



Volume 11 Issue 1 Year 2026 | Page 55-64 ISSN: 2527-9866
Received: 16-12-2025 | Revised: 25-12-2025 | Accepted: 31-12-2025

Classification of Diabetic Retinopathy Using ShuffleNet V2 and Real-ESRGAN with CLAHE Image Enhancement

Nicholas Edison¹, Tinaliah²

Multi Data Palembang University, Palembang, South Sumatra, Indonesia, 30113

e-mail: nicholasedison_2226250003@mhs.mdp.ac.id¹, tinaliah@mdp.ac.id²

*Correspondence: nicholasedison_2226250003@mhs.mdp.ac.id

Abstract: Diabetic retinopathy (DR) is a microvascular complication of diabetes that can lead to blindness if not detected and treated early. Manual DR grading from fundus images is time-consuming and highly dependent on expert availability, motivating the need for automated and efficient decision-support systems. This study proposes a lightweight DR severity classification model using ShuffleNet V2 combined with a preprocessing pipeline consisting of Contrast Limited Adaptive Histogram Equalization (CLAHE) and Real-ESRGAN-based super-resolution. Unlike prior works that mainly employ these enhancement techniques with deeper or computationally expensive networks, this study explicitly investigates their synergistic integration with ShuffleNet V2 to improve lesion visibility while preserving computational efficiency for resource-constrained environments. Experiments conducted on the APTOS 2019 dataset demonstrate that the proposed combination significantly improves classification performance, achieving a best accuracy of 90.70%, with balanced precision, recall, and F1-score when optimized using Adam. Comparative analysis with SGD optimizer further reveals a trade-off between accuracy and inference speed. The results confirm that combining CLAHE and Real-ESRGAN with ShuffleNet V2 offers an effective and efficient solution for automated diabetic retinopathy grading, highlighting its suitability for large-scale screening and low-resource clinical deployment.

Keywords: Diabetic Retinopathy, ShuffleNet V2, CLAHE, Real-ESRGAN

1. Introduction

Diabetes mellitus is a metabolic disorder characterized by defective insulin activity, which results in raised blood sugar levels, leaving above the normal range and causing several serious complications, one of which is diabetic retinopathy [1], [2]. Diabetic retinopathy itself is a microvascular complication that affects the retinal blood vessels, with epidemiological studies projecting that by 2030 there will be approximately 154.9 million patients, 30% of whom are at risk of blindness, although around 90% of new cases can actually be prevented with early and appropriate management [3], [4], [5].

In clinical practice, diabetic retinopathy is commonly classified into five severity levels (No DR, Mild, Moderate, Severe, and Proliferative DR), and manual assessment of fundus images is time-consuming and highly dependent on the expertise of ophthalmologists, highlighting the need for automated decision-support systems that can provide consistent, efficient, and accurate classification [4]. Recent studies have demonstrated that deep learning-based artificial intelligence, particularly Convolutional Neural Network (CNN), achieves competitive to superior performance in various medical image classification tasks compared with traditional methods such as Support Vector Machines and even with more recent architectures such as Vision Transformers in certain scenarios [6], [7], [8].

Among lightweight CNN architectures, ShuffleNet V2 has been reported to offer a favorable balance between accuracy and computational efficiency, making it suitable for deployment on resource-constrained devices such as mobile platforms and low-power clinical equipment, and therefore highly relevant for large-scale, rapid screening of diabetic retinopathy [9], [10]. Prior work applying ShuffleNet V2 to diabetic retinopathy severity classification has shown promising results, but its evaluation was restricted to a narrow set of metrics, namely Cohen's kappa and loss, and has not comprehensively assessed standard classification metrics including accuracy, precision, recall, and F1-score, while comparative studies also indicate that ShuffleNet V2 can match or outperform other lightweight models such as MobileNetV2 and even deeper architectures like ResNet50 in terms of accuracy [11], [12].

Image pre-processing has also been shown to play a crucial role in improving model performance for fundus image analysis, with combinations such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and Enhanced Super-Resolution GAN (ESRGAN) substantially enhancing image quality and classification accuracy, although the effectiveness of such pipelines has not yet been systematically evaluated on ShuffleNet V2 and can be sensitive to the ordering of operations due to risks of synthetic artefacts and over-enhancement [13], [14]. The APTOS 2019 fundus dataset employed in this study consists of relatively high-resolution images that still suffer from common degradations such as blur, sensor noise, uneven illumination, compression artefacts, and unclear microstructures, as well as imbalanced class distribution, particularly in the severe and proliferative categories, which motivates the use of super-resolution techniques like Real-ESRGAN and targeted strategies for handling class imbalance [15].

In response to a gap identified in prior ShuffleNet V2-based work on diabetic retinopathy severity classification, which evaluated performance using only Cohen's kappa and loss, this study conducts a comprehensive performance assessment using standard classification metrics, including accuracy, precision, recall, and F1-score. Moreover, this research incorporates CLAHE and Real-ESRGAN, modifications of ESRGAN to be used in the preprocessing pipeline. These modifications have received strong evaluations in previous CNN diabetic retinopathy studies, yet they have not been systematically deployed in conjunction with the ShuffleNet V2 model. The objective of the methods introduced in this study include the ability to augment lesion visibility and the detail of textures in the retina, along with limited parameter over-boosting, through model fine-tuning and ablation studies. Moreover, the studies allow for the development of accurate, stable, and efficient models that can be used for early diabetic retinopathy detection in health-relating establishments with limited computational capabilities.

2. Literature Review

The current literature points out that models based on CNN-based models such as EfficientNet and its variants, RegNet, ResNet, DenseNet and Xception have shown to possess a great level of accuracy and performance when it comes to the automatic detection and grading of the severity of Diabetic Retinopathy from fundus images in both the binary and multi-class levels [16], [17]. Additional studies have also pointed out that EfficientNet-Lite and ShuffleNet are among the most lightweight models that can be used when there are resource limitations on the devices to be used, while ensuring that accuracy would still remain at a competitive level [18], [19].

Many researchers have investigated different CNN backbones and transfer learning methods concerning the classification of retinal images from Diabetes Mellitus (DM) complications, achieving and reporting classification accuracy greater than 90% and with notable precision, recall, and F1-score across a variety of publicly available datasets, such as APTOS 2019, EyePACS, Messidor, DDR, and IDRiD [16], [20]. Recently, the focus of research has shifted to lightweight models with architectures based on ShuffleNet, demonstrating equal or greater performance than other lightweight compact networks at all levels of accuracy and F1-score, confirming these

models are appropriate for the analysis of DR complications [12], [19]. Before the work on ShuffleNet V2 concerning DR classification, other studies primarily evaluated the use of kappa scores and the loss only, leaving a void in the exhaustive evaluation of standard multiclass metrics such as accuracy, precision, recall, and F1-score [11].

Numerous studies show that leveraging specific image processing techniques such as contrast adjustment and super resolution impacts model training and performance positively (particularly for fundus images). Many CNN-based studies show notable improvements in performance for various metrics after implementing CLAHE image enhancement techniques. Additionally, studies that integrate CLAHE and ESRGAN (Enhanced Super Resolution Generative Adversarial Networks) show significant performance improvements in classification tasks that assess various severity levels of Diabetic Retinopathy (DR), as compared to models trained and tested on raw images (especially the DenseNet-121 on the APTOS 2019 dataset) [4], [13]. However, the performance of CLAHE-ESRGAN pipelines have yet to be systematically explored on compact architectures such as ShuffleNet V2.

This study seeks to address this gap. In this context, we have implemented the ShuffleNet V2 architecture as the lightweight CNN backbone for the DR classification task, where the target outcome entails the classification of the images in the APTOS 2019 dataset into five different levels of DR severity, during deployment where the hardware is sufficiently constrained. Moreover, the study proposes to include a preprocessing technique that brings together the CLAHE and Real-ESRGAN in order to make possible the enhancement of the lesions in the images and the microscopic details of the images during the study, and to incorporate deliberate design choices that reduce the possibility of over enhancement. Finally, the study reports on a comprehensive evaluation that incorporates several performance measures such as multiclass accuracy, precision, recall and F1 score in an imbalanced classification, and thus contributes to the understanding of performance of studies that incorporate ShuffleNet V2 and the classification of DR and enhancement of fundus images.

3. Methods

This study is based on the APTOS 2019 Blindness Detection dataset (<https://www.kaggle.com/competitions/aptos2019-blindness-detection>), which is accessible on Kaggle and consists of 3,662 retinal fundus photographs, which are sorted into five categories: No DR (1805), Mild (370), Moderate (999), Severe (193), and Proliferative DR (295). A sample of every category is shown in Figure 1.

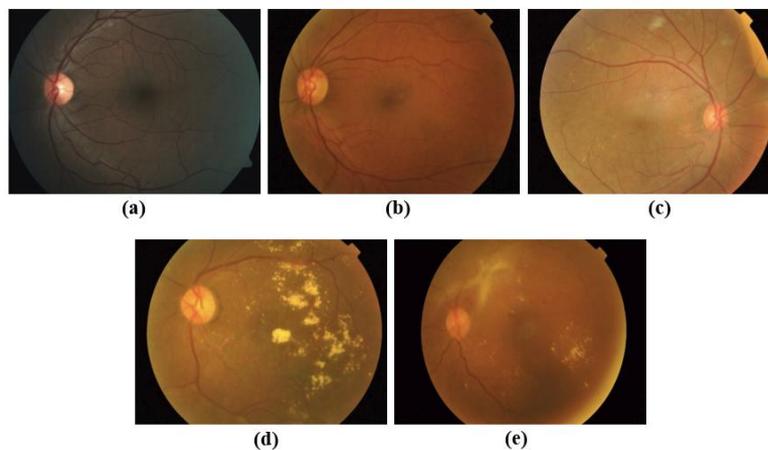


Figure 1. Sample from each class: (a) No DR, (b) Mild, (c) Moderate, (d) Severe, (e) Proliferative DR

This study's first step is to obtain fundus images from the APTOS 2019 Blindness Detection dataset, which is then subjected to multiple preprocessing steps. These steps include changing the size of the images to 224x224 pixels, normalizing the pixel intensity values, contrast correction using CLAHE, and resolution improvement with Real-ESRGAN. This stage produces fundus images with improved quality and contrast, and the CLAHE parameters are kept to a reasonable clip limit while Real-ESRGAN is configured using safe defaults to avoid excessive artifacts and over-enhancement.

For addressing the class imbalance, this study uses data augmentation via oversampling methods because they raise the number of samples without changing the features of the original images, which makes it safer than Synthetic Minority Over-sampling Technique (SMOTE) or GAN-based oversampling methods that produce samples that may have pathological structure differences [21]. A recent study that classifies dermoscopic images has shown that augmentation-based balancing improves model performance by almost 1–2% relative to the use of GAN samples, most likely because the spatial resolution and details of the lesions are better preserved [22]. Rather than using focal loss or class weighting, augmentation is preferred because the latter methods only change the penalty weights without adding any new images to the training set [23]. For the most extreme and proliferative classes, which have the fewest samples, visual diversity is the most critical for robust generalization. This is supported by the empirical evidence that focal loss and class-weighted losses that were designed to improve performance for minority classes actually resulted in performance degradation for major classes, and that augmentation-based oversampling improves balanced accuracy for the minority class by almost 3–6% [24].

Overfitting risks are dealt with in advance with a variety of techniques other than controlled pre-processing. The ShuffleNet V2 architecture inherently allows for more stability with lower overfitting tendencies on smaller datasets. To prevent the training algorithm from gaining an accuracy on training data through overfitting while neglecting the validation data, performance on the training data is tracked through the validation dataset.

Class augmentation is applied to the non-majority classes through random rotations and random horizontal and vertical flips until class sizes are equal to the largest class (No DR). The data set is then split into training, validation, and test set in the ratio of 80%:10%:10%. ShuffleNet V2 is configured with the 1× width multiplier to balance accuracy and computational efficiency for deployment on resource-constrained hardware. Model training is conducted using a batch size of 64 over 20 epochs. Two optimization strategies are evaluated: Adam (learning rate = 0.001) and SGD (learning rate = 0.01, momentum = 0.9). After training, the model is evaluated on the test set using standard classification metrics, as illustrated in Figure 2.

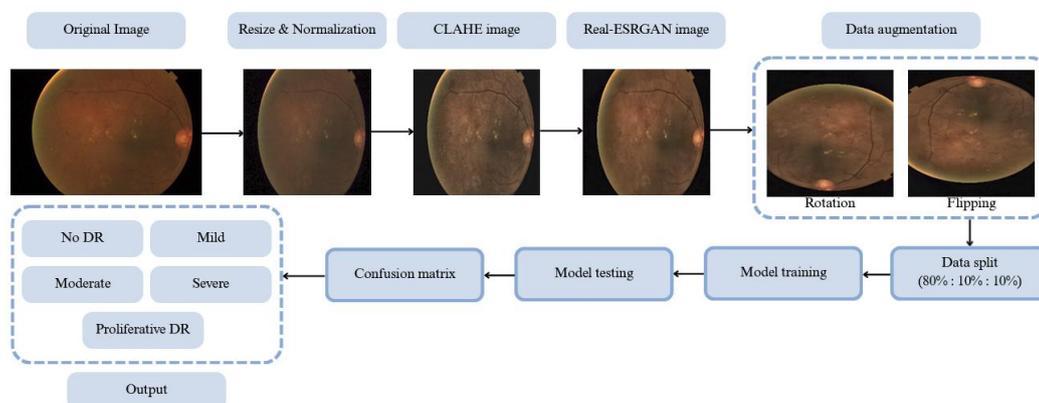


Figure 2. Model flowchart

In addition to the main design, this study also investigates several comparative scenarios which performance will be evaluated against the primary configuration. The complete set of experimental scenarios considered in this research is presented in Table I.

Table I. Comparative Scenarios

No.	Scenarios	CLAHE	Real-ESRGAN	Data Augmentation	Optimizer
1	Scenario 1	✗	✗	✗	Adam
2	Scenario 2	✓	✗	✗	
3	Scenario 3	✗	✓	✗	
4	Scenario 4	✗	✗	✓	
5	Scenario 5	✓	✓	✗	
6	Scenario 6	✓	✗	✓	
7	Scenario 7	✗	✓	✓	
8	Scenario 8	✓	✓	✓	
9	Scenario 9	✗	✗	✗	
10	Scenario 10	✓	✗	✗	
11	Scenario 11	✗	✓	✗	
12	Scenario 12	✗	✗	✓	SGD
13	Scenario 13	✓	✓	✗	
14	Scenario 14	✓	✗	✓	
15	Scenario 15	✗	✓	✓	
16	Scenario 16	✓	✓	✓	

The performance of the model for all scenarios is evaluated and presented using confusion matrices, from which the values of accuracy, precision, recall, and F1-score are derived for each configuration. In addition, the computational cost is assessed by measuring the inference time for every scenario to quantify the model’s efficiency.

4. Results and Discussion

As an initial step in reporting the experimental findings, the confusion matrices and the evaluation metrics table are presented in Figure 3, Table II and Table III, to quantitatively illustrate the model’s classification outcomes.

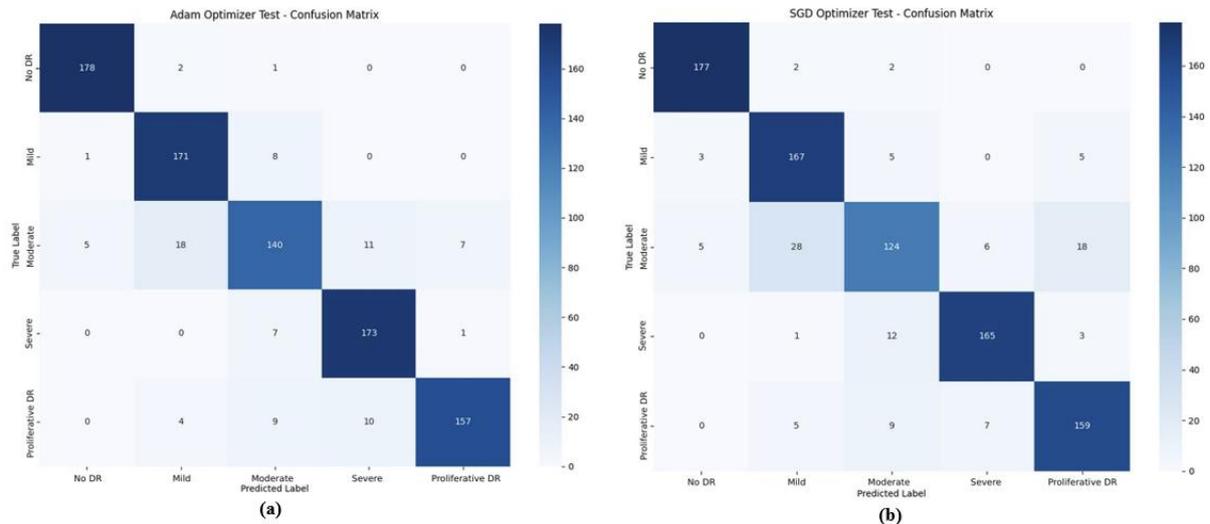


Figure 3. Confusion matrix of: (a) Scenario 8, (b) Scenario 16

In reference to the confusion matrices presented in Figure 3, both models appear to have captured the correct overall classification behavior, as exemplified by the overrepresentation of correct classification along all classes of the confusion matrices' main diagonal. Generally, the Adam Optimizer is seen to yield results that are slightly better and comparatively more balanced across the classes than the results from the SGD Optimizer. This is evident from the greater values in the diagonal of the confusion matrices of the Adam Optimized models from most of the classes, especially the Mild, Moderate, and Severe classes, in which Adam exhibited less classification errors than SGD to those classes.

For the Adam Optimized models, No DR and Severe classes were classified with very high accuracy and with no confusion with other classes. However, the Moderate class did experience some confusion whereby it was incorrectly classified into the Mild or Severe classes. This observation indicates that the features representing these classes might have some underlying gaps or have high similarity that could lead to the classes having some degree of confusion. Nevertheless, the scope of classifying errors was relatively narrow in relation to the classes as a whole, which indicates that the models did exhibit some good generalization on the dataset.

In contrast, the SGD-based model shows a greater level of uncertainty for the Moderate and Proliferative DR classes, where a considerable number of samples are misclassified from adjacent classes. While No DR and Severe also maintained high rates of correct classification, the greater off-diagonal scores suggest SGD has greater difficulty in capturing the discriminative features for certain classes. This contrast states that, in the context of the current study, the Adam optimizer is associated with more stable learning with greater separation of classes than SGD.

Table II. Comparative Scenarios using Adam Optimizer

No	Scenarios	Accuracy	Precision	Recall	F1-score	Inference time
1	Scenario 1	81.20%	65.46%	60.00%	61.51%	1.6170 s
2	Scenario 2	81.20%	70.20%	64.02%	63.04%	1.5969 s
3	Scenario 3	81.74%	71.21%	61.26%	64.18%	1.6373 s
4	Scenario 4	83.17%	83.46%	83.17%	82.47%	1.6659 s
5	Scenario 5	82.56%	65.40%	61.50%	61.13%	1.6087 s
6	Scenario 6	87.71%	87.75%	87.71%	87.65%	1.6528 s
7	Scenario 7	88.48%	88.42%	88.48%	88.35%	1.6779 s
8	Scenario 8	90.70%	90.72%	90.70%	90.59%	1.7642 s

Table II captures the performance improvement of the Adam Optimizer over eight different experimental designs. Characteristic of such performance evaluation is the improvement of most performance metrics. The performance metrics of the different Adam Experimental Settings show an upward trend in accuracy, precision, recall, and F1-score. The trend indicates that the system is achieving better metrics in the latter experimental settings because of better design. The overall best performance is observed in Scenario 8, which had the best accuracy (90.70) and other performance metrics including precision (90.72), recall (90.70), and F1-score (90.59) which indicates improvement in classification.

The earlier scenarios (Scenarios 1-3) show appreciable accuracy levels but lowish precision and recall, demonstrating an inability to distinguish all classes correctly. However, Scenarios 4 to 7 show significant improvements to precision and recall, demonstrating better class discrimination and consequently, low misclassification. While slightly different across the scenarios, inference times are all low, and the differences are too limited to justify the performance improvements as a trade-off. Overall, the results show that Scenario 8 is the best possible compromise in all respects classification accuracy, consistency, and computational efficacy.

Table III. Comparative Scenarios using SGD Optimizer

No	Scenarios	Accuracy	Precision	Recall	F1-score	Inference time
1	Scenario 9	79.84%	54.94%	54.45%	53.98%	0.0593 s
2	Scenario 10	78.47%	52.38%	57.40%	54.35%	0.0788 s
3	Scenario 11	77.38%	50.20%	47.39%	45.89%	0.0563 s
4	Scenario 12	81.17%	82.86%	81.16%	81.50%	0.0972 s
5	Scenario 13	79.29%	55.24%	51.42%	51.32%	0.0639 s
6	Scenario 14	83.39%	84.57%	83.37%	83.46%	0.1923 s
7	Scenario 15	85.71%	85.99%	85.71%	85.79%	0.1691 s
8	Scenario 16	87.71%	87.63%	87.71%	87.49%	0.2339 s

Table III contains the comparison among the eight models trained using the SGD optimizer where the results show a degree of disparity among the models in terms of their classification performances. This also showcases how SGD is impacted by the model configurations. Generally, it is observed that SGD shows some inconsistency in achieving the classification goals than the Adam optimizer over the models in precision, recall, and F1 score.

Preliminary configurations (Scenarios 9–11) demonstrate low precision and recall with a moderate degree of accuracy, suggesting that the model fails to produce reliable class-specific predictions within this configuration. Starting with Scenario 12, an improvement is evident, with precision, recall, and F1-score increasing considerably, indicative of enhanced learning stability and better discriminatory learning. Scenario 16 has the highest accuracy of 87.71%, precision of 87.63%, recall of 87.71%, and an F1-score of 87.49%, thus portraying a more even classification performance, among the SGD based experiments.

Simultaneous application of CLAHE, Real-ESRGAN, and data augmentation, all enhanced through the Adam optimizer, has resulted in the highest levels of performance, confirming the metrics and the confusion matrices. In this best case, the model attained the metrics of 90.70% accuracy, 90.72% precision, 90.70% recall, and 90.59% F1-score, indicating that the model achieved good levels of balance in the performance metrics with regard to the positively predicted samples and the true positives across all the five levels of diabetic retinopathy disease severity. The confusion matrix also confirmed this by showing the high population of the correctly predicted sample that was lined along the main diagonal and the reduced misclassification of severity levels that were adjacent to each other. In diabetic retinopathy grading, this diagonal dominance is

especially important and even critical because most of the residual errors are between the clinically adjacent levels and not the more distant classes which in turn implies that the learned weak features representation maintained the ordinal and pathological progression of the disease. ShuffleNet V2 along with the other proposed class balancing and preprocessing techniques meets the resolution of the research problem of this thesis by demonstrating the ability to classify all diabetic retinopathy grades with high accuracy and with a significantly reduced cost of computation.

Active learning is a highly effective sampling strategy; however, when using only a subset of the proposed components, performance declines across the board due to the deterioration of the pipeline as evidenced by the lower values of the metrics calculated. In data unaugmented scenarios, even if CLAHE and Real-ESRGAN are applied alone a performance drop is exhibited (e.g. a model Adam with 82.56% accuracy) which also results in a higher panel of the confusion matrices regarding class closure in these minority and adjacency of class severity. Such behaviour is a consequence of the class imbalance during model training, where classes such as Severe and Proliferative DR, being minority and underrepresented, provide fewer gradient updates and as a consequence, the model learns to concentrate its decisions on the hyperplanes of majority classes. This issue is addressed by data augmentation, whereby the added effective size and spatial variation of the pathological patterns allow adaptive feature learning to be less biased and considerably stable. This is augmentation's impact in tackling class imbalance that has underprovided the generalisation of severely and proliferative cases that are of great clinical importance yet scarce in the original dataset.

The role of image enhancement techniques was confirmed through the scenario-based assessments as well as the classification error distribution per class. Those configurations which included CLAHE and/or Real-ESRGAN tended to be more accurate than those without enhancements, suggesting that better contrast and resolution are beneficial to the detection of microvascular lesions and nuanced textural variations between adjacent stages. Contrarily, Real-ESRGAN restores high-frequency structural details that are typically lost from native blur or sub-optimal resolution. This combination ensures that diagnostically important information is added to before feature extraction, especially important for a lightweight architecture like ShuffleNet V2 which is heavily dependent on the quality of those early layers.

The results presented in Tables II and III demonstrate that, from an optimisation standpoint, the Adam optimisation technique performs significantly better than Stochastic Gradient Descent (SGD) in terms of accuracy, precision, recall, and F1 score; this pattern is also validated with confusion matrices that are smaller and are positively biased diagonal-wise. While it is true that SGD is much faster (in the order of 0.06–0.23 seconds per batch compared to 1.6–1.8 seconds for Adam), this explains why the overlapping of classes in the results is also greater and hence the lower reliability of classification is observed in certain cases. If higher speeds are required, then SGD could be the better option. However, for consistently better results in diabetic retinopathy classification, Adam performs better in terms of classification and convergence.

The inference time analysis corroborates that even when integrated with super-resolution enhancement methods, ShuffleNet V2 retains time efficiency across all instances. Overall, the scenario-based analysis combined with confusion matrix analysis indicates that the proposed system is capable of overcoming the difficulties of low-quality images and class imbalance discussed in the introduction with reliable and efficient automatic classification of diabetic retinopathy that is valid and can be scaled to screen large populations.

5. Conclusions

The area of research under consideration deals with the design of lightweight diabetic retinopathy screening models based on ShuffleNet V2 with optimised pre-processing arguments such as CLAHE, Real-ESRGAN, and data augmentation which tackles the issues of image degradation and class imbalance across the APTOS 2019 dataset. Configuring ShuffleNet V2 1× with full pre-processing and augmentation optimised by the Adam recall of 90.70% and F1 of 90.56% confirms the model is highly accurate and can classify the 5-tier DR grading schema while also being resource savvy. The results also confirm the models can be effectively utilised across DR screening exercises with the models concentrating on screening deployment across resource optimised environments and limited on sophisticated devices. Having a lightweight backbone allows for use in edge devices and cost-efficient screening systems. This also allows for faster severe and proliferative case classifications and quicker referrals for important clinical interventions needed for vision loss prevention. However, this study has some promising results but several shortcomings as well. Our experiments are limited to one public dataset, which could create boundaries for generalisability to images collected from different cameras, populations, or clinical protocol systems. Furthermore, this iteration centres on classifying images without detailed documentation of lesions or specific interpretability systems, which makes our results less trusted in a clinical use-case and could hinder adoption. Future efforts will validate our suggested model on datasets from multiple centres, add techniques for explainability, and refine our model for faster work. Additionally, combining severity grading and referral recommendations with patient data over time is an ideal and promising workflow for automated diabetic retinopathy screening systems.

References

- [1] S. F. Muzhaffarah, R. S. Simamora, dan Roulita, "HUBUNGAN DUKUNGAN KELUARGA DENGAN KEPATUHAN KONTROL GULA DARAH PADA PENDERITA DIABETES MELLITUS (DM)," *J. Penelit. Perawat Prof.*, vol. 6, no. 4, hal. <http://jurnal.globalhealthsciencegroup.com/index.p>, 2024.
- [2] Y. Y. Pasaribu dan Antony, "State Accountability and Legal Responsiveness in Preventing Diabetes Melitus From an Early Age: Between Moving Towards a Golden Indonesia or an Anxious Indonesia," *Sci. Midwifery*, vol. 12, no. 3, 2024.
- [3] R. F. N. Purnama, I. K. Nintyastuti, dan M. Rizki, "Retinopati Diabetik : Manifestasi Klinis, Diagnosis, Tatalaksana dan Pencegahan," *Lomb. Med. J.*, vol. 2, no. 2, hal. 1–4, 2023.
- [4] M. Hayati *et al.*, "Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning," *Procedia Comput. Sci.*, vol. 216, no. 2022, hal. 57–66, 2023, doi: 10.1016/j.procs.2022.12.111.
- [5] M. Elveny, T. Anjulina, B. Siregar, dan R. Syah, "Identification of Diabetic Retinopathy with Retinal Fundus Imagery Using Probabilistic Neural Network," *J. Phys. Conf. Ser.*, vol. 1641, no. 1, 2020, doi: 10.1088/1742-6596/1641/1/012055.
- [6] A. I. Kushartanto, F. Fauziah, dan R. T. Aldisa, "Comparison of CNN and SVM Methods on Web-based Skin Disease Classification Process," *Sinkron*, vol. 8, no. 2, hal. 778–788, 2024, doi: 10.33395/sinkron.v8i2.13349.
- [7] H. Wang, Y. Ji, K. Song, M. Sun, P. Lv, dan T. Zhang, "ViT-P: Classification of Genitourinary Syndrome of Menopause from OCT Images Based on Vision Transformer Models," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, doi: 10.1109/TIM.2021.3122121.
- [8] Tedyyana, A., Ahmad, A. A., Idrus, M. R., Mohd Shabli, A. H., Abu Seman, M. A., Ghazali, O., & Abd Razak, A. H. (2024). Enhance Telecommunication Security Through the Integration of Support Vector Machines. *International Journal of Advanced Computer Science & Applications*, 15(3).
- [9] H. Li, X. Yue, Z. Wang, W. Wang, H. Tomiyama, dan L. Meng, "A survey of Convolutional Neural Networks —From software to hardware and the applications in measurement," *Meas. Sensors*, vol. 18, 2021, doi: 10.1016/j.measen.2021.100080.
- [10] Y. Huang, Y. Lu, L. Fan, K. Xiang, dan H. Ma, "Fault Arc Detection Method Based on Improved

- ShuffleNet V2 Network,” *Processes*, vol. 13, no. 1, hal. 1–13, 2025, doi: 10.3390/pr13010135.
- [11] R. Gambhir, S. Bhardwaj, A. Kumar, dan R. Agarwal, “Severity Classification of Diabetic Retinopathy using ShuffleNet,” *2021 Int. Conf. Intell. Technol. CONIT 2021*, hal. 1–5, 2021, doi: 10.1109/CONIT51480.2021.9498569.
- [12] Z. Lu, J. Miao, J. Dong, S. Zhu, X. Wang, dan J. Feng, “Automatic classification of retinal diseases with transfer learning-based lightweight convolutional neural network,” *Biomed. Signal Process. Control*, vol. 81, hal. 104365, 2023, doi: <https://doi.org/10.1016/j.bspc.2022.104365>.
- [13] G. Alwakid, W. Gouda, dan M. Humayun, “Enhancement of Diabetic Retinopathy Prognostication Using Deep Learning, CLAHE, and ESRGAN,” *Diagnostics*, vol. 13, no. 14, 2023, doi: 10.3390/diagnostics13142375.
- [14] N. Nishio, J. Nishio, dan A. Nishio, “Early Detection and Severity Classification of Diabetic Retinopathy Through Image Processing and Deep Learning,” in *2024 IEEE International Conference on Future Machine Learning and Data Science (FMLDS)*, 2024, hal. 383–388. doi: 10.1109/FMLDS63805.2024.00074.
- [15] L. Zhao, H. Chi, T. Zhong, dan Y. Jia, “Perception-oriented generative adversarial network for retinal fundus image super-resolution,” *Comput. Biol. Med.*, vol. 168, hal. 107708, 2024, doi: <https://doi.org/10.1016/j.compbimed.2023.107708>.
- [16] M. Youldash *et al.*, “Early Detection and Classification of Diabetic Retinopathy: A Deep Learning Approach,” *AI MDPI*, vol. 5, no. 4, hal. 2586–2617, 2024, doi: 10.3390/ai5040125.
- [17] Y. K. Bintang dan H. Imaduddin, “Pengembangan Model Deep Learning Untuk Deteksi Retinopati Diabetik Menggunakan Metode Transfer Learning,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 9, no. 3, hal. 1442–1455, 2024, doi: 10.29100/jupi.v9i3.5588.
- [18] M. Muhammad dan I. Susilawati, “Model Cnn Berbasis Efficientnet-Lite Untuk Deteksi Retinopati Diabetik Menggunakan Tensorflow Lite,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 5, hal. 8754–8761, 2025, doi: 10.36040/jati.v9i5.15192.
- [19] D. Saproo, Seema, A. N. Mahajan, dan A. Balara, “Employing Lightweight Computational Models To Analyse and Classify Diabetic Retinopathy Images,” *ICTACT J. Data Sci. Mach. Learn.*, vol. 6, no. 2, hal. 772–776, 2025, doi: 10.21917/ijdsml.2025.0158.
- [20] B. Menaouer, Z. Dermene, N. El Houada Kebir, dan N. Matta, “Diabetic Retinopathy Classification Using Hybrid Deep Learning Approach,” *SN Comput. Sci.*, vol. 3, no. 5, 2022, doi: 10.1007/s42979-022-01240-8.
- [21] S. Alanazi dan R. Alanazi, “Enhancing diabetic retinopathy detection through federated convolutional neural networks : Exploring different stages of progression,” *Alexandria Eng. J.*, vol. 120, no. February, hal. 215–228, 2025, doi: 10.1016/j.aej.2025.02.026.
- [22] M. Alsaidi, M. Tanveer, J. Ahmed, dan A. Hanqi, “Tackling the class imbalanced dermoscopic image classification using data augmentation and GAN,” *Multimed. Tools Appl.*, vol. 83, no. 16, hal. 49121–49147, 2024, doi: 10.1007/s11042-023-17067-1.
- [23] N. Mahmoodi, H. Shirazi, M. Fakhredanesh, dan K. DadashtabarAhmadi, “Automatically weighted focal loss for imbalance learning,” *Neural Comput. Appl.*, vol. 37, no. 5, hal. 4035–4052, Des 2024, doi: 10.1007/s00521-024-10323-x.
- [24] R. Escobar, D. Guerrero, L. Carvalho, T. Bocklitz, J. Popp, dan J. L. Oliveira, “A Data Augmentation Methodology to Reduce the Class Imbalance in Histopathology Images,” *J. Imaging Informatics Med.*, vol. 37, no. 4, hal. 1767–1782, 2024, doi: 10.1007/s10278-024-01018-9.

Acknowledgements

The authors would like to express their sincere gratitude to Universitas Multi Data Palembang for the academic support and facilities provided throughout this research. The authors also acknowledge the providers of the APTOS 2019 Blindness Detection dataset (Kaggle) for making the retinal fundus images publicly available, enabling this study to be conducted and evaluated in a reproducible manner. In addition, appreciation is extended to all lecturers, colleagues, and reviewers who provided constructive feedback that helped improve the quality of this manuscript.