**Ocular Disease Recognition**

*A project report submitted*

*to*

**MANIPAL ACADEMY OF HIGHER EDUCATION**

*For Partial Fulfillment of the Requirement for the Award of the Degree*

*of*

**Bachelor of Technology**

*in*

**Information Technology**

*by*

**Sneha Dharne**

**Reg. No. 190911088**

*Under the guidance of*

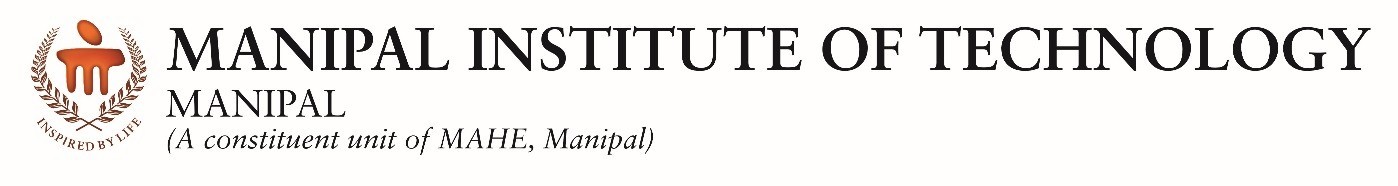
Mrs. Veena K. M.

Assistant Professor – Sr. Scale

Department of I & CT

Manipal Institute of Technology

Manipal, India



**July 2023**

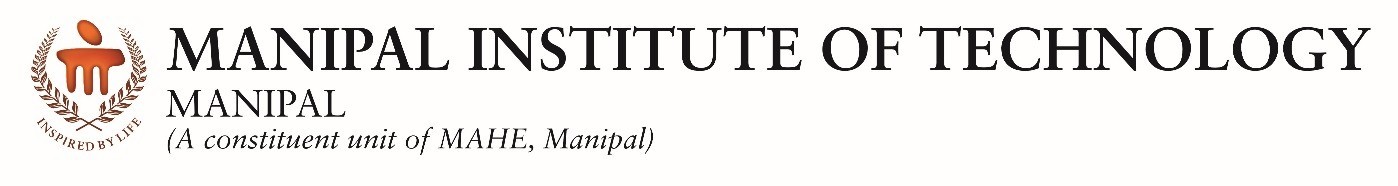
I dedicate my thesis to my friends and family.

**DECLARATION**

I hereby declare that this project work entitled **Ocular Disease Recognition** is original and has been carried out by me in the Department of Information and Communication Technology of Manipal Institute of Technology, Manipal, under the guidance of **Mrs. Veena K. M.** , **Assistant Professor, Sr. Scale**, Department of Information and Communication Technology, M. I. T., Manipal. No part of this work has been submitted for the award of a degree or diploma either to this University or to any other University.

Place: Manipal

Date : 13-07-2023 Sneha Dharne



**CERTIFICATE**

This is to certify that this project entitled **Ocular Disease Recognition** is a bonafide project work done by **Miss Sneha Dharne (Reg.No.:190911088)** at Manipal Institute of Technology, Manipal, independently under my guidance and supervision for the award of the Degree of Bachelor of Technology in Information Technology.

Mrs. Veena K. M. Dr. Smitha N Pai

Assistant Professor, Sr. Scale Professor & Head

Department of I & CT Department of I & CT

Manipal Institute of Technology Manipal Institute of Technology Manipal, India Manipal, India

**ACKNOWLEDGEMENTS**

I express my heartfelt gratitude to Mrs. Veena K. M, Assistant Professor (Senior Scale), Department of Information and Communication Technology, Manipal Institute of Technology, Manipal, for her invaluable guidance throughout this project. I would also like to extend my thanks to the Head Of the Department and Dr. Smitha N. Pai for providing me with the opportunity to undertake this project within the department. I would also like to sincerely thank the faculty of the Department of Information and Communication Technology for their unwavering support and guidance. Their trust in my abilities, as demonstrated through previous project opportunities, has been instrumental in developing my skills and making me capable enough to successfully undertake this project.

Finally, I would like to acknowledge the presence and unwavering support of my friends and colleagues. Their constant motivation and encouragement were vital in keeping me focused and determined, enabling me to complete this project within the designated timeframe.

**ABSTRACT**

This study focuses on the recognition of ocular diseases from fundus images using a combination of generic image preprocessing techniques and state-of-the-art convolutional neural network (CNN) models with transfer learning. The aim is to accurately identify and classify various ocular diseases, including Normal, Diabetes, Glaucoma, Cataract, AMD, Hypertension, Myopia, and Others.

To improve the quality of fundus images and facilitate the identification of disease-specific features, image enhancement techniques such as histogram equalization and Gaussian filtering are employed. These techniques enhance contrast, details, and local contrast, improving image visibility and quality.

In addition, a deep learning architecture called MIRNet is utilized to further enhance the images by capturing global and local features while preserving both details and textures. MIRNet utilizes invertible building blocks, residual connections, attention mechanisms, and a perceptual loss function to produce visually pleasing images with improved brightness, contrast, and details while reducing noise and artifacts associated with low-light conditions.

Transfer learning is employed to leverage pre-trained CNN models, specifically ResNet50 and VGG16, which have been trained on the ImageNet dataset. The pre-trained models' weights are utilized as an initialization point, and the final layers or some intermediate layers are modified or replaced to adapt to the ocular disease recognition task. This approach reduces the training time and data requirements while benefiting from the pre-trained models' learned representations and features.

To optimize the performance of the models, hyperparameter tuning is conducted using the Keras Tuner library. The library facilitates the search for optimal hyperparameters using techniques such as random search, hyperband, or Bayesian optimization. Area

under ROC curve (or AUC) is the performance metric evaluated to further assess the model.

The ultimate goal of this study is to develop an accurate and efficient system for ocular disease recognition from fundus images. The findings and insights gained from this research can contribute to advancements in medical imaging and assist healthcare professionals in the early detection and treatment of ocular diseases, potentially improving patient outcomes and quality of life.

**Contents**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Acknowledgments | | | iv | |
| Abstract | | | v | |
| List of Figures | | | viii | |
| List of Tables | | | ix | |
| 1 | Introduction |  | |
|  | 1.1 Importance of Early and Accurate Detection of Chronic Ocular Diseases…………………………………………………. | 1 | |
|  | 1.2 Leveraging Transfer Learning for Ocular Disease Detection | 1 | |
| 2 | Literature Review | 3 | |
| 3 | Problem Statement and Objectives |  | |
|  | 3.1 Literature Insights | 4 | |
|  | 3.2 Objectives | 4 | |
| 4 | Methodology |  | |
|  | 4.1 Preprocessing Techniques | 6 | |
|  | 4.2 Class Imbalance | 8 | |
|  | 4.3 Transfer Learning | 9 | |
| 5 | Results | 15 | |
| 6 | Conclusion and Future Scope | 17 | |
| References | | | 19 | |
| Annexure | | | 20 | |
| Project Details | | | 24 | |
| Originality Report | | | 25 | |
| CO-PO Mapping | | | 26 | |
|  | | |  | |

**List of Abbreviations**

|  |  |
| --- | --- |
| AMD | Age-Related Macular Degeneration |
| AUC | Area Under Curve |
| CLAHE | Contrast Limited Adaptive Histogram Equalization |
| COD | Chronic Ocular Diseases |
| CNN | Convolutional Neural Network |
| ILSVRC | ImageNet Large-Scale Visual Recognition Challenge |
| MSR | Multi-Scale Retinex |
| RESNET50 | Residual Network (50 Layers) |
| ROC | Receiver Operating Characteristic |
| SMOTE | Synthetic Minority Over-sampling Technique |
| VGG16 | Visual Geometry Group (16 Layers ) |

**List of Tables**

|  |  |  |
| --- | --- | --- |
| 5.1 | Comparison of AUC and loss function for state-of-the-art pre-trained models | 16 |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| 4.1 | Image Preprocessing Results | 8 |
| 4.2 | Image Preprocessing | 9 |
| 4.3 | Transfer Learning | 11 |
| 4.4 | RESNET50 Model Architecture | 12 |
| 4.5 | VGG16 Model Architecture | 12 |
| 4.6 | Tuner Search Loop | 13 |
| 4.7 | RESNET50 Trainable Parameters | 13 |
| 4.8 | VGG16 Trainable Parameters | 13 |

**Chapter 1**

**Introduction**

**1.1 Importance of Early and Accurate Detection of Chronic Ocular Diseases**

Chronic ocular diseases, such as age-related macular degeneration, glaucoma, and diabetic retinopathy, pose significant challenges in the field of ophthalmology. Early and accurate detection of these diseases is crucial for successful treatment and effective management, as it allows for timely interventions that can potentially preserve visual function and prevent irreversible damage. Fundus images, which provide high-resolution views of the back of the eye, have become a widely utilized diagnostic tool in the field. However, interpreting these images can be intricate and demanding, requiring extensive training and expertise from ophthalmologists and eye care professionals.

**1.2 Leveraging Transfer Learning for Ocular Disease Detection**

In recent years, transfer learning has emerged as a powerful technique in the field of deep learning for ocular disease detection. Transfer learning allows researchers to leverage the knowledge and representations learned by pre-trained models on large-scale datasets, such as ImageNet, and apply them to related tasks with smaller medical image datasets. By utilizing pre-trained convolutional neural network (CNN) models as feature extractors, researchers can benefit from the rich hierarchical representations learned during the pre-training phase, which can aid in the accurate detection and classification of ocular diseases.

In this study, we explore the application of transfer learning using two pre-trained models for the detection of chronic ocular diseases at an early stage. Our focus lies on leveraging the power of pre-trained models to extract meaningful features from fundus images and using these features for disease detection. By employing appropriate pre-processing techniques to enhance the quality and visibility of fundus images, we aim to improve the performance of the pre-trained models in detecting ocular diseases.

Through this study, we seek to assess the effectiveness of transfer learning with pre-trained models on the detection of ocular disease. By evaluating the performance of the models on a dataset of fundus images, we aim to gain insights into the potential of these approaches for improving the early diagnosis and treatment of chronic ocular diseases.

The early detection of ocular diseases holds tremendous potential for transforming clinical practice and patient outcomes. By identifying signs of ocular diseases at their earliest stages, healthcare professionals can initiate timely interventions, such as medical treatments or lifestyle modifications, to prevent or slow down disease progression.

Transfer learning, combined with appropriate pre-processing techniques, presents a promising approach for the automated detection of chronic ocular diseases at an early stage. By leveraging the knowledge learned from pre-trained models and improving the visibility of fundus images, we can potentially enhance the accuracy and efficiency of disease detection. This research contributes to the growing body of knowledge in the field of transfer learning and automated ocular disease detection, offering insights into potential advancements that can revolutionize the field of ophthalmology.

**Chapter 2**

**Literature Review**

# In recent years, the importance of preprocessing techniques for color fundus images in the accurate detection of chronic ocular diseases (COD) has gained significant attention among researchers. It has been widely recognized that the application of appropriate transformations to input color fundus images before feeding them into deep neural models plays a crucial role in enhancing the diagnostic performance of deep learning-based COD diagnosis. These preprocessing techniques effectively address various challenges encountered in fundus images, such as irregular illumination, low contrast, and the presence of unimportant features. By mitigating these challenges, these techniques lead to improved reliability in the diagnosis of COD.

# Mayya et al. [1] referred to the problem of accurate detection of chronic ocular diseases using fundus images as a multilabel binary classification task in their paper. They proposed the use of multilabel binary classification to assign multiple disease labels to each image, allowing for a more comprehensive understanding of the underlying conditions. This approach aimed to capture the diverse range of chronic ocular diseases that can be present in a single patient simultaneously.

# However, not all approaches have yielded optimal results. For instance, Islam et al. [2] proposed a shallow convolutional neural network (CNN) model for COD prediction. They employed a technique where two fundus images from the left and right eyes were merged and fed as inputs to the network. Although this approach seemed promising, it encountered challenges. The merging of multiple fundus images increased the complexity of the problem and led to overfitting issues. As a result, the model's performance suffered, as relevant information was lost due to the reduced image dimensions.

# To address this limitation, subsequent studies, such as those conducted by Li et al. [3] and He et al. [4], focused on fusing features from separate CNN networks

# trained on left and right eye fundus images.

# By leveraging the distinctive information captured by each eye, these studies aimed to improve the accuracy of COD classification. Additionally, He et al. [5] further enhanced the fusion approach by training a teacher network that incorporated features from both eye images and diagnostic keywords, providing a more comprehensive representation for accurate COD detection.

# To summarize, the existing literature highlights the significance of preprocessing techniques for color fundus images in the accurate detection of chronic ocular diseases. Mayya et al. emphasized the use of multilabel binary classification to capture the diversity of chronic ocular diseases present in a single patient. While some approaches, such as merging multiple fundus images, encountered challenges and resulted in decreased performance, subsequent studies focused on fusing features from separate eye images to improve classification accuracy. These advancements aim to enhance the accuracy and reliability of COD diagnosis using fundus images.

# Chapter 3

**Problem Statement and Objectives**

**3.1 Literature Insights**

While substantial progress has been made in ocular disease recognition from color fundus images using deep learning, there is a specific gap in exploring the effectiveness of generic preprocessing techniques for broad classification of fundus images into normal and abnormal categories. Existing research primarily focuses on disease-specific preprocessing techniques, but there is a need to assess the performance of generic preprocessing techniques in providing a quick and general classification of fundus images.

Previous approaches, such as merging multiple fundus images or employing complex architectures, have encountered overfitting issues or high computational complexity. There is a need to address these challenges effectively to develop models that generalize well to unseen data while being computationally efficient.

**3.2 Objectives**

1. To evaluate and compare the effectiveness of generic preprocessing techniques, such as contrast enhancement, noise reduction, and image normalization, in broadly classifying fundus images into normal and abnormal categories.
2. To develop and implement a classification model that utilizes the output of the generic preprocessing techniques as input and provides a binary classification of fundus images as either normal or abnormal.
3. To mitigate overfitting issues and reduce computational complexity by utilizing transfer learning, leveraging pre-trained models such as ResNet50 or VGG16, as initialization points for the ocular disease recognition task.
4. To optimize the hyperparameters of the transfer learning models using appropriate techniques such as Keras Tuner, in order to find the best configuration that maximizes the models' performance while minimizing computational requirements.
5. To analyze and interpret the results to gain insights into the performance and limitations of the generic preprocessing techniques for rapid and preliminary classification of fundus images.

By accomplishing these objectives, this study aims to advance the understanding of effective preprocessing techniques, mitigate overfitting issues, reduce computational complexity, and develop accurate and efficient models for ocular disease recognition. The study also aims to explore the effectiveness of generic preprocessing techniques in broadly classifying fundus images into normal and abnormal categories. The findings will provide valuable insights into the potential of these techniques as a quick and preliminary classification tool, allowing healthcare professionals to prioritize and focus on abnormal cases, thus optimizing the efficiency of ocular disease diagnosis and treatment.

**Chapter 4**

**Methodology**

This study focuses on comparing results shown by the generic image preprocessing techniques with two state of the art CNN models using Transfer Learning. Image enhancement techniques, such as histogram equalization and Gaussian filtering, can help improve the quality of fundus images and make it easier to identify features of interest. Images could be enhanced further using a deep learning architecture (MIRNet) dedicated to enhance details and textures of the image while preserving both global and local features. We will study the effectiveness of these techniques in improving the performance of the model.

* 1. **Preprocessing Techniques**

1. CLAHE

CLAHE (Contrast Limited Adaptive Histogram Equalization) is an image enhancement technique that improves contrast and details by dividing the image into tiles, applying adaptive histogram equalization independently to each tile, and incorporating contrast limiting to prevent noise amplification. It is a valuable tool in various fields like medical imaging and photography for enhancing image visibility and contrast.

1. Gaussian Blur and MultiScale Retinex

MSR with Gaussian blur is an algorithm that enhances image appearance by decomposing it into multiple scales, applying Retinex processing independently to each scale to enhance local contrast, and incorporating a Gaussian blur step to reduce noise and high-frequency details. It combines the processed scales using a weighting scheme to produce an enhanced image with improved visual quality and preservation of essential details.

1. MIRNET

MIRNet is a deep learning architecture designed to enhance low-light images by

|  |  |
| --- | --- |
| Figure 4.11: Original Image, cropped and resized | Figure 4.12 RGB CLAHE |
| Figure 4.13 MSR Enhancement | Figure 4.14 RGB CLAHE + MSR |
| Figure 4.15 MIRNET Model | Figure 4.16 RGB CLAHE + MIRNET |

Figure 4.1 Image preprocessing results

capturing global and local features, utilising invertible building blocks and residual

connections, incorporating attention mechanisms, and optimising results using a perceptual loss function. It produces visually pleasing images with improved brightness, contrast, and details, while reducing noise and artefacts commonly associated with low-light conditions.

These preprocessing techniques, CLAHE, Gaussian Blur with MSR, and MIRNet, were sequentially applied to the fundus images, as depicted in Figure 4.2. This preprocessing pipeline aimed to enhance the overall quality of the images, making it easier to identify important features and leading to improved performance in subsequent classification tasks. The results of individual preprocessing techniques are shown in figure 4.1

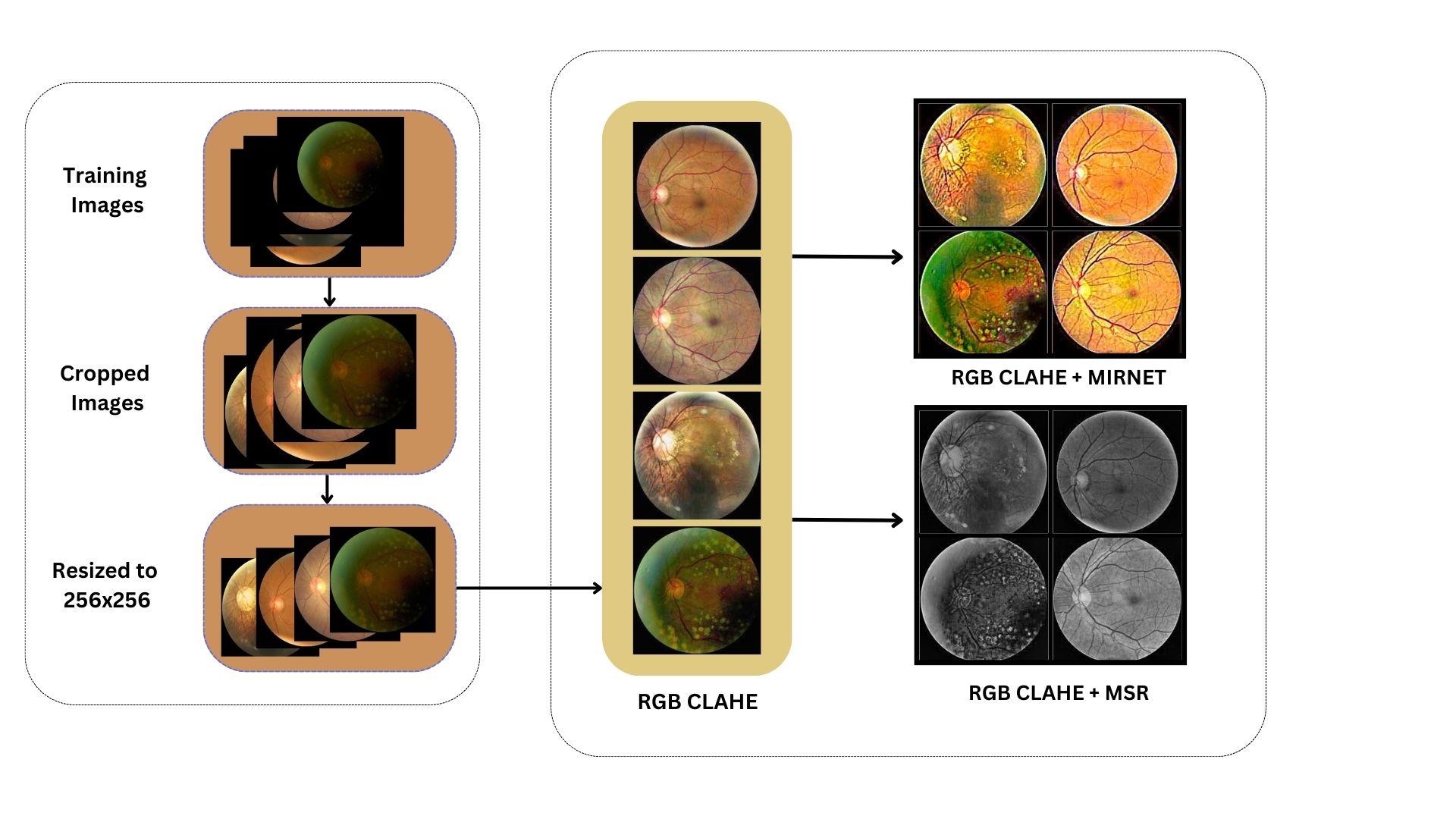


Figure 4.1: Image Preprocessing

* 1. **Class Imbalance**

**4.2.1 SMOTE (Synthetic Minority Over-sampling Technique):**

SMOTE is an oversampling technique used to address class imbalance in machine learning datasets. It focuses on boosting the representation of the minority class by generating synthetic samples. SMOTE works by interpolating feature vectors of existing minority class samples to create new synthetic instances. This process involves selecting a sample from the minority class and identifying its k nearest

neighbors Synthetic samples are then generated along the line connecting the selected sample and its neighbors. By introducing these synthetic instances, SMOTE rebalances the class distribution and provides the learning model with more diverse examples from the minority class. This helps improve the model's ability to learn and generalize from the minority class, reducing the risk of biased predictions.

**4.2.2 TomekLinks:**

TomekLinks is an undersampling technique that complements oversampling methods such as SMOTE. Its primary goal is to remove noisy and borderline samples that may introduce ambiguity and misclassification. TomekLinks identifies pairs of instances known as Tomek pairs, where each instance is the nearest neighbor of the other and belongs to a different class. These pairs typically exist near the decision boundary between classes. By eliminating instances belonging to the majority class from Tomek pairs, TomekLinks enhances the separation between classes and reduces potential overlap. This leads to a clearer decision boundary, improving the model's ability to discriminate between different classes and reducing the risk of misclassification.

SMOTE and TomekLinks are applied sequentially to balance the representation of eye diseases in the dataset. SMOTE is first utilized to generate synthetic samples for the minority class, ensuring a more equitable distribution of eye disease instances. This helps prevent bias and allows the learning model to learn effectively from both the majority and minority classes. Then, TomekLinks is employed to further refine the dataset by removing instances from Tomek pairs, promoting better class separation and reducing the potential for misclassification. By integrating both SMOTE and TomekLinks, the code aims to create a more balanced and representative dataset, enhancing the model's performance and reliability in the task of eye disease classification.

**4.3 Transfer Learning**

Transfer learning is a machine learning technique where a pre-trained model, trained on a large dataset for a specific task, is used as a starting point for solving a different but related task. Instead of training a model from scratch on the new task, transfer learning leverages the knowledge and representations learned from the pre-trained model.

1. Pre-training: A deep neural network model, such as a Convolutional Neural Network (CNN), is trained on a large dataset for a specific task, often on a large-scale dataset like ImageNet. This training process requires significant computational resources and time. The pre-trained model learns to extract meaningful features and representations from the data, capturing general patterns and knowledge.
2. Fine-tuning: The pre-trained model's weights and parameters are used as an initialization point for the new task. The final layers or some intermediate layers of the pre-trained model are modified or replaced to adapt to the specific requirements of the new task. These modified layers are typically randomly initialized, while the remaining layers retain the pre-trained weights.
3. Training on the new task: The new model, consisting of the pre-trained layers and the modified layers, is trained on a smaller dataset specific to the new task. Since the initial layers have already learned meaningful representations from the pre-training, the training process on the new task requires fewer training samples and less time compared to training a model from scratch. The modified layers are adjusted to specialize in the new task, while the pre-trained layers act as feature extractors, providing a head start.

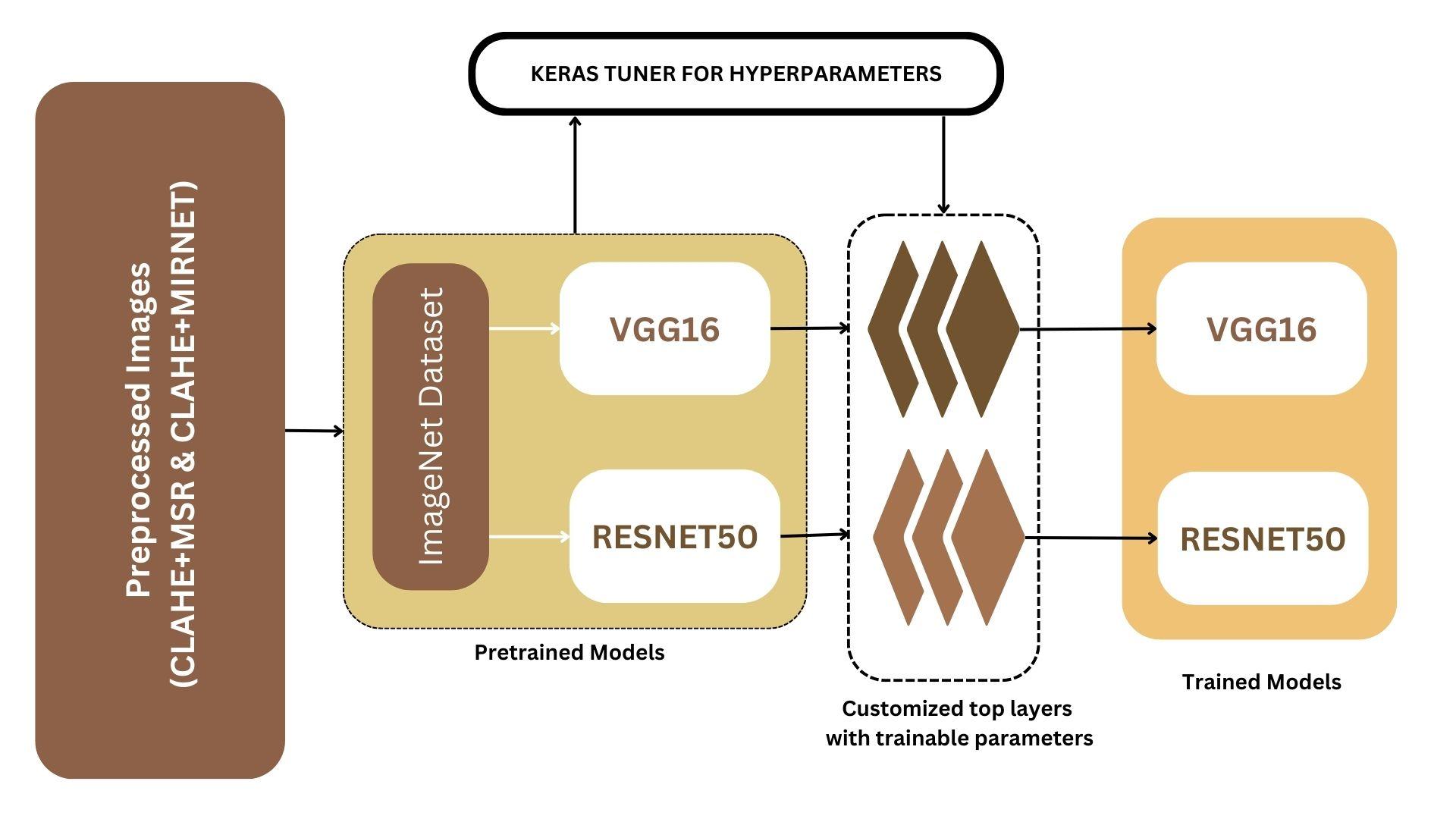


Figure 4.2: Transfer Learning

The transfer learning process, as described above, is illustrated in Figure 4.3. It visually represents the flow of knowledge transfer from pre-training to fine-tuning and training on the new task. By leveraging transfer learning, the study aims to enhance the performance of the pre-trained models, ResNet50 and VGG16, for the ocular disease recognition task on the ODIR dataset

**4.3.1 RESNET50**

ResNet (short for Residual Network) is a deep convolutional neural network architecture proposed by Microsoft Research in 2015. It was designed to address the problem of vanishing gradients in very deep networks by introducing skip connections or residual connections. It achieved remarkable performance in the ILSVRC 2015 competition, surpassing human-level performance on the ImageNet dataset.

ResNet-50 has 50 layers, combining convolutional, pooling, and fully connected layers.

The key innovation of ResNet is the introduction of residual blocks, where shortcut connections bypass one or more convolutional layers. These shortcut connections allow the model to learn residual mappings, enabling the network to focus on learning the residual (difference) between the input and output of each block. The RESNET50 model too, pretrained on the ImageNet dataset is used as the initialization point for the classification task.

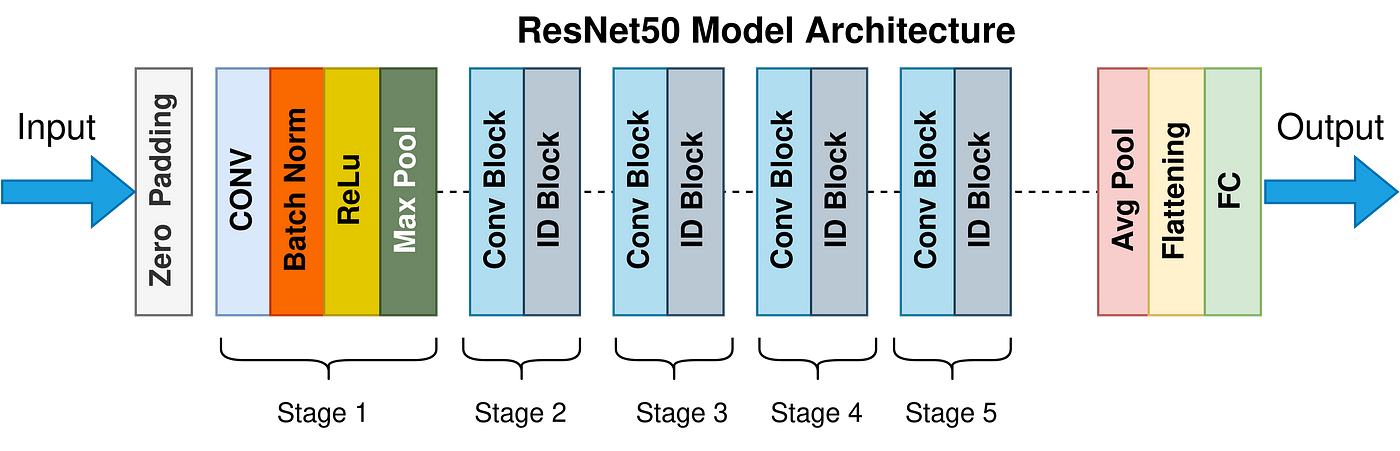


Figure 4.3: RESNET50 model Architecture [6]

In the context of this study, the pre-trained ResNet-50 model, initialized with weights learned from the ImageNet dataset, is utilized as the starting point for the ocular disease

recognition task. The architecture of ResNet-50, as depicted in Figure 4.3, guides the configuration and organization of convolutional, pooling, and fully connected layers, facilitating feature extraction and classification for the classification task at hand.

**4.3.2 VGG16**

Visual Geometry Group (VGG) at the University of Oxford proposed VGG16, a convolutional neural network architecture. It has gained popularity for its deep architecture and remarkable performance in various computer vision tasks. The name "VGG16" refers to the number of layers in the network, specifically 16 layers.

The core of the VGG16 architecture is a stack of 13 convolutional layers, followed by 3 fully connected layers. The convolutional layers are responsible for learning and extracting features from input images. VGG16 adopts a uniform configuration of using 3x3 filters throughout these convolutional layers. This choice of filter size enables the network to capture finer details and enhances its ability to learn complex hierarchical representations of visual features.

Using multiple convolutional layers stacked on each other allows VGG16 to learn increasingly abstract features as the network goes deeper. The network gradually progresses from low-level features such as edges and textures to higher-level features representing more complex patterns and objects. This depth and layer-wise organization contribute to the network's ability to effectively model and understand visual data.



Figure 4 4: VGG16 Model Architecture [7]

Following the stack of convolutional layers, VGG16 includes three fully connected layers responsible for the final classification or prediction task. These fully connected

layers combine the high-level features learned by the preceding convolutional layers to make predictions.

During training, VGG16 learns the weights of its layers using a large labeled dataset, such as the ImageNet dataset. During the training process of VGG16, the model utilizes the backpropagation algorithm. This algorithm computes the gradients of the loss function concerning the model's weights. These gradients are then used to update the weights, iteratively fine-tuning the network's parameters and optimizing its overall performance. By iteratively adjusting the weights based on the computed gradients, VGG16 can learn to make more accurate predictions and improve its ability to classify and understand visual data.

Figure 4.5 provides a visual representation of the VGG16 architecture, illustrating the organization and arrangement of the convolutional and fully connected layers.

**4.3.3 Keras Tuning**

Keras Tuner is a hyperparameter tuning library that automates the process of selecting optimal hyperparameters for deep learning models. It provides a flexible and efficient framework for hyperparameter search using techniques like random search, hyperband, or Bayesian optimization.

For evaluating the performance of the models during the search, Keras Tuner allows you to specify metrics such as accuracy, precision, recall, F1 score, or AUC. By selecting AUC as the metric, you can focus on optimising the area under the ROC curve, which is particularly useful for binary classification tasks.



Figure 4.6 Tuner Search Loop [8]

The usage of Keras Tuner with specific models like ResNet50 and VGG16 remains the same. One defines the model architecture, sets up the hyperparameter search space, specifies the objective metric (in this case, AUC), and lets Keras Tuner search for the best hyperparameters. (See Figure 4.6 for the workflow of the Tuner Search Loop) Keras Tuner is model-agnostic and can be used with any model architecture supported by the underlying deep learning framework, such as TensorFlow or Keras. Figures 4.7 and 4.8 are the results in the form of trainable parameters and total parameters (hyperparameters of number of neurons per layer and number of layers for optimum complexity and generalizability ) of the pretrained model attained using keras tuning.

**4.3.4 Evaluation Metrics**

The ROC (Receiver Operating Characteristic) Curve represents the trade-off between the true positive rate and the false positive rate. AUC ( Area Under ROC Curve) represents the probability that a randomly chosen positive instance will be ranked higher than a randomly chosen negative instance by the model. An AUC value of 1 indicates a perfect classifier, while an AUC of 0.5 suggests a random or ineffective classifier.

Loss, on the other hand, measures the error or discrepancy between the predicted outputs of the model and the actual ground truth labels. The loss function, typically defined during the model training process, is optimized to minimize this error. Lower values of model loss indicate a better fit between the predicted and actual values, signifying improved performance.

Analyzing both model AUC and model loss provides comprehensive insights into the performance of a machine-learning model. A high AUC indicates the model's ability to accurately classify instances, while a low model loss signifies better alignment between the predicted and true labels.

Since, our problem statement is defined as a multi-label binary classification, analyzing AUC and Loss functions would provide better analytical insights.

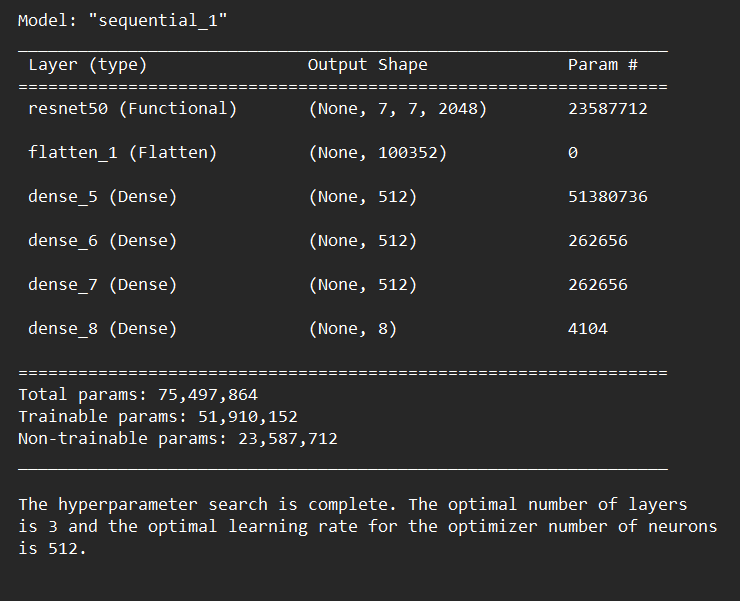


Figure 4.7 RESNET50 Trainable Parameters

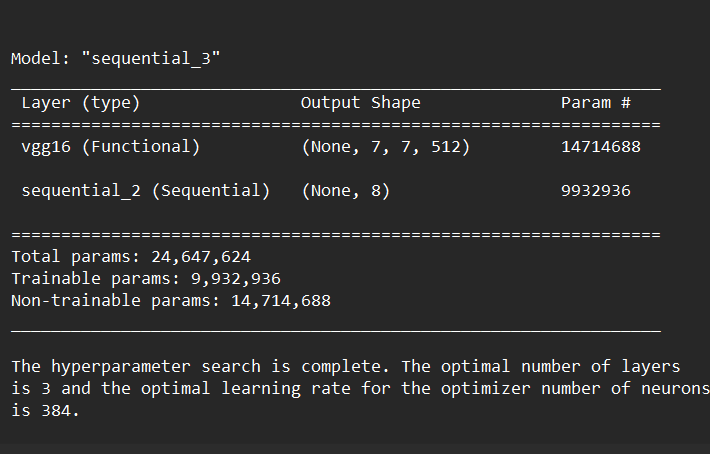


Figure 4.8 VGG16 Trainable Parameters

**Chapter 5**

**Results**

Table 5.1 depicts the comparison of AUC and Loss functions of the RESNET50 and the VGG16 models. We can evaluate the following factors of the models:

**Model Performance:**

The performance of the trained machine learning model was evaluated using the Area Under the Curve (AUC) metric. The AUC values obtained during training and validation converged at 0.8 (for the RESNET50 model) and 0.9 (for the VGG16 model), indicating a good discrimination ability of the model. This suggests that the model can effectively distinguish between positive and negative instances. Both the ResNet50 and VGG16 models have achieved significant success in capturing relevant features for ocular disease recognition. Despite having more trainable parameters, the VGG16 model exhibits a slightly higher AUC value, suggesting that it has a stronger discriminatory power in classifying ocular diseases.

**Generalization:**

The convergence of the loss and validation loss at 0.2 indicates that the model is generalizing well to unseen data. The low values of loss and val\_loss suggest that the model is not overfitting to the training data and can make accurate predictions on new data.

**Model Stability:**

The convergence of both AUC and val\_AUC, as well as loss and val\_loss, indicates that the model is stable during the training process. The consistent improvement and lack of significant fluctuations suggest that the model is finding a good set of weights and is not trapped in a suboptimal solution.

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Preprocessing On Images | AUC | LOSS |
| RESNET50 | CLAHE+MIRNET |  |  |
| CLAHE+MSR |  |  |
| VGG16 | CLAHE+MIRNET |  |  |
| CLAHE+MSR |  |  |

Table 5.1: Comparison of AUC and loss function for state-of-the-art pre-trained models

**Chapter 6**

**Conclusion and Future Scope**

Based on the evaluation of the trained machine learning models, the following conclusions can be drawn:

1. The trained models, using RESNET50 and VGG16 architectures, demonstrated good discrimination ability, as evidenced by the Area Under the Curve (AUC) values of 0.8 and 0.9, respectively. This indicates that the models can effectively distinguish between fundus images belonging to different categories, such as normal and abnormal.
2. The convergence of the loss and validation loss at 0.2 suggests that the models generalize well to unseen data. The low values of loss and val\_loss indicate that the models have not overfit to the training data and can make accurate predictions on new, unseen data.
3. The stability of the models is demonstrated by the convergence of both AUC and val\_AUC, as well as loss and val\_loss. The consistent improvement and absence of significant fluctuations during the training process indicate that the models are finding a robust set of weights and are not trapped in suboptimal solutions.

While the ResNet50 model performs slightly lower in terms of AUC, it still demonstrates considerable discriminative ability. This highlights the effectiveness of the residual connections and skip connections employed in the ResNet architecture, which help mitigate the vanishing gradient problem and improve feature learning.

Overall, the trained machine learning models show promising performance, generalization ability, and stability in classifying fundus images into their respective categories.

**Future Scope:**

* + - 1. Validation on External Datasets: To further validate the performance and generalizability of the trained models, it is essential to evaluate them on external datasets. This will provide a more comprehensive assessment of the models'

ability to classify fundus images accurately.

* + - 1. Refinement of Preprocessing Techniques: While the generic preprocessing techniques used in this study showed promising results, further refinement and exploration of different preprocessing approaches can be undertaken. This may include investigating novel techniques for contrast enhancement, noise reduction, and image normalization, leading to improved model performance and robustness.
      2. Integration of Clinical Data: Incorporating additional clinical data, such as patient demographics, medical history, and clinical examination findings, into the models can enhance their diagnostic accuracy and provide more comprehensive disease recognition. The fusion of clinical data with image-based features may contribute to a more holistic approach in ocular disease diagnosis.
      3. Multi-Class Classification: Extending the binary classification approach to multi-class classification, where fundus images are categorized into specific ocular diseases, can provide valuable insights into disease identification and aid in personalized treatment strategies. Exploring techniques such as one-vs-rest or softmax activation for multi-class classification can be considered.
      4. Explainability and Interpretability: Investigating methods for interpreting the decisions made by the trained models can enhance the trust and acceptance of the models in clinical practice. Techniques such as feature importance analysis, attention mechanisms, or saliency mapping can be explored to provide insights into the model's decision-making process.

In summary, this study lays the foundation for accurate ocular disease recognition from fundus images using machine learning models. The future scope encompasses validation on external datasets, refinement of preprocessing techniques, integration of clinical data, multi-class classification, and exploring explainability methods. These advancements will contribute to the development of robust and clinically applicable models for ocular disease diagnosis and treatment.

**References**

1. Mayya, V., Kamath S, S., Kulkarni, U., Kaiyoor Surya, D., & Acharya, U. R. (2021). An empirical study of preprocessing techniques with convolutional neural networks for accurate detection of chronic ocular diseases using fundus images. Computer Methods and Programs in Biomedicine, 206, 106267.<https://doi.org/10.1016/j.cmpb.2021.106267>.
2. Islam MT, Imran SA, Arefeen A, Hasan M, Shahnaz C (2019) Source and camera independent ophthalmic disease recognition from fundus image using neural network. In: 2019 IEEE International conference on signal processing, information, communication systems (SPICSCON).
3. Li N, Li T, Hu C, Wang K, Kang H (2021) A benchmark of ocular disease intelligent recognition: One shot for multi-disease detection. In: Wolf F, Gao W (eds) Benchmarking, measuring, and optimizing. Springer International Publishing, Cham, pp 177–193.
4. He J, Li C, Ye J, Qiao Y, Gu L (2021) Multi-label ocular disease classification with a dense correlation deep neural network. Biomedical Signal Processing and Control, pp 63.
5. He J, Li C, Ye J, Qiao Y, Gu L (2021) Self-speculation of clinical features based on knowledge distillation for accurate ocular disease classification. Biomed Sig Process Control 67:102491.
6. <https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758>
7. <https://medium.com/@mygreatlearning/everything-you-need-to-know-about-vgg16-7315defb5918>
8. https://www.researchgate.net/figure/Hyperparameter-tuning-process-with-keras-tuner\_fig3\_358360497

**Annexure**

**Annexure A: Hyperparameter tuning code**

import tensorflow as tf

from tensorflow import keras

from tensorflow.keras import layers

from tensorflow.keras.applications import VGG16

import keras\_tuner

from keras\_tuner import Hyperband

from tensorflow.keras.metrics import AUC

# Build the model

def build\_model(hp):

# Load pre-trained VGG16 model

vgg\_base = VGG16(weights='imagenet', include\_top=False, input\_shape=(224, 224, 3))

# Freeze the VGG16 base layers

vgg\_base.trainable = False

# Create the custom top layer

top\_model = keras.Sequential()

top\_model.add(layers.Flatten(input\_shape=vgg\_base.output\_shape[1:]))

# Define the hyperparameters to tune

num\_layers = hp.Int('num\_layers', min\_value=2, max\_value=4, step=1)

num\_neurons = hp.Int('num\_neurons', min\_value=128, max\_value=512, step=128)

for \_ in range(num\_layers):

top\_model.add(layers.Dense(num\_neurons, activation='relu'))

top\_model.add(layers.Dense(8, activation='sigmoid'))

# Combine the base and top layers

model = keras.Sequential([vgg\_base, top\_model])

# Compile the model with the Adam optimizer and custom metrics

model.compile(optimizer='adam', loss='binary\_crossentropy', metrics=[AUC(name='auc')])

return model

# Create the tuner

tuner = keras\_tuner.tuners.Hyperband(

build\_model,

objective=keras\_tuner.Objective("auc", direction="max"),

max\_epochs=10,

)

data\_dir\_preprocessed = '/content/drive/MyDrive/ODIR-5K/clahe+mirnet' # Replace with the path to your preprocessed images

test\_image\_dir = '/content/drive/MyDrive/ODIR-5K/test-clahe-mirnet' # Replace with the path to your test images

df['filename'] = ""

df['filename'] = df.apply(lambda row: os.path.join(data\_dir\_preprocessed, str('clahe\_mirnet\_'+row['ID'])), axis=1)

batch\_size = 64

epochs = 10

# Create an ImageDataGenerator for data augmentation

train\_datagen = ImageDataGenerator(rescale=1./255,

rotation\_range=20,

width\_shift\_range=0.2,

height\_shift\_range=0.2,

horizontal\_flip=True)

# Load the training and validation data

# Replace with the path to your training data file

train\_set = df.sample(frac=0.8, random\_state=42)

val\_set = df.drop(train\_set.index)

# Create the flow from dataframe for training and validation sets

train\_generator = train\_datagen.flow\_from\_dataframe(dataframe=train\_set,

directory=data\_dir\_preprocessed,

x\_col="filename",

y\_col=["Normal", "Diabetes", "Glaucoma", "Cataract",

"AMD", "Hypertension", "Myopia", "Others"],

target\_size=(224, 224),

batch\_size=batch\_size,

class\_mode="raw",

shuffle=True)

val\_generator = train\_datagen.flow\_from\_dataframe(dataframe=val\_set,

directory=data\_dir\_preprocessed,

x\_col="filename",

y\_col=["Normal", "Diabetes", "Glaucoma", "Cataract",

"AMD", "Hypertension", "Myopia", "Others"],

target\_size=(224, 224),

batch\_size=batch\_size,

class\_mode="raw",

shuffle=False)

# Just pass the generator directly to tuner.search

tuner.search(train\_generator, epochs=10, validation\_data=val\_generator)

best\_hps = tuner.get\_best\_hyperparameters(num\_trials=1)[0]

model = tuner.hypermodel.build(best\_hps)

model.summary()

**Project Details**

*Student Details*

|  |  |  |  |
| --- | --- | --- | --- |
| **Student Name** | Sneha Dharne |  |  |
| Registration  Number | 190911088 | Section/Roll No. | B/31 |
| Email Address | snehadattadharne@gmail.com | Phone N0. (M) | 8451037225 |

*Project Details*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Project Title** |  | Ocular Disease Recognition | |  |
| Project Duration | 6 Months |  | Date of Reporting | 07/02/2023 |

*Internal Guide Details*

|  |  |
| --- | --- |
| **Faculty Name** | Mrs Veena K. M. |
| Full Contact  Address with PIN  Code | Department of Information and Communication Technology,  Manipal Institute of Technology,  Manipal-576104 |
| Email Address | veena.gv@manipal.edu |



**CO and PO Mapping**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CLOs** | | **PO1** | **PO2** | **PO3** | **PO4** | **PO5** | **PO6** | **PO7** | **PO8** | **PO9** | **PO10** | **PO11** | **PO12** |
| ICT 4299.1 | Assess the work available in the literature related to the project to identify the limitations and risks. | 3 | 3 | - | 3 | 3 | 2 | 2 | - | - | - | - | 3 |
| ICT 4299.2 | Practice planning and time management in solving the problem. | 3 | 3 | - | - | - | - | - | 3 | 3 | 2 | 3 | 2 |
| ICT 4299.3 | Demonstrate professional skills to work effectively in a team or individually. | - | - | - | - | - | - | - | 3 | 3 | 3 | 2 | - |
| ICT 4299.4 | Develop the ability to adopt a methodological approach to solve societal problems.. | 3 | 2 | 3 | 2 | 2 | 3 | 2 | 3 | - | - | - | 3 |
| ICT 4299.5 | Conduct experimentation and testing to achieve the  defined objectives through  computing/coding/statistical  analysis | 3 | 3 | 3 | 3 | 3 | 3 | - | - | - | - | - | 3 |
| ICT 4299.6 | Compose the technical report with effective communication on  incorporating ethical practices. | - | 3 | - | - | - | - | - | 3 | - | 3 | - | - |
| **ICT 4299 (Avg. correlation level)** | | **3** | **2.8** | **3** | **2.67** | **2.34** | **2.67** | **2** | **3** | **3** | **2.67** | **2.5** | **3** |

**PROGRAM OUTCOMES (PO)**

Engineering Graduates will be able to:

1. **Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

1. **Problem analysis:** Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

1. **Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

1. **Conduct investigations of complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

1. **Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

1. **The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

1. **Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

1. **Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

1. **Individual and team work:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

1. **Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

1. **Project management and finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one’s own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

1. **Life-long learning:** Recognize the need for and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CLOs** | | **PSO1** | **PSO2** | **PSO3** | **PSO4** | **PSO5** | **PSO6** | **PSO7** | **PSO8** | **PSO9** |
| ICT 4299.1 | Assess the work available in the literature related to the project to identify the limitations and risks. | 3 | - | - | 2 | - | 3 | - | - | - |
| ICT 4299.2 | Practice planning and time management in solving the problem. | - | 3 | 2 | - | - | - | - | - | 2 |
| ICT 4299.3 | Demonstrate professional skills to work effectively in a team or individually. | - | - | 3 | - | - | - | 3 | - | 3 |
| ICT 4299.4 | Develop the ability to adopt a methodological approach to solve societal problems.. | 2 | - | 3 | 3 | 3 | 3 | 3 | 3 | - |
| ICT 4299.5 | Conduct experimentation and testing to achieve the  defined objectives through  computing/coding/statistical  analysis | - | - | 3 | 2 | 3 | 3 | 2 | 2 | - |
| ICT 4299.6 | Compose the technical report with effective  communication on  incorporating ethical practices. | - | - | - | - | - | 3 | - | - | - |
| **ICT 4299 (Avg. correlation level)** | | **2.5** | **3** | **2.75** | **2.34** | **3** | **3** | **2.67** | **2.5** | **2.5** |

* 1. To identify, analyse and develop software systems using appropriate techniques and concepts related to information technology
  2. To design an algorithm or process within realistic constraints to meet the desired needs through analytical, logical and problem-solving skills.
  3. To apply state of the art IT tools and technologies, IT infrastructure management abilities in treading innovative career path as a prospective IT engineer
  4. Apply the principles of science, maths and computer programming to solve complex problems related to information technology.
  5. Apply knowledge of programming, computational intelligence, computer graphics and visualization, data analytics, software system design, cyber security to arrive at solutions to real world problems.
  6. Apply IT knowledge to design and develop systems with respect to societal, user, customer needs, health and safety, diversity, inclusion, societal, environmental codes of practise and industry standard.
  7. Integrate and interface industry relevant hardware and software components and technology to come up with innovative and creative solutions.
  8. Use of industry standard software tools and platform to design and analyze IT systems.
  9. Learn to function collaboratively as a member of leader in diverse teams in multidisciplinary settings to manage the process effectively and document, present and communicate with the engineering community.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| COURSE  Code | Course Title | PO 1 | PO2 | PO3 | PO4 | PO5 | PO6 | PO7 | PO8 | PO9 | PO10 |
| ICT 4299 | Project Work | 2.5 | 2.6 | 3 | 2.67 | 2.34 | 2.67 | 2 | 3 | 3 | 2.67 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| COURSE  Code | Course Title | PO11 | PO12 | PSO1 | PS02 | PSO3 | PS O4 | PSO5 | PSO6 | PSO7 | PS O8 | PSO9 |
| ICT 4299 | Project Work | 2.5 | 3 | 2.5 | 3 | 2.75 | 2.34 | 3 | 3 | 2.67 | 2.5 | 2.5 |

**IET (AHEP Mapping):**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **CLOs** | | **C1** | **C2** | **C3** | **C4** | **C5** | **C6** | **C13** | **C16** | **C17** |
| ICT 4299.1 | Assess the work available in the  literature related to the project to identify the limitations and risks. | - | 3 | 2 | 3 | - | - | 2 | - | - |
| ICT 4299.2 | Practice planning and time management in solving the problem. | - | - | 3 | - | - | 2 | - | 3 | 3 |
| ICT 4299.3 | Demonstrate professional skills to work effectively in a team or individually. | - | - | - | - | - | - | - | 3 | 3 |
| ICT 4299.4 | Develop the ability to adopt a methodological approach to solve societal problems.. | - | - | 3 | - | 3 | 3 | - | 2 | - |
| ICT 4299.5 | Conduct experimentation and testing to achieve the defined objectives through  computing/coding/statistical  analysis | 3 | 2 | 2 | - | - | - | 3 | 3 | 2 |
| ICT 4299.6 | Compose the technical report with effective communication on incorporating ethical practices. | - | - | - | - | - | - | - | 3 | 3 |
| **ICT 4299 (Avg. correlation level)** | | **3** | **2.5** | **2.5** | **3** | **3** | **2.5** | **2.5** | **2.8** | **2.75** |

|  |  |  |
| --- | --- | --- |
| **Category** | **AHEP**  **LO number** | **AHEP LO Statements** |
| Science & Maths | C1 | Apply knowledge of mathematics, statistics, natural science and engineering principles to the solution of complex problems. Some of the knowledge will be at the forefront of the particular subject of study |
| Engineering Analysis | C2 | Analyse complex problems to reach substantiated conclusions using first principles of mathematics, statistics, natural science and engineering principles |
| C3 | Select and apply appropriate computational and analytical techniques to model complex problems, recognising the limitations of the techniques employed |
| C4 | Select and evaluate technical literature and other sources of information to address complex problems |
| Design & Innovation | C5 | Design solutions for complex problems that meet a combination of societal, user, business and customer needs as appropriate. This will involve consideration of applicable health & safety, diversity, inclusion, cultural, societal, environmental and commercial matters, codes of practice and industry standards |
| C6 | Apply an integrated or systems approach to the solution of complex problems |
| The  Engineer & Society | C7 | Evaluate the environmental and societal impact of solutions to complex problems and minimise adverse impacts |
| C8 | Identify and analyse ethical concerns and make reasoned ethical choices informed by professional codes of conduct |
| C9 | Use a risk management process to identify, evaluate and mitigate risks (the effects of uncertainty) associated with a particular project or activity |
| C10 | Adopt a holistic and proportionate approach to the mitigation of security risks |
| C11 | Adopt an inclusive approach to engineering practice and recognise the responsibilities, benefits and importance of supporting equality, diversity and inclusion |
| Engineering Practice | C12 | Use practical laboratory and workshop skills to investigate complex problems |
| C13 | Select and apply appropriate materials, equipment, engineering technologies and processes, recognising their limitations |
| C14 | Discuss the role of quality management systems and continuous improvement in the context of complex problems |
| C15 | Apply knowledge of engineering management principles, commercial context, project and change management, and relevant legal matters including intellectual property rights |
| C16 | Function effectively as an individual, and as a member or leader of a team |
| C17 | Communicate effectively on complex engineering matters with technical and non-technical audiences |
| C18 | Plan and record self-learning and development as the foundation for lifelong learning/CPD |