International Journal of Arts, Sciences and Education
Volume 5 Issue 3 | 2025

ISSN: 2799 - 1091
Page No. 1-18

Enhanced Robustness in Early Multi-Class Detection of Diabetic Retinopathy Across Datasets Using EfficientNet and CLAHE

Ageng Cahyo Widjaya , Irmawati
Department of Information System, Universitas Multimedia Nusantara, Banten
15810, Indonesia
*Corresponding author: irmawati@umn.ac.id

Abstract

Diabetic retinopathy (DR) is a severe complication of diabetes mellitus (DM) that can lead to vision loss. DR has a high prevalence among DM patients and is one of the leading causes of preventable blindness. Early detection of DR is crucial to prevent adverse effects on vision in later stages. This study utilizes EfficientNet and CLAHE methods to detect DR levels based on fundus images. Experiments were conducted on three datasets: DDR, IDRiD, and Messidor. These datasets contain five stages of DR: normal, mild, moderate, severe, and proliferative diabetic retinopathy (PDR). Our proposed model achieved a validation accuracy of 85.97% and a testing accuracy of 86.16%, outperforming other models such as Inception-ResNet-v2 (82.18%), Inception-V3 (78.79%), and ResNet-50 (74.32%). The validation and testing accuracy values above 85% indicate that the model can accurately predict labels. Interestingly, when a mixed dataset was used, the testing accuracy decreased. This decline in accuracy may be due to increased data variability, inconsistent preprocessing, and differences in image quality. Nonetheless, this model significantly contributes to preventing severe complications and vision loss due to DR.

Keywords: Diabetic retinopathy detection, EfficientNet, CLAHE, Fundus images, Transfer learning.

Introduction

Diabetes mellitus (DM) is a metabolic disease that causes inappropriate elevations in blood glucose levels, Sapra and Bhandari (2023). Diabetic Retinopathy (DR) is a condition characterized by damage to the small blood vessels in the retina of individuals with diabetes mellitus, which can lead to vision loss due to macular complications and the growth of new blood vessels in the retina and iris. DR is classified into two categories: Non-Proliferative Diabetic Retinopathy (further subdivided into mild, moderate, and severe) and Proliferative Diabetic Retinopathy (PDR) by Raja Memon, Lal, and Aziz Sahto (2017).

According to estimates by the International Diabetes Federation Saeedi et al. (2019), the global population suffering from DM is projected to reach 578 million by 2030. It is predicted to increase to 700 million by 2045. Diabetic retinopathy remains a common complication arising from DM and is the leading cause of avoidable blindness among working-age adults. According to an article by Teo et al. (2021), the global prevalence of diabetic retinopathy is approximately 22.27% or about 103 million individuals, and the prevalence of vision-threatening diabetic retinopathy is around 28 million.

Halting the progression of non-proliferative diabetic retinopathy (NDPR) at an early stage is crucial to reducing the risk of severe vision loss. However, current treatments are targeted at the advanced stages of DR when vision has already been significantly affected, highlighting the importance of making the arrest of DR progression more feasible. Additionally, most treatments for advanced stages, such as conventional laser therapy, intravitreal anti-VEGF or corticosteroid injections, and vitreoretinal surgery, are expensive and invasive and can have serious complications by Park and Roh (2016). In the early phase of DR, patients exhibit no symptoms, but in more advanced stages, the condition leads to floaters, blurred vision, distortion, and a gradual decline in visual acuity. Therefore, despite the challenges, it is crucial to detect DR early to prevent adverse outcomes in later stages by Qummar et al. (2019).

Color fundus images are used to detect DR. This method can only be performed by experts, thus requiring considerable time and cost. Therefore, an automatic computer vision detection method is needed to aid in the early detection of DR and reduce the number of cases of poor vision and blindness. This study aims to develop a deep-learning model capable of detecting the stages of DR in fundus images. We used the DDR dataset, which contains five levels of DR. These five levels are normal, mild, moderate, severe, and proliferative diabetic retinopathy. This dataset will train a deep learning model that can detect and classify the stages of DR based on the given fundus images.

DR is a leading cause of vision impairment and blindness among working-age adults globally, necessitating early and accurate detection methods to mitigate its impact. Traditional methods of DR detection rely heavily on manual examination by ophthalmologists, which is both time-consuming and prone to human error. Consequently, automated detection systems leveraging deep learning techniques have garnered significant attention in recent years. Various models, including convolutional neural networks (CNN), have enhanced the accuracy and efficiency of DR detection from retinal images.

Recent advancements have seen the application of sophisticated architectures such as EfficientNet, which have demonstrated superior performance in image classification tasks due to their optimized scaling techniques by Tan and Le (2019). EfficientNet models, when combined with image enhancement techniques like Contrast Limited Adaptive Histogram Equalization (CLAHE), have shown promise in improving the robustness and accuracy of DR detection across diverse datasets by Tan and Le (2019) and Li et al. (2021). The use of CLAHE specifically addresses issues related to varying lighting conditions and contrast in retinal images, further enhancing the performance of deep learning models by Alwakid, Gouda, and Humayun (2023).

Several studies have reported notable success in multi-class DR classification using deep learning models. For instance, Gulshan et al. (2016) demonstrated the effectiveness of a deep CNN in detecting DR and diabetic macular edema, achieving performance comparable to that of ophthalmologists. Moreover, research by Qummar et al. (2019) has shown that integrating deep learning with robust preprocessing techniques can significantly reduce misclassification rates, leading to more reliable diagnostic outcomes.

The study by Lin and Wu (2023) aims to demonstrate the impact of implementing Standard Operating Procedures (SOPs) in DR data processing using the ResNet-50 model on the EyePACS dataset. Their research provides standardized procedures for the preprocessing of fundus images. The proposed ResNet-50 model in this study achieved a test accuracy of 74%. Similarly, El Houby (2021) employs a transfer learning approach with the VGG-16 model for DR detection. Experiments were repeatedly conducted by altering the number of DR grade categories. Based on the evaluation results, the model without augmentation yielded better accuracy, although it exhibited overfitting. In addition to VGG-16, the study tested other models, namely ResNet50 and Inception. All three models achieved the same accuracy of 75% in the five-class DR classification.

Suedumrong et al. (2022) utilized CNN architecture and transfer learning methods using GoogLeNet and VGG-16 models to address the challenges in DR detection. These models were trained on the EyePACS dataset which consists of 88,702 images. The experiments revealed that using the GoogLeNet model with the Sigmoid function as the activation function and a learning rate 0.01 achieved an accuracy of 61.58%. In contrast, the VGGNet-16 model with fine-tuning reached an accuracy of 71.65%. This study highlights the potential of new techniques to improve accuracy while considering training time, suggesting that experiments with blue or red-coloured images as input could further enhance model performance.

Gangwar and Ravi (2021) employed a transfer learning approach with the Inception-ResNet-v2 model for DR detection. They added a custom CNN block to the pre-trained model to build a hybrid model evaluated on the Messidor-1 and APTOS 2019 datasets. The hybrid model achieved an accuracy 6.3% higher than that of the GoogLeNet model, with accuracies of 72.33% on the Messidor-1 dataset and 82.18% on the APTOS dataset. The study underscores the importance of balancing data distribution and further exploring transfer learning to improve overall performance.

These advancements underscore the critical role of innovative image processing and machine learning techniques in advancing the field of DR detection. The integration of EfficientNet with CLAHE represents a promising avenue for enhancing the robustness and accuracy of early multi-class DR detection, paving the way for improved patient outcomes and more efficient clinical workflows.

The latest innovation in deep learning is the transfer learning model. This method utilizes a pre-trained neural network and reduces the number of parameters by applying some parts of the existing model to the new model.Transfer learning models, such as EfficientNet and ConvNeXt, used by Irmawati et al. (2023) to detect potato leaf disease, and Inception-v3, used by Irmawati, Basari, and Gunawan (2021) to assess human blastocyst quality, enable the efficient application of deep learning techniques to address new problems.

Objective of the Study

This study provides several significant benefits. To develop a model with optimal accuracy in detecting diabetic retinopathy. The research also contributes to reducing misclassification of diabetic retinopathy stages experienced by patients, leading to more accurate and reliable diagnoses. The results of this study are valuable references for future research in the development of diabetic retinopathy detection models, allowing efforts in this field to continue to evolve and deliver broader benefits to the medical community and patients.

Research Methodology

Dataset

Based on the literature review, several datasets have been used in previous research. The datasets include Indian Diabetic Retinopathy Image Dataset (IDRiD), REHLC, Messidor-1, Messidor-2, DDR, STARE, and HRF. The datasets used in this study are the DDR, IDRiD, and Messidor datasets, which comprise images labelled into five categories: normal, mild, moderate, severe, and proliferative diabetic retinopathy. These datasets include many fundus images, ensuring a diverse representation of different stages of DR. After obtaining the three datasets, the next step is to understand the datasets. This step is performed to comprehend the datasets before preparing them for the modelling stage.

The DDR dataset by Emma Dugas Jared (2015) originates from EyePACS and contains high-resolution retinal images captured under diverse conditions. This dataset uses a grading scale from 0 to 4, where 0 represents a normal eye and 4 indicates a proliferative DR condition. In this study, not all of the dataset will be used, considering the balance of the data. The IDRiD dataset by Porwal et al. (2018) comprises 516 images and is divided into five classes, similar to the DDR dataset. The Messidor dataset by Decencière et al. (2014) is a database created to facilitate research on computer-aided diagnosis of diabetic retinopathy. It comprises 1,744 images obtained using a 3CCD colour video camera mounted on a Topcon TRC NW6 non-mydriatic retinograph with a 45-degree field of view. These images were captured using 8 bits per colour channel at resolutions of 1440x960, 2240x1488, or 2304x1536 pixels. Table 1 and Table 2 below provides details of the datasets used in this study and the number of images in each class.

Table 1. Description of the dataset used

ISSN: 2799 - 1091
Page No. 1-18

No	*Dataset*	*Format*	*Dimensions*	*Number of Images*
1	DDR	.jpg	512 *x* 512	11.571
2	IDRiD	.jpg	4288 *x* 2848	516
3	Messidor	.png	512 *x* 512	1.744

Table 2. Number of images in each dataset class

No	*Dataset*	*Normal*	*Moderate*	*Severe*	*Proliferate DR*
1	DDR	6261	600	3598	200
2	IDRiD	168	25	168	93
3	Messidor	1017	270	347	75

3.2 Data preprocessing

Fig.1 illustrates the steps to prepare the datasets for developing the DR detection model. Before preprocessing, each dataset's dimensions are resized to 224 x 224, and the black areas are removed, ensuring that only the eye region is used in the images.

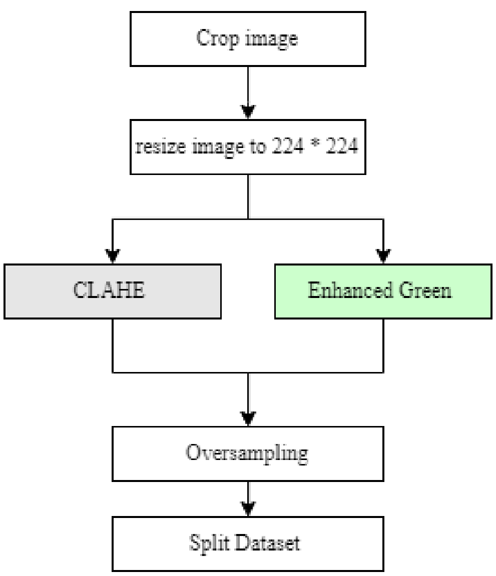

Fig. 1. The flow of data preprocessing

The image adjustment is performed by detecting the black areas. Subsequently, the images are cropped into a circular shape, retaining only the eye region. The images are then resized to 224 x 224 and saved in .jpg format. This step aims to standardize each dataset used in model development. The images below show the images before and after the adjustment process.

ISSN: 2799 - 1091
Page No. 1-18

This study implemented two image preprocessing techniques: CLAHE and Enhanced Green. Fig.2 below presents sample images resulting from the preprocessing methods.

Original Image CLAHE Image Enhanced green Image

Fig.2 below presents sample images resulting from the preprocessing methods.

Original Image CLAHE Image Enhanced green Image

Fig. 2. Samples of image enhancement techniques: original, CLAHE, and Enhanced Green

Fig. 2. Samples of image enhancement techniques: original, CLAHE, and Enhanced Green

The first dataset used in this study is the DDR dataset. The DDR dataset was balanced by reducing the amount of data in excessive classes. Subsequently, classes with insufficient data were supplemented using images from the IDRiD and Messidor datasets. The oversampling method was performed for classes with too few samples to ensure a balanced dataset across all classes. The following Table 3 presents the results of the data split into training, validation, and test sets.

Table 3. The distribution data of training, validation and testing

No	*Dataset*	*Training*	*Validation*	*Testing*
1	DDR	2213	948	351
2	DDR (Oversampled)	3015	1290	477
3	DDR + IDRiD + Messidor	2507	1072	396
4	DDR + IDRiD + Messidor (Oversampled)	2939	1257	465

Table 3 below shows the number of images per class in the dataset used for modelling. In the DDR dataset, the number of images per class is limited to 900 to reduce training time and maintain class balance. Classes with fewer images were supplemented using oversampling and dataset-merging techniques. If a class still had fewer than 900 images after merging, it was further increased using oversampling method. In Table 4, the number of images is presented for each class.

Table 4. The number of images in each class

No	*Dataset*	*Normal*	*Mild*	*Moderate*	*Severe*	*PDR*
1	DDR	900	600	900	200	912
2	DDR (Oversampled)	900	1121	900	949	912
3	DDR + IDRiD + Messidor	900	895	900	368	912
4	DDR + IDRiD + Messidor (Oversampled)	900	895	900	1054	912

3.3 Modeling

The next stage after preprocessing is the modeling stage. This study developed the deep learning model using the EfficientNetB0 architecture from the Keras library. This model is utilized to classify images of five classes to detect diabetic retinopathy. The model's architecture starts with an input layer sized 224 * 224 pixels, consisting of a GlobalAveragePooling2D layer and three Dense layers, each with a ReLU activation function. These Dense layers, also known as fully connected layers, connect every neuron from the preceding layer to every neuron in the subsequent layer. Subsequently, a Dropout layer with a rate of 0.3 is implemented after the third Dense layer to prevent overfitting. The output of the Dropout layer is then flattened and passed through another Dense layer with 64 units and ReLU activation. The final layer serves as the output layer with five units and softmax activation, which is used to classify the stages of diabetic retinopathy into five classes. The optimizer used is Adam, with a learning rate of 0.0001.

The evaluation phase involves comparing the outcomes of each model trained using the generated dataset. This comparison encompasses various aspects, including model performance metrics such as validation and testing accuracy (1). Fig.2 illustrates the proposed framework for DR-level classification.

$$\text{Accuracy}=(TP+TN)/(TP+TN+FP+FN) \qquad (1)$$

International Journal of Arts, Sciences and Education
Volume 5 Issue 3 | 2025
ISSN: 2799 - 1091
Page No. 1-18

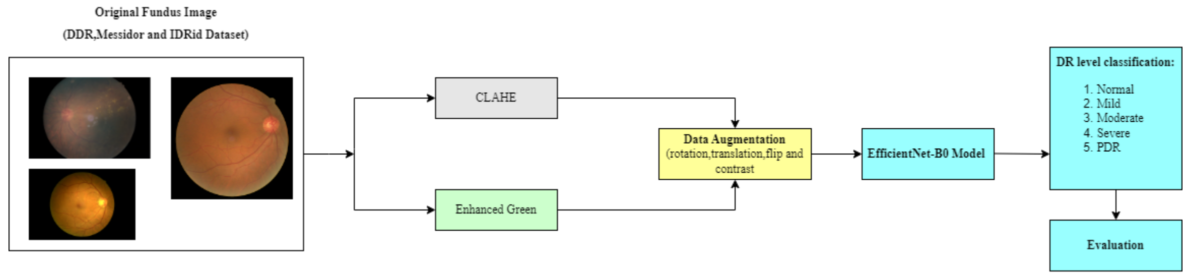

Fig. 2. The proposed framework for DR-level classification

Results and Discussion

4.1 DDR with CLAHE

This training is carried out on the DDR dataset by applying CLAHE. Fig.3 displays loss and accuracy in training. It can be seen that the model is overfitting. Evaluation of training data produces a loss of 0.6447 and an accuracy of 74.88%. The model is evaluated using validation data and test data. In the validation data, the resulting loss was 1.0064, with an accuracy of 63.40%.

Meanwhile, in the test data, the resulting loss was 0.9421, with an accuracy of 62.68%. The evaluation results show that the model experiences overfitting on the training data because the accuracy of the test data is lower than that of the training data. However, it can still perform exceptionally well in predicting classes on the test data.

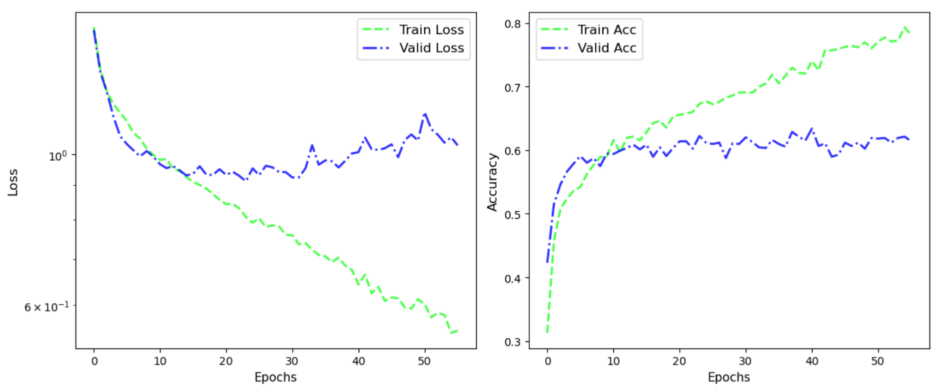

Fig. 3. Accuracy and loss graph on DDR with CLAHE

In Fig.4 and and Fig.5, it is evident that the model predicts the PDR class with high precision and recall, achieving values of 0.77 and 0.89, respectively. However, the model exhibits lower performance for the Mild and Severe classes, with precision and recall both below 0.5. This indicates that the model tends to predict these classes with lower accuracy. Therefore, the evaluation results show that while the model performs well in predicting the PDR class, there is still room for improvement, particularly in enhancing the performance for the Mild and Severe classes.

ISSN: 2799 - 1091
Page No. 1-18

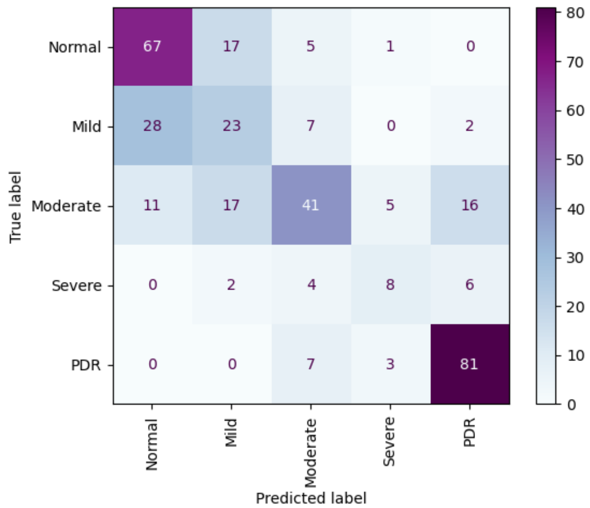

Fig. 4. The Confusion matrix results for validation data on DDR with CLAHE

Classification Report

	precision	recall	f1-score	support
Normal	0.63	0.74	0.68	90
Mild	0.39	0.38	0.39	60
Moderate	0.64	0.46	0.53	90
Severe	0.47	0.40	0.43	20
PDR	0.77	0.89	0.83	91
accuracy			0.63	351
macro avg	0.58	0.57	0.57	351
weighted avg	0.62	0.63	0.62	351

Fig. 5. The classification report for DDR with CLAHE

4.2 DDR with CLAHE and Oversampling

In the DDR dataset, each class is limited to a maximum of 900 images to reduce the time required for training. The oversampling technique is carried out on each class with a small number so that the dataset can reach at least 900. Fig. 6 below shows the accuracy and loss results for training and validation data in each epoch. This model training uses a patience level of 15 and stops at the 55th epoch.

ISSN: 2799 - 1091
Page No. 1-18

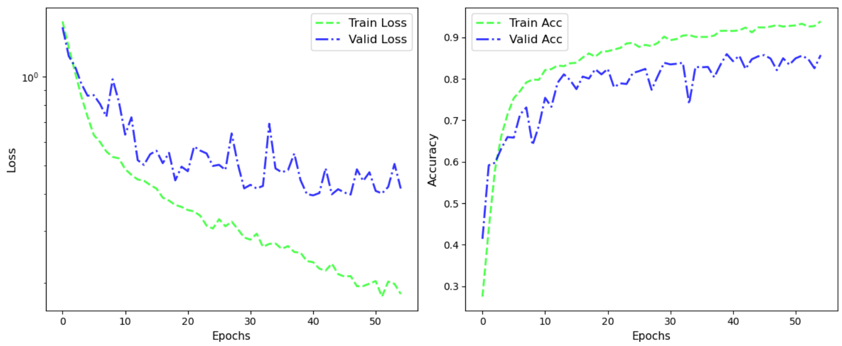

Fig. 6. Accuracy and loss graph on DDR and CLAHE with oversampling

Fig.7 below presents the confusion matrix for this dataset. Overall, the model demonstrates a high level of accuracy. However, it encounters difficulties in accurately classifying the DR level within the moderate category. A similar pattern is observed in Fig.8, which applies the model to the test dataset, indicating persistent challenges in the classification of moderate-level DR.

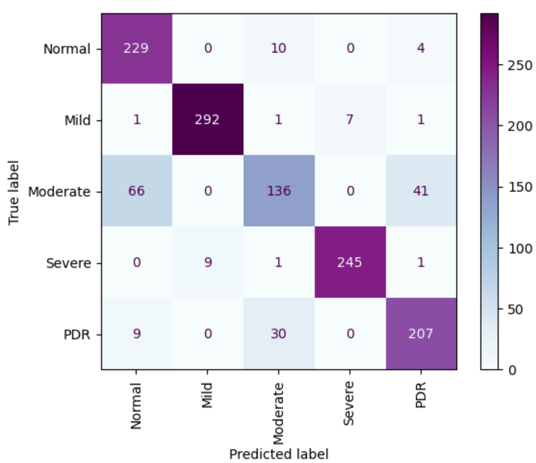

Fig. 7. Confusion matrix results for validation on DDR with CLAHE and oversampling

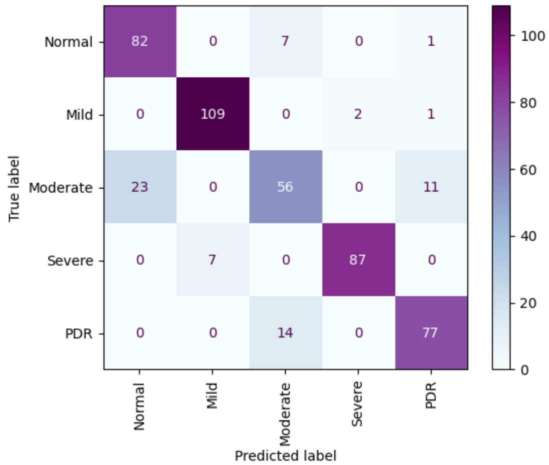

Fig. 8. Confusion matrix results for testing on DDR with CLAHE and oversampling

4.2 Mixed Three Datasets with Oversampling

In this experiment, we combined three datasets, DDR, IDRiD, and Messidor, using an oversampling technique. The results are presented in Fig.9, which displays the accuracy and loss for both training and validation data using the mixed dataset. Training of this model concluded at the 45th epoch. The loss graph indicates that the model is experiencing overfitting.

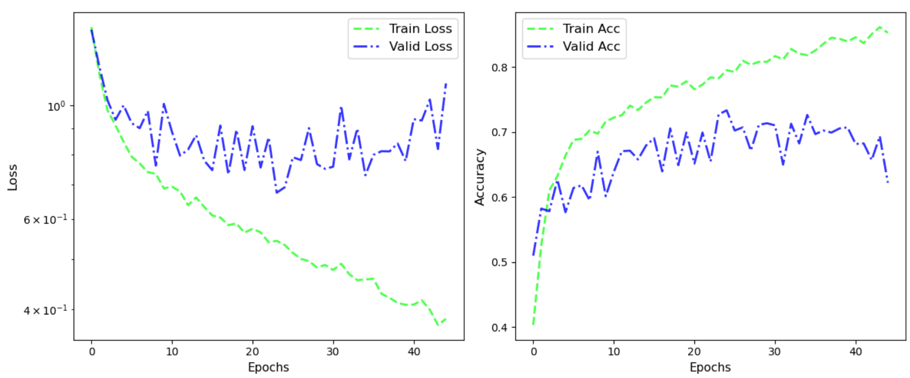

Fig. 9. Accuracy and loss graph of mixed dataset with oversampling

The Fig.10 and Fig.11 below illustrates the confusion matrix resulting from the validation and testing process using a mixed dataset with oversampling. It is evident that training with this dataset causes the model to have difficulty distinguishing between the normal, mild, and moderate classes.

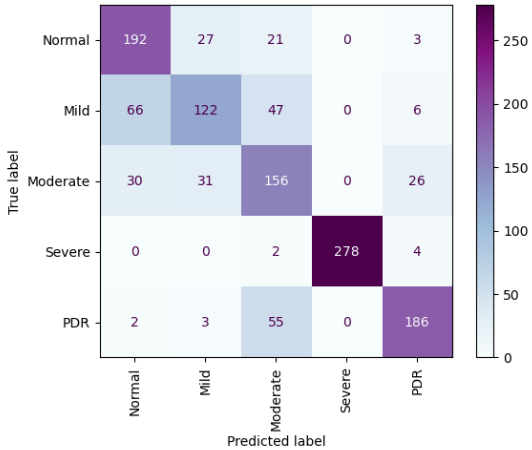

Fig. 10. Confusion matrix results for validation data on mixed dataset and oversampling

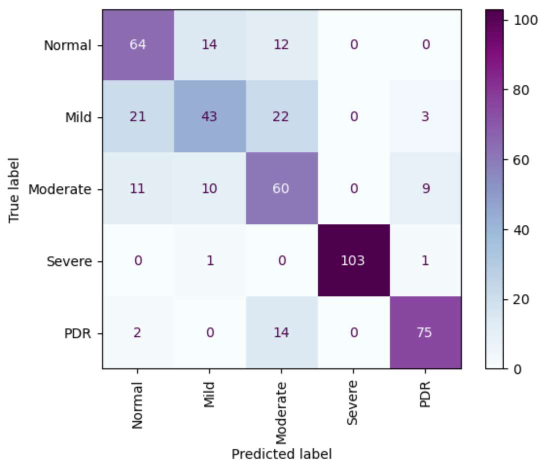

Fig. 11. Confusion matrix results for testing data of mixed dataset with oversampling

4.3 Mixed Three Datasets with CLAHE and Oversampling

The loss in mixed datasets processed with CLAHE is relatively smaller compared to datasets without CLAHE. However, this model's accuracy is also lower, at 74.19%. This model has an overall accuracy of 74%, with balanced precision, recall, and F1-score for each class. The confusion matrix indicates that the model predicts the Severe class very well, achieving precision and recall values of 100% and 98%, respectively. However, for the Normal, Mild, and Moderate classes, the model shows slightly lower precision and recall values. For the PDR class, although precision and recall are high (85% and 82%), there is still room for improvement. The Fig.12 and Fig.13 below illustrates the confusion matrix resulting from the training process using a mixed dataset with CLAHE and oversampling.

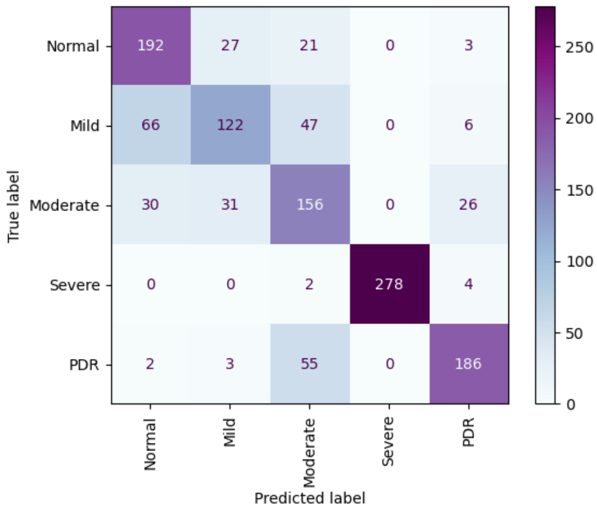

Fig. 12. Confusion matrix for validation of mixed dataset with CLAHE and oversampling

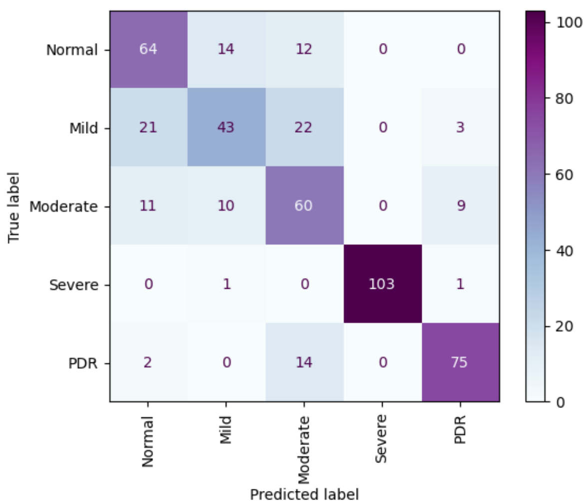

Fig. 13. Confusion matrix for testing of mixed dataset with CLAHE and oversampling

Fig. 14 shows the DDR dataset with the CLAHE technique and oversampling providing the highest testing accuracy of 86.16%. Table.6 presents the test accuracy and loss results obtained for each dataset using the same EfficientNet-B0 model. To reduce the training time, the DDR dataset limits the number of images per class to 900. An oversampling technique is applied to classes with fewer images to ensure that each class contains at least 900 images.

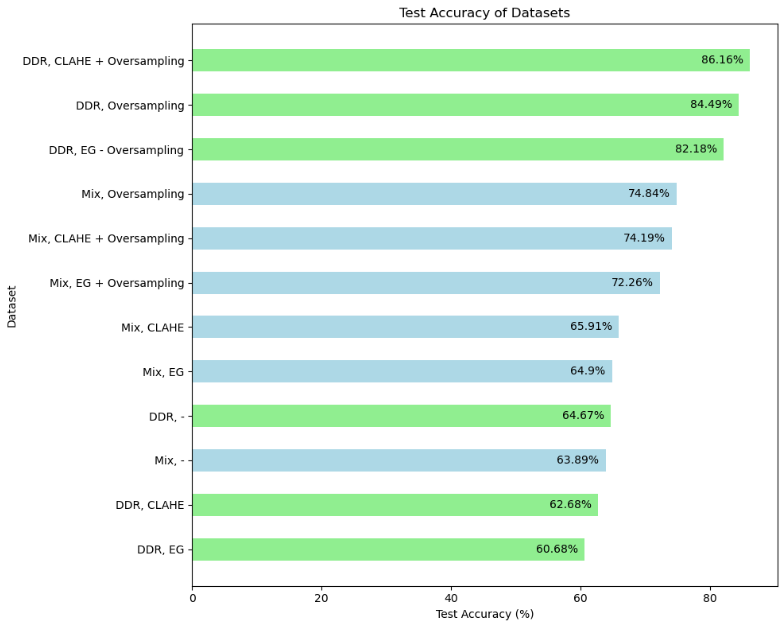

Fig. 14. Comparison of test result accuracy across all experimental scenarios

Table 5. Accuracy and loss results for all experiments

Dataset	Pre-processing	Loss			Accuracy		
		Train	Val	Test	Train	Val	Test
DDR	-	0,5101	1,2434	1,0944	81,07%	62,76%	64,67%
	Oversampling	0,2161	0,5521	0,5266	92,64%	84,03%	84,49%
	Enhanced Green	0,8658	0,9403	0,9273	64,71%	61,92%	60,68%
	EG + Oversampling	0,2092	0,4314	0,4887	91,41%	84,57%	82,18%
	CLAHE	0,6447	1,0064	0,9421	74,88%	63,40%	62,68%
	CLAHE + Oversampling	**0,2213**	**0,3994**	0,3936	91,91%	**85,97%**	**86,16%**
Mix Dataset (DDR + Messidor + IDRiD)	-	0,6176	0,9497	0,9625	74,75%	64,09%	63,89%
	Oversampling	0,4955	0,6917	0,6979	79,76%	73,35%	74,84%
	Enhanced Green	0,6332	0,9155	0,9628	74,51%	66,14%	64,90%
	EG + Oversampling	0,4545	0,6331	0,6593	81,32%	74,62%	72,26%
	CLAHE	0,4790	0,9303	0,8976	81,65%	67,16%	65,91%
	CLAHE + Oversampling	**0,4168**	**0,6646**	0,6317	83,12%	**74,30%**	**74,19%**

The evaluation results of each experiment in this research demonstrate that training with the DDR dataset, utilizing preprocessing techniques such as CLAHE and oversampling, yields the highest accuracy compared to other datasets. The model produced from this experiment will be compared with previous research, as shown in Table 6 below.

Table 6. Comparison with previous research

Model	Dataset	Pre-processing	Test Accuracy
CNN Inception-ResNet-v2 hybrid (Gangwar & Ravi, 2021)	APTOS 2019	Standard	82,18%
Inception-V3 (Elisha Widyaya et al., 2021)	APTOS 2019	*Enhanced green*	78,79%
ResNet-50 (C. L. Lin & Wu, 2023)	Kaggle DDR	*Standard Operation Procedure*	74,32%
VGG-16 (Houby, 2021)	Kaggle DDR	-	73,70%
CNN Inception-ResNet-v2 hybrid (Gangwar & Ravi, 2021)	Messidor-1	-	72,33%
VGG-16 (Suedumrong et al., 2022)	Kaggle DDR	-	71,65%
EfficientNet-B0 (model yang diusulkan)	**Kaggle DDR**	**CLAHE, *Oversampling***	**86,16%**
	DDR+IDRiD+Messidor		**74,19%**

Table 6 above presents the models used in the research, datasets, preprocessing techniques, and test accuracy. The CNN Inception-ResNet-v2 hybrid model achieved a high accuracy of 82.18% on the APTOS dataset. The table also shows that the Inception-V3 model achieved an accuracy of 78.79% on the APTOS 2019 dataset with the enhanced green preprocessing technique. The ResNet-50 model applied to the Kaggle DDR dataset with Standard Operation Procedure reached an accuracy of 74.32%. Additionally, the VGG-16 model was used on the Kaggle DDR dataset without specified preprocessing and achieved an accuracy of 73.70%. The CNN Inception-ResNet-v2 hybrid model applied to the Messidor-1 dataset achieved an accuracy of 72.33%. Furthermore, the VGG-16 model used on the Kaggle DDR dataset without other specified preprocessing techniques achieved an accuracy of 71.65%. Based on this comparison, it can be seen that the EfficientNet-B0 model with the proposed CLAHE preprocessing and oversampling techniques is relatively superior, with an accuracy of 86.16%.

Conclusion

This research was conducted to develop a diabetic retinopathy detection model using EfficientNet-B0 transfer learning, trained on multiple datasets. The datasets used in the training were divided into several types for subsequent training and evaluation. Various preprocessing techniques were applied, including CLAHE, enhanced green, and oversampling. The evaluation results of the trained models indicate that applying CLAHE can reduce the loss generated by the model. Models using the DDR dataset also performed better. However, models trained with mixed datasets were more capable of detecting various images. Additionally, dataset balance significantly impacted the model's accuracy, comparing models trained with and without oversampling.

In this study, the method used was transfer learning with the EfficientNet-B0 model. Future research will compare this model with its latest versions. The datasets used were also adjusted to be smaller to reduce training time. In future research, the datasets can be used to improve the model's accuracy.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to thank the support provided by Universitas Multimedia Nusantara during this study and special thanks to this journal for publishing this paper.

References

Alwakid, G., Gouda, W., & Humayun, M. (2023). Deep learning-based prediction of diabetic retinopathy using CLAHE and ESRGAN for enhancement. *Healthcare (Switzerland), 11*(6), 1–17. https://doi.org/10.3390/healthcare11060863

Decencière, E., Zhang, X., Cazuguel, G., Lay, B., Cochener, B., Trone, C., Gain, P., Ordonez, R., Massin, P., & Erginay, A. (2014). Feedback on a publicly distributed image database: The Messidor database. *Image Analysis and Stereology, 33*(3), 231–234.

Dugas, E., & Cukierski, J. W. (2015). *Diabetic retinopathy detection*. Kaggle. https://kaggle.com/competitions/diabetic-retinopathy-detection

Gangwar, A. K., & Ravi, V. (2021). Diabetic retinopathy detection using transfer learning and deep learning. In *Evolution in Computational Intelligence: Frontiers in Intelligent Computing: Theory and Applications (FICTA 2020), Volume 1* (pp. 679–689). Springer.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., & Cuadros, J. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA, 316*(22), 2402–2410.

Houby, E. M. F. E. (2021). Using transfer learning for diabetic retinopathy stage classification. *Applied Computing and Informatics*. https://doi.org/10.1108/ACI-10-2020-0136 *(Add DOI if available)*

Irmawati, I., Basari, B., & Gunawan, D. (2021). Classification of human blastocyst quality using wavelets and transfer learning. In *Advances in Computer, Communication and Computational Sciences: Proceedings of IC4S 2019* (pp. 965–974). Springer.

Irmawati, I., Shabrina, N. H., Indarti, S., & Prastomo, N. (2023). Early detection of potato leaf pest and disease using EfficientNet and ConvNeXt architecture. In *2023 7th International Conference on New Media Studies (CONMEDIA)* (pp. 167–172). IEEE.

Li, T., Bo, W., Hu, C., Kang, H., Liu, H., Wang, K., & Fu, H. (2021). Applications of deep learning in fundus images: A review. *Medical Image Analysis, 69*, 101971.

Lin, C.-L., & Wu, K.-C. (2023). Development of revised ResNet-50 for diabetic retinopathy detection. *BMC Bioinformatics, 24*(1), 157.

Park, Y. G., & Roh, Y.-J. (2016). New diagnostic and therapeutic approaches for preventing the progression of diabetic retinopathy. *Journal of Diabetes Research, 2016*, Article 1753584. https://doi.org/10.1155/2016/1753584

Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., & Meriaudeau, F. (2018). Indian diabetic retinopathy image dataset (IDRiD). *IEEE Dataport*. https://doi.org/10.21227/H25W98

Qummar, S., Khan, F. G., Shah, S., Khan, A., Shamshirband, S., Rehman, Z. U., Khan, I. A., & Jadoon, W. (2019). A deep learning ensemble approach for diabetic retinopathy detection. *IEEE Access, 7*, 150530–150539.

Raja Memon, W., Lal, B., & Sahto, A. A. (2017). Diabetic retinopathy: Frequency at level of HbA1C greater than 6.5%. *The Professional Medical Journal, 24*(02), 234–238. https://doi.org/10.17957/tpmj/17.3616

Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Unwin, N., Colagiuri, S., Guariguata, L., Motala, A. A., & Ogurtsova, K. (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas. *Diabetes Research and Clinical Practice, 157*, 107843.

Sapra, A., & Bhandari, P. (2023). Diabetes. *Continuing Education Activity*. *(Add publisher if available)*

Suedumrong, C., Leksakul, K., Wattana, P., & Chaopaisarn, P. (2022). Application of deep convolutional neural networks VGG-16 and GoogLeNet for level diabetic retinopathy detection. In *Proceedings of the Future Technologies Conference (FTC) 2021, Volume 2* (pp. 56–65). Springer.

Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning* (pp. 6105–6114). PMLR.

Teo, Z. L., Tham, Y.-C., Yu, M., Chee, M. L., Rim, T. H., Cheung, N., Bikbov, M. M., Wang, Y. X., Tang, Y., & Lu, Y. (2021). Global prevalence of diabetic retinopathy and projection of burden through 2045: Systematic review and meta-analysis. *Ophthalmology, 128*(11), 1580–1591.