

# Retina Damage Detection and Classification Based On OCT Images

## Graduation Thesis: The Interim Report

**Yueju Han**

Student ID: 50079687

Program: AI

u11yh21@abdn.ac.uk

*Aberdeen-SCNU Joint Institute of Data Science and AI,  
University of Aberdeen, Aberdeen AB24 3UE, UK*

## Introduction

Retina diseases are a primary concern in ophthalmology due to their high prevalence and challenging treatment options. The primary symptom of these illnesses is the distortion of any portion of the retina, which causes visual problems and can lead to irreversible visual impairment further [1, 2]. Moreover, the medical causes of these diseases are complex, as evidenced by different status of retina damages. Therefore, early detection and diagnosis of retina diseases is critical for accurately inferring and preventing their progression to more severe stages.

In recent years, the introduction of Optical Coherence Tomography (OCT) has largely facilitated the diagnosis and treatment retina diseases. This innovative imaging technique captures high-resolution cross-sectional images at a microscopic level using low-coherence light [3], making it the most commonly used method for imaging the retina today. Notably, the convenient availability of OCT images also created significant demand for certified professionals to interpret images and diagnose the diseases correctly. Furthermore, some lesions may elude direct observation which make manual interpretation prone to errors [2]. Thus, the automated diagnosis system is of great interest in the field of ophthalmology, as it can help streamline the process of interpreting OCT images and improve diagnostic accuracy.

Recently, the advance of artificial intelligence has provided a solution to this challenge. Based on past research, the accuracy of traditional approaches for classifying retinal disorders has ranges from 80% to 91% [1]. Moreover, deep learning models can be more effective at the classification of OCT images because they can extract patterns and features from labeled images and can attain even higher accuracy score [4]. Hence, in this project, I aim to develop a professional deep learning model that achieves high accuracy in classifying the retina diseases and is efficient in making diagnosis decision.

## Goals

The primary goal of this project is to develop an advanced deep learning model that enhances the accuracy and efficiency of diagnosing retinal diseases using OCT images. This main goal can be further divided into the following sub-goals:

- **Enhance Classification Accuracy:** Improve the model's ability to accurately classify the four retina diseases by leveraging the state-of-the-art architecture and making appropriate modification.
- **Optimize Decision-Making Efficiency:** Develop a system that not only achieves high diagnostic accuracy but also ensures fast decision-making, aiding clinicians in delivering timely diagnoses and treatment plans.
- **Better Clinical Interpretation :** Resolve the interpretation challenge by delivering transparent and enlightening interpretability outcomes, which could guide physicians' focus toward critical areas, helping them double-check and verify results.

If time permitted, I plan to further improve my deep learning skills and challenge towards more sophisticated task and problems. They are listed in the following points:

- **Exploring Few-Shot Learning:** Investigate the potential of few-shot learning techniques to classify retinal diseases with limited labeled data. This goal aims to address the challenge of acquiring sufficient labeled data for training deep learning models, which is often a constraint in medical imaging tasks.
- **Domain Adaptation Capabilities:** Explore domain adaptation strategies to improve the model's performance across different datasets, ensuring that the model can generalize well even when trained on data from one domain and applied to another.

## Methodology

The methodology of this research aims to develop a deep learning model for the classification of retinal diseases using Optical Coherence Tomography (OCT) images. The process involves multiple stages, and below are the key components of the methodology:

- **Data Preprocessing:** The open source dataset OCT2017 [5] is utilized as the main dataset. In this dataset, OCT images are already labelled by the professionals in a well-structured form. However, I observe that the separation of the dataset is not in the standard way. There are approximately up to 120,000 images in training set, while only about 1,000 images in the test set and only 32 images in the validation set. So it requires more standard way to separate the dataset.
- **Data Augmentation:** Applying data augmentation techniques such as rotation, flipping, and zooming to artificially increase the dataset and improve model robustness.

- **Model Development:** This research will utilize deep learning models, particularly convolutional neural networks (CNNs), known for their effectiveness in image classification tasks. A pre-trained model (e.g., ResNet, VGG) may be used as a baseline, with transfer learning applied to adapt the model to OCT image data. To achieve the better outcomes, some state-of-the-art general models will also be investigated and modified to adopt to this task
- **Model Training:** Using cross-entropy loss function for the multi-class classification. Applying optimizers like Adaw or SGD to minimize the loss function. In this stage, I will use grid search to find the most optimal hyperparameters for each model.
- **Model Evaluation:** Key performance indicators (KPIs) will include accuracy, precision, recall, F1 score. Confusion Matrix would be utilized to visualize the model's performance and identify areas of improvement.
- **Visualization Using GRAD-CAM:** To improve the transparency and interpretability of the deep learning model, GRAD-CAM (Gradient-weighted Class Activation Mapping) [6] will be applied for visualizing the regions of the OCT images that the model focuses on when making its predictions.
- **Optional Extensions:** If time allows, additional strategies such as few-shot learning and domain adaptation will be explored to improve the model's performance under conditions of limited labeled data and varying data distributions.

## Resources Required

To carry out this research, Python and the PyTorch framework are essential tools for developing and training the deep learning model. The operating system I will use locally is Windows, with Visual Studio Code (VSCode) chosen as the integrated development environment (IDE).

For model training, the availability of a GPU is crucial. To meet this requirement, I plan to rent computational resources from online cloud service providers such as Alibaba Cloud, Autodl, etc. However, these platforms typically operate on Linux-based systems. As a result, some modification are necessary to carry out training process with the code writing on Windows.

## Risk Assessment

## Timetable

<b>Risk Name</b>	<b>Cause</b>	<b>Impact</b>	<b>Mitigation Strategy</b>
<b>Computational Resource Limitations</b>	The need for GPUs for training large models, with potential cost or availability issues.	If resources are unavailable or too expensive, training could be delayed.	Optimize the model to use fewer resources or switch to smaller models, explore cost-effective GPU rental options.
<b>Model Performance and Overfitting</b>	Small datasets or lack of regularization could lead to overfitting.	Poor generalization to unseen data, resulting in low performance in real-world scenarios.	Employ dropout, data augmentation, and regularization techniques. Monitor validation performance and implement early stopping.
<b>Time Pressure and Project Progress</b>	Tight deadlines and the need for quick results.	Project could face delays or incomplete results, especially during model tuning and data processing.	Properly allocate time, focus on critical milestones first, and adjust project scope if necessary. Prioritize model optimization and core functionality.

Table 1: Risk Assessment

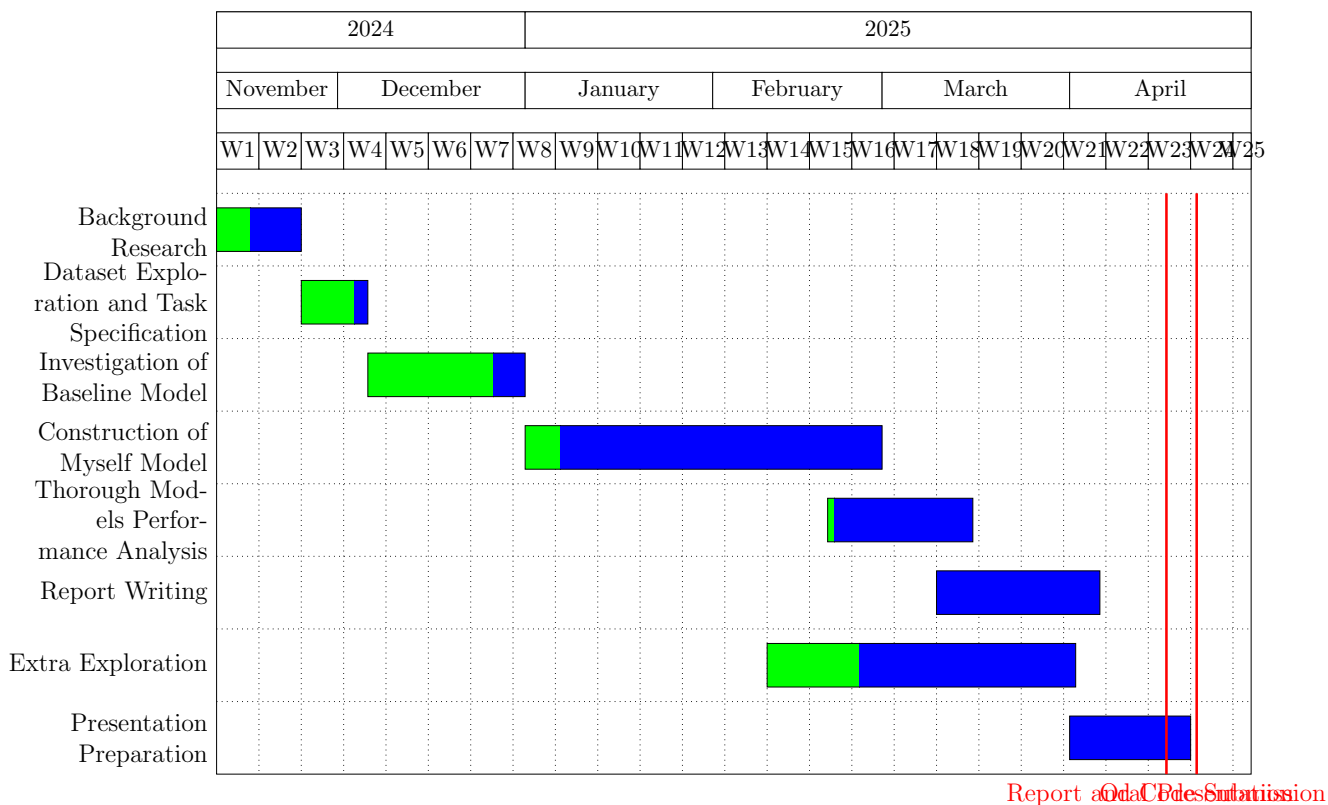


Figure 1: Main Project Activities.

## References

- [1] Malliga Subramanian, Kogilavani Shanmugavadivel, Obuli Sai Naren, K Premkumar, and K Rankish. Classification of retinal oct images using deep learning. In *2022 international conference on computer communication and informatics (ICCCI)*, pages 1–7. IEEE, 2022.
- [2] Umer Sadiq Khan and Saif Ur Rehman Khan. Boost diagnostic performance in retinal disease classification utilizing deep ensemble classifiers based on oct. *Multimedia Tools and Applications*, pages 1–21, 2024.
- [3] Lehan Wang, Weihang Dai, Mei Jin, Chubin Ou, and Xiaomeng Li. Fundus-enhanced disease-aware distillation model for retinal disease classification from oct images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 639–648. Springer, 2023.
- [4] GR Hemalakshmi, M Murugappan, Mohamed Yacin Sikkandar, S Sabarunisha Begum, and NB Prakash. Automated retinal disease classification using hybrid transformer model (svit) using optical coherence tomography images. *Neural Computing and Applications*, pages 1–18, 2024.
- [5] Daniel S Kermany, Michael Goldbaum, Wenjia Cai, Carolina CS Valentim, Huiying Liang, Sally L Baxter, Alex McKeown, Ge Yang, Xiaokang Wu, Fangbing Yan, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *cell*, 172(5):1122–1131, 2018.
- [6] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626, 2017.