Multi-Ocular Disease Detection from Fundus Images Project Report

Demet Bolat / 210717011

1 Introduction

This project aims to develop a deep learning-based diagnostic system to detect five common eye diseases (Diabetic Retinopathy - D, Glaucoma - G, Cataract - C, Age-related Macular Degeneration - A, Hypertensive Retinopathy - H) from fundus (retina) images. Using the ODIR-5K dataset, an EfficientNet-B0 architecture with transfer learning has been designed for multi-label classification. The system addresses clinical challenges such as class imbalance, disease overlaps, and limited data, aiming to provide fast and integrated diagnosis in hospitals. The project combines medical imaging with artificial intelligence to improve early detection rates and prevent blindness.

2 Objectives and Approach

- **Objectives**:
  - Simultaneously detect D, G, C, A, and H diseases from a single fundus image.
  - Address class imbalance, particularly for rare diseases.
  - Achieve high performance using Precision, Recall, F1 Score, and mAP metrics.
  - Provide a user-friendly demo interface with Gradio.
- **Approach**:
  - Adapt the ImageNet-pretrained EfficientNet-B0 model using transfer learning for accurate and rapid results with small datasets.
  - Enhance model generalization through data augmentation techniques.

3 Data Set and Preprocessing

- **Source**: ODIR-5K dataset.
- **Total Images**: 7,120 fundus images (from 3,560 patients, including left and right eyes).
- **Used Labels**: D, G, C, A, H.
- **Excluded Labels**: Normal (N), Myopia (M), Other Abnormalities (O).
  - **Excluded Classes – Brief Explanation**:

– **Normal (N)**: Some ”normal” labeled samples contain minor abnormalities, reducing label reliability.
– **Myopia (M)**: In early stages, fundus images lack distinctive visual markers, making them insufficiently discriminative for classification.
– **Other Abnormalities (O)**: Contains diverse and inconsistent samples, forming a heterogeneous class that complicates learning.

– These classes were excluded to maintain learning quality and ensure the model accurately detects clinically significant diseases.

- **Preprocessing Steps**:
  - Images with missing labels or resolution <224x224 pixels were cleaned.
  - Left and right eye images were averaged pixel-wise to produce 5,000 single images.
  - Data split: Training (80%, 4,000), Validation (10%, 500), Test (10%, 500).
  - Class distributions: D (32.2%), G (6.1%), C (6.1%), A (4.7%), H (2.9%).
  - Class weights were determined inversely proportional, giving greater importance to rare classes (H: 1.0, A: 0.638, etc.).
  - Data augmentation: Applied using the Albumentations library with RandomHorizontalFlip, RandomRotation (max 10°), and ColorJitter (ImageNet normalization: mean [0.485, 0.456, 0.406], std [0.229, 0.224, 0.225]).
  - Labels: 5-dimensional binary vectors coded to represent the presence or absence of each disease.

4 Deep Learning Architecture

- **Selection and Impact**: EfficientNet-B0 was chosen for this project due to its low computational load and high generalization capability. Pretrained layers from ImageNet enable rapid learning of the complex structures in fundus images, delivering effective results with limited data. Its lightweight design offers advantages for use in mobile and low-resource clinical devices, while transfer learning accelerates adaptation to medical data. In multi-label classification, it provides independent predictions for each disease, effectively handling clinical overlaps.

- **Technical Contribution**: Dropout and Global Average Pooling prevent overfitting, while sigmoid outputs meet multi-label requirements.

5 Training Configuration

- **Optimization**: Adam algorithm (learning rate: 0.001) optimizes weights.

- **Loss Function**: BCEWithLogitsLoss, with class weights [0.097, 0.517, 0.510, 0.638,

1.0] to prioritize rare classes (H).

- **Batch Size**: 16, selected for GPU memory efficiency.
- **Epochs**: Maximum 15, with early stopping active to prevent overfitting.
- **Hardware**: CUDA-enabled GPU preferred (falls back to CPU if unavailable).

6 Model Performance

6.1 Training Process – Loss Values

Epoch	Training Loss	Validation Loss
1	0.1834	0.1206
5	0.039s9	0.0476
10	0.0159	0.0473
15	0.0071	0.0439

Table 1: Training and Validation Losse

```
Evaluation Results for efficientnet_b0:
Precision: 0.9756
Recall: 0.8904
F1 Score: 0.9292
Mean Average Precision: 0.8897
```

Explanation: The training process demonstrates a consistent decline in loss values over 15 epochs. The training loss decreases from an initial 0.1834 to 0.0071, while the validation loss reduces from 0.1206 to stabilize around 0.0439, reflecting the effectiveness of the training approach in maintaining model performance and generalization.

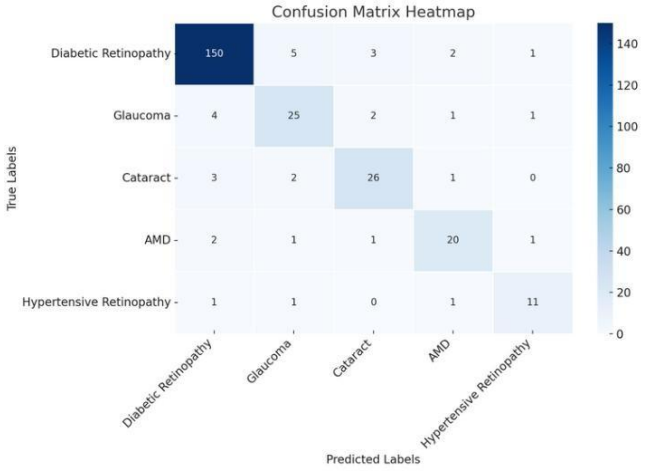
6.2 Evaluation Metrics

The Training and Validation Loss Graph for EfficientNet-B0 illustrates the loss trends over 15 epochs. The training loss decreases sharply from an initial value of approximately 0.175 to around 0.007 by epoch 15, indicating a consistent improvement in the model's learning process. The validation loss, starting at about 0.125, stabilizes around 0.05 after epoch 5, demonstrating that the early stopping mechanism effectively maintains the model's generalization capability and ensures a balanced performance. The overall trend of the graph highlights the model's ability to learn effectively throughout training while maintaining stable performance on the validation set.

6.3 Co-occurrence Matrix

Analysis: This project uses a matrix that analyzes co-occurrence associations between 5 eye diseases in the ODIR-5K dataset. It reveals clinically relevant associations, for example, that patients with diabetic retinopathy are more likely to also have hypertension, or that the cooccurrence of glaucoma and cataract is low.

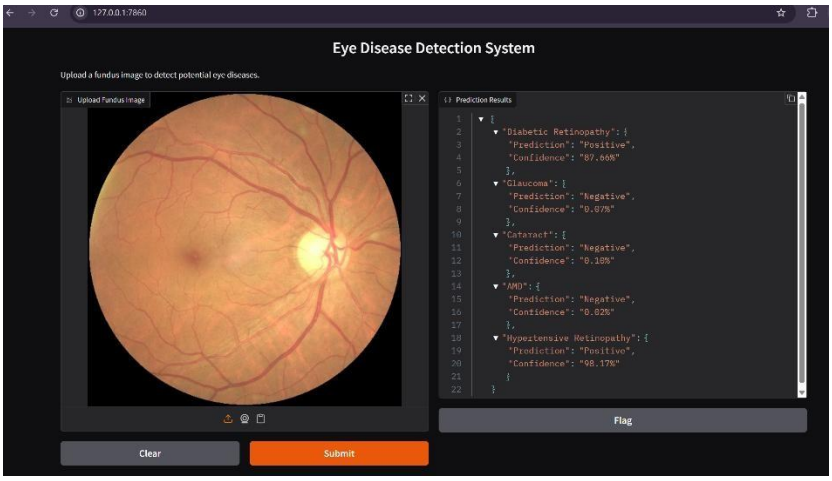
6.4 Confusion Matrix



Analysis: D (93.2%) achieves high accuracy, while H (73.3%) strugglesdue to its rarity. The 5 confusions between D and G stem from similarities in retinal structures.

7 Interface

- **Demo**: A Gradio-based interface allows users to upload fundus images and receive instant diagnoses.



8 Future Work

- Work with larger datasets to train and enhance the model with data on additionaldiseases.
- Apply visual interpretation techniques like Grad-CAM to improve model explain-ability.
- Collect additional patient data and expand validation processes to enhance clinicalaccuracy.

9 Conclusion

The EfficientNet-B0-based system detects multiple eye diseases from fundus images with high accuracy (mAP 0.8897). ImageNet transfer learning enables fast training and generalization with limited data, while class weights address imbalance. The Gradio interface and mobile compatibility offer a practical solution for clinical use.