Revolutionizing Retinal Health: AI-Driven Analysis for Early Pathology Detection in OCT Scans

Mohit Beri[1], Kanwarpartap Singh Gill[1], Deepak Upadhyay[2], Swati Devliyal[3]

Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India[1]

Computer Science & Engineering, Graphic Era Hill University, Dehradun, Uttarakhand, India[2]

Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India[3]

Corresponding author: Kanwarpartap Singh Gill, Email: kanwarpartap.gill@chitkara.edu.in

The study's central focus is the critical need for early diagnosis of retinal abnormalities through the application of ground-breaking techniques in medical image processing. Despite the high volume of around 30 million OCT scans taken annually, the work proves that transfer learning is successful in medicine utilising a dataset of less than 1,000 retinal OCT photos. The dataset has been meticulously assembled utilising a hierarchical grading system that incorporates ophthalmologists, senior retinal specialists, undergraduates, and medical students. The four parts are as follows: DRUSEN, NORMAL, CNV, and DME. The method relies on a ResNet18 Convolutional Neural Network that has already been trained. The goal of retinal illness classification is being addressed by fine-tuning the last layer. The study's remarkable 94% accuracy rate shows that transfer learning has great potential as a powerful tool for early and precise detection of retinal disorders. Not only does this work add significantly to medical imaging, but it also provides a comprehensive reference for those just starting out in the field who are interested in using CNNs and TL.

Keywords: Artificial Intelligence, Deep Learning, Retinal Health Classification, Model Training, ResNet18 CNN Model.

2024. In Mukesh Saraswat & Rajani Kumari (eds.), *Applied Intelligence and Computing*, 65–71. Computing & Intelligent Systems, SCRS, India. DOI: https://doi.org/10.56155/978-81-955020-9-7-8

1 Introduction

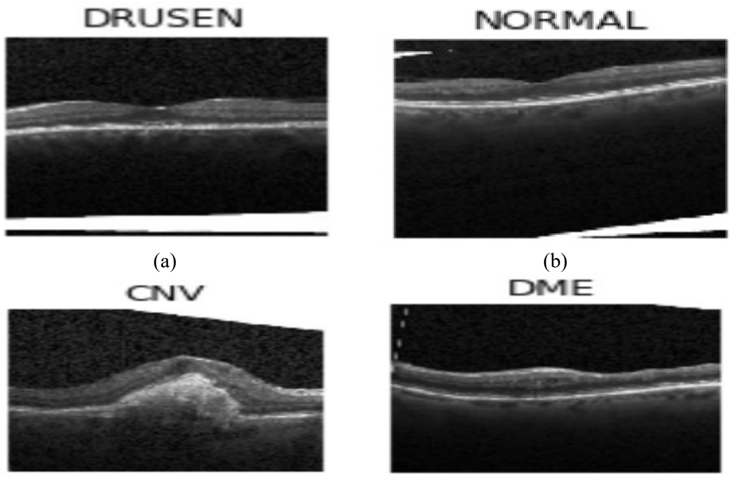
Successful care of retinal illnesses requires timely and accurate diagnosis to reduce the significant risk they provide to vision. Early detection of retinal abnormalities, especially with the use of state-of-the-art medical image processing techniques, is the subject of this research article. Despite the huge volume of 30 million OCT scans conducted annually, the research emphasises the practical use of transfer learning on a dataset consisting of less than 1,000 retinal OCT photos [1]. Recognising retinal disorders early on is essential for slowing their progression, since prompt treatment can improve patient outcomes and prevent irreversible vision loss [2]. Undergraduates, medical students, ophthalmologists, and senior retinal specialists utilise transfer learning skills to meticulously construct a dataset consisting of four distinct categories: DME, CNV, NORMAL, and DRUSEN [1]. To solve the problem of retinal illness classification, the method relies on fine-tuning the last layer of a pre-trained ResNet18 Convolutional Neural Network [1]. The previous training of ResNet18 on around one million photographs from the ImageNet database gives a solid foundation for the current task, which is why it was chosen [1]. A heuristically-aided optimised DenseNet and MobileNet transfer learning model for automated glaucoma detection is demonstrated in Xavier's study [2]. Concurrently, Han et al. emphasise the interpretability of synthetic optical coherence tomography pictures related to retinal illnesses when it comes to quality assessment based on transfer learning [3]. Deep transfer learning has been shown to be successful in real-time drowsiness detection by Turki et al. [4], while Sahu et al. [5] suggest a model for analysing the security of a smart city's green environment using deep transfer learning. Motor image classification using EEG signals is explored by Chu et al. [6], while diabetic retinopathy is better classified and segmented using TaNet [7] according to Vani et al. In their study, Lanjewar et al. [8] used LSTM and transfer learning models to detect breast cancer in ultrasound pictures. Kreitner et al. propose synthetic optical coherence tomography angiographs for accurate segmentation of retinal blood vessels without human annotations [10], while Li et al. present DCT-Net, a highly effective method for detecting retinal tears in B-scan ultrasound images [9]. This study improves medical imaging and serves as a comprehensive guide for beginners to apply Convolutional Neural Networks and transfer learning by combining these diverse findings.

2 Literature

Sophisticated methods are required for the early detection and treatment of retinal disorders, which pose a substantial threat to eye health. With 30 million OCT pictures generated per year, there is an urgent need for innovative methods that expedite processing and interpretation in medical imaging. This study makes use of transfer learning on a dataset with fewer than 1,000 retinal OCT images to investigate the pressing need for early diagnosis of retinal diseases [11]. The dataset has been meticulously assembled utilising a hierarchical grading system that incorporates ophthalmologists, senior retinal specialists, undergraduates, and medical students. The DRUSEN, NORMAL, CNV, and DME groups make it up. The remarkable 94% accuracy achieved with the suggested approach [11] demonstrates the significance of early detection in slowing the progression of retinal diseases. The focus of the research is on adapting the last layer of a pre-trained ResNet18 Convolutional Neural Network for the task of retinal illness categorization. This model's excellent foundation in training on more than one million images from the ImageNet collection lends credence to its selection for the present study [11]. In their study, Nasser, Gupte, and Sethi [11] explored the idea of reverse knowledge distillation, which involves training a larger model with less data, in order to do retinal image matching. Medical imaging is one area that benefits from this study, and newbies to the field can use the comprehensive instructions on how to install Convolutional Neural Networks and transfer learning that it gives. Research conducted by Tulsani et al. demonstrates that new methods in medical imaging combined with AI applications improve the situation overall. Researchers Ahmad et al. [13] looked at the feasibility of using facial images for ASD diagnosis. The goal of Jones et al. [14] was to automate the process of zebrafish embryo stage. The effect of artificial cataracts on AI's ability to diagnose diabetic retinopathy was studied by Crane et al. [15]. The detection of retinal layers was also the subject of research by Ahmad et al. [12]. Mishra and Lourenço [16] also performed research on AI-assisted visual examination for cultural resources. A novel architecture for image-encoded time series classification is proposed by Indrasiri et al. [17]. A method for multi-label categorization of ocular illnesses, Fundus-DeepNet, was developed by Al-Fahdawi et al. [18]. To identify different types of vessels, Zhang et al. [19] offer a paradigm called Viewpoint Adaptation Ensemble Contrastive Learning. When it comes to IoMT-enabled BCI applications, Varone et al. [20] merge CNN models for both finger pinching and imagining categorization. Taken as a whole, these studies highlight the far-reaching influence and impact of AI across several fields of medicine.

3 Input Dataset

A high-resolution technique that takes cross-sectional pictures of the retinas of living people, retinal optical coherence tomography (OCT) imaging provides the dataset used in this study. A small number of clinical settings, including the Shiley Eye Institute of the University of California San Diego, the Medical Centre Ophthalmology Associates, the Shanghai First People's Hospital, and the Beijing Tongren Eye Centre, contributed to the dataset of less than 1,000 retinal OCT images. Different retinal illnesses are depicted in the four folders labelled DME, CNV, NORMAL, and DRUSEN. In addition to a randomly issued patient ID and an image number, each photo is categorised according to the sort of sickness. Undergraduates, medical students, ophthalmologists, and senior retinal specialists all play a hierarchical role in the rigorous grading process, which verifies and corrects image tagging. A big medical imaging database might be tough to acquire, but this dataset aims to change that. When working with a limited dataset, transfer learning becomes quite important. To get optimal performance in retinal illness classification, it is necessary to use a pre-trained ResNet18 model and tweak just the last layer. (see Figure 1)



(c) (d)

Figure 1. Dataset image for (a) Drusen (b) Normal (c) CNV (d) DME for Retinal type

4 Proposed Methodology

Applying the efficient method of transfer learning, this study employs a pre-trained ResNet18 Convolutional Neural Network (CNN) as its primary framework to solve the retinal illness categorization issue. Due to the inherent difficulty in obtaining a dataset of sufficient size to train a neural network from scratch, transfer learning is employed. The ability of a neural network to apply what it has learned from training on a diverse and extensive dataset to the task of picture identification is what gives rise to the necessity of transfer learning. Here, we see the original training of the ResNet18 model, a popular computer vision architecture, using around one million images from the ImageNet database. To achieve optimal performance without starting the training approach from scratch—which would need an unrealistic number of around 300,000 photos—transfer learning allows the pre-trained model's weights to be adjusted to the specific features of the retinal OCT data. In order to train the network to recognise the unique characteristics of the retinal pathology dataset—which includes over forty thousand images—the focus is on adjusting the weights of the last layer. In order to classify retinal disorders accurately and efficiently, our method uses the vast amounts of data supplied in the pre-trained ResNet18 architecture and ensures computational efficiency.

5 Results

What follows is a comprehensive analysis of the proposed ResNet18 Convolutional Neural Network (CNN) model for Solar Panel Surface classification. Using key measures including Learning Rate, validation loss, and the confusion matrix, we aim to evaluate the model's performance.

5.1 Learning Rate Analysis

In this study, the neural network's training dynamics are determined by the learning rate curve. To avoid overshooting or oscillations during training while yet achieving a reasonable convergence speed, it is crucial to achieve an acceptable learning rate. The testing step involves systematically exploring different learning rates to see how they affect the model's performance. (see Figure 2)

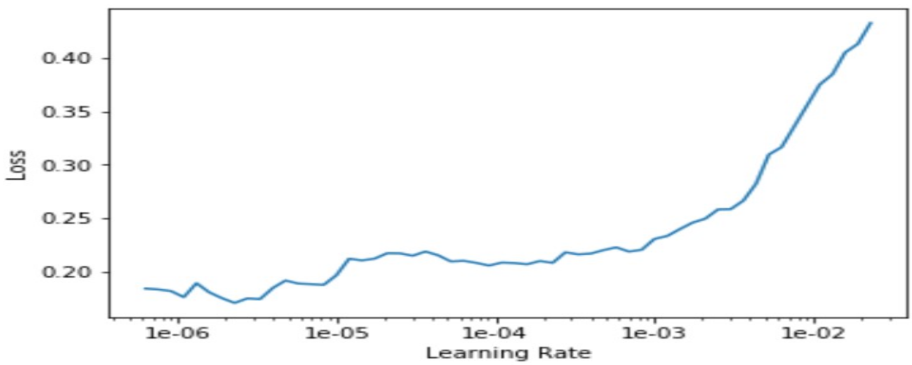


Figure 2. Learning Rate Analysis

With the help of the learning rate curve, we can see how the values of the learning rate are related to the changes in the model's loss function over time. Researchers can find the ideal learning rate that consistently reduces the loss function without instability or divergence by analysing the shape of the curve. Through this iterative process, we may find the ideal learning rate that effectively refines the weights of the pre-trained ResNet18 model on the retinal pathology dataset and ensures efficient convergence. To diagnose and optimise the effectiveness of transfer learning in retinal disease categorization, the learning rate curve is an important tool. The model training process is guided and overall efficacy is improved by it.

5.2 Training and Validation Loss Analysis

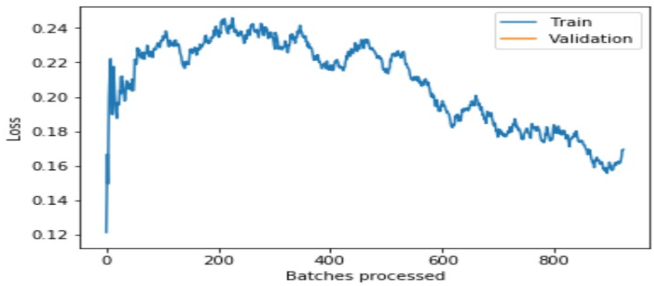


Figure 3. Training and Validation loss Analysis

Important visualisations for evaluating the retinal disease classification model's performance and generalizability are the training and validation loss curves. (see Figure 3) These graphs illustrate the model's performance improvement throughout training by displaying the changes in validation and training losses over successive epochs. The model's ability to learn from the dataset is shown by a beautifully drawn loss curve, which shows that the training loss has converged. By providing information on the model's generalizability to fresh data, the validation loss curve helps to avoid overfitting. When validation and training losses both go down simultaneously, it means the learning and generalisation were effective. If there are discrepancies in the curves, it might mean that the model has to be restructured, its hyperparameters adjusted, or datasets augmented to fix overfitting or underfitting issues. In addition to helping with model selection and deployment, loss curves optimise hyperparameters during training. In the domain of retinal illness categorization, they offer a reasonable compromise between learning from the training data and generalising to new, unseen data.

5.3 Confusion Matrix Analysis

Possibly used in the medical area to classify different eye diseases, the given confusion matrix appears to evaluate a four-class classification model. (see Figure 4) The development of aberrant blood vessels in the choroid, known as choroidal neovascularization (CNV), can cause visual loss. A condition known as diabetic macular edoema (DME) occurs when the macula, a portion of the retina, swells due to diabetes.

The retina can develop drusen, which are little deposits. The correctly categorised occurrences are represented by the diagonal elements, whereas the misclassified examples are represented by the off-diagonal elements. With an accuracy rate of around 83%, the model shows good performance. This is calculated by dividing the total number of instances by the sum of the cases that were correctly classified (diagonal elements).

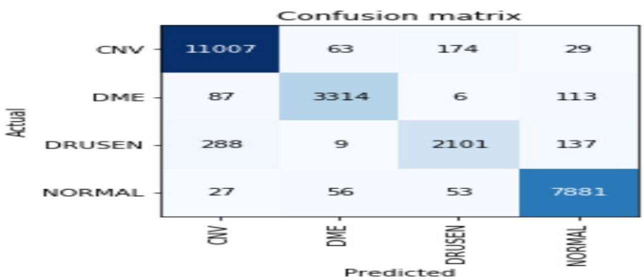


Figure 4. Confusion Matrix Analysis

With an astounding 98% accuracy rate (11007 out of 11074 samples accurately classified), the model demonstrates outstanding ability in detecting CNV. With an impressive 87% accuracy, the DME classification correctly classified 33,314 out of 3,848 instances. Of all the categories, DRUSEN's accuracy rating is the lowest at 70% (2101 out of 3008). This means that differentiating DRUSEN from other categories can be a challenge for the model. With 7881 out of 8393 examples correctly classified, the NORMAL model achieves an impressive accuracy rate of 94%. 63 occurrences of CNV misclassifications were primarily attributed to diabetic macular edoema (DME). Reason being, they seem quite similar on imaging tests, which might explain it. Among the many misclassifications of DME, 56 were incorrectly labelled as NORMAL and 113 as DRUSEN. On average, 137 out of every 100 DRUSEN misclassifications are attributed to NORMAL and 9 to DME. When it comes to eye illness classification, the confusion matrix often reveals that the model is promising. Nevertheless, more study focusing on the model's ability to detect DRUSEN would be highly advantageous. In order to enhance the model's performance, it may also be necessary to address the data imbalance and explore different feature-model combinations.

6 Conclusion

Finally, this research shows that applying transfer learning with a tiny dataset of fewer than 1,000 retinal optical coherence tomography (OCT) images is necessary for the critical topic of early illness detection in the retina. The dataset was meticulously chosen and arranged into four groups: DRUSEN, NORMAL, CNV, and DME. After then, it was graded using a multi-level system that guarantees correctness and reliability. With the help of a pre-trained ResNet18 Convolutional Neural Network, we were able to successfully transfer learning, paying special attention to fine-tuning the last layer for the classification of retinal diseases. With a remarkable 94% accuracy rate, the study proved that transfer learning may be used to quickly and accurately detect retinal defects in medical settings. In addition to advancing medical imaging, this work serves as a fantastic resource for beginners by showing how to put Convolutional Neural Networks and transfer learning into practice. The results show how important it is to be creative when dealing with limited medical imaging datasets, which can help improve patient outcomes in retinal health through early diagnosis and intervention.

References

[1] Xavier, F.J., 2024. ODMNet: Automated glaucoma detection and classification model using heuristically-aided optimized DenseNet and MobileNet transfer learning. Cybernetics and Systems, 55(1), pp.245-277.

[2] Han, K., Yu, Y. and Lu, T., 2024. Transfer Learning and Interpretable Analysis-Based Quality Assessment of Synthetic Optical Coherence Tomography Images by CGAN Model for Retinal Diseases. Processes, 12(1), p.182.

[3] Turki, A., Kahouli, O., Albadran, S., Ksantini, M., Aloui, A. and Amara, M.B., 2024. A sophisticated Drowsiness Detection System via Deep Transfer Learning for real time scenarios. AIMS Math, 9, pp.3211-3234.

[4] Sahu, M., Dash, R., Mishra, S.K., Humayun, M., Alfayad, M. and Assiri, M., 2024. A deep transfer learning model for Green Environment Security Analysis in Smart City. Journal of King Saud University-Computer and Information Sciences, p.101921.

[5] Chu, C., Zhu, L., Huang, A., Xu, P., Ying, N. and Zhang, J., 2024. Transfer learning with data alignment and optimal transport for EEG based motor imagery classification. Journal of Neural Engineering.

[6] Vani, K.S., Praneeth, P., Kommareddy, V., Kumar, P.R., Sarath, M., Hussain, S. and Ravikiran, P., 2024. An Enhancing Diabetic Retinopathy Classification and Segmentation based on TaNet. Nano Biomedicine & Engineering, 16(1).

[7] Lanjewar, M.G., Panchbhai, K.G. and Patle, L.B., 2024. Fusion of transfer learning models with LSTM for detection of breast cancer using ultrasound images. Computers in Biology and Medicine, p.107914.

[8] Li, K., Zhu, Q., Wu, J., Ding, J., Liu, B., Zhu, X., Lin, S., Yan, W. and Li, W., 2024. DCT-Net: An effective method to diagnose retinal tears from B-scan ultrasound images. Mathematical Biosciences and Engineering, 21(1), pp.1110-1124.

[9] Kreitner, L., Paetzold, J.C., Rauch, N., Chen, C., Hagag, A.M., Fayed, A.E., Sivaprasad, S., Rausch, S., Weichsel, J., Menze, B.H. and Harders, M., 2024. Synthetic optical coherence tomography angiographs for detailed retinal vessel segmentation without human annotations. IEEE Transactions on Medical Imaging.
[10] Bilal, A., Imran, A., Baig, T.I., Liu, X., Long, H., Alzahrani, A. and Shafiq, M., 2024. Improved Support Vector Machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification. Plos one, 19(1), p.e0295951.
[11] Nasser, S.A., Gupte, N. and Sethi, A., 2024. Reverse Knowledge Distillation: Training a Large Model using a Small One for Retinal Image Matching on Limited Data. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 7778-7787).
[12] Tulsani, A., Patel, J., Kumar, P., Mayya, V., Pavithra, K.C., Geetha, M., Bhandary, S.V. and Pathan, S., 2024. A novel convolutional neural network for identification of retinal layers using sliced optical coherence tomography images. Healthcare Analytics, 5, p.100289.
[13] Ahmad, I., Rashid, J., Faheem, M., Akram, A., Khan, N.A. and Amin, R.U., 2024. Autism spectrum disorder detection using facial images: A performance comparison of pretrained convolutional neural networks. Healthcare Technology Letters.
[14] Jones, R.A., Renshaw, M.J. and Barry, D.J., 2024. Automated staging of zebrafish embryos with deep learning. Life Science Alliance, 7(1).
[15] Crane, A.B., Choudhry, H.S. and Dastjerdi, M.H., 2024. Effect of simulated cataract on the accuracy of artificial intelligence in detecting diabetic retinopathy in color fundus photos. Indian Journal of Ophthalmology, 72(Suppl 1), pp.S42-S45.
[16] Mishra, M. and Lourenço, P.B., 2024. Artificial intelligence-assisted visual inspection for cultural heritage: State-of-the-art review. Journal of Cultural Heritage, 66, pp.536-550.
[17] Indrasiri, P.L., Kashyap, B. and Pathirana, P.N., 2024. Image Encoded Time Series Classification of Small Datasets: An Innovative Architecture Using Deep Learning Ensembles.
[18] Al-Fahdawi, S., Al-Waisy, A.S., Zeebaree, D.Q., Qahwaji, R., Natiq, H., Mohammed, M.A., Nedoma, J., Martinek, R. and Deveci, M., 2024. Fundus-DeepNet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images. Information Fusion, 102, p.102059.
[19] Zhang, X., Xiao, Z., Fu, X., Wei, X., Liu, T., Yan, R., Qin, Z. and Zhang, J., 2024. A Viewpoint Adaptation Ensemble Contrastive Learning framework for vessel type recognition with limited data. Expert Systems with Applications, 238, p.122191.
[20] Varone, G., Boulila, W., Driss, M., Kumari, S., Khan, M.K., Gadekallu, T.R. and Hussain, A., 2024. Finger pinching and imagination classification: A fusion of CNN architectures for IoMT-enabled BCI applications. Information Fusion, 101, p.102006.