I AI

Duygu AKMAN, Ali Emre YENİHAYAT, Berk ÜLKER

1. Description of the Project

This project aims to develop an artificial intelligence model capable of predicting potential ocular diseases by utilizing individuals' eye images (such as retinal and fundus photographs, as well as optic disc images), along with demographic data including age and gender. The model will analyze visual data through deep learning techniques while incorporating demographic information as auxiliary inputs. Various model architectures will be explored and compared in terms of accuracy, computational efficiency, and overall performance. Additionally, the project will address key challenges in multi-class and multi-label classification, data imbalance, and asymmetry between left and right eye images.

2. Literature Titles and Links

1. **Optimising deep learning models for ophthalmological disorder classification:** https://www.nature.com/articles/s41598-024-75867-3

2. **Recognition of eye diseases based on deep neural networks for transfer learning and improved D-S evidence theory:** https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-023-01176-2

3. **Deep learning-based classification of eye diseases using CNN for OCT images:** https://www.frontiersin.org/journals/computerscience/articles/10.3389/fcomp.2023.1252295/full

4. **Ocular Diseases Diagnosis in Fundus Images using Deep Learning:** https://arxiv.org/abs/1905.02544

3. Description of the Dataset

Within the scope of the project, the ODIR-5K (Ocular Disease Intelligent Recognition) dataset will be used. (https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k)
Data Type: Visual + Tabular (CSV)
Total Records: Approximately 5,000 patients (10,000 eye images)
Image Format: JPEG (RGB)
Labeling: Multi-labeled — each patient may be associated with more than one disease.

Attributes:

Column Name	Description
ID	Unique identifier assigned to each patient
Patient Age	Age of the patient
Patient Sex	Gender of the patient
Left-Fundus / Right-Fundus	Image files of the left and right eyes
N, D, G, C, A, H, M, O	Eight disease categories
filepath	Path to the corresponding image file
target	Numerical label vector (e.g., [1, 0, 0, 0, 0, 0, 0, 0])
filename	Name of the image file

4. Platforms, Libraries, and Hardware

Platform / Environment:

- Python
- Kaggle Notebook / Visual Studio Code / PyCharm
- Git / GitHub
- Experiment tracking tools: Weights & Biases (wandb), MLflow, TensorBoard, etc.

Python Libraries:

- PyTorch, scikit-learn, pandas, numpy, OpenCV, matplotlib, etc.

Framework / Approach:

- Custom and transfer learning models based on Convolutional Neural Networks (CNNs)
- Multimodal (fusion) models integrating visual and tabular data
- Hyperparameter optimization
- Multi-input and multi-label model design

Hardware:

- GPU: NVIDIA RTX 4080
- RAM: 32 GB
- CPU: Intel i9-14900HX

5. Model Design and Training Strategy

- **Image Processing:** CNN-based architectures will be employed for visual feature extraction.
- **Tabular Data:** Age and gender information will be processed through dedicated dense layers.
- **Fusion Layer:** Visual and tabular features will be integrated within a multimodal fusion layer.
- **Classification Type:** Multi-label classification, allowing simultaneous prediction of multiple ocular diseases.
- **Training Parameters:** Optimization will be performed with respect to batch size, learning rate, and optimizer selection.
- **Performance Evaluation:** Metrics such as Accuracy, Precision, Recall, F1-score, and ROC-AUC will be used to assess model performance.
- **Explainability:** Model decision-making will be visualized and interpreted using Grad-CAM.
- **Data Augmentation:** Techniques including rotation, flipping, color jittering.
- **Regularization:** Dropout, weight decay, and early stopping methods will be implemented to prevent overfitting.

6. Task Matrix

	Duygu AKMAN	Ali Emre YENİHAYAT	Berk ÜLKER
Selection of Database	X	X	X
Selection of Models			X
Image Preprocessing	X		
Data Cleaning			X
Preparing Dataloader	X	X	
Hyperparameter Optimization		X	X
Model Checkpointing	X		X
Augmentation	X	X	
Model Testing	X		
Imputation		X	X
GRAD-CAM		X	
Confusion Matrix	X		X
Graphics		X	X
Report	X	X	