Indonesian Journal of Electrical Engineering and Computer Science
Vol. 36, No. 1, October 2024, pp. 509~516
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v36.i1.pp509-516

Eye disease detection using transfer learning based on retinal fundus image data

Helmi Imaduddin, Alivia Rahma Sakina
Department of Informatics, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia

Article Info

Article history:

Received Dec 2, 2023
Revised Jun 6, 2024
Accepted Jun 25, 2024

Keywords:

Deep learning
DenseNet-121
Eye disease
ResNet-152
Retinal image
Transfer learning

ABSTRACT

The escalating global prevalence of blindness remains a pressing concern, with eye diseases representing the primary culprits behind this issue. Vision is integral to various aspects of human life, underscoring the significance of effective eye disease detection. Presently, disease detection relies largely on manual methods, which are susceptible to misdiagnosis. However, the advent of technology has paved the way for disease detection through the application of deep learning methodologies. Deep learning exhibits substantial potential in disease detection, particularly when applied to image data, as attested by its accuracy in algorithmic assessments. This research introduces a novel approach to disease detection, specifically transfer learning-based deep learning. The study seeks to evaluate and compare the performance of various models, including EfficientNetB3, DenseNet-121, VGG-16, and ResNet-152, in identifying three prevalent eye diseases: cataract, diabetic retinopathy, and glaucoma, utilizing retinal fundus image data. Extensive experimentation reveals that the DenseNet-121 model achieves the highest accuracy levels, boasting precision, recall, F1-score, and accuracy values of 96.5%, 96%, 96.25%, and 96.20%, respectively. These results demonstrate the superior performance of the employed transfer learning model, signifying its efficacy in detecting eye diseases.

Corresponding Author:

Helmi Imaduddin
Department of Informatics, Universitas Muhammadiyah Surakarta
Surakarta, Indonesia
Email: helmi.imaduddin@ums.ac.id

1. INTRODUCTION

Vision plays a pivotal role throughout one's lifespan. Visual impairment or eye diseases can manifest across various age groups, significantly impacting the quality of life. Notably, eye diseases stand as a primary cause of blindness, further exacerbating the challenges faced by individuals [1]. The global prevalence of visual impairment has seen a concerning rise, with data from the World Health Organization (WHO) in 2019 revealing that over 2.2 billion people worldwide suffer from visual impairment or blindness. Alarmingly, Indonesia ranks third globally in terms of blindness prevalence, with a rate of 1.47%, and within Southeast Asia, it possesses the highest rate at 3% [2].

Analyzing data spanning from 1990 to 2015, cataracts emerge as the leading cause of blindness, accounting for 36.58% of cases. Following closely is undercorrected refractive error at 36.43%, while other conditions such as glaucoma (5.81%), age-related macular degeneration (2.44%), corneal disease (2.43%), diabetic retinopathy (0.16%), and trachoma (0.04%) constitute the remaining causes of blindness [3]. These statistics underscore cataracts as the predominant contributor to blindness. Cataracts involve a degenerative process characterized by opacity in the eye's lens fibers. Additionally, conditions like glaucoma, characterized by optic disc cupping and resulting from elevated intraocular pressure, play a significant role in vision loss. Diabetic retinopathy, marked by damage to the eye's blood vessels, is another prominent cause of blindness [4]. These high incidence rates highlight a concerning issue: many individuals affected by eye diseases remain unaware of their condition, leading to delayed diagnosis and treatment, often worsening the prognosis.

Conversely, the process of diagnosing eye disorders can be time-consuming, particularly when relying on manual observations, which may lead to misdiagnosis. However, with the advancement of technology, disease identification can be facilitated through the use of technological means, with deep learning emerging as a prominent approach. Deep learning, a subset of machine learning, is characterized by its focus on artificial intelligence algorithms inspired by the neural structure and functioning of the brain, known as artificial neural networks [5]. Within the realm of deep learning, one method that stands out is transfer learning. This approach harnesses the potential to accurately identify diseases, thus offering significant assistance in the field of medicine. Transfer learning serves as a technique to expedite the training process within convolutional neural networks (CNNs), a key component of deep learning, designed to address limitations inherent in previous methods [6].

In a study conducted by Sarki *et al.* [7], the application of transfer learning was employed to diagnose diabetic eye diseases by analyzing retinal fundus images using CNN architectures, specifically VGG-16. The findings indicated that there was an accuracy rate of 83.43%. In addition, Pin conducted a study wherein the ResNet50 model was employed for the purpose of detecting eye problems [8]. The results indicated an accuracy rate of 85.79% when applied to a dataset consisting of 1,304 fundus images. In a separate study, Sugeno did research on the application of EfficientNetB3 for the identification of eye illnesses [9]. The findings of this inquiry revealed an accuracy rate of 84.42%. Additionally, Taşar employed the Transfer Learning methodology to discern various medical conditions, such as skin cancer, through the examination of dermoscopy images [10]. In this particular case, the DenseNet-121 architecture was employed, resulting in a remarkable accuracy rate of 94.29%.

Based on the aforementioned findings, this study aims to identify specific eye disorders, namely cataract, glaucoma, and diabetic retinopathy. This research proposes to conduct a comparative analysis of various algorithms using transfer learning methodology, specifically focusing on EfficientNetB3, DenseNet-121, VGG-16, and ResNet-152, using retinal fundus image data. The ultimate goal is to determine the algorithm that exhibits the highest level of accuracy in detecting various eye diseases. This work presents a novel transfer learning approach using different hyperparameters. We have conducted 12 experiments by modifying different learning rate and epoch values for each algorithm, then the results of the classification will be compared to find out which transfer learning method has the best performance for eye disease image data. So that the best results are obtained for the classification of eye disease images.

2. METHOD

In this study, the transfer learning technique was utilized to classify image data pertaining to eye diseases. This approach involves utilizing a pre-trained model and adjusting its parameters to cater to the specific characteristics of the new case, which, in this context, relates to eye disease classification. This study applies the transfer learning process to compare the performance of various models with parameter adjustments according to new cases in eye diseases. This research will identify three eye diseases namely cataract, diabetic retinopathy, and glaucoma. Le *et al.* [11] mentioned that using transfer learning will result in stronger classification. This is one of the references for researchers to use transfer learning in eye disease classification.

The employed methodology can be described as follows: It begins with the gathering of the necessary datasets. Afterwards, data preprocessing is carried out to prepare the data for integration with the selected model. There are three subsets within the dataset: training data, testing data, and validation data. In the training phase, the designated training dataset is used to train the model. During each training iteration, the validation dataset is used to evaluate the performance of the model. After the training process is complete, the model is tested using the testing dataset [12]. Figure 1 is a visual representation of the overall workflow for the method used in this study.

2.1. Data collection

The dataset utilized in this study comprises publicly available retinal fundus images of eye diseases, sourced from Kaggle. This dataset amalgamates data from diverse origins, including the Indian Diabetic Retinopathy Image Dataset (IDRiD), ocular recognition, and high-resolution fundus (HRF) datasets. The dataset employed in this research encompasses a total of 4,217 images, which have been categorized into four distinct classes. These classes comprise 1,038 images of cataract cases, 1,098 images depicting diabetic retinopathy, 1,007 images showcasing glaucoma, and 1,074 images representing normal eye conditions. A comprehensive breakdown of the dataset distribution across these classes is presented in Table 1.

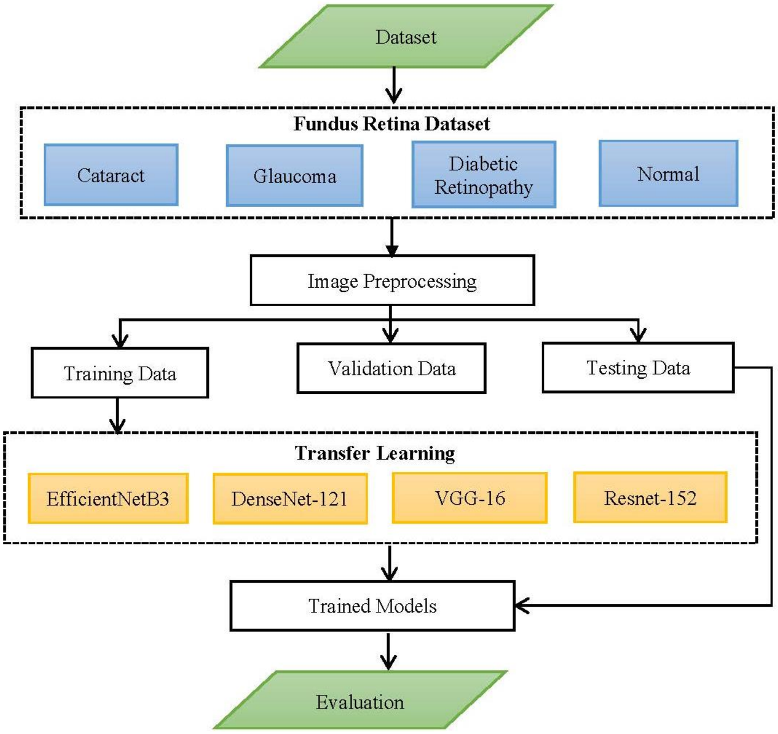

Figure 1. Research method workflow

Table 1. Division of eye disease dataset

Class	Data
Cataracts	1,038
Diabetic retinopathy	1,098
Glaucoma	1,007
Normal	1,074

2.2. Image preprocessing

The image preprocessing stage encompasses a series of techniques applied to the dataset to enhance its usability for subsequent processes. A critical aspect of this stage involves noise removal to ensure data quality. Additionally, since the loaded data exhibits variations in resolution sizes, standardization is achieved by resizing all images to a consistent dimension of 224×224×3. In the realm of deep learning, a substantial volume of data is typically essential for optimal model performance. However, image classification methods often encounter limitations due to insufficient data availability for model training. Consequently, image augmentation emerges as an effective technique, particularly when dealing with datasets lacking ample data points. This method proves valuable by augmenting the training dataset without the necessity of acquiring additional data, thereby circumventing the need for extra storage capacity. In the context of this research, the Keras library is harnessed to leverage the image data generator function, which plays a pivotal role in preventing overfitting. This function encompasses various graphical parameters designed to generate synthetic images [13]. Within this study, specific parameters are employed, including the preprocessing function=scalar and horizontal flip=true.

2.3. Data splitting

During the data processing stage, a data splitting process, commonly referred to as data division, is conducted in order to acquire a dataset that is proportionally representative. The utilization of this stage is also employed to mitigate bias during the assessment of model performance. The data division ratio utilized in this study involves allocating 90% of the data for training purposes, while 5% is allocated for testing and another 5% for validation. The training data consisted of 3,795 instances, while the test and validation datasets each contained 211 instances. The data will subsequently be partitioned randomly utilizing the random state function set to 123.

2.4. Transfer learning

Transfer learning is a widely recognized method that utilizes the structural framework of a pre-existing model to improve the precision of models, especially when faced with datasets that are not large enough for extensive training. Transfer learning leverages feature extraction techniques derived from pre-existing models. The transfer learning models have been trained on extensive datasets such as ImageNet, which consists of 1.2 million images belonging to 1,000 different classes. This training process enables the models to acquire significant features that can be applied to new target data [14], [15].

One of the main benefits of utilizing transfer learning is its capacity to optimize the duration of training and reduce generalization errors, as evidenced by previous studies [16]. Within the scope of this research, the utilization of transfer learning methodology is observed, employing four pre-trained CNN architecture models: EfficientNetB3, DenseNet-121, VGG-16, and ResNet-152. These models have been specifically customized for the purpose of classifying eye diseases. The application of transfer learning models in a strategic manner greatly enhances the efficiency and effectiveness of the classification task being performed.

2.4.1. EfficientNetB3

EfficientNetB3 is a CNN model that includes three critical elements in its architecture: width, depth, and resolution. This combination is strategically designed to achieve higher levels of accuracy while simultaneously minimizing both the optimal parameter size and the number of floating-point operations (FLOPs) [17]. EfficientNetB3's architecture includes two convolution layers, seven mobile bottleneck convolution layers, one pooling layer, and one fully connected layer. Within each convolution layer, the process begins with a convolutional operation with a 3×3 kernel size and a filter value of 24. Following that, the rectified linear unit (ReLU) activation function is used, which is followed by the max pooling operation, which is also followed by the ReLU activation function. This operation sequence continues until the network reaches the 7th convolution layer. Following a series of convolutions, ReLU activation functions, and max pooling, the next step involves creating a fully connected layer and applying the SoftMax function. This final step is critical in classifying the extracted features, yielding an output indicating the image class under consideration [18].

2.4.2. DenseNet-121

DenseNet represents a CNN architecture distinguished by its unique approach of connecting each layer to all subsequent layers in a feed-forward manner. This connectivity ensures that each subsequent layer receives input feature-maps from all preceding layers. The utilization of DenseNet offers several advantages, including the mitigation of gradient-related challenges, fortification of feature propagation throughout the network, and reduction in the overall number of parameters required for effective operation [19]. Furthermore, a prominent characteristic of DenseNet is the sequential connection density of layers within the network. In the context of this study, DenseNet-121 is the chosen variant, characterized by four dense blocks, three transition layers, and a total of 121 layers. This layer composition encompasses 117 convolution layers, three transition layers, and one classification layer. Each convolution layer is associated with a composite operation sequence, consisting of batch normalization (BN), ReLU, and convolution operations. The classification subnetwork of DenseNet-121 encompasses global average pooling at 77%, followed by a fully-connected layer with 1000 dimensions, concluding with the SoftMax function for classifying the extracted features and producing the desired image class output [20].

2.4.3. VGG-16

VGG-16 stands as a prominent CNN architecture model that has demonstrated remarkable performance on the ImageNet dataset. Notably, it secured the top position in the ImageNet 2014 visual vision challenge. The architecture of VGG-16 is characterized by a specific convolutional filter specification, employing 3×3 filters. VGG-16 is designed to process RGB images with dimensions of 224×224 pixels. As an initial step, it normalizes the pixel values, which typically range from 0 to 255, to the normalized scale of 0 to 1. Subsequently, the image undergoes a series of convolutional layers and fully connected layers. VGG-16 boasts a total of 13 convolution layers and 3 fully connected layers within its structure. To manage image size reduction and augment filter depth, VGG-16 incorporates 2D MaxPool layers strategically. The number of filters progressively escalates with the depth of the model, commencing at 64 filters and progressively increasing to 128, 256, and 512 filters as features are extracted from the input image. The ultimate output of the VGG-16 model consists of a feature representation derived from the input image, rendering it well-suited for a variety of classification or detection tasks.

2.4.4. RestNet-152

ResNet, short for residual networks, represents a CNN architecture meticulously engineered to address the notorious problem of gradient loss encountered in deep neural networks. This innovation is achieved through the implementation of skip connections between layers, a technique referred to as residual learning. Such architectural enhancement culminates in a network that is considerably more tractable during the training process. This, in turn, facilitates the design of deeper networks, and it exerts a favorable impact on overall model accuracy. Among the various ResNet architectures available, ResNet-152 stands out, boasting the lowest recorded top-1 error and top-5 error rates, approximately measuring 21.43% and 5.71%, respectively [21]. The ResNet-152 architecture itself unfolds with specific specifications: it commences with a 7×7 convolution operation, featuring a stride of 2. Subsequently, a 3×3 max-pooling operation with a stride of 2 is executed. Furthermore, the model incorporates batch normalization (1×1), followed by a ReLU activation (1×1) and a 3×3 convolution operation within each of the conv2_x, conv3_x, conv4_x, and conv5_x stages. The final steps involve average pooling and the Softmax function to yield classification results.

2.5. Training model

The primary frameworks for developing the model in this study are Python and TensorFlow. The runtime environment used is GPU, which provides significant benefits in terms of improved performance and reduced training time, especially when dealing with complex neural network models [22]. In the experimental phase, the model is pretrained by using the ImageNet public dataset to initialize weight values. Comprehensive hyperparameter tuning is performed to improve the model's efficiency. This entails fine-tuning a variety of parameters such as learning rate, momentum, epoch count, batch size, and others [23]. A custom Keras callback subclass is created to ensure efficient performance monitoring throughout the training process. This callback system is critical in monitoring the model's training progress and facilitating parameter adjustments to improve the training monitoring process. Table 2 shows the hyperparameter tuning used.

Table 2. Tuning hyperparameter

Class	Data
Loss function	Categorical cross-entropy
Optimizer	Adamax
Activation	Softmax
Learning rate	[(0.01), (0.001), (0.0001)]
Epoch	[(10), (20), (30)]
Batch size	40
Momentum	0.99

2.6. Evaluation

Evaluating a model's performance is an important step in the process, especially when determining the best model. The confusion matrix is a valuable technique for evaluating classification model performance. This matrix represents the model's predictions in relation to the actual data conditions in detail. Two fundamental metrics, accuracy and the F1-score, can be derived from the confusion matrix. The proportion of correctly predicted samples in relation to the total sample count is referred to as accuracy. The F1-score, on the other hand, provides a combined measure that balances precision and recall. Precision is the ratio of correctly predicted positive cases to total positive predictions, whereas recall is the ratio of true positive predictions to total true positive data instances.

The outcomes of these computations serve as a means to assess the algorithm's efficacy. Accuracy values fall within a range of 0 to 1 or 0 to 100% when expressed as a percentage. A higher accuracy value is indicative of a more effective utilization of the algorithm [24].

3. RESULTS AND DISCUSSION

The primary objective of employing transfer learning models, in conjunction with rigorously tested hyperparameter values, is to facilitate a comparative evaluation of various model architectures. This comparative analysis investigates the performance of EfficientNetB3, DenseNet-121, VGG-16, and ResNet-152, with a specific emphasis on assessing accuracy and overall efficacy. In order to accomplish this objective, modifications are made to the learning rate function, taking into consideration the substantial influence of the learning rate on accuracy. An increased learning rate accelerates the training procedure but may potentially undermine the accuracy of the network, whereas a reduced learning rate results in a slower yet potentially more precise or stable training process. In addition, a range of diverse epoch values is utilized in this study, acknowledging the significant impact that epoch settings can have on the performance of the model. Epochs are of utmost importance in the context of machine learning, serving as iterative stages that indicate the level of training the algorithm has undergone on the complete dataset. Insufficient numbers of training iterations can potentially lead to the model's inability to capture intricate data patterns, resulting in the phenomenon known as underfitting. On the contrary, an excessive number of epochs can lead to overfitting of the model, wherein it memorizes the training set.

Therefore, the evaluation process assumes significant importance in this research. The evaluation method chosen for this study involves the implementation of a confusion matrix, which incorporates various metrics such as precision, recall, F1-score, and accuracy. Every individual value within the matrix plays a significant role in assessing the model's effectiveness and facilitates comprehension of the specific errors made by the model. The comprehensive presentation of the research findings, which includes the matrix values for each model, can be found in Table 3.

Table 3. Model comparison results

Model	Learning rate	Epoch	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
EfficientNetB3	0.01	10	89	88.75	88.5	88.63
	0.001	20	**93.75**	**93.75**	**94**	**93.84**
	0.0001	30	92.5	92.5	92.5	92.42
DenseNet-121	0.01	10	77.75	66.5	66.75	68.24
	0.001	20	95.5	95.75	95.75	95.73
	0.0001	30	**96.5**	**96**	**96.25**	**96.20**
VGG-16	0.01	10	74.25	49.75	47.25	51.18
	0.001	20	91.25	91.25	91	91
	0.0001	30	**96.25**	**95.75**	**96**	**95.73**
ResNet-152	0.01	10	21.5	24.7	11.28	29.85
	0.001	20	95	94.75	94.75	**94.78**
	0.0001	30	**95**	**95**	**95**	93.84

The results of the comparative analysis demonstrate that the EfficientNetB3 model achieves superior performance when trained with a learning rate of 0.001 and an epoch count of 20. DenseNet-121 exhibits superior performance compared to other models, especially when utilizing a learning rate of 0.0001 and an epoch value of 30. Furthermore, it is worth noting that both VGG-16 and ResNet-152 exhibit remarkable performance when a learning rate of 0.0001 and an epoch count of 30 are employed. The aforementioned results highlight the significant influence of different learning rates and epoch values on the resultant accuracy measure. It is imperative to acknowledge that there exists a positive correlation between accuracy and epoch count, suggesting a unidirectional association. Put simply, when the number of epochs is increased, the accuracy values for both the training and validation datasets also increase. On the other hand, a negative correlation can be observed between the number of epochs and the loss metric. This suggests that as the number of epochs increases, there is a corresponding decrease in the loss value observed in the training data.

Upon evaluating the accuracy values obtained from the trained models, it is evident that the DenseNet-121 model achieved the highest accuracy, reaching an impressive 96.20%. This exceptional performance was achieved with a learning rate set at 0.0001 and 30 epochs. Following closely, VGG-16 attained an accuracy of 95.73% under the same configuration of a learning rate of 0.0001 and 30 epochs. ResNet-152 secured an accuracy of 94.76% using a learning rate of 0.001 and 20 epochs, while EfficientNetB3 yielded an accuracy of 93.84% with the same learning rate and epoch settings. It is essential to acknowledge that the architectural design of these models can exert a profound influence on their accuracy. Variables such as layer size, number of layers, layer type, and the interplay between layers significantly impact the models' ability to discern and interpret patterns within the data.

Previous studies in this domain have also yielded noteworthy results. For instance, Islam *et al.* [25] reported an accuracy of 96.25% in detecting glaucoma disease using EfficientNetB3. Similarly, Paradisa *et al.* [26] utilized DenseNet-121+Inception-ResNetV2 architecture to detect diabetic eye disease, garnering an accuracy of 91%. Notably, these studies often focused on single-eye disease detection or carried out the classification separately. In contrast, this research encompassed the simultaneous classification of three eye disease categories cataract, glaucoma, and diabetic retinopathy yielding an average accuracy of 96.20% with the DenseNet-121 architecture, surpassing other models, including VGG-16, EfficientNetB3, and ResNet-152.

These accuracy results across various models affirm their effectiveness in classifying retinal fundus image data for the detection of the aforementioned eye diseases. Several factors can influence the accuracy outcomes, including dataset size, data processing stages like image augmentation and data splitting, and the configuration of hyperparameters during the model training process. These factors will affect the accuracy produced by each model. The better the data preparation done before model training. Then the better the accuracy produced by each model. Nagpal *et al.* [27] states that data preprocessing or data preparation before data is further processed will reduce noise in the image to be processed, this will result in better classification.

4. CONCLUSION

The process of detecting eye diseases in our research faced several challenges, with limited data availability being a significant impediment that could potentially affect the accuracy of our models. Transfer learning, a fundamental concept employed in this study, involves leveraging pre-trained models that have been trained on extensive datasets to facilitate the understanding of new data. This approach was adopted to enhance the accuracy of our models. The proposed method performs well with several transfer learning models, including EfficientNetB3, DenseNet-121, VGG-16, and ResNet-152, coupled with the exploration of various hyperparameters, we achieved the highest accuracy with the DenseNet-121 model. It demonstrated precision, recall, and F1-score values of 96.5%, 96%, and 96.25%, respectively, along with an overall accuracy of 96.20%. Furthermore, our investigation highlighted the significant impact of different learning rates and epochs on accuracy. Notably, the use of a learning rate set at 0.0001 and an epoch count of 30 consistently yielded higher accuracy values compared to other configurations. Consequently, our research demonstrates the effectiveness of the selected transfer learning model in classifying retinal fundus image data for the detection of cataract, glaucoma, and diabetic retinopathy diseases. This study serves as a valuable benchmark for future research endeavours, offering insights into potential avenues for improving accuracy. Future investigations could explore a broader spectrum of transfer learning models and leverage larger datasets to further enhance the classification performance in this critical domain of eye disease detection.

ACKNOWLEDGEMENTS

The authors express gratitude to the Universitas Muhammadiyah Surakarta for providing research support, enabling the completion of this research.

REFERENCES

[1] S. R. Flaxman *et al.,* "Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis," *Lancet Glob. Heal*., vol. 5, no. 12, pp. e1221–e1234, Dec. 2017, doi: 10.1016/S2214-109X(17)30393-5.

[2] E. Mardalena and F. Hayati, "Prevalence of blindness due to cataracts in RSUD dr. Zainoel Abidin Banda Aceh," *Jurnal Riset Sains dan Teknologi JRST*, vol. 11, no. 3, pp. 597–602, 2021, doi: 10.47647/jsr.v11i3.794

[3] V. Nangia *et al.,* "Prevalence and causes of blindness and vision impairment: magnitude, temporal trends and projections in south and Central Asia," *British Journal of Ophthalmology,* vol. 103, no. 7, pp. 871–877, Nov. 2018, doi: 10.1136/bjophthalmol-2018-312292

[4] R. S. Salvi, S. R. Labhsetwar, P. A. Kolte, V. S. Venkatesh, and A. M. Baretto, "Predictive analysis of diabetic retinopathy with transfer learning," in *2021 International Conference on Nascent Technologies in Engineering, ICNET 2021* - Proceedings, 2021, doi: 10.1109/ICNTE51185.2021.9487789.

[5] W. Supriyanti and D. A. Anggoro, "Classification of pandavas figure in shadow puppet images using convolutional neural networks," *Khazanah Informatika : Jurnal Ilmu Komputer dan Informatika*, vol. 7, no. 1, pp. 18–24, 2021, doi: 10.23917/khif.v7i1.12484.

[6] I. D. Apostolopoulos and T. A. Mpesiana, "COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, Jun. 2020, doi: 10.1007/s13246-020-00865-4.

[7] R. Sarki *et al.,* "Image preprocessing in classification and identification of diabetic eye diseases," *Data Science and Engineering*, vol. 6, no. 4, pp. 455–471, Aug. 2021. doi: 10.1007/s41019-021-00167-z.

[8] K. Pin, J. Ho Chang, and Y. Nam, "Comparative study of transfer learning models for retinal disease diagnosis from fundus images," *Computers, Materials, Continua*, vol. 70, no. 3, pp. 5821–5834, 2022, doi: 10.32604/cmc.2022.021943.

[9] A. Sugeno, Y. Ishikawa, T. Ohshima, and R. Muramatsu, "Simple methods for the lesion detection and severity grading of diabetic retinopathy by image processing and transfer learning," *Computers in Biology and Medicine*, vol. 137, p. 104795, Oct. 2021, doi: 10.1016/j.compbiomed.2021.104795.

[10] B. Taşar, "SkinCancerNet: Automated classification of skin lesion using deep transfer learning method," *Traitement Du Signal*, vol. 40, no. 1, pp. 285–295, Feb. 2023, doi: 10.18280/ts.400128.

[11] D. Le *et al.,* "Transfer learning for automated OCTA detection of diabetic retinopathy," *Translational Vision Science and Technology,* vol. 9, no. 2, p. 35, Jul. 2020, doi: 10.1167/tvst.9.2.35.

[12] H. Imaduddin and B. A. Hermansyah, "Transfer learning for detecting COVID-19 on x-ray using deep residual network," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 6, 2022, doi: 10.11591/eei.v11i6.4334.

[13] S. Gu, M. Pednekar, and R. Slater, "Improve image classification using data augmentation and neural networks," *SMU Data Science Review*, vol. 2, no. 2, Aug. 2019, Accessed: Apr. 17, 2023, [Online]. Available: https://scholar.smu.edu/datasciencereview/vol2/iss2/1.

[14] M. Huh, P. Agrawal, and A. A. Efros, "What makes ImageNet good for transfer learning?," *Berkeley Artif. Intell. Res. Lab. UC Berkeley*, Aug. 2016, Accessed: Apr. 17, 2023. [Online]. Available: https://arxiv.org/abs/1608.08614v2.

[15] R. Godasu, D. Zeng, and K. Sutrave, “Transfer learning in medical image classification: challenges and opportunities opportunities,” *MWAIS 2020 Proceedings*., pp. 5–28, 2020, Accessed: Apr. 18, 2023.

[16] A. Jaradat *et al.*, “Automated monkeypox skin lesion detection using deep learning and transfer learning techniques,” *International Journal of Environmental Research and Public Health*, vol. 20, no. 5, p. 4422, Mar. 2023, doi: 10.3390/ijerph20054422.

[17] A. Bhatt, A. Ganatra, and K. Kotecha, “COVID-19 pulmonary consolidations detection in chest X-ray using progressive resizing and transfer learning techniques,” *Heliyon*, vol. 7, no. 6, Jun. 2021, doi: 10.1016/j.heliyon.2021.e07211.

[18] F. Zaelani and Y. Miftahuddin, “Comparison of EfficientNetB3 and MobileNetV2 methods for identification of fruit types using leaf features,” *JITTER*, vol. 9, no. 1, pp. 1–11, 2022, doi: 10.33197/jitter.vol9.iss1.2022.911.

[19] G. Huang, Z. Liu, and L. van der Maaten, “Densely connected convolutional networks,” In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4700-4708. 2017, doi: 10.48550/arXiv.1608.06993.

[20] X. Li, X. Shen, Y. Zhou, X. Wang, and T. Q. Li, “Classification of breast cancer histopathological images using interleaved DenseNet with SENet (IDSNet),” *PLoS One*, vol. 15, no. 5, May 2020, doi: 10.1371/journal.pone.0232127.

[21] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Dec. 2016, doi: 10.1109/CVPR.2016.90.

[22] R. D. Nurfita, G. Ariyanto, “Tensorflow-based deep learning implementation for fingerprint recognition,” *Journal Emit*, vol. 18, no. 01, pp. 22–27, 2018, doi: 10.23917/emitor.v18i01.6236.

[23] C. M. Sharma, K. Tomar, R. K. Mishra, and V. M. Chariar, “Indian sign language recognition using fine-tuned deep transfer learning model,” *Proceedings of the 1st International Conference on Innovation in Computer and Information Science*, 2021. doi:10.5220/0010790300003167

[24] E. F. Ohata, J. V. S. das Chagas, G. M. Bezerra, M. M. Hassan, V. H. C. de Albuquerque, and P. P. R. Filho, “A novel transfer learning approach for the classification of histological images of colorectal cancer,” *Journal Supercomput*, vol. 77, no. 9, pp. 9494–9519, 2021, doi: 10.1007/s11227-020-03575-6.

[25] M. T. Islam, S. T. Mashfu, A. Faisal, S. C. Siam, I. T. Naheen and R. Khan, “Deep learning-based glaucoma detection with cropped optic cup and disc and blood vessel segmentation,” in *IEEE Access*, vol. 10, pp. 2828-2841, 2022, doi: 10.1109/ACCESS.2021.3139160.

[26] R. H. Paradisa, A. Bustamam, W. Mangunwardoyo, A. A. Victor, A. R. Yudantha, and P. Anki, “Deep feature vectors concatenation for eye disease detection using fundus image,” *Electronics*, vol. 11, no. 1, 2022, doi: 10.3390/electronics1101002.

[27] D. Nagpal, N. Alsubaie, B. O. Soufiene, M. S. Alqahtani, M. Abbas, and H. Almohiy, “Automatic detection of diabetic hypertensive retinopathy in Fundus images using transfer Learning,” *Applied Sciences*, vol. 13, no. 8, p. 4695, Apr. 2023, doi: 10.3390/app13084695.

BIOGRAPHIES OF AUTHORS

Helmi Imaduddin is currently a lecturer at the Faculty of Communication and Informatics, Universitas Muhammadiyah Surakarta. His primary areas of interest in research include natural language processing, image processing, and classification. In 2015, he successfully completed his Bachelor of Science in Informatics Engineering from Universitas Muhammadiyah Surakarta. Subsequently, in 2020, he obtained his Master of Science in Information Technology from Gadjah Mada University in Yogyakarta. He can be contacted at email: helmi.imaduddin@ums.ac.id.

Alivia Rahma Sakina is an informatics engineering student at Muhammadiyah University of Surakarta. She has research interests in data engineering, computer vision and machine learning. Recently She has been actively working as a data engineer in the telecommunications network sector. She can be contacted at email: aliviarahma56@gmail.com.