A Multimodal Approach for Ophthalmic Disease Diagnosis

Minor Project-II Presentation (B.Tech 3rd A.I.)
AIC 3950

Presented By:	Faculty Number:	Group No. :	Supervisor:
Faisal Ali Khan	21AIB196		Dr. Junaid Ali Reshi
Imaad Hasan	21AIB120	14	

Contents

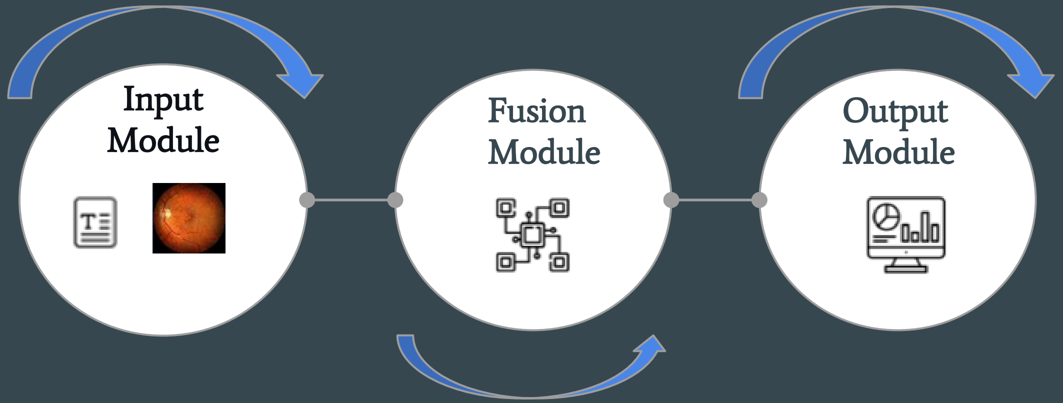
Introduction

What is Multimodal Learning?

It is a subfield of artificial intelligence that seeks to effectively process and analyze data from multiple modalities.

These systems constitute three fundamental components :

- An input module
- A fusion module
- An output module



Problem

- **Incompleteness:** Single-source fundus imaging often results in incomplete diagnostics, missing the full complexity of ocular conditions.
- **Accuracy Issues:** Reliance on visual fundus images alone can lead to inaccuracies and diagnostic errors.
- **Diagnostic Bias:** Traditional methods are subject to biases influenced by image quality and interpreter experience.
- **Lack of Integration:** Current diagnostics do not adequately incorporate crucial patient-specific data like demographics and clinical history.
- **A Need for Better Methods:** There is a critical need for a multimodal diagnostic approach that integrates various data types to improve reliability and precision.

Related Work

- **Al-Fahdawi et al., (2023) :** Developed Fundus-DeepNet, a sophisticated multi-label deep learning system that integrates enhanced image preprocessing and deep feature extraction techniques to accurately detect multiple ocular diseases.
- **Ran Xiao et al., (2023) :** Developed a multimodal deep learning model that integrates ECG waveforms and patient demographics to enhance diagnostic accuracy for myocardial infarction, achieving significant improvements in performance.
- **Keerthiveena et al., (2020):** Focused on computer-aided diagnosis using feature fusion, underlining the benefit of integrating multiple data sources for enhanced diagnostic accuracy.
- **Chelaramani et al., (2020):** Explored multi-task learning to enhance fine-grained eye disease prediction by developing methods for generating textual diagnoses, facilitating a deeper understanding of diagnostic outcomes.

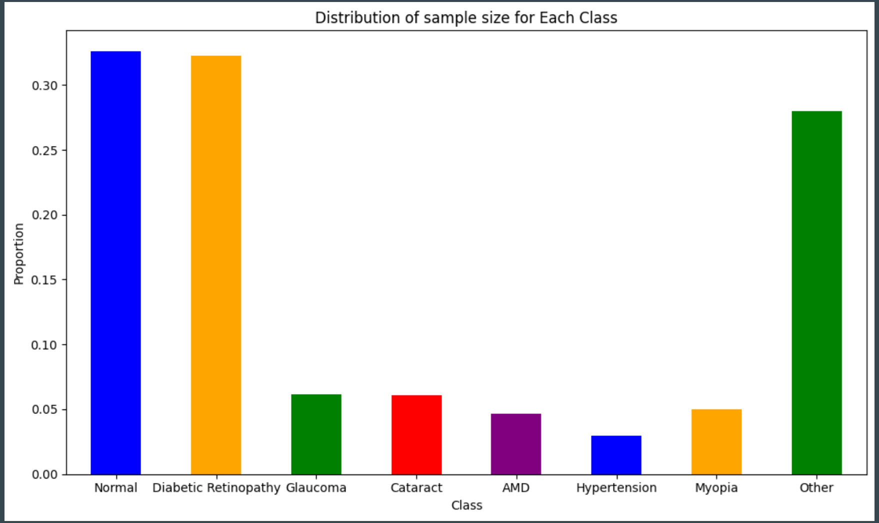
System Design - The Dataset

A structured ophthalmic database of **5,000 patients with age, color fundus photographs from left and right eyes and doctors' diagnostic keywords** from doctors (in short, ODIR). They classify patient into *eight labels* including **normal (N), diabetes (D), glaucoma (G), cataract (C), AMD (A), hypertension (H), myopia (M) and other diseases/abnormalities (O)** based on both eye images and additionally patient age.

Training Set : 7000 images (3500 samples)

On-site Test(Validation Set) : 1000 images (500 samples)

Off -site Test(Test Set) : 2000 images (1000 samples)



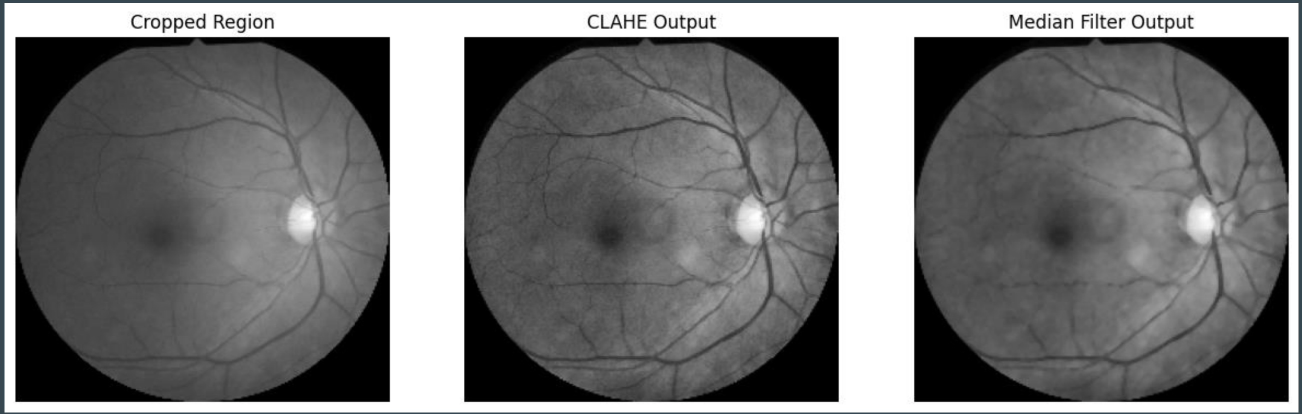
System Design - Data Preparation

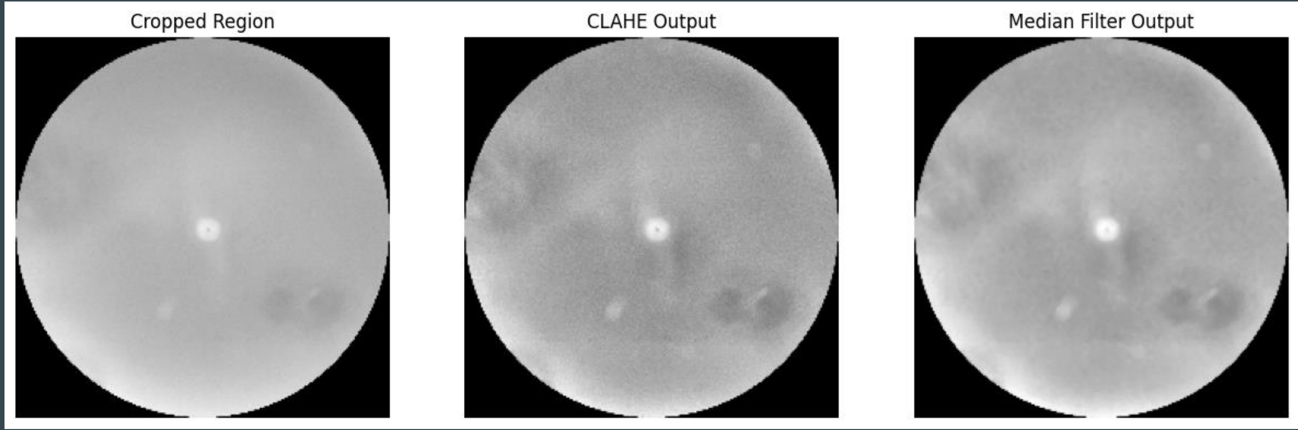
Preprocessing:

- ❑Circular border cropping to remove black borders around the eye images.
- ❑Image resizing to a standard (224x224) pixels.
- ❑Contrast enhancement using the CLAHE method.
- ❑Median filter applied for noise reduction.
- ❑Conversion of images to grayscale to standardize the input.
- ❑Generation of a clipping mask to extract only the relevant part of the images.

Enhancement Techniques:

- Use of CLAHE (Contrast Limited Adaptive Histogram Equalization):
  1. Operates on small regions in the image (tiles).
  2. Improves the contrast of the images while preserving natural appearance.
- CLAHE parameters:
  1. **Clip limit (CL):** 2 for increased brightness.
  2. **Block size (BS):** (8×8) for enhanced contrast and wider intensity range.





System Design - Original Model Architecture

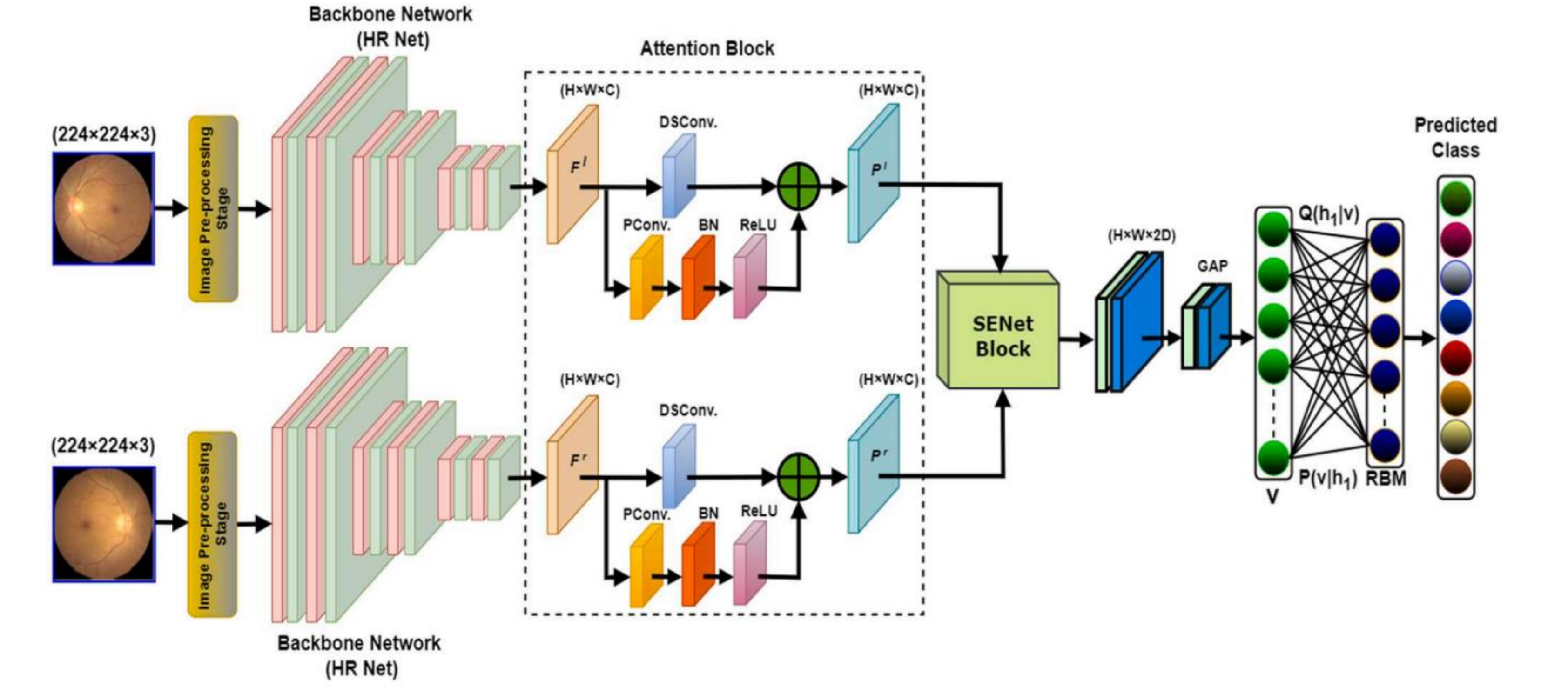
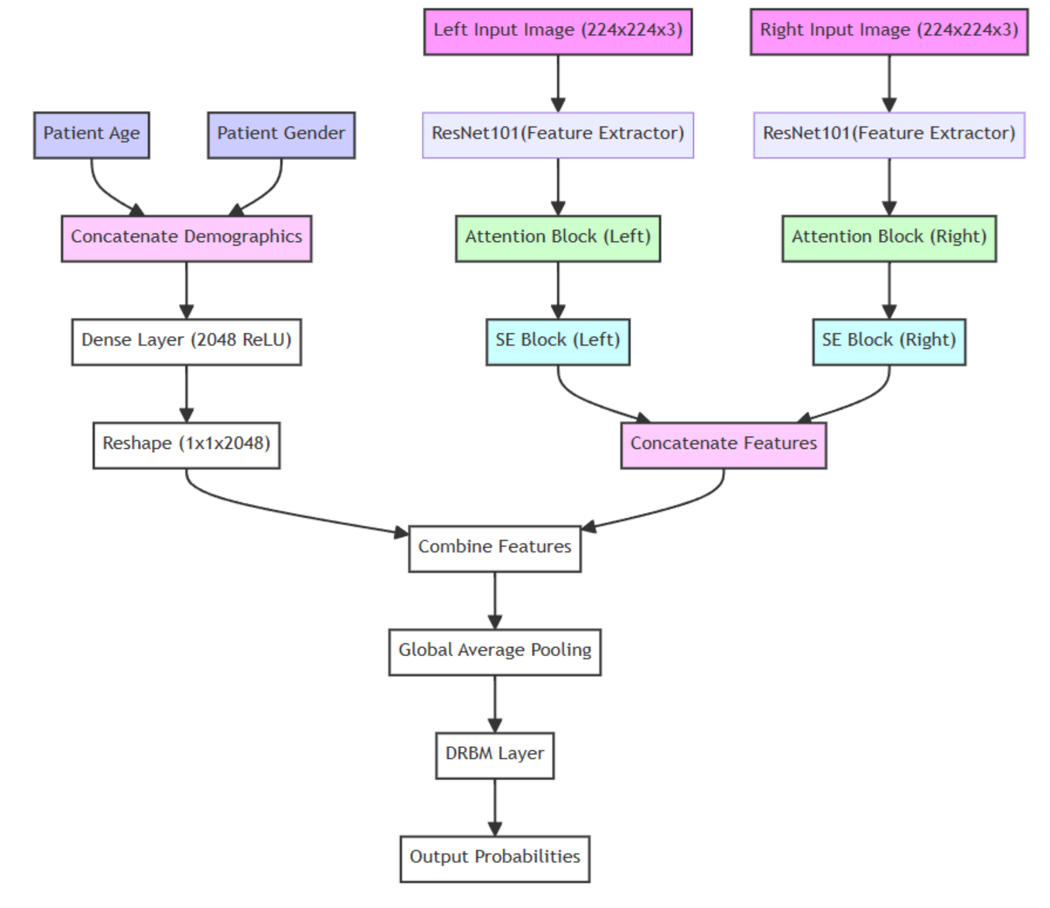


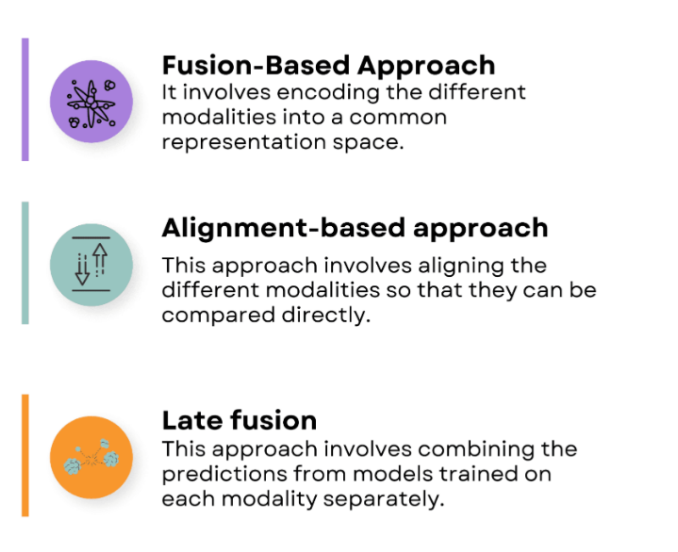
Fig. 1. The diagram illustrating the structure of the proposed Fundus-DeepNet system.

Source: **Al-Fahdawi, S., Al-Waisy, A. S., Zeebaree, D. Q. et al.(2024).** Fundus-DeepNet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images. Information Fusion, 102, 102059

System Design - Proposed Model Architecture



System Design - Late Feature Fusion

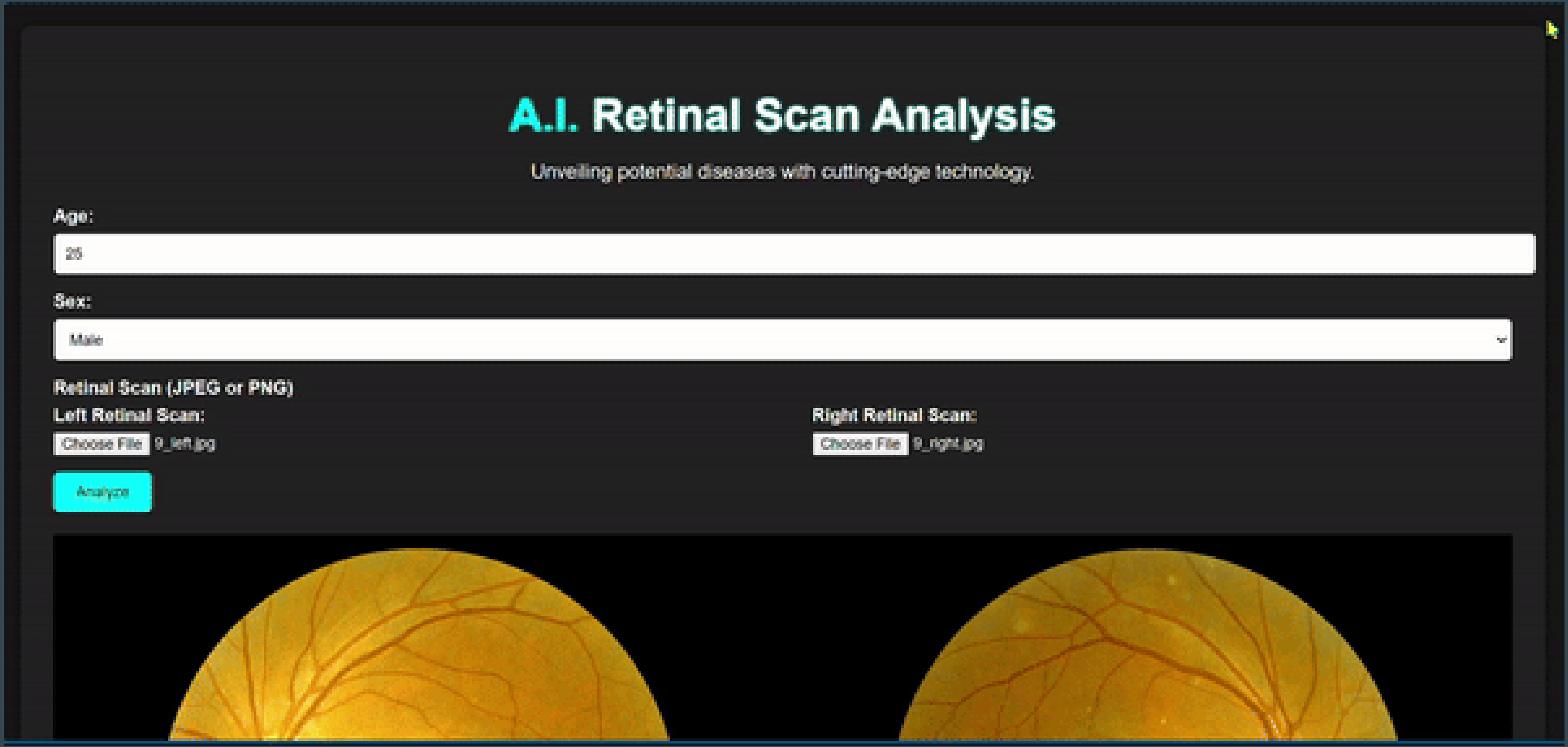


Source: https://www.kdnuggets.com

The type of feature fusion used in our model is **Late Fusion.** The fusion approach is implemented:

- **Separate Feature Extraction:** Features are extracted independently from both the left and right inputs using possibly the same feature extraction method .
- **Attention and SE Blocks:** Each set of features from the left and right inputs are further processed separately through attention blocks and SE blocks.
- **Concatenation of Features:** The features processed by the attention and SE blocks from both inputs are then concatenated. This is the step of **late fusion** where the combination of features happens late in the process, just before the final decision-making layers.

System Implementation



Results

Testing:

Accuracy Score : 70%

AUC Score : 83%

Training:

Accuracy Score : 91%

AUC Score : 100%

Original Results:

Accuracy : 74%

AUC Score : 98%

References

- **Al-Fahdawi, S., Al-Waisy, A. S., Zeebaree, D. Q. et al.(2024).**Fundus-DeepNet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images. Information Fusion, 102, 102059. [DOI: 10.1016/j.inffus.2023.102059]

- 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 Lecture notes in computer science (pp. 177–193). https://doi.org/10.1007/978-3-030-71058-3_11

- https://www.kdnuggets.com/2023/03/multimodal-models-explained.html

Bibliography

- **Al-Fahdawi, S., Al-Waisy, A. S., Zeebaree, D. Q. et al.(2024).**Fundus-DeepNet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images. Information Fusion, 102, 102059. [DOI: 10.1016/j.inffus.2023.102059]

- **Xiao, R., et al. (2023).**Integrating multimodal information in machine learning for classifying acute myocardial infarction. Physiological Measurement, 44(4), 044002. Published online 2023 Apr 18. [DOI: 10.1088/1361-6579/acc77f]

- **Gour, N., & Khanna, P. (2021).**Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network. Biomedical Signal Processing and Control, 66, 102329. [DOI: 10.1016/j.bspc.2020.102329]

- **Chelaramani, S., Gupta, M., Agarwal, V., Gupta, P., & Habash, R. (2020).**Multi-task learning for fine-grained eye disease prediction. In S. Palaiahnakote, G. Sanniti di Baja, L. Wang, & W.Q. Yan (Eds.), ACPR 2019 (pp. 734–749). Springer, Cham. [DOI: 10.1007/978-3-030-41299-9 57]

Thank you for your attention!

Any Questions?