset is conducted through the implementation of Augmentation techniques for both the two-class and five-class scenarios. Additionally, all images are resized to 224×224 pixels as part of the preprocessing steps to ensure consistency and compatibility with the model architecture. Augmentation involves the adjustment of brightness and contrast, strategically employed to address potential bias and enrich the dataset. The outcomes of Augmentation for the two-class and five-class scenarios are presented in Table II.

This strategic refinement aims to mitigate the imbalanced class distribution within the DR dataset, a common concern that could compromise the model's generalizability. Augmentation, specifically through brightness and contrast adjustments, serves as a nuanced technique to introduce diversity to the dataset, thus minimizing the risk of overfitting. The rationale behind this methodology stemmed from the acknowledgment that the quality of images profoundly influences the outcomes of detection, with fluctuations in brightness and contrast being prevalent in practical settings. Through systematic manipulation of these variables, our objective was to replicate a spectrum of environmental conditions and fortify the adaptability of our model to fluctuations in image quality. This augmentation procedure facilitated the expansion of our dataset, consequently enhancing the model's capacity for generalization. The augmentation outcomes, delineated in Table II, underscore the tangible impact of this method on enhancing the dataset's representativeness and, by extension, fortifying the robustness of the subsequent model. Additionally, we utilized augmentation three times to ensure balance across classes and to mitigate significant disparities in dataset size between classes. To enhance detection quality, we applied Gaussian blur to the images. This method aims to reduce noise and smooth out irrelevant details in the images, allowing the detection algorithm to focus on more significant features. Consequently, the use of Gaussian blur helps improve the precision and accuracy of detection in the processed images.

TABLE II. AUGMENTATION OUTCOMES FOR 2 AND 5-CLASSES CONFIGURATIONS

DR Classes	Original Sample	2-classes	5-classes
Normal	1805	2x Augmented	Not Augmented
Mild	370	3x Augmented	3x Augmented
Moderate	999	Not Augmented	Not Augmented
Severe	193	3x Augmented	3x Augmented
PDR	295	3x Augmented	3x Augmented

C. CNN Architecture

In recent times, the emergence of Deep Learning has significantly transformed various domains, particularly in the areas of pattern recognition and image processing. Deep Learning, renowned for its capacity to autonomously acquire layered representations of data, has become a formidable asset in addressing intricate challenges. Among the array of techniques within this framework, CNNs have emerged as a pivotal method, demonstrating impressive capabilities in extracting pertinent features from raw data, particularly in the domain of image analysis. In this investigation, we capitalize on the potential of CNNs to discern nuanced patterns within images. Through the utilization of CNNs, we not only enhance the precision and efficiency of our model but also uncover fresh perspectives that transcend conventional approaches. By meticulously scrutinizing CNN architectures and their application within our research domain, we underscore the

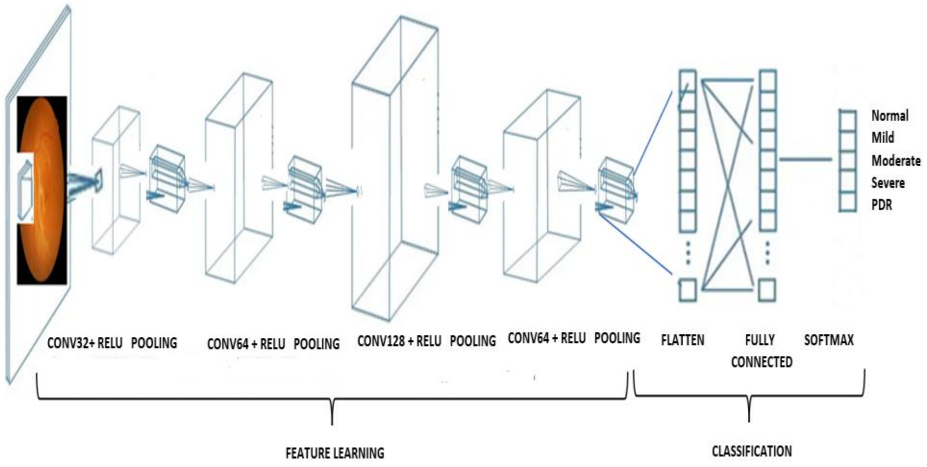

Figure 1. The proposed CNN architecture

unique merits of Deep Learning, specifically CNNs, in propelling progress and offering comprehensive resolutions to complex issues.

Within this investigation, the architecture of a CNN assumes the role of a discerning model dedicated to the identification of DR. Depicted in Figure 1, the intricate CNN architecture unfolds through a series of strategically orchestrated layers. This model, a CNN, is meticulously structured with sequential layers designed to systematically process intricate image data. Commencing the convolutional journey is the inaugural Conv2D layer boasting 32 filters of dimensions 3×3, invoking ReLU activation to adeptly discern intricate features embedded within the visual data. Significantly, this layer incorporates L2 regularization, employing a judicious coefficient of 0.001, to safeguard against potential overfitting. Subsequent to this, a MaxPooling2D layer with dimensions 2×2 is introduced, serving the pivotal purpose of diminishing the spatial dimensions of the processed image.

The convolutional cascade further unfolds through the incorporation of the second, third, and fourth Conv2D layers, each distinct in filter quantity and configuration, yet consistently adhering to the principles of ReLU activation and L2 regularization. Notably, every convolutional stratum is succeeded by a congruent MaxPooling2D layer, ensuring a seamless progression. Following this convolutional continuum, the images undergo a transformative shift, transmogrifying into one-dimensional vectors through the agency of the Flatten layer. This heralds the entrance of a subsequent Dense layer, boasting 32 neurons and invoking ReLU activation to facilitate profound learning from the gleaned features. Culminating in the architectural design is the imposition of a Dropout layer, strategically set at a 0.5 dropout rate, serving as a prudent measure against overfitting. The model's denouement materializes in the form of a concluding Dense layer, featuring 2 neurons and deploying softmax activation, thus effectuating the classification of images into the desired dual classes. This intricate and sophisticated CNN architecture stands as a testament to the meticulous approach undertaken in the realm of image classification.

D. Detection Performance Evaluation

In the evaluation stage, the meticulously crafted CNN model from the antecedent phase will undergo rigorous testing. This critical phase aims to furnish insights through the presentation of a confusion matrix, delineating pivotal metrics including accuracy, precision, recall, and F1-score. The scrutiny of these metrics serves as a comprehensive assessment, providing a nuanced understanding of the model's performance and its aptitude in discerning patterns within the dataset. This phase serves as a decisive juncture in gauging the efficacy of the CNN model, elucidating its proficiency in classification tasks and contributing to the elucidation of its practical viability.

1) Accuracy

The metric of accuracy serves as an indicator of the model's precision in classification, characterized by Equation (1).

$$Accuracy = \frac{TN+TP}{FP+TN+TP+FN} \tag{1}$$

2) Precision

Precision serves as a metric to elucidate the precision exhibited by a model in forecasting positive events amidst a sequence of predictive endeavors as indicated by Equation (2).

$$Precision = \frac{TP}{FP+TP} \tag{2}$$

3) Recall

In tandem with precision and accuracy metrics, a comprehensive evaluation of a system's efficacy necessitates the consideration of recall or its sensitivity to a specific class, as elucidated by Equation (3):

$$Recall = \frac{TP}{FN+TP} \tag{3}$$

4) F1-Score

The F1-Score, or F-Measure, is an evaluation metric that combines precision and recall into a single value as indicated by Equation (4). In the context of classification, the F1-Score provides a comprehensive overview of the balance between a model's ability to make accurate positive predictions (precision) and its ability to identify all true positive cases (recall). By amalgamating these two metrics, the F1-Score offers a holistic measure of the overall quality of the model, with the highest value reflecting an optimal balance between precision and recall. Therefore, the F1-Score is particularly valuable in situations where the need to optimize both precision and recall is equally critical, such as in disease detection or classifying imbalanced data scenarios.

$$F1 - Score = \frac{2\ X\ Precision\ X\ Recall}{Precision+Recall} \tag{4}$$

III. Results and Discussion

A. Two-classes Classification

In this study, a CNN model has been successfully developed and tested to classify fundus images into two categories: normal and DR. The model underwent training on a significantly large dataset with optimized training parameters, including a test size of 0.15, Adam optimizer with a learning rate of 0.001, 75 epochs, and a batch size of 32.

The testing results revealed that the model achieved an accuracy rate of 98.67%, accompanied by precision and recall rates of 98.68% and 98.64%, respectively, and an F1-Score of 98.66%. The Confusion Matrix presented in Table III illustrates minimal errors in the model's identification of normal and DR conditions, showcasing strong potential for early and accurate detection of DR in clinical applications. This underscores the model's robust performance and underscores its viability for clinical implementation in the realm of DR diagnosis.

TABLE III. The Results of Confusion Matrix for 2-Classes

Predicted	**Actual**	
	Normal	**DR**
Normal	533	9
DR	7	657

Accuracy	98.67%
Precision	98.68%
Recall	98.64%
F1-Score	98.66%

The graphical representation of the CNN model's performance, as delineated in Figure 2, provides a nuanced insight into its efficacy in distinguishing between normal ocular conditions and instances of DR. During the initial epochs of training, the model displayed a commendable efficiency in adapting its weight parameters, as evidenced by the rapid descent of its loss function. This signifies a proficient learning process, where the model swiftly aligns itself with the inherent complexities of the dataset. Notably, this trend persevered until the loss approached a state of stability, indicating a careful convergence towards an optimal model configuration.

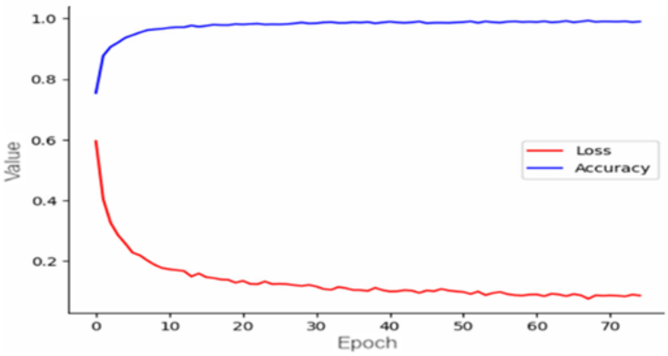

Figure 2. Loss and accuracy in DR 2-classes classification

Moreover, the discernible ascent in accuracy throughout the training process is a noteworthy observation. The model's ability to sharply improve its accuracy towards nearly flawless values, nearing the theoretical maximum of 100%, underscores its adeptness in correctly classifying instances of normal and DR conditions. The swift and consistent enhancement in accuracy, coupled with a steady reduction in loss, suggests that the model successfully captured pertinent patterns within the data.

Importantly, the absence of signs of overfitting, as manifested by the flatlining of the loss curve in the latter epochs, is a crucial facet. This implies that the model's learning process was well-balanced, avoiding undue adaptation to the training data at the expense of generalizability. The plateau observed in the loss curve signifies a point of stability, further substantiating the model's potential for optimal learning and robust performance on previously unseen data.

B. Five-Classes Classification

Within the context of this investigation, a meticulously designed and rigorously tested CNN model has been developed to classify fundus images into five distinct categories: Normal, Mild, Moderate, Severe, and PDR. A crucial aspect of the model's development strategy involved thorough training on a comprehensive dataset. The training parameters, which were meticulously adjusted, included a test size ratio of 0.15, utilization of the Nadam optimizer with a learning rate of 0.001, a training duration of 75 epochs, and a batch size of 32.

Importantly, the model demonstrates remarkable proficiency in accurately identifying various stages of retinopathy, as evidenced by its commendable performance metrics. The model achieves an accuracy rate of 87.81%, indicating its precision in making correct classifications. Additionally, the precision score of 86.60%, recall rate of 86.08%, and F1-score of 86.30% collectively highlight the model's effectiveness in navigating the complexities of this intricate classification system. A comprehensive analysis of the classification outcomes for each specific class is meticulously provided in Table IV, offering a detailed understanding of the model's discriminative capabilities.

The graphical representation of the model's training trajectory, spanning 67 epochs, offers valuable insights into the dynamic learning process, as depicted in Figure 3. Notably, the graph unveils an initial sharp descent in loss, transitioning into a more gradual descent, indicative of the model's convergence and sustained progression over time. Concurrently, accuracy experiences a noteworthy surge in the early stages, eventually reaching a stabilized state at elevated levels, albeit not achieving full saturation. This discernible pattern implies the model's adept parameter adaptation to the dataset, effectively assimilating crucial features from retinal imagery for precise DR classification. While the model's performance demonstrates considerable promise for medical diagnostic applications, prudent adjustments and additional validation endeavors may be imperative to mitigate cross-classification errors and fortify the predictive reliability.

TABLE IV. The Results of the Confusion Matrix for 5-Classes.

Predicted	**Actual**				
	Normal	**Mild**	**Moderate**	**Severe**	**PDR**
Normal	344	6	9	0	2
Mild	7	273	10	1	4
Moderate	14	19	136	12	19
Severe	0	1	11	134	9
PDR	1	7	18	2	208

Accuracy	87.81%
Precision	86.60%
Recall	86.08%
F1-Score	86.30%

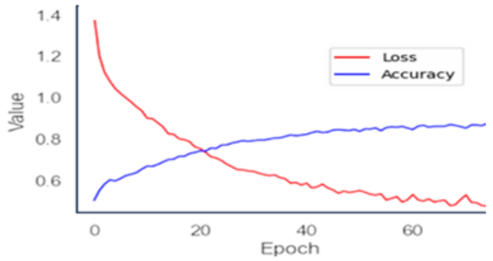

Figure 3. Loss and accuracy in 5-classes DR classification

C. Discussions

This research endeavors to draw a comprehensive comparison between our proposed methodology and established machine learning algorithms. Table V serves as a visual representation, delineating the nuances between antecedent studies and the innovative approach we put forth, all within the confines of a shared dataset.

In this examination, we systematically evaluate the efficacy of our proposed method by placing it side by side with conventional machine learning algorithms that have found application in analogous studies. The intention is to discern the unique attributes and advancements that our methodology introduces to the field. Table V emerges as a succinct visual aid, elucidating the distinctive features and comparative performance metrics across various research endeavors. This juxtaposition not only contributes to the validation of our approach but also underscores its potential as a noteworthy augmentation within the broader landscape of machine learning applications. Moving beyond binary classification, Table VI showcases the robustness of our methodology in the more challenging multi-class setting, reflecting our model's proficiency in handling a nuanced classification landscape and its promising application in complex medical diagnostic tasks.

TABLE V. COMPARISON OF THE PROPOSED RESULTS WITH THE PREVIOUS RESEARCH FOR 2-CLASSES.

Method	Year	Accuracy (%)
Composite DNN [16]	2021	97.82
Xception+VGG16+ DNN [17]	2020	97,92
DenseNet121 [18]	2020	94.44
VGG16 [19]	2021	97.05
Supervised Contrastive learning [20]	2022	98.36
This study	2023	**98.67**

TABLE VI. COMPARISON OF THE PROPOSED RESULTS WITH THE PREVIOUS RESEARCH FOR 5-CLASSES

Method	Year	Accuracy (%)
Composite DNN [16]	2021	82,54
Xception+VGG16+DNN [17]	2020	80,96
Deep Convolution Features+SVM [21]	2020	77.90
Inception-ResNet-v2+CNN [22]	2021	82.18
VGG16 [19]	2021	75.50
Supervised Contrastive Learning [20]	2022	84.36
This Study	2023	**87.81**

In the comparative analysis of binary-class classification for the identification of DR within the APTOS 2019 dataset, findings reveal that the method propounded in 2023, denoted as "our method," achieved an apex accuracy of 98.67%. This supremacy is conspicuous when contrasted with antecedent investigations such as Xception+VGG16+DNN [17] 97.92%, DenseNet121 [18] 94.44%, Composite DNN [16] 97.82%, and VGG16 [19] 97.05%. Notably, the accuracy of our method surpasses that of the study conducted by Supervised Contrastive learning [20] in 2022, registering at 98.36%. This nuanced advancement underscores the superior discriminative prowess of our proposed methodology in effecting the categorization of ocular states into the dichotomy of Normal and DR, vis-à-vis antecedent methodologies.

Within the intricate domain of multi-class classification pertaining to the nuanced stratification of DR severity, our proposed method in 2023 attains a commendable accuracy of 87.81%. This substantial achievement markedly transcends the outcomes of prior investigations including Deep Convolution Features+ SVM [21] 77.90%, Xception+VGG16+DNN [17] 80.96%, VGG16 [19] 75.50%, and Supervised Contrastive Learning [20] 84.36%. Although Inception-ResNet-v2+CNN [22] study in 2021 yielded an accuracy of 82.18%, our method exhibits a sustained trajectory of improvement. The pinnacle accuracy achieved by our method accentuates the model's proficiency in discerning ocular conditions across a spectrum of severity levels, thereby bearing consequential implications for clinical praxis.

The consistent augmentation in performance, particularly discernible in the realm of multi-class classification, positions the proposed method as an invaluable contributory milestone in the evolution of DR detection methodologies [23]. These empirical findings substantiate the methodological reliability and discriminatory acuity of the proposed framework in the context of DR detection utilizing the APTOS 2019 dataset [24]. The inference derived is that this methodological innovation may serve as a seminal underpinning for subsequent strides in the diagnosis and therapeutic intervention for ocular pathologies concomitant with DR [25].

In the domain of DR detection, our approach stands as a pioneering instance of deep learning methodology, characterized by a specifically tailored CNN architecture. This unique architecture, deliberately configured to find an equilibrium between effectiveness and computational efficiency [26], played a pivotal role in achieving a noteworthy accuracy of 98.67%. This surpasses the performance of established methodologies in binary-class classification, underscoring the efficacy of a meticulously designed CNN architecture in distinguishing between Normal and DR conditions [27].

To contend with imbalanced class classification challenges, our method strategically integrates advanced data augmentation techniques. Notably, the utilization of contrast and brightness adjustments, chosen for their ability to mimic real-world image variations, contributes to the model's adaptability [28]. Furthermore, preprocessing methodologies, including Gaussian blurring, enhance the clarity of retinal objects, thereby amplifying the model's discernment capabilities. This comprehensive methodological framework has yielded a commendable accuracy of 87.81% in the nuanced landscape of multi-class DR severity classification, surpassing precedent studies.

Our approach succeeds through the intelligent integration of a bespoke CNN architecture—balancing accuracy and efficiency—with advanced data preprocessing for real-world applicability [29]. Together, these elements enhance robustness in binary and multi-class DR classification [30], evidencing our method's effectiveness and conceptual depth.

IV. Conclusion

In conclusion, our study demonstrates a significant leap in DR detection and severity grading with a CNN model, achieving remarkable accuracy rates of 98.67% in binary classifications and 87.81% in multi-class scenarios, outperforming conventional machine learning approaches. This advancement is bolstered by effective data augmentation and preprocessing strategies, enhancing the model's adaptability and precision in medical imaging contexts. However, the challenge of model interpretability remains, suggesting a need for future research focused on improving transparency through methods like attention mechanisms and feature visualization. Future studies should also explore refining the CNN architecture, expanding data augmentation techniques, and extending applications to diverse datasets, alongside real-world validation with healthcare professionals. Our findings advocate for the potential of deep learning in ophthalmology, setting the stage for further advancements in AI-driven diagnostic tools that could revolutionize clinical practices and patient care.

Acknowledgment

The research was partly supported by PUTI Q2 2022 Grant No. NKB-1323/UN2.RST/HKP.05.00/2022 from Universitas Indonesia.

References

[1] M. Deore, A. Kshirsagar, V. Patel, V. Patel, and V. Phaltankar, "Diabetes Retinopathy Detection," *IJSREM*, vol. 07, no. 03, Mar. 2023, doi: 10.55041/IJSREM18164.

[2] S. M. Mohammed, Z. J. Ali, and S. S. Najam, "Diagnosis of Retinopathy in Patients Diabetes," in *2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)*, Bhilai, India: IEEE, Jan. 2023, pp. 1–4. doi: 10.1109/ICAECT57570.2023.10118153.

[3] J. Mathias, S. Gadkari, M. Payapilly, and A. Pansare, "Categorization of Diabetic Retinopathy and Identification of Characteristics to Assist Effective Diagnosis," in *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India: IEEE, Mar. 2021, pp. 801–806. doi: 10.1109/ESCI50559.2021.9396908.

[4] T. H. Fung, B. Patel, E. G. Wilmot, and W. M. Amoaku, "Diabetic retinopathy for the non-ophthalmologist," *Clin Med*, vol. 22, no. 2, pp. 112–116, Mar. 2022, doi: 10.7861/clinmed.2021-0792.

[5] S. H. Sinclair and S. S. Schwartz, "Diabetic Retinopathy–An Underdiagnosed and Undertreated Inflammatory, Neuro-Vascular Complication of Diabetes," *Front. Endocrinol.*, vol. 10, p. 843, Dec. 2019, doi: 10.3389/fendo.2019.00843.

[6] G. Alwakid, W. Gouda, and M. Humayun, "Enhancement of Diabetic Retinopathy Prognostication Using Deep Learning, CLAHE, and ESRGAN," *Diagnostics*, vol. 13, no. 14, p. 2375, Jul. 2023, doi: 10.3390/diagnostics13142375.

[7] C. Mohanty *et al.*, "Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy," *Sensors*, vol. 23, no. 12, p. 5726, Jun. 2023, doi: 10.3390/s23125726.

[8] A. K. Nivedha, Rahini. T, S. Banu. M, Sowmiya. V, and Sreelakshmi. T, "Detection and Classification of Diabetic Retinopathy Using Deep Learning," *IJRASET*, vol. 11, no. 5, pp. 856–860, May 2023, doi: 10.22214/ijraset.2023.51626.

[9] T. Swapna, D. Akhila, P. Srija, T. Shivani, and P. Srinidhi, "Diabetic Retinopathy Classification using Transfer learning," *IJATCSE*, vol. 12, no. 3, pp. 110–116, Jun. 2023, doi: 10.30534/ijatcse/2023/021232023.

[10] Payal Nannewar and Dr. Sanjay. L. Haridas, "Diabetic Retinopathy Detection Through Deep Learning Techniques," *IJSRST*, pp. 729–734, Apr. 2023, doi: 10.32628/IJSRST523102105.

[11] M. Saini and S. Susan, "Diabetic retinopathy screening using deep learning for multi-class imbalanced datasets," *Computers in Biology and Medicine*, vol. 149, p. 105989, Oct. 2022, doi: 10.1016/j.compbiomed.2022.105989.

[12] M. Bhavya, M. Anush, R. Gagan, H. Gowtham, and U. Yashwanth, "Detection of Diabetic Retinopathy Using Deep Learning," *IJRASET*, vol. 10, no. 7, pp. 1880–1886, Jul. 2022, doi: 10.22214/ijraset.2022.45597.

[13] T. K. Shivaprasad, Sameeksha, K. S. Sathvik, and Y. P. Yashawantha, "Early Diabetic Retinopathy Detection using Deep Learning," *IJRASET*, vol. 11, no. 4, pp. 3967–3971, Apr. 2023, doi: 10.22214/ijraset.2023.51209.

[14] N. Duraichi, S. Jalaja, C. D. Merlin, J. S. Meena, R. N. Kamali, and K. Manoj, "Detection and Classification of Diabetic Retinopathy using Deep Learning," *CM*, no. 26, pp. 808–813, Mar. 2023, doi: 10.18137/cardiometry.2023.26.808813.

[15] "APTOS 2019 Blindness Detection." Accessed: Nov. 15, 2023. [Online]. Available: https://kaggle.com/competitions/aptos2019-blindness-detection

[16] J. D. Bodapati, N. S. Shaik, and V. Naralasetti, "Composite deep neural network with gated-attention mechanism for diabetic retinopathy severity classification," *J Ambient Intell Human Comput*, vol. 12, no. 10, pp. 9825–9839, Oct. 2021, doi: 10.1007/s12652-020-02727-z.

[17] J. D. Bodapati *et al.*, "Blended Multi-Modal Deep ConvNet Features for Diabetic Retinopathy Severity Prediction," *Electronics*, vol. 9, no. 6, p. 914, May 2020, doi: 10.3390/electronics9060914.

[18] S. S. Chaturvedi, K. Gupta, V. Ninawe, and P. S. Prasad, "Automated Diabetic Retinopathy Grading using Deep Convolutional Neural Network," *ARXIV*, 2020, doi: 10.48550/ARXIV.2004.06334.

[19] G. Kumar, S. Chatterjee, and C. Chattopadhyay, "DRISTI: a hybrid deep neural network for diabetic retinopathy diagnosis," *SIViP*, vol. 15, no. 8, pp. 1679–1686, Nov. 2021, doi: 10.1007/s11760-021-01904-7.

[20] M. R. Islam *et al.*, "Applying supervised contrastive learning for the detection of diabetic retinopathy and its severity levels from fundus images," *Computers in Biology and Medicine*, vol. 146, p. 105602, Jul. 2022, doi: 10.1016/j.compbiomed.2022.105602.

[21] V. Dondeti, J. Bodapati, S. Shareef, and V. Naralasetti, "Deep Convolution Features in Non-linear Embedding Space for Fundus Image Classification," *RIA*, vol. 34, no. 3, pp. 307–313, Jun. 2020, doi: 10.18280/ria.340308.

[22] A. K. Gangwar and V. Ravi, "Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning," in *Evolution in Computational Intelligence*, V. Bhateja, S.-L. Peng, S. C. Satapathy, and Y.-D. Zhang (Eds.), Advances in Intelligent Systems and Computing, vol. 1176, Singapore: Springer Singapore, 2021, pp. 679–689. https://doi.org/10.1007/978-981-15-5788-0_64

[23] M. Vijayan and V. S, "A Regression-Based Approach to Diabetic Retinopathy Diagnosis Using Efficientnet," *Diagnostics*, vol. 13, no. 4, p. 774, Feb. 2023, doi: 10.3390/diagnostics13040774.

[24] S. Haghighi, M. Jasemi, S. Hessabi, and A. Zolanvari, "PyCM: Multiclass confusion matrix library in Python," *JOSS*, vol. 3, no. 25, p. 729, May 2018, doi: 10.21105/joss.00729.

[25] Z. Xu, J. Rittscher, and S. Ali, "SSL-CPCD: Self-supervised learning with composite pretext-class discrimination for improved generalisability in endoscopic image analysis," *ARXIV*, 2023, doi: 10.48550/ARXIV.2306.00197.

[26] M. Dhouibi, A. K. Ben Salem, A. Saidi, and S. Ben Saoud, "Acceleration of convolutional neural network based diabetic retinopathy diagnosis system on field programmable gate array," *IJ-ICT*, vol. 12, no. 3, p. 214, Dec. 2023, doi: 10.11591/ijict.v12i3.pp214-224.

[27] L.-Y. Chen, S.-M. Hsu, J.-C. Wang, T.-H. Yang, and H.-S. Chuang, "Photonic crystal enhanced immunofluorescence biosensor integrated with a lateral flow microchip: Toward rapid tear-based diabetic retinopathy screening," *Biomicrofluidics*, vol. 17, no. 4, p. 044102, Jul. 2023, doi: 10.1063/5.0158780.

[28] C. Plateau-Holleville and Y. Benezeth, "Class-aware data augmentation by GAN specialisation to improve endoscopic images classification," in *2022 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, Ioannina, Greece: IEEE, Sep. 2022, pp. 1–7. doi: 10.1109/BHI56158.2022.9926846.

[29] M. Li, P. Peng, H. Sun, M. Wang, and H. Wang, "An Order-Invariant and Interpretable Dilated Convolution Neural Network for Chemical Process Fault Detection and Diagnosis," *IEEE Trans. Automat. Sci. Eng.*, pp. 1–11, 2023, doi: 10.1109/TASE.2023.3290202.

[30] H. Louati, S. Bechikh, A. Louati, C.-C. Hung, and L. Ben Said, "Deep convolutional neural network architecture design as a bi-level optimization problem," *Neurocomputing*, vol. 439, pp. 44–62, Jun. 2021, doi: 10.1016/j.neucom.2021.01.094.