Applied Mathematics and Computational Intelligence
Volume 14, No. 3, 2025 [29-36]

Detection of Diabetic Retinopathy Using a Transfer Learning Approach with DarkNet19

Michael Chi Seng Tang*, Abdulwahab Funsho Atanda, Huong Yong Ting

Design and Technology Centre, University of Technology Sarawak, 96000 Sibu, Sarawak, Malaysia

* Corresponding author : michaeltang@uts.edu.my

Received: 2 July 2024
Revised: 6 February 2025
Accepted: 24 February 2025

ABSTRACT

The early detection of diabetic retinopathy (DR) in fundus images is crucial for preventing vision loss in diabetic patients. Deep learning models have played a significant role in advancing DR detection. This paper explores the relatively unexplored DarkNet19 model's performance in comparison to well-established models like ResNet18 and ResNet50 for DR detection. A balanced dataset of healthy and DR images was created through standardization and augmentation techniques. The models underwent binary classification training and testing, and their performance was evaluated using accuracy and precision metrics. DarkNet19 outperformed the other models, achieving higher accuracy (0.7347) and precision (0.9233), demonstrating its potential to enhance early DR diagnosis and reduce the risk of vision loss. This research contributes to the field of DR detection, highlighting the effectiveness of DarkNet19.

Keywords: Diabetic Retinopathy, Transfer Learning, DarkNet19, Deep Learning, Medical Imaging.

1 INTRODUCTION

Deep learning has been widely used in many areas for automated detection, such as in the medical field [1], [2], [3], [4], [5], [6], [7], vehicle recognition [8], and waste material classification [9]. The prevalence of diabetic retinopathy (DR) which is a serious eye condition affecting people with diabetes and a leading cause of blindness, coupled with the projected increase in diabetic patients, calls for concerted efforts by health practitioners, researchers, and the Government. The field of DR detection in fundus images has witnessed significant advancements with the integration of deep learning models. Early detection of DR is paramount in preventing vision loss among diabetic patients, and the utilization of artificial intelligence (AI) and deep learning (DL) has shown promise in achieving this objective. Previous research has explored various deep learning models such as ResNet18, ResNet50, and others for the detection of DR, highlighting the potential of these technologies in enhancing early diagnosis and treatment.

Khalifa et al. [10], Sambyal et al. [11], Caicho et al. [12], and Lin et al. [13] have all contributed to this growing body of knowledge, emphasizing the importance of early DR detection and the role of deep learning models in achieving this goal. These studies have introduced modified models and innovative techniques to improve the accuracy and efficiency of DR detection, setting the stage for further exploration in this domain.

Despite the extensive exploration of various pre-trained models, one notable model, DarkNet19, has remained relatively unexplored in the context of DR detection. This paper aims to bridge this gap by evaluating the performance of DarkNet19 and comparing it to well-tested pre-trained models. The assessment of DarkNet19's effectiveness in DR detection is a key contribution of this research.

In this paper, we present the methodology employed for the evaluation, including data preprocessing, model training, and testing procedures. The fundus image dataset used in this study, comprising both healthy and DR images, underwent standardization and augmentation to create a balanced dataset. Various deep learning models, including DarkNet19, ResNet18, and ResNet50, were trained and evaluated on this dataset for binary classification of DR.

The results and discussions section provides insights into the performance of these models, with a focus on accuracy and precision metrics. Notably, the Proposed Model (DarkNet19) demonstrates promising results, indicating its potential as an effective tool for DR detection. The comparison with established models highlights DarkNet19's superior performance, particularly in correctly identifying positive instances within the classification task with high accuracy.

This paper contributes to the ongoing efforts in the field of DR detection by shedding light on the capabilities of DarkNet19 and its potential to enhance the early diagnosis of DR, ultimately minimizing the risk of vision loss among diabetic patients. The remaining sections of this paper are divided into several subsections as follows: Section 2.0 discusses the related work of this study. Section 3.0 presents the methodology employed in this study. Section 4.0 discusses the results obtained from the experiment. Finally, Section 5.0 concludes this paper.

2 RELATED WORKS

Several prior papers have been published that utilized deep learning models for the detection of DR in fundus images. For example, Khalifa et al. [10] implemented ResNet18 for the detection of DR. They address the significance of early detection of DR., Leveraging advancements in computer science, particularly artificial intelligence (AI) and deep learning (DL), the research explores deep transfer learning models for medical DR detection, utilizing the Asia Pacific Tele-Ophthalmology Society (APTOS) 2019 dataset. The selected models, including AlexNet, Res-Net18, SqueezeNet, GoogleNet, VGG16, and VGG19, were chosen for their efficiency with fewer layers compared to larger models. Data augmentation techniques were employed to enhance model robustness. The study demonstrates the potential of DL in improving early DR detection and minimizing vision loss in diabetic patients, with AlexNet model achieving the highest accuracy of 97.9%.

Sambyal et al. [11] have also tested the ResNet18 model for DR detection. They address the issue of diabetic retinopathy (DR) diagnosis, a significant complication of diabetes and a leading cause of global blindness. Manual diagnosis by ophthalmologists is time-consuming and challenging. The paper introduces modified deep residual networks for both binary and multistage classification of DR. The models were evaluated using the MESSIDOR dataset, achieving high accuracy levels. For binary classification, the modified ResNet18, ResNet34, and ResNet50 models demonstrated accuracies of 99.47%, 99.47%, and 99.87%, respectively. In multistage classification, the modified

ResNet18 achieved an accuracy of 99.37%, modified ResNet34 had 99.16%, and modified ResNet50 reached 99.37%. A comparison with existing models revealed significant accuracy improvements of at least 0.83% for binary and over 5% for multistage classification. Overall, the proposed DR classification models outperformed existing methods on this benchmark dataset.

Caicho et al. [12] tested the ResNet50 model for DR detection. They focused on the importance of early detection and treatment of DR, a leading cause of blindness in diabetes patients. It highlighted the challenges of detecting DR in its early stages, even for highly trained healthcare professionals, and emphasized the potential of artificial intelligence, particularly Convolutional Neural Networks (CNN), as a more efficient and cost-effective solution. The study utilized three different CNN models, including AlexNet, GoogleNet, and ResNet50, to classify the five stages of DR.

Lin et al. [13] also used the ResNet50 model for DR detection. They address the issue of DR, a condition that can lead to vision loss or blindness if not detected and treated early. The study leverages the advancements in deep learning, particularly the ResNet-50 model, to enhance the accuracy of DR prediction. By employing visualization techniques and preprocessing, the study aims to calibrate the model effectively. Comparing its performance with other common CNN models such as Xception, AlexNet, VggNet-s, VggNet-16, and ResNet-50, the research identifies an overfitting problem in the examined models. However, the modified ResNet-50 model (with a training accuracy of 0.8395 and a test accuracy of 0.7432) outperforms the others, showcasing its ability to mitigate overfitting, reduce loss, and minimize fluctuations. The study also introduces two approaches: a standard procedure for preprocessing fundus images and a revised ResNet-50 structure with adaptive learning rate adjustments, regularization, and structural changes. While the study doesn't aim to create the most accurate DR screening network, it highlights the significance of preprocessing and the visualization of modified CNN structures, providing insights for future improvements in CNN architecture design.

With numerous pre-trained models already implemented for DR detection through transfer learning, primarily the ResNet variants, one pre-trained model remains untested: DarkNet19. DarkNet19 is among the unexplored pre-trained models in the field of research. Hence, this study aims to assess the performance of the DarkNet19 model in comparison to other well-tested pre-trained models. The paper's contribution lies in the evaluation of DarkNet19, an untested model, for its effectiveness in DR detection.

3 METHODOLOGY

The fundus images utilized in this study were sourced from various datasets, including Messidor [14], private datasets from Sibu Hospital, and private datasets from the Department of Ophthalmology at the Health Campus of Universiti Sains Malaysia. Each of these datasets presented images with different resolutions. Consequently, a standardization process was applied to resize all images to 224×224 pixels, enabling compatibility with pre-trained models such as ResNet18 [15] and ResNet50 [15]. However, DarkNet19 [16] necessitated a specific image size of 256×256 pixels for loading. In total, the dataset comprised 26,348 healthy images (classified as Negative) and 9,980 images depicting DR (classified as Positive). Figure 1 shows several images obtained from the datasets.

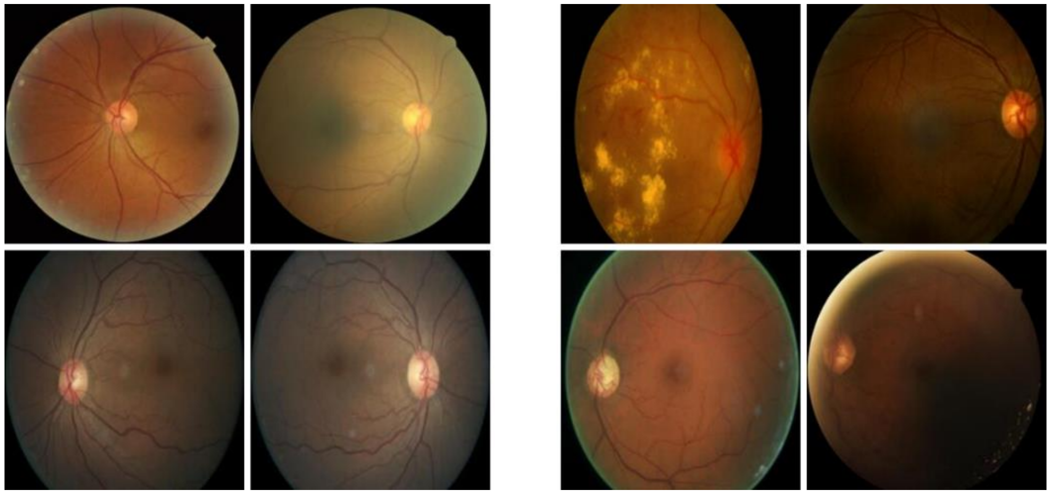

Figure 1 : Several images obtained from the datasets. (a) Negative images. (b) Positive images.

Notably, the dataset exhibited an imbalance between these two classes, with an abundance of negative instances and a scarcity of positive ones. To rectify this imbalance, image augmentation techniques were employed exclusively on the positive images, which included horizontal and vertical flipping. This augmentation process increased the count of positive images to 29,940, thus creating a more balanced dataset.

Subsequently, the dataset was partitioned into a training set and a testing set, with 75% of the images allocated to training and the remaining 25% reserved for testing. The training set was employed to train various deep learning models, while the testing set served as an evaluation benchmark to determine the model's performance and identify the best-performing one.

The deep learning models initially underwent pre-training to recognize multiple classes. However, for this study, only two classes were relevant: negative and positive. To accommodate this binary classification task, modifications were introduced. The final fully connected layer of each pre-trained model was removed and replaced with a new one designed to classify images into the two specified classes. Training was carried out using Stochastic Gradient Descent with Momentum (SGDM) with a mini-batch size of 128, a learning rate of 0.01, and over a span of 10 epochs. Once the models completed training, they were subjected to testing using the independent testing set, which comprised images that had not been seen during the training process.

Each image in the testing set was classified as either positive or negative by the model. Following this classification, two performance metrics were computed, including accuracy and precision. The specific equations for these performance metrics can be found in [7]. Ultimately, the study compared the performance of the DarkNet19, ResNet18, and ResNet50 models, shedding light on their efficacy in this binary classification task. Figure 2 shows the flowchart of the methodology.

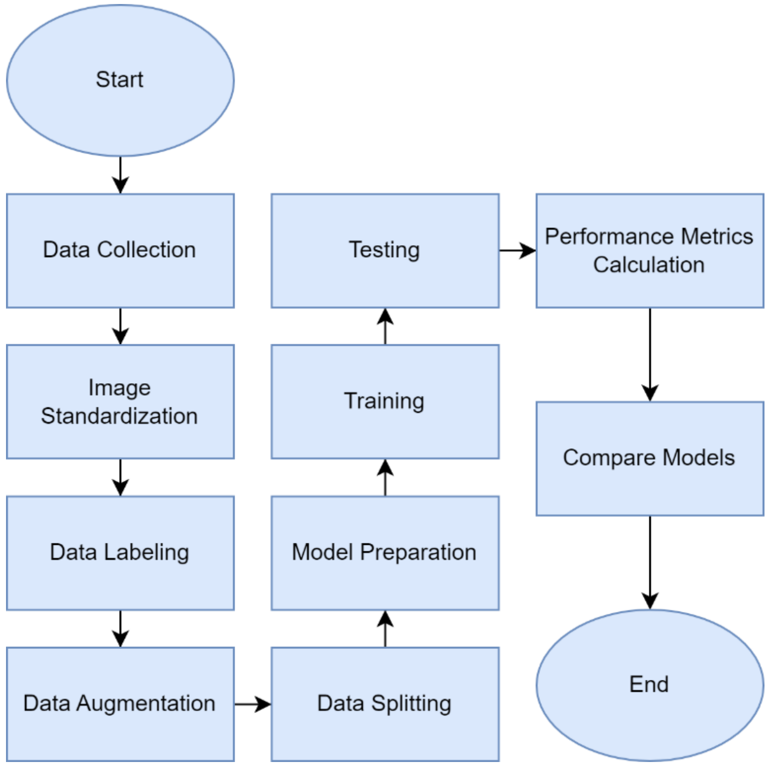

Figure 2 : Flowchart of the methodology.

4 RESULTS AND DISCUSSIONS

Once the performance metrics are obtained from each deep learning model, they are compared to determine which one performs the best. The performance metrics, including accuracy and precision, are displayed in Table 1.

Table 1 : Performance of different models used for DR detection. The best results are highlighted in bold.

Model	Accuracy	Precision
ResNet18	0.7320	0.7738
ResNet50	0.7324	0.7718
Proposed Model (DarkNet19)	**0.7347**	**0.9233**

The table depicts a comparison of performance among different machine learning models in the context of a classification task. Three models were evaluated: ResNet18, ResNet50, and a Proposed Model referred to as DarkNet19.

The accuracy metric was utilized to assess the models' overall classification correctness. ResNet18 was found to achieve an accuracy of 0.7320, signifying that approximately 73.20% of instances were correctly classified. Similarly, ResNet50 yielded an accuracy of 0.7324, indicating that approximately 73.24% of instances were correctly categorized. These results highlight the close similarity in accuracy between ResNet18 and ResNet50.

Remarkably, the Proposed Model (DarkNet19) exhibited superior performance, boasting an accuracy of 0.7347, signifying that it correctly classified approximately 73.47% of instances. This elevated accuracy score suggests that DarkNet19 may offer a more robust and accurate solution for the given classification task when compared to the other models under consideration.

The precision metric, which is another essential measure of a model's performance, was examined in addition to accuracy. Precision assesses the ability of a model to make correct positive predictions and is calculated as the ratio of true positive predictions to the total number of positive predictions made by the model.

For ResNet18, the precision was found to be 0.7738, indicating that approximately 77.38% of the positive predictions made by the model were correct. Similarly, ResNet50 exhibited a precision of 0.7718, signifying that approximately 77.18% of its positive predictions were accurate. These precision scores indicate a relatively high level of correctness in identifying positive cases for both ResNet18 and ResNet50.

In contrast, the Proposed Model (DarkNet19) demonstrated a significantly higher precision of 0.9233. This suggests that DarkNet19 excelled in making accurate positive predictions, with approximately 92.33% of its positive predictions being correct. This substantial increase in precision implies that DarkNet19 is particularly proficient at correctly identifying positive instances within the classification task.

5 CONCLUSIONS

In conclusion, this study thoroughly assessed the performance of the DarkNet19 model in DR detection, comparing it to ResNet18 and ResNet50. The results showcased DarkNet19's superior accuracy and precision, highlighting its potential as a robust tool for early DR diagnosis. While ResNet models performed credibly, DarkNet19's exceptional performance underscores its significance in enhancing DR detection, offering promising prospects for improving diabetic patient care. This research contributes to the evolving field of DR detection, emphasizing the need for continued exploration and integration of deep learning models in ophthalmology to mitigate the risk of vision loss among diabetic individuals.

ACKNOWLEDGEMENT

The authors would like to thank University of Technology Sarawak, UTS for funding this work under an internal grant UTS/RESEARCH/4/2023/08.

REFERENCES

[1] N. Tariq *et al.*, "Ground Truth from Multiple Manually Marked Images to Evaluate Blood Vessel Segmentation," in *Proceedings of the 12th International Conference on Robotics, Vision, Signal Processing and Power Applications*, 2024, pp. 531–536. doi: 10.1007/978-981-99-9005-4_67.

[2] M. C. S. Tang and N. H. Rashidi, "Enhanced Brain Tumor Classification using Modified ResNet50 Architecture," *Borneo J. Sci. Technol.*, vol. 6, no. 1, pp. 11–18, Jan. 2024, doi: 10.35370/bjost.2024.6.1-02.

[3] M. C. S. Tang and S. S. Teoh, "Brain Tumor Detection from MRI Images Based on ResNet18," in *2023 6th International Conference on Information Systems and Computer Networks (ISCON)*, IEEE, Mar. 2023, pp. 1–5. doi: 10.1109/ISCON57294.2023.10112025.

[4] M. C. S. Tang and S. S. Teoh, "Blood vessel segmentation in fundus images using hessian matrix for diabetic retinopathy detection," in *2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, IEEE, Nov. 2020, pp. 0728–0733. doi: 10.1109/IEMCON51383.2020.9284931.

[5] M. C. S. Tang, S. S. Teoh, and H. Ibrahim, "Retinal vessel segmentation from fundus images using DeepLabv3+," in *2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA)*, IEEE, May 2022, pp. 377–381. doi: 10.1109/CSPA55076.2022.9781891.

[6] M. C. S. Tang, S. S. Teoh, H. Ibrahim, and Z. Embong, "Neovascularization detection and localization in fundus images using deep learning," *Sensors*, vol. 21, no. 16, p. 5327, Aug. 2021, doi: 10.3390/s21165327.

[7] M. C. S. Tang, S. S. Teoh, H. Ibrahim, and Z. Embong, "A deep learning approach for the detection of neovascularization in fundus images using transfer learning," *IEEE Access*, vol. 10, pp. 20247–20258, 2022, doi: 10.1109/ACCESS.2022.3151644.

[8] M. C. S. Tang and H. Y. Ting, "Evaluating the Performance of ResNet Variant Models for Car Model Detection Using a Transfer Learning Approach," in *SpringerBriefs in Applied Sciences and Technology*, 2024, pp. 125–132. doi: 10.1007/978-3-031-68751-8_15.

[9] M. Tang, K. C. Ting, and N. H. Rashidi, "DenseNet201-Based Waste Material Classification Using Transfer Learning Approach," *Appl. Math. Comput. Intell.*, vol. 13, no. 2, pp. 113–120, Jun. 2024, doi: 10.58915/amci.v13i2.555.

[10] N. E. M. Khalifa, M. Loey, M. H. N. Taha, and H. N. E. T. Mohamed, "Deep Transfer Learning Models for Medical Diabetic Retinopathy Detection," *Acta Inform. Medica*, vol. 27, no. 5, pp. 327–332, 2019, doi: 10.5455/aim.2019.27.327-332.

[11] N. Sambyal, P. Saini, R. Syal, and V. Gupta, "Modified residual networks for severity stage classification of diabetic retinopathy," *Evol. Syst.*, vol. 14, no. 1, pp. 17–35, Feb. 2023, doi: 10.1007/s12530-022-09427-3.

[12] J. Caicho *et al.*, "Diabetic Retinopathy: Detection and Classification Using AlexNet, GoogleNet

and ResNet50 Convolutional Neural Networks," in *Communications in Computer and Information Science*, vol. 1532 CCIS, 2022, pp. 259–271. doi: 10.1007/978-3-030-99170-8_19.

[13] C.-L. Lin and K.-C. Wu, "Development of revised ResNet-50 for diabetic retinopathy detection," *BMC Bioinformatics*, vol. 24, no. 1, p. 157, Apr. 2023, doi: 10.1186/s12859-023-05293-1.

[14] E. Decencière *et al.*, "Feedback on a publicly distributed image database: The Messidor database," *Image Anal. Stereol.*, vol. 33, no. 3, pp. 231–234, Aug. 2014, doi: 10.5566/ias.1155.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[16] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 6517–6525. doi: 10.1109/CVPR.2017.690.