Eye Disease Recognition

KAVYA M
e5222044@sret.edu.in

***Abstract*—** Early detection of eye diseases through fundus images has always posed a challenge for clinicians due to the complexity and potential for human error in manual diagnosis. An automated computer system for detecting ocular diseases is essential to improve accuracy and efficiency in identifying various eye conditions using fundus images. This research proposes a deep learning-based approach to facilitate ocular disease detection. We employed Vision Transformer, VGG-19, and ResNet50 image classification algorithms to classify the ODIR dataset, which contains 6,000 fundus images across eight distinct classes. These classifications represent different ocular conditions. Given the imbalance in the dataset, this study addressed the issue by equalizing the number of images per category, converting the multiclass classification problem into a binary classification task. Additionally, we enhanced the image data using Local Binary Pattern (LBP) transformations to improve the models' performance. Our results indicate that this approach provides an effective and accurate method for eye disease detection.

Introduction

Nearly 2.2 billion people around the world experience vision problems. The World Health Organization (WHO) estimates that there may have been a reduction in at least 1 billion of these incidents. Over time, there has been an increase in eye illnesses, with changes being one of the causes of a change in human behavior caused by technology and the creation of technological apparatus. That impact caused ocular diseases have had a significant impact on modern human life. Common eye diseases include glaucoma, diabetes, and hypertension. Blindness is brought on by cataracts, pathological myopia, etc. could occur. Even though eye diseases' effects can be blindingly severe and result in early detection of the disorders can lessen the impact of the disease. Growing older and contact with certain substances, UV radiation, and genetic issues. With the prediction being digitized the model could be utilized for the analysis of the ocular disease. The disease should be accurately identified at all times. and effectiveness. One of the most important human organs is the eye. primarily vision aids in the recognition and detection of 3D objects. loss of one eyesight or both eyesight may cause a person to live an unsettling way of life because people make decisions based on what they observe in their daily lives. The effect of vision unsettling health can have an economic and personal impact. These symptoms are linked to a variety of eye illnesses. These include excruciating eye pain, an abrupt loss of vision in one or both eyes, fuzzy vision, red eyes, and droopy eyelids.

Given the serious effects of eye problems on people, to find eye diseases, this research was done on lifestyle with greater precision when using the fundus image. This study is primarily concerned with accurate recognition taking into account the Local Binary Pattern features (LBP) of ocular disease. It sights to create a variety of extracting features techniques and Neural Networks to distinguish typical visual disorders involving fundus photographs.

We discovered that eye illness identification using deep learning focuses on just one aberration in comparison to the prior methods. For various challenges, several designs were utilised in this paper, and the outcomes are fairly good. Thus, with accuracy more than 90% for each activity, illnesses including glaucoma, diabetic retinopathy, and cataract are adequately managed.

Dataset

We made use of the kaggle data collection for our research. In the dataset, we can find Diabetes, Glaucoma, Cataract, Related Macular Degeneration, Myopia, Hypertensive Retinopathy, and others.

The lens of your eye, which is typically straightforward, becomes clouded by a cataract.. Most cataracts grow slowly and don't initially impair your vision. However, cataracts will eventually obstruct your vision over time. Because of the white film it produces, it is one of the simplest abnormalities to spot. Figure 1 shows how a realistic depiction of a cataract appears in fundus images.

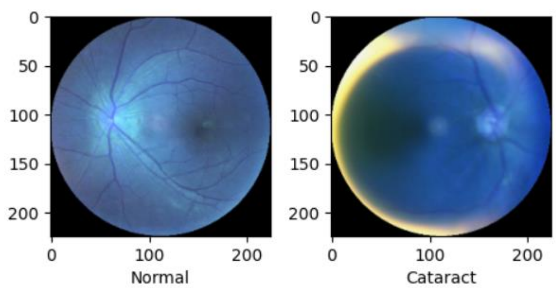

Fig.1: Normal Eye vs Cataract Eye

Image Preprocessing The overall performance of the model is significantly influenced by the quality of the cornea's fundus analysis, making it crucial to focus on this area. However, noise in the images can negatively impact the accuracy of the classification model. When converting the image to grayscale, the fundus color often closely resembles the background color, which can cause confusion for the classifier, especially with images that have white backgrounds. To address this issue and make it easier to identify points of interest, the backgrounds of all images were manually removed as the first step in the pre-processing phase.

Feature Extraction The texture elements of the image are extracted using a technique called the local binary pattern. Each pixel in the picture has a value which the LBP receives, and the form of the LBP histogram distribution is used to determine the regularity of the image texture. The LBP is more efficient in computer vision operations due to its strong discriminative strength and computational simplicity. The field of medical image processing makes extensive use of LBP.

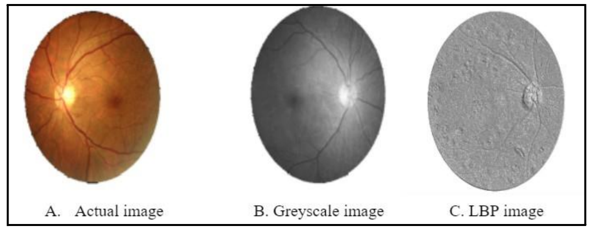

Fig.2: Actual image vs Greyscale image vs LBP image

Proposed Architecture

VGG-19

The advanced CNN model, VGG-19, consists of layers that have been pre-trained and possess a strong understanding of the shape, color, and structural features of an image. VGG-19 is a very deep network that has been trained on a vast and diverse set of images, making it highly effective for complex classification tasks.

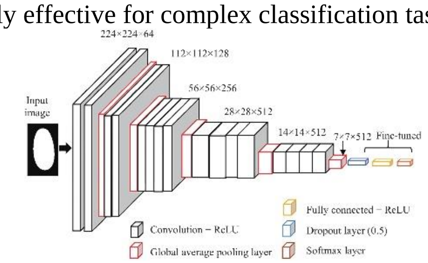

Fig.3: Architecture of VGG19

ResNet50

Instead of attempting to learn specific characteristics, ResNet or Residual Network makes use of residual learning. Residual may be easily regarded as the feature learned from that layer's input subtracted. More than ResSNet50, there are other ResNet variations. Creating a shortcut connection that omits one or more levels is the fundamental concept behind ResNet.

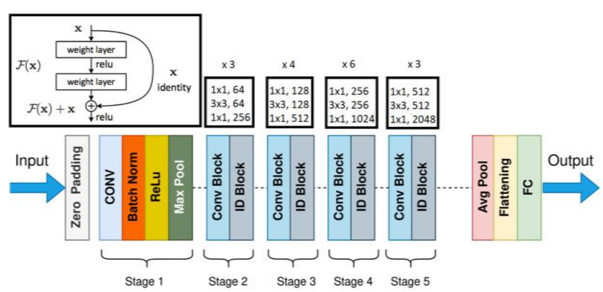

Fig.4: Architecture of ResNet50

Vision Transformer

The Vision Transformer (ViT) is an image classification model that employs a Transformer-like architecture. It operates by dividing the input image into fixed-size patches, then linearly embedding each patch and adding positional embeddings to form a sequence of vectors. These vectors are processed through a standard Transformer encoder. For classification, an additional "classification token" is appended to the sequence, which the model learns during training, allowing it to perform effective image classification.

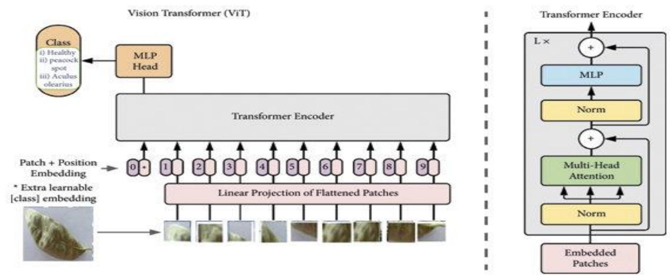

Fig.5: Architecture of Vision Transformer

Result

- **Without LBP**

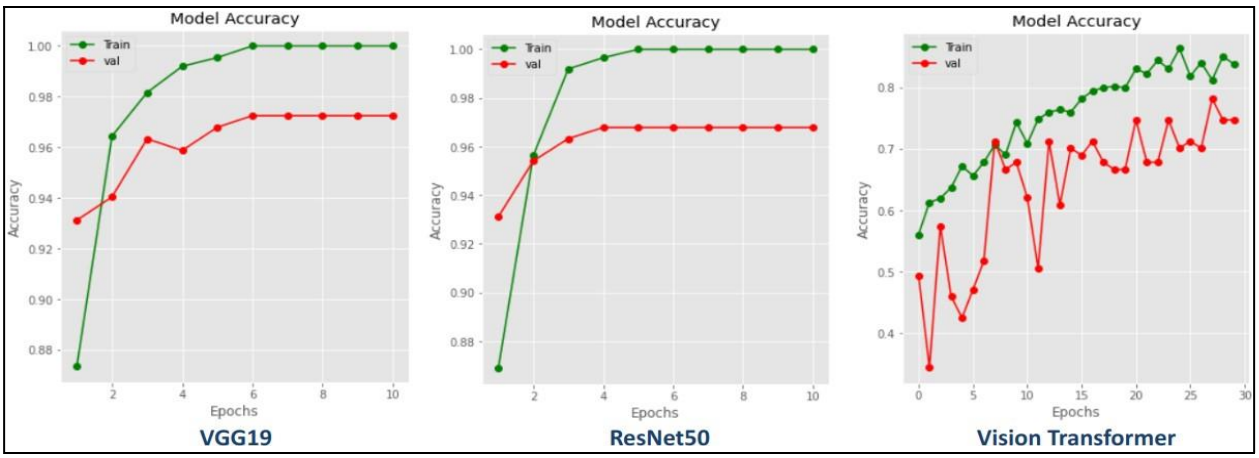

Fig.6 Model Accuracy vs Loss Curve

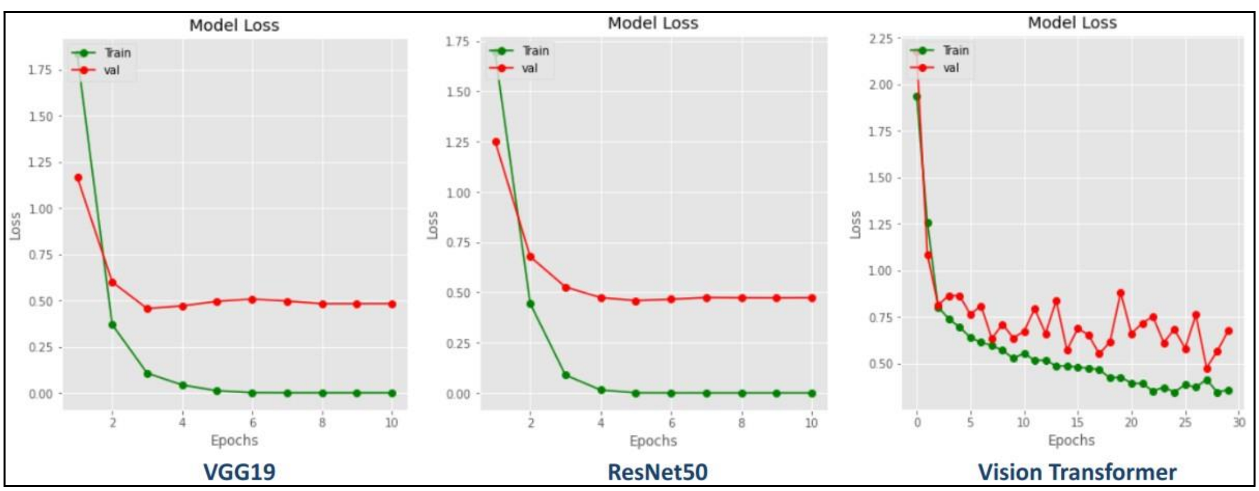

Fig.7: Model Loss vs epoch curve

Comparison table

The table shows that VGG19 performs very well with a validation accuracy of 99.08% and very less 0.08 validation loss. The vision transformer model performs very poorly compared to VGG19 and ResNet50.

	Model	Training_Accuracy	Training_Loss	Validation_Accuracy	Validation_Loss
1.	VGG19	0.996552	0.014352	0.990826	0.084891
2.	ResNet50	1.000000	0.000006	0.977064	0.131539
3.	Vision Transformer	0.863346	0.344338	0.781609	0.476366

Fig.8: Result Comparison of all models before using LBP

- **With LBP**

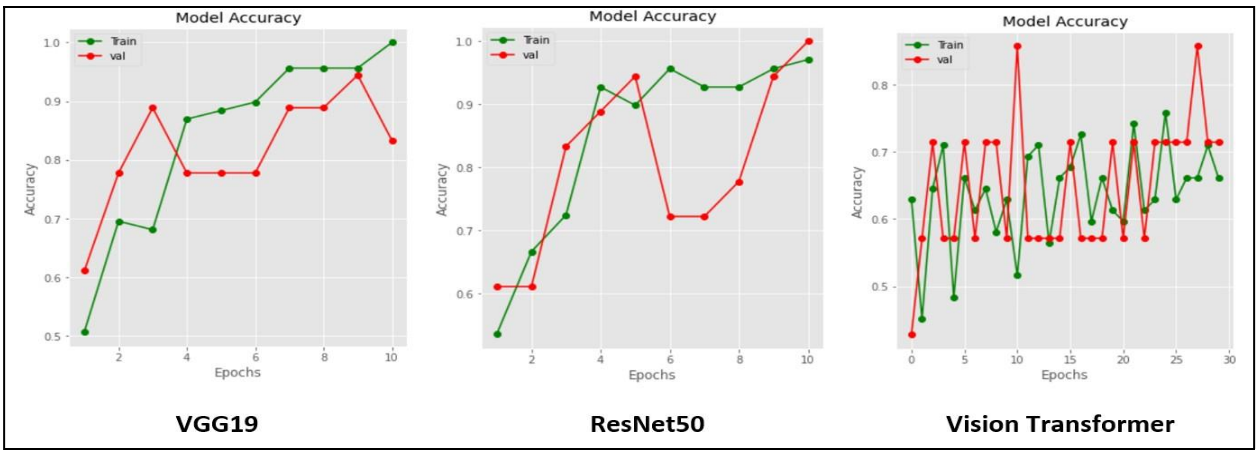

Fig.9: Model Accuracy vs Epoch Curve

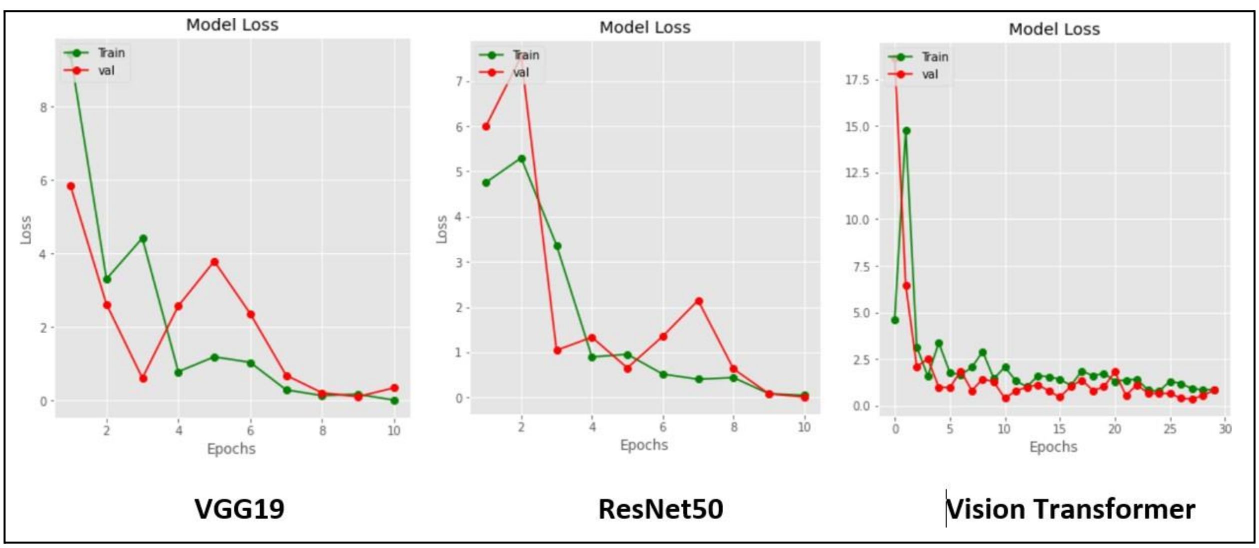

Fig.10: Model Loss vs epoch curve

Comparison table

Here VGG19 training accuracy is very good. ResNet50 model provided a validation accuracy of 100% and a validation loss of 0.004. The vision transformer here also performs very badly.

	Model	Training_Accuracy	Training_Loss	Validation_Accuracy	Validation_Loss
0	VGG19	1.000000	0.000014	1.00	0.000021
1	ResNet50	1.000000	0.000002	1.00	0.004034
2	Vision Transformer	0.842857	0.516162	0.75	1.442586

Fig.11: Result Comparison of all models after using LBP

When we compare the model with the dataset after applying LBP and without LBP, then in both cases we cannot see any significant changes. One thing that we can notice is that the minimum validation loss in ResNet50 with LBP is very less as compared to models trained with the LBP image dataset.

Conclusion

In this study, the VGG-19, ResNet50, and Vision Transformer models were employed to classify various ocular disorders. These models demonstrated the capability to accurately predict whether a fundus is healthy or affected by an eye disease. Our findings revealed that using the LBP-enhanced dataset resulted in high accuracy and minimal loss, indicating the models' effectiveness. To further address the dataset imbalance, generative adversarial networks (GANs) could be utilized to generate synthetic images of eye diseases, enhancing model training. Such advancements have the potential to revolutionize the field of eye disease diagnosis, providing significant support to medical professionals. While the current models show promise, there is ample opportunity for further improvement through continued research and exploration in the future.

References

[1] A. Burlacu, A. Iftene, E. Busoiu, D. Cogean, and A. Covic. Challenging the supremacy of evidence-based medicine through artificial intelligence: the time has come for a change of paradigms, Nephrology Dialysis Transplantation (ndt), gfz203, Oxford Academic, 7 November, 2019, pp. 1–4.

[2] A. Burlacu, A. Iftene, D. Jugrin, I.V. Popa, P.M. Lupu, C. Vlad, and A. Covic. Using Artificial Intelligence Resources in Dialysis and Kidney Transplant Patients: A Literature Review, BioMed Research International, vol. 2020, article ID 9867872, 2020, 14 pages.

[3] H.-C. Shin, H. R. Roth, M. Gao et al., “Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1285– 1298, 2016.

[4] L. Abhishek, “Optical character recognition using ensemble of svm, mlp and extra trees classifier,” in 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1–4.

[5] .Varghese, R. R. Varghese, K. Balakrishnan, and J. S. Paul, “Level identification of brain mr images using histogram of a lbp variant,” in 2012 IEEE International Conference on Computational Intelligence and Computing Research, 2012, pp. 1–4