成都理工大學中英合作办学
CDUT Sino-British Collaborative Education

OXFORD BROOKES UNIVERSITY

Design of Deep Convolution Neural Network for Image Classification

Chinese Name: Zhong Wentao

English Name：Franklin

Student ID：202018010409

Supervisor Name: Dr Grace Ugochi Nneji

成都理工大學 中英合作办学
CDUT Sino-British Collaborative Education

Contents

01 Abstract

Abstract

- This study uses convolutional neural networks (CNNs) with spatial attention to improve image classification.
- By integrating InceptionV2, SE Networks, and ResNet architectures, the research enhances image recognition.
- Validation on standard datasets shows superior performance, highlighting model versatility and effectiveness.
- The innovation lies in combining deep learning with spatial attention for a more efficient and accurate classification system.

Keywords: Image Classification, Convolutional Neural Networks, Spatial Attention Mechanism, Deep Convolutional Neural Networks

02 Introduction

Introduction

Current Challenges

- Varying lighting conditions
- Low data resources
- Fluctuating image quality
- Low performance of the recognition model

Background

Image classification is a key area in computer vision and image processing, essential for various applications like security, human-computer interaction, social media analysis, and advertisement targeting.

Solutions

- Data Augmentation
- Data Pre-processing
- Attention Mechanism
- Machine Learning using CNNs

03 Datasets

Human Face Datasets

Figure 1: Sample display of human face dataset [1]

Split ratio of Human Face Datasets

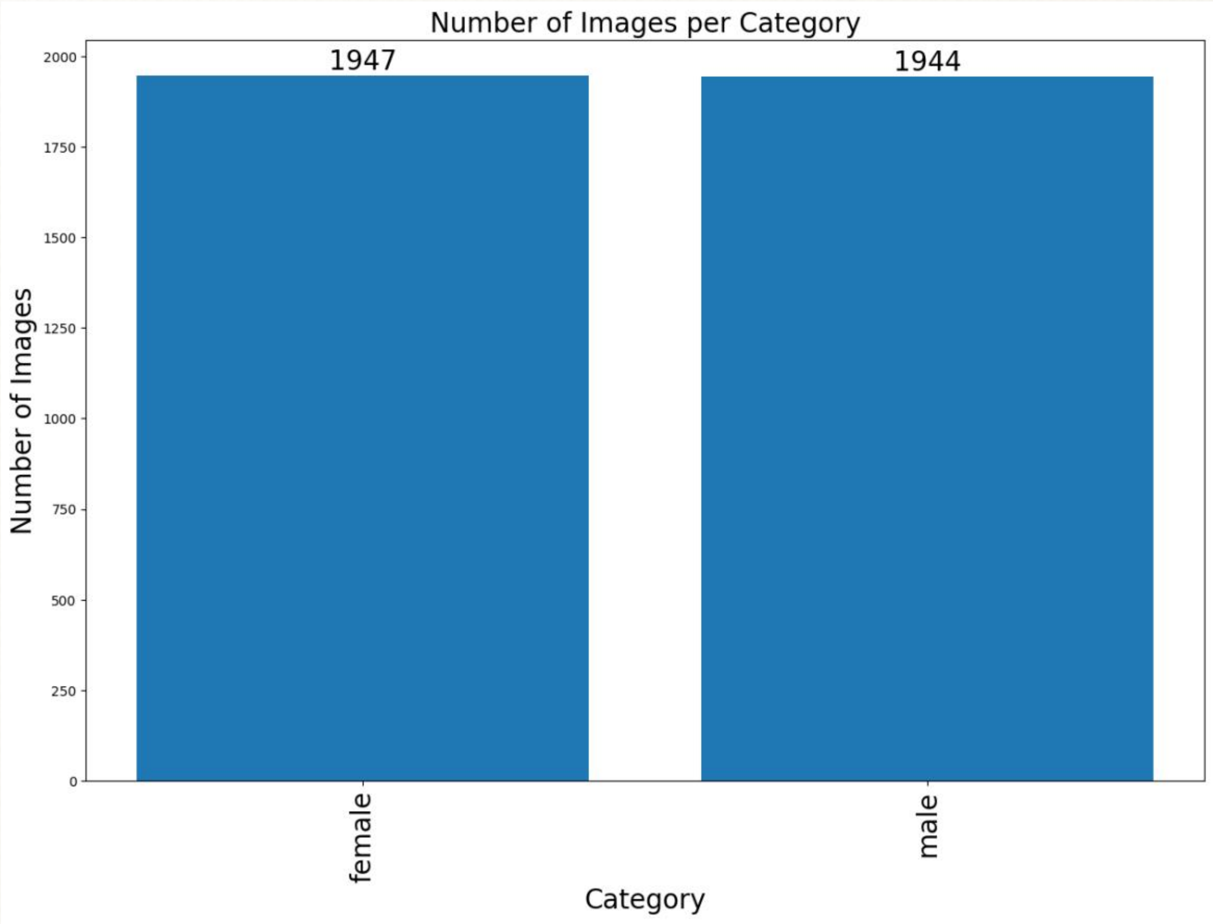


Figure 2: Distribution of human face datasets

Diseased Plant Leaf Datasets

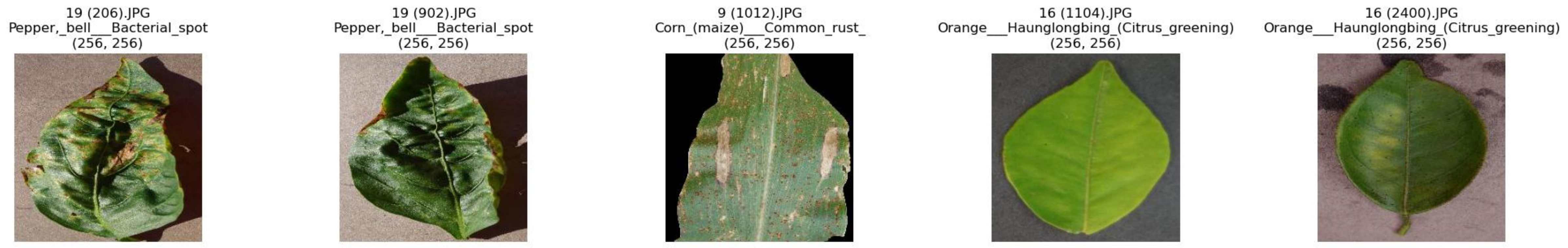


Figure 3: Sample display of plant disease dataset [2]

Split ratio of Diseased Plant Leaf Datasets

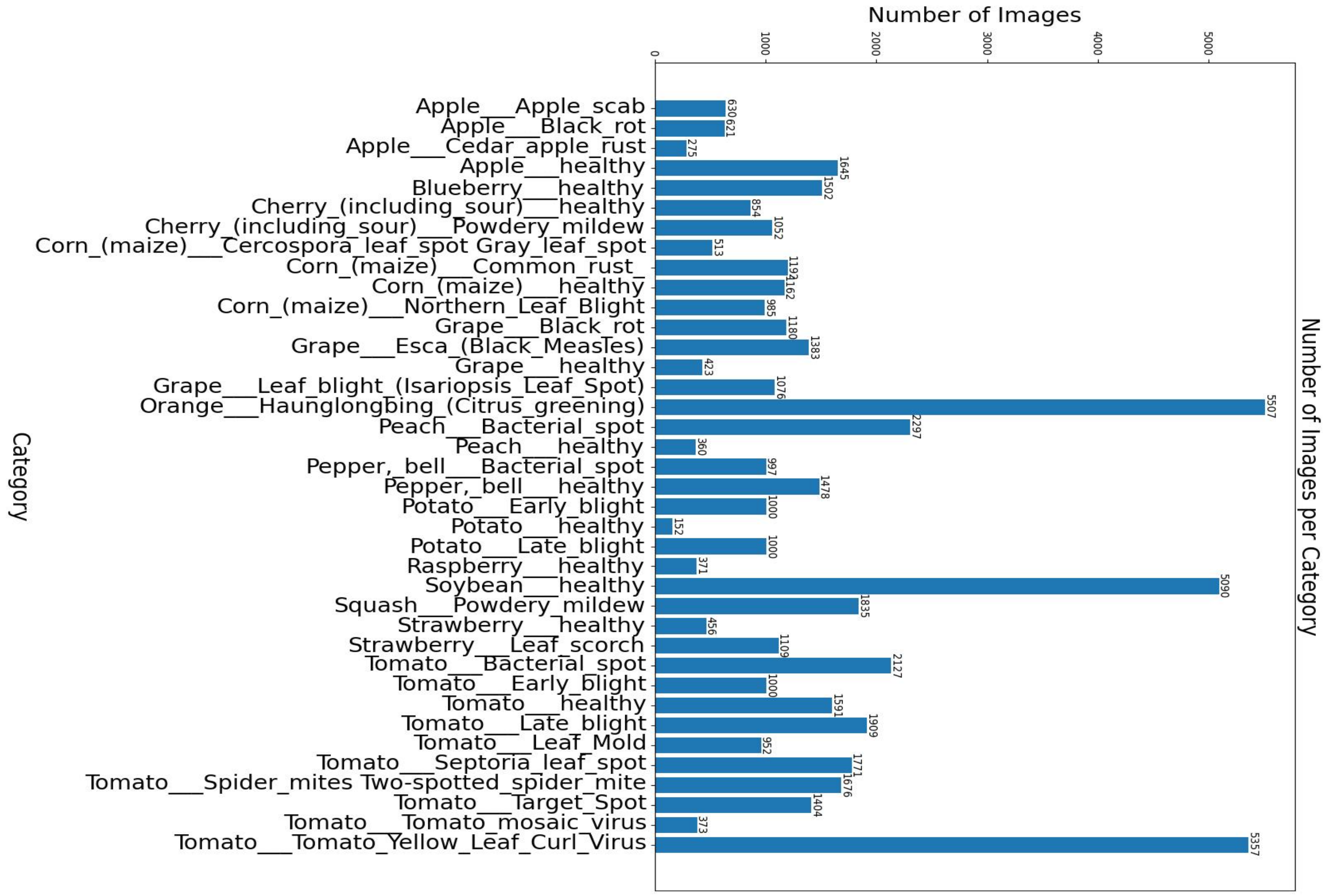


Figure 4: Distribution of plant disease datasets

Skin Cancer Datasets

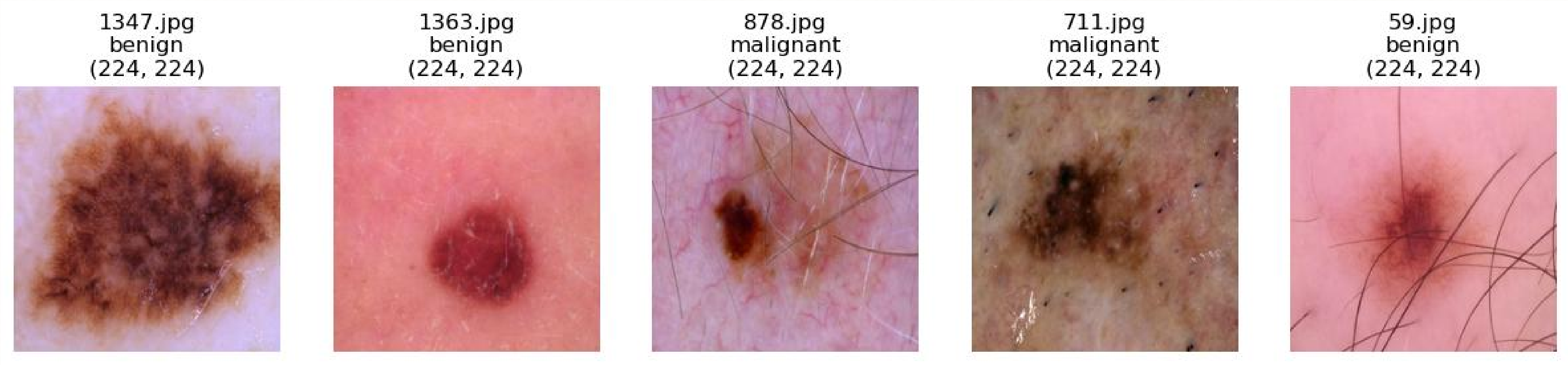


Figure 5: Sample display of skin cancer dataset [3]

Split ratio of Skin Cancer Datasets

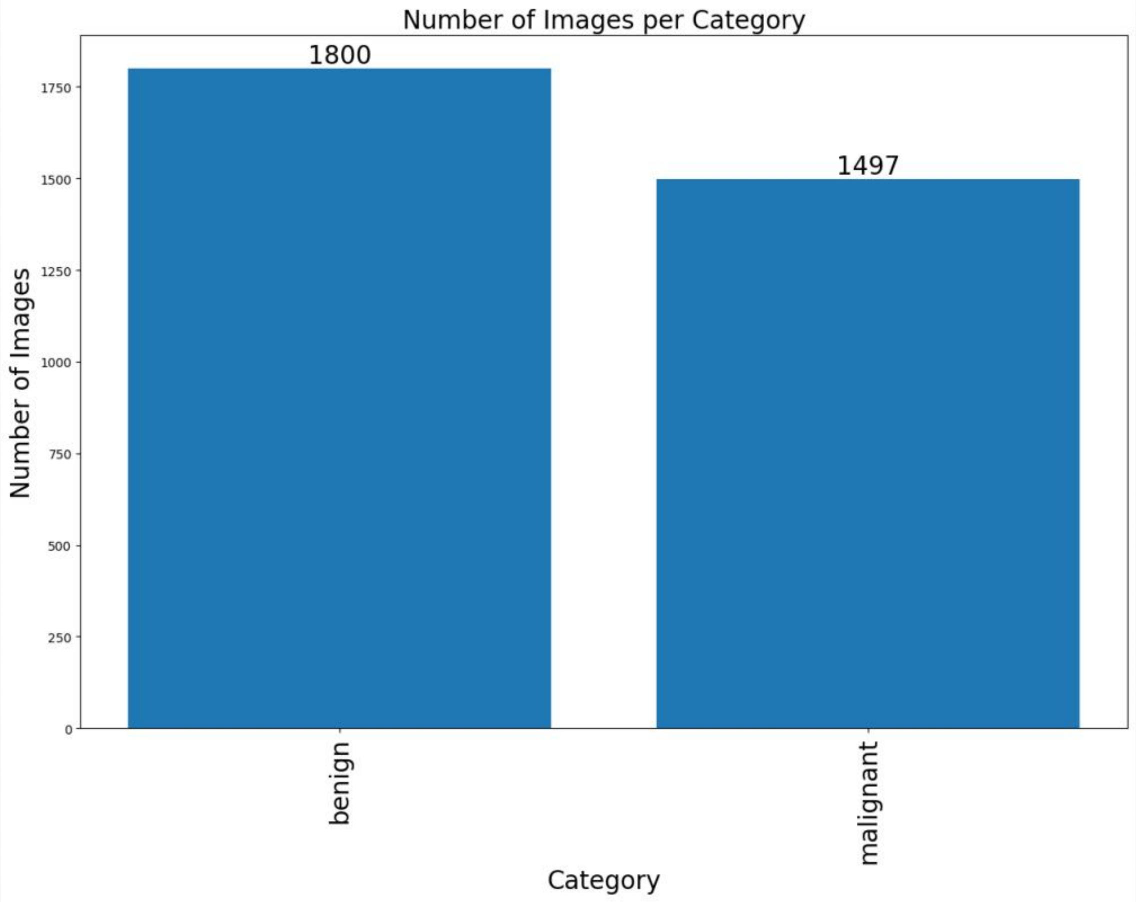


Figure 6: Distribution of skin cancer datasets

04 Proposed Model

Architecture of FranklinNet

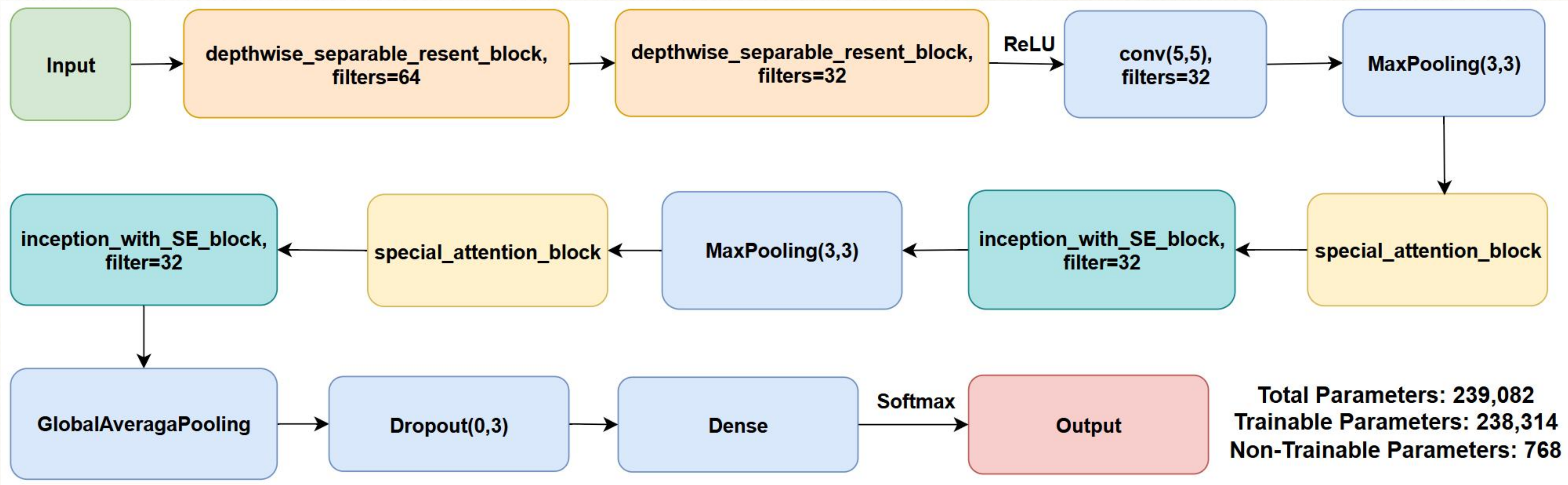


Figure 7: Ensemble model architecture diagram for FranklinNet

05 Heat Map

Heat map of model focus after using the attention mechanism

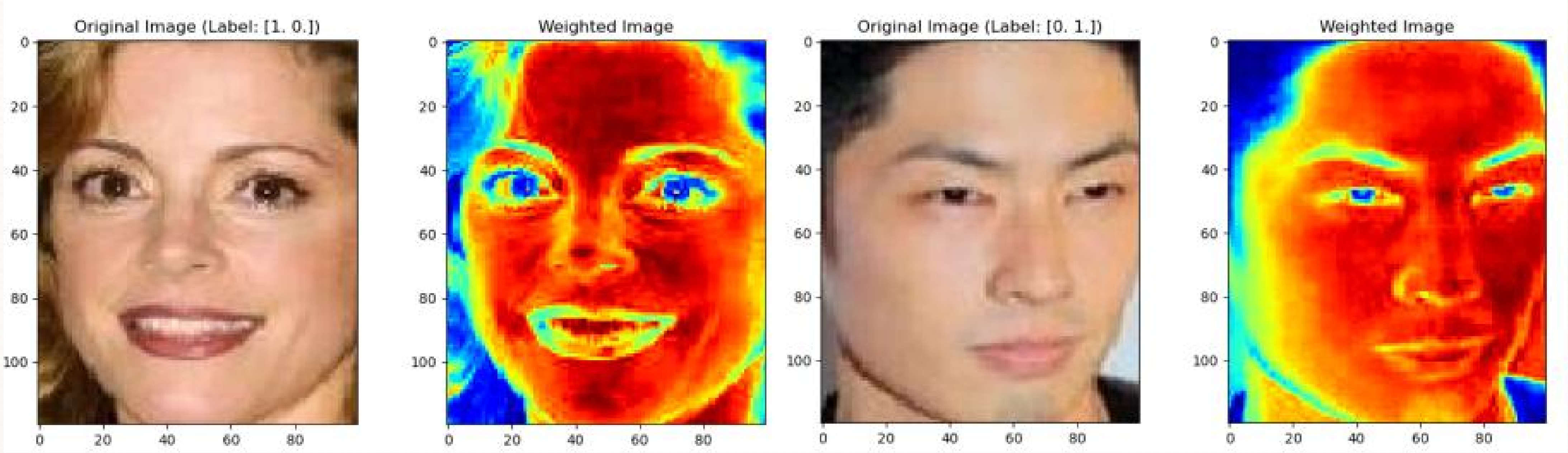
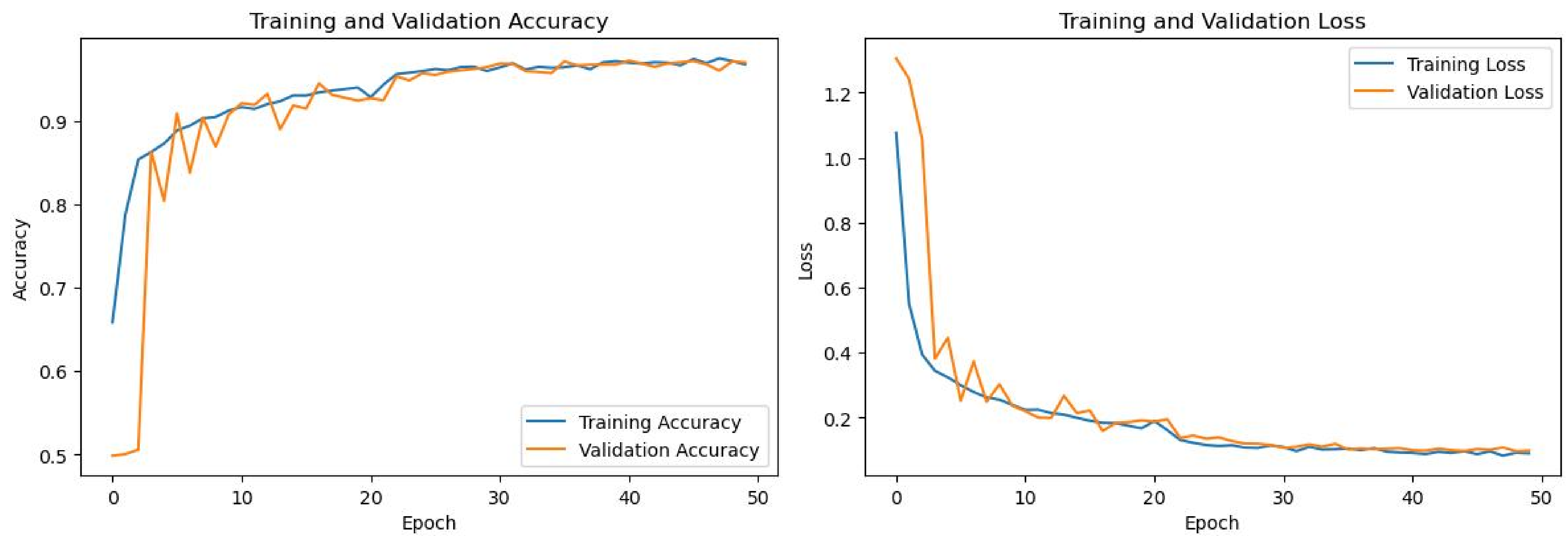


Figure 8: Facial Feature Focus Visualization via Grad-CAM

06 Results

Result of Human Face



9-1): Train_Acc=0.9676, Val_Acc=0.9799

9-2): Train_Loss=0.0891, Val_Loss=0.0516

Figure 9: Performance evaluation of FranklinNet using Human face dataset

Result of Human Face

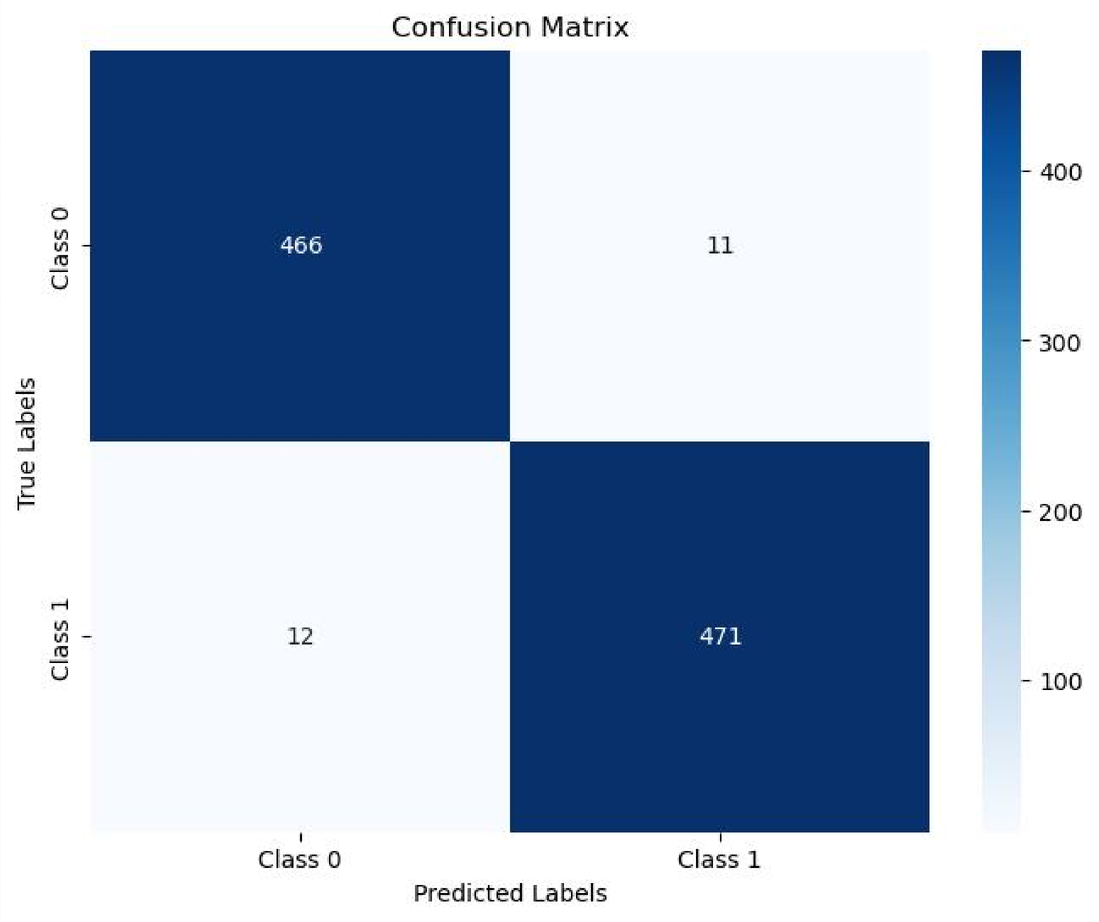


Figure 10: Confusion Matrix of FranklinNet using Human face dataset

Result of Human Face

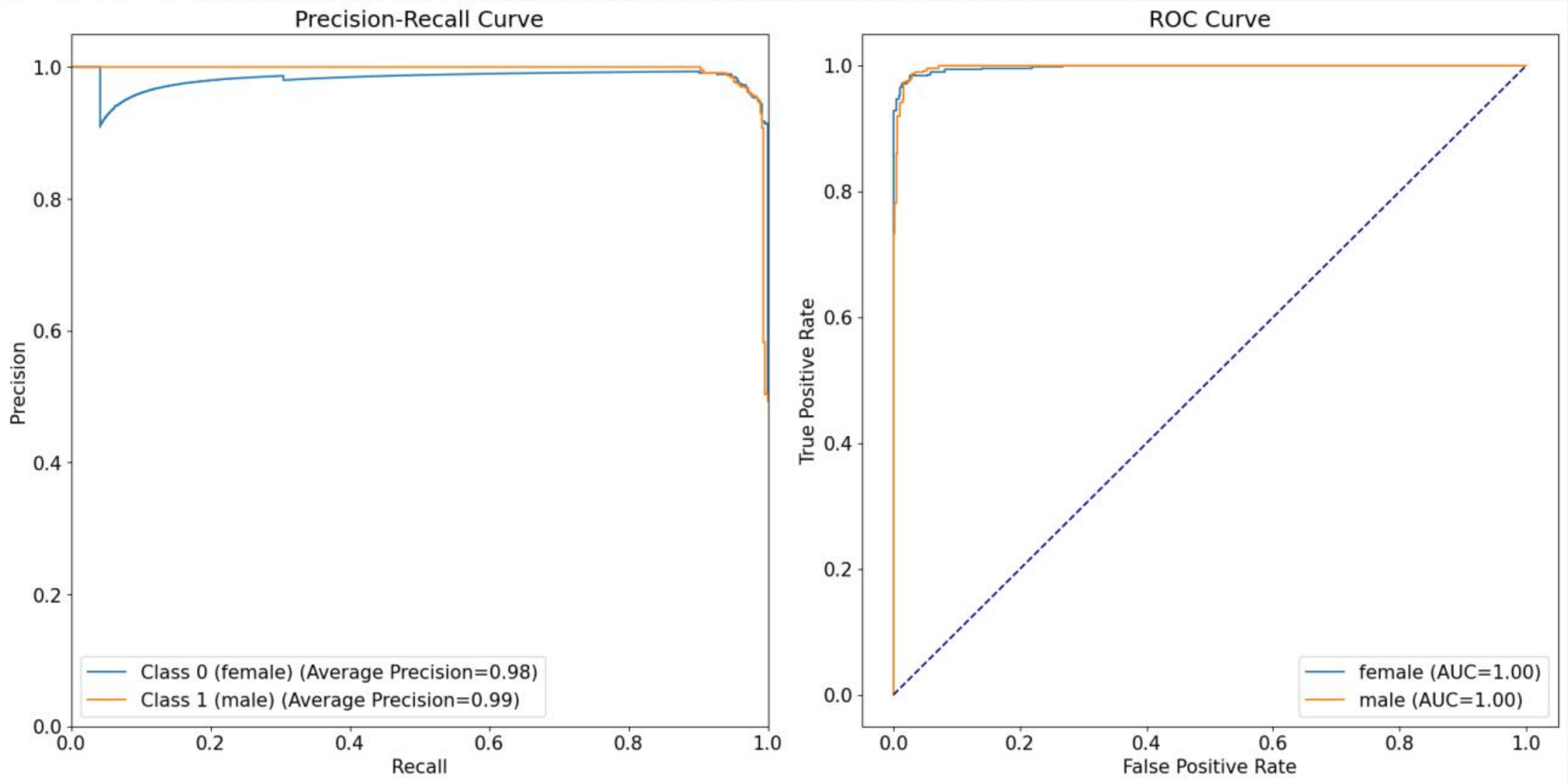
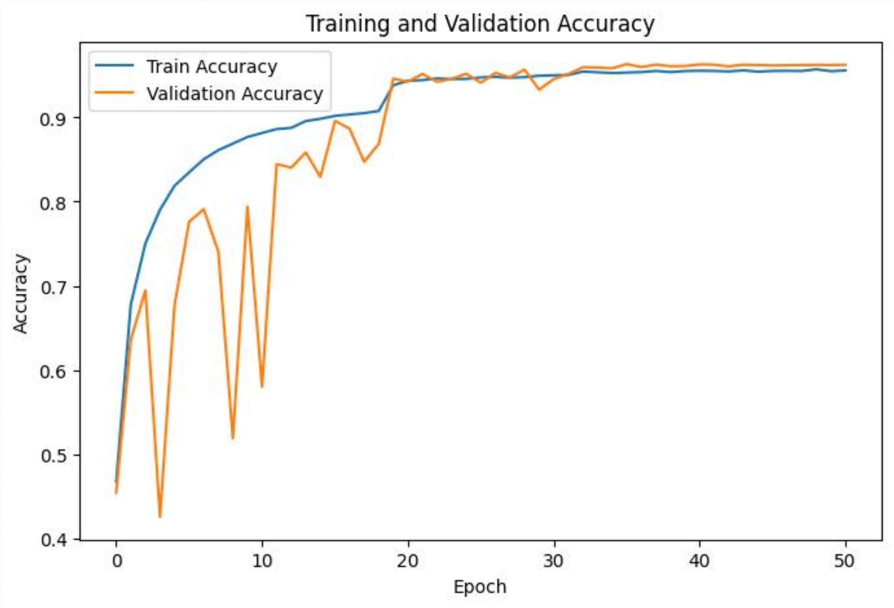
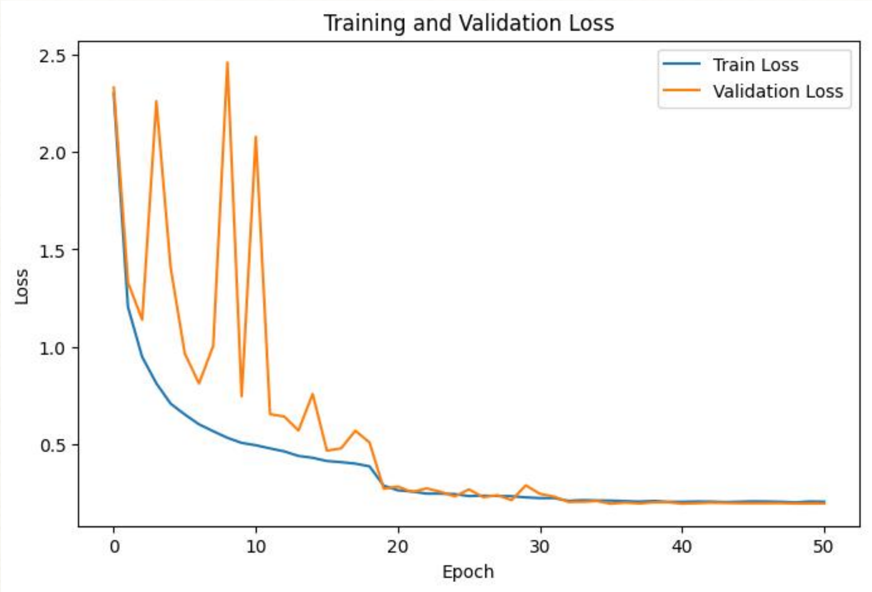


Figure 11: PRC and ROC of FranklinNet using Human face dataset

Result of Diseased Plant Leaf



12-1): Train_Acc=0.9558, Val_Acc=0.9625



12-2): Train_Loss=0.2054, Val_Loss=0.1965

Figure 12: Performance evaluation of FranklinNet using Plant disease dataset

Result of Diseased Plant Leaf

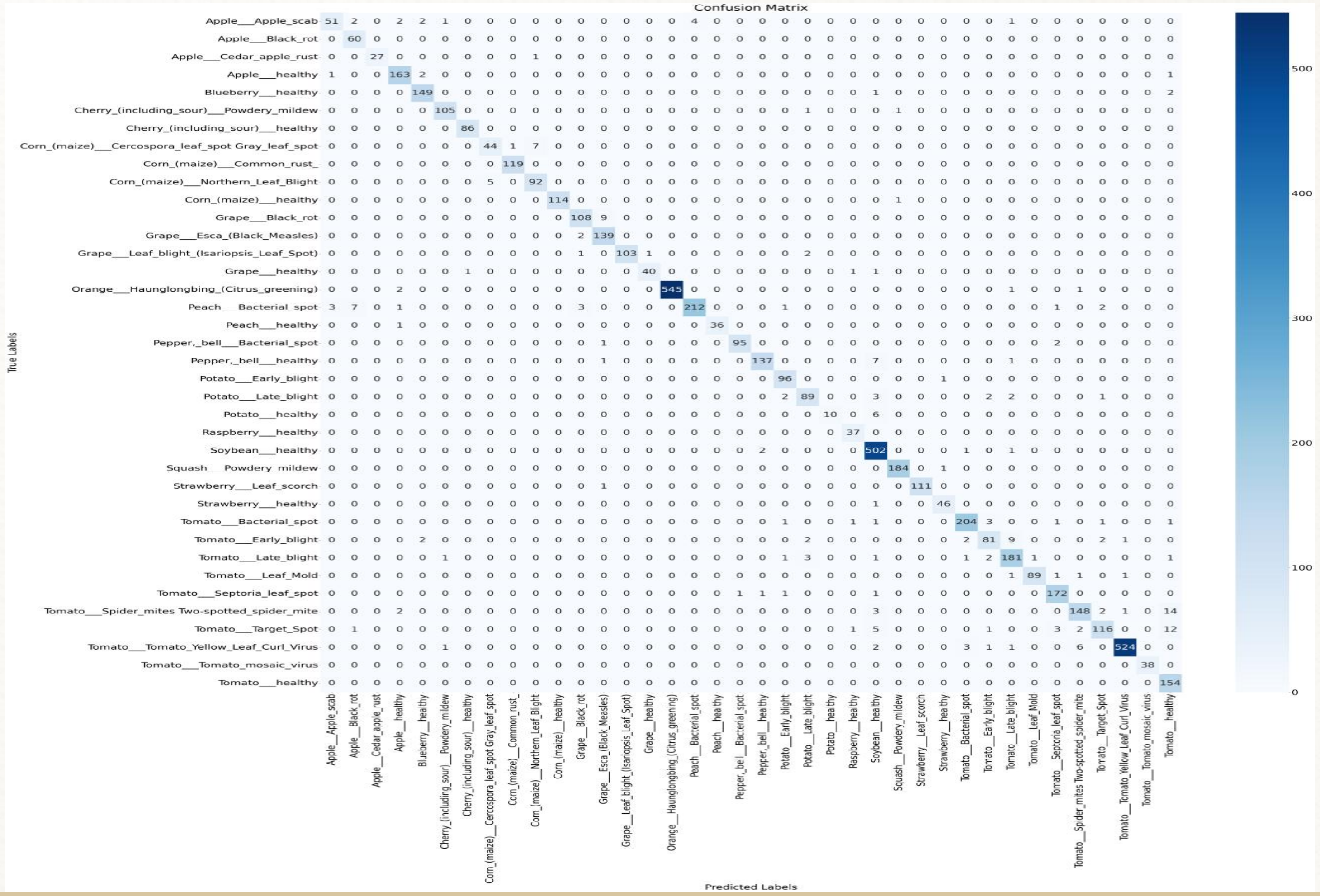


Figure 13: Confusion Matrix of FranklinNet using Plant disease dataset

Result of Diseased Plant Leaf

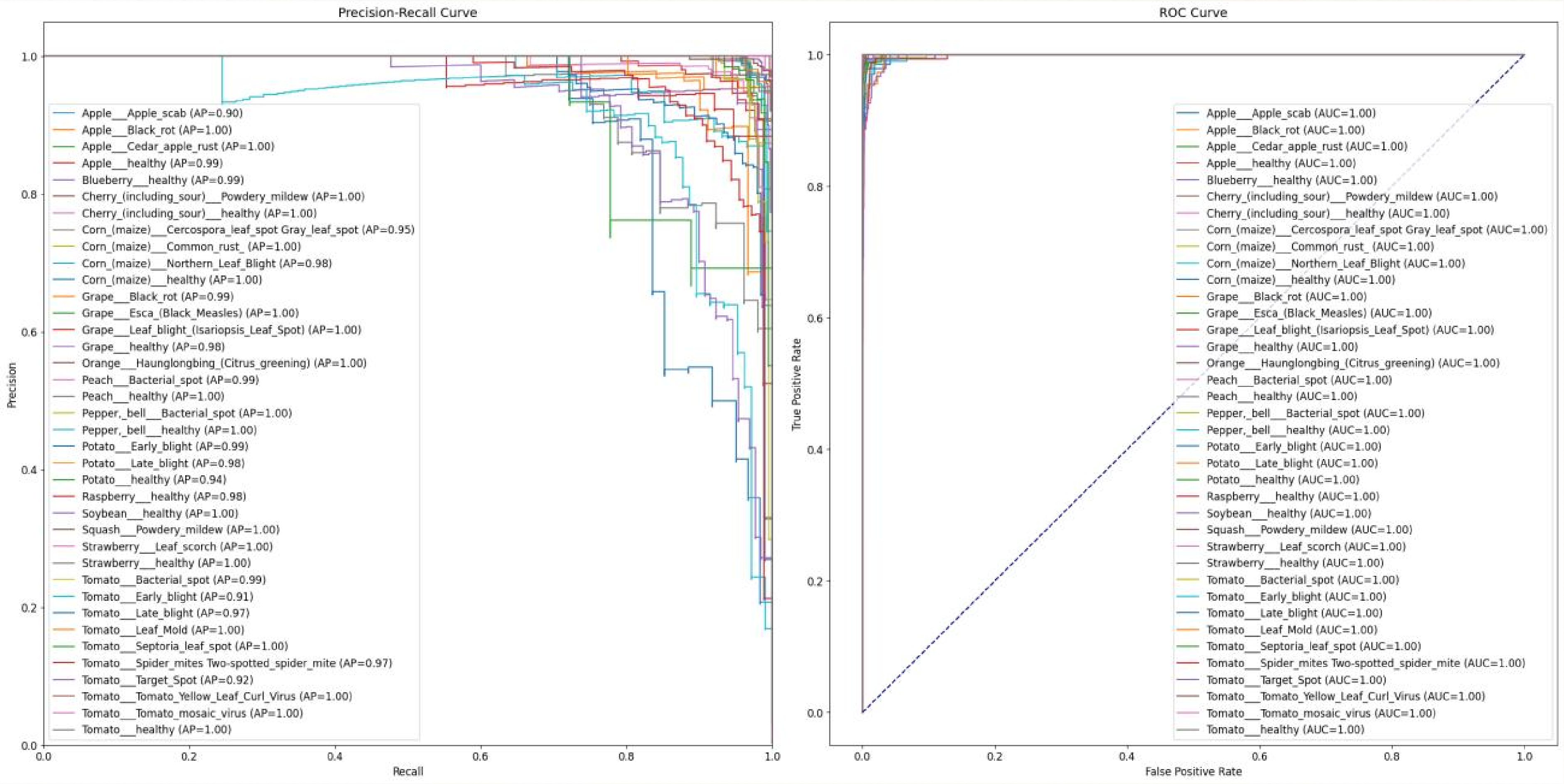
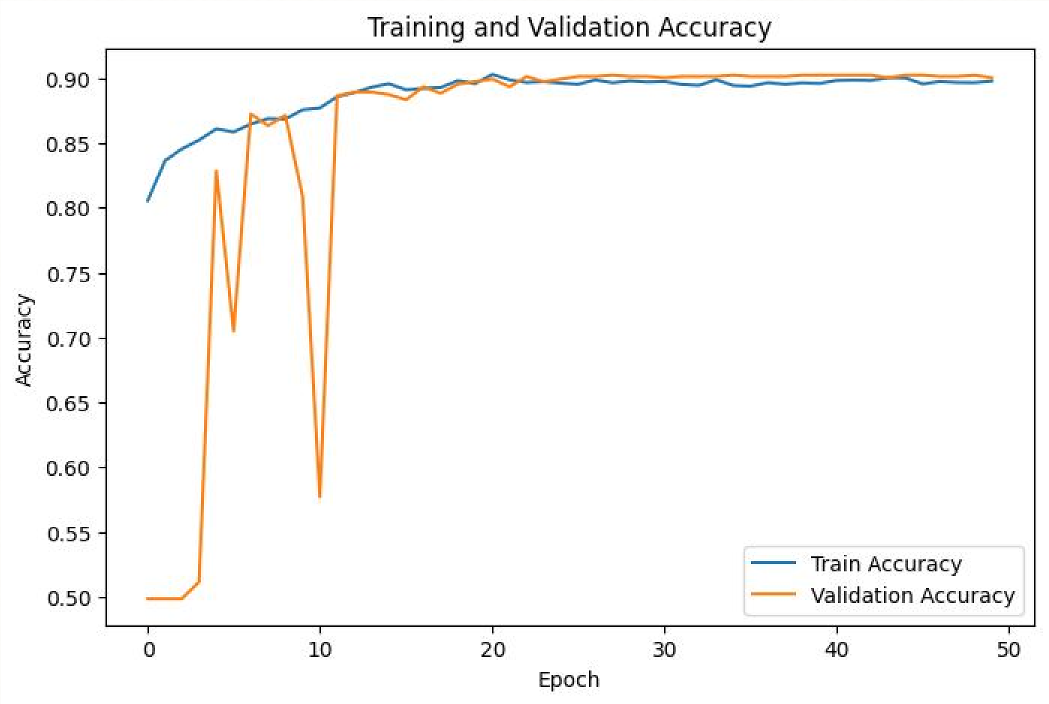
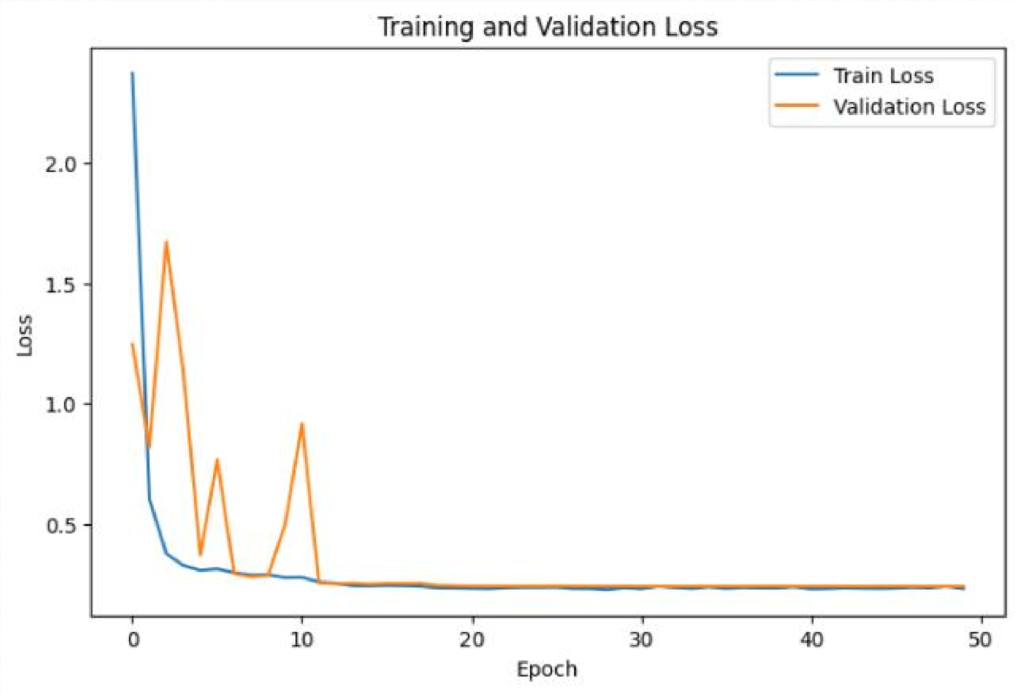


Figure 14: PRC and ROC of FranklinNet using Plant disease dataset

Result of Skin Cancer



15-1): Train_Acc=0.8977, Val_Acc=0.9002



15-2): Train_Loss=0.2335, Val_Loss=0.2425

Figure 15: Performance evaluation of FranklinNet using Skin cancer dataset

Result of Skin Cancer

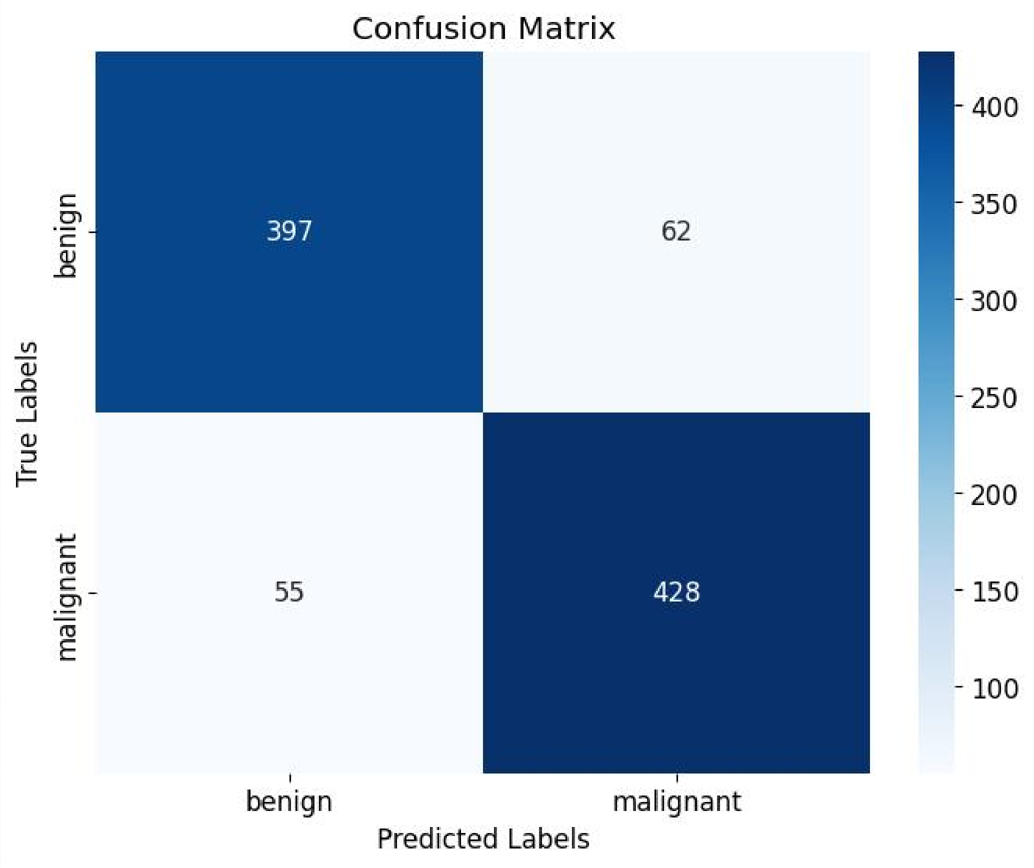


Figure 16: Confusion Matrix of FranklinNet using Skin cancer dataset

Result of Skin Cancer

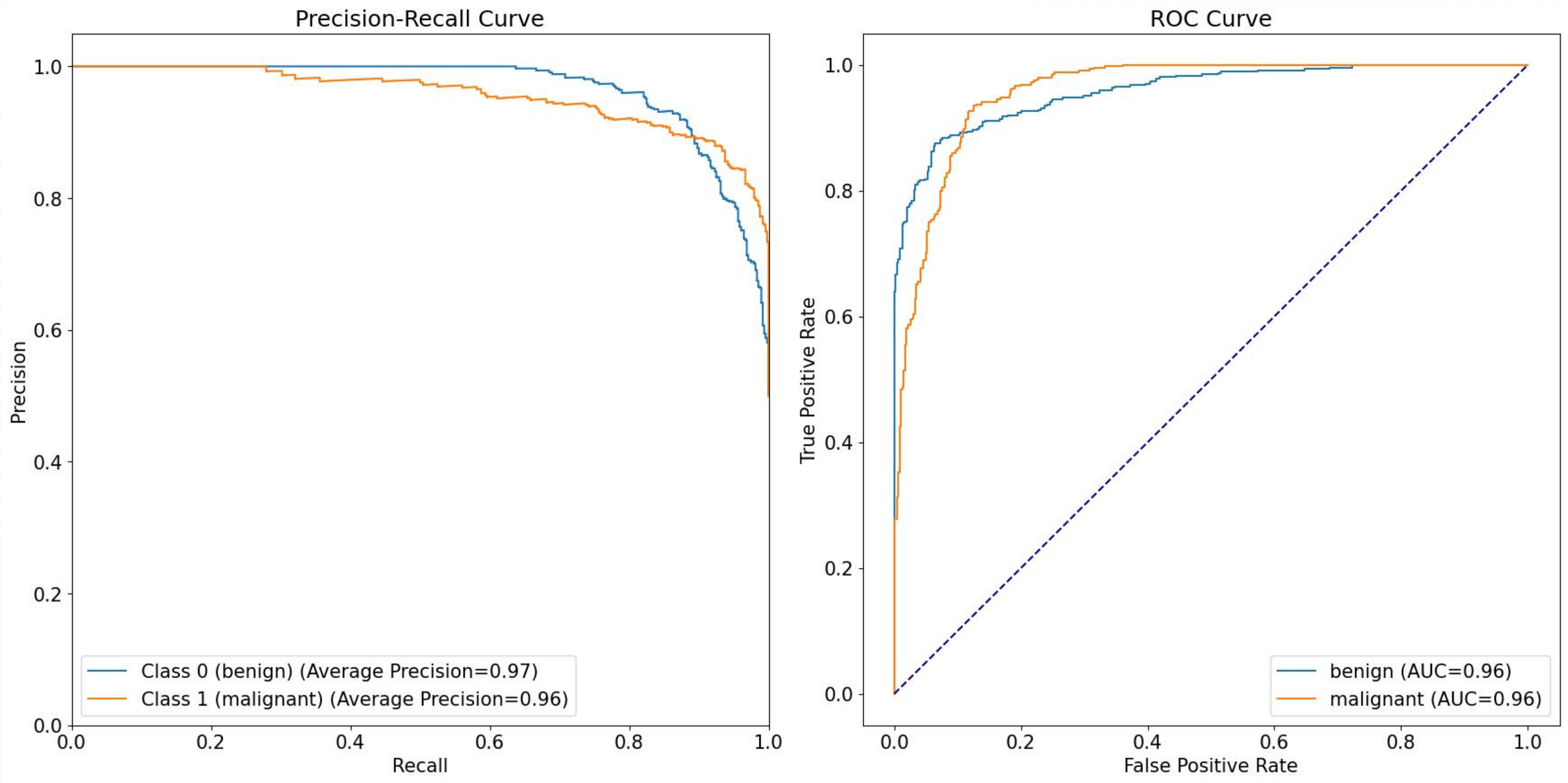


Figure 17: PRC and ROC of FranklinNet using Skin cancer dataset

Results and Analyses of the different datasets Experimented on Using FranklinNet

Performance results for FranklinNetusing the test dataset.

Dataset \ Metrics	Loss	Accuracy	Precision	Recall	Sensitivity	Specificity	F1 score	AUC
Human Face	0.05	0.98	0.98	0.98	0.98	0.98	0.98	0.99
Plant disease	0.19	0.96	0.96	0.96	0.96	0.96	0.96	0.99
Skin cancer	0.26	0.89	0.89	0.89	0.89	0.89	0.88	0.96

Table 1: Performance comparison of FranklinNet with three datasets.

07 Graphic User Interface

GUI

Figure 18: Main page

GUI

Figure 19: Choose model and upload image

GUI

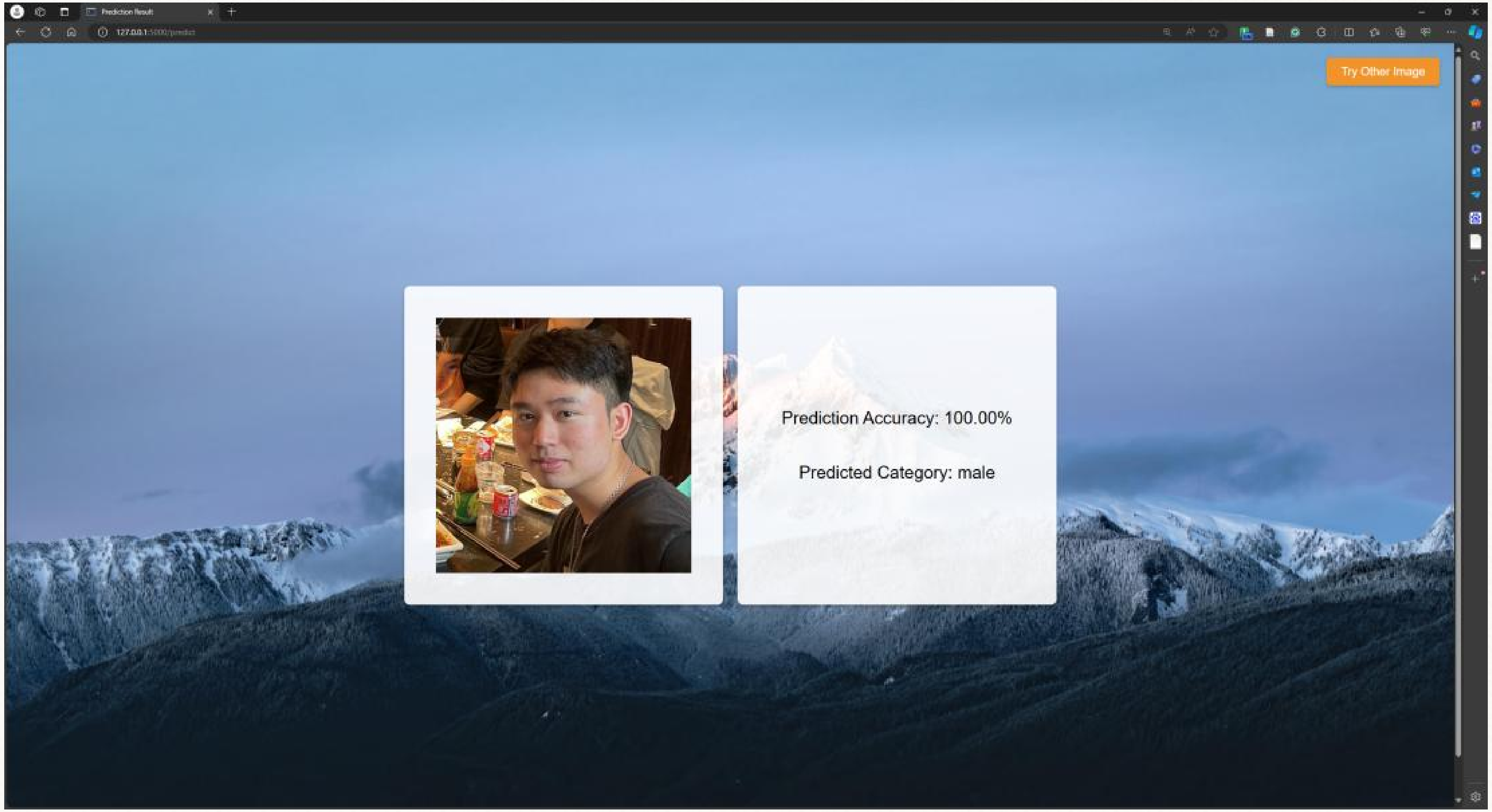


Figure 20: Prediction result page

08 Limitation and Challenges Faced

Limitation

The model may not fully adapt to the diverse application scenarios, leading to decreased performance in real-world applications.

Generalization Ability

Although various CNN architectures and attention mechanisms enhance performance, they also increase model complexity and computational demands, limiting their use in resource-constrained environments.

Computational Resources

The long-tail distribution in datasets may cause the model to overfit to common categories while underperforming on rare ones.

Sample Imbalance

Challenges

Real-time requirements:

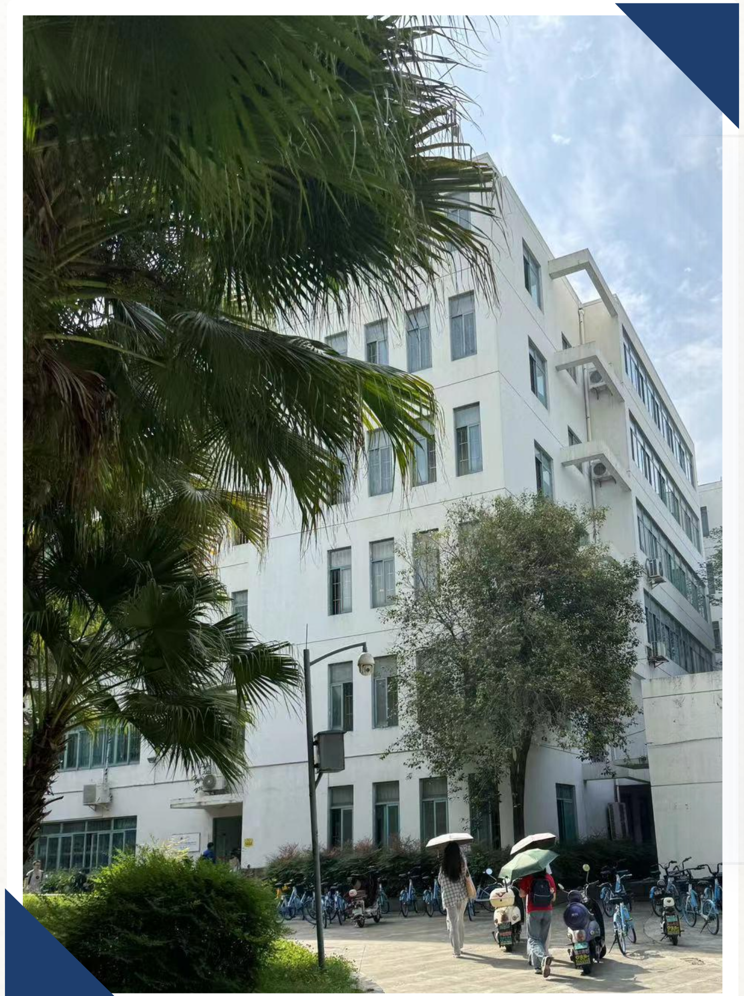
The model may not fully adapt to the diverse application scenarios, leading to decreased performance in real-world applications.

Multi-task learning and model generalization:

Although various CNN architectures and attention mechanisms enhance performance, they also increase model complexity and computational demands, limiting their use in resource-constrained environments.

09 Conclusion and Future work

Conclusion



This study demonstrates that the ensembled use of various CNN architectures and the incorporation of spatial attention mechanisms can significantly enhance the performance of image classification tasks.

Conclusion

This approach shows remarkable robustness in dealing with:

- Varying lighting conditions
- Low data resources
- Fluctuating image quality
- Low performance of the recognition model

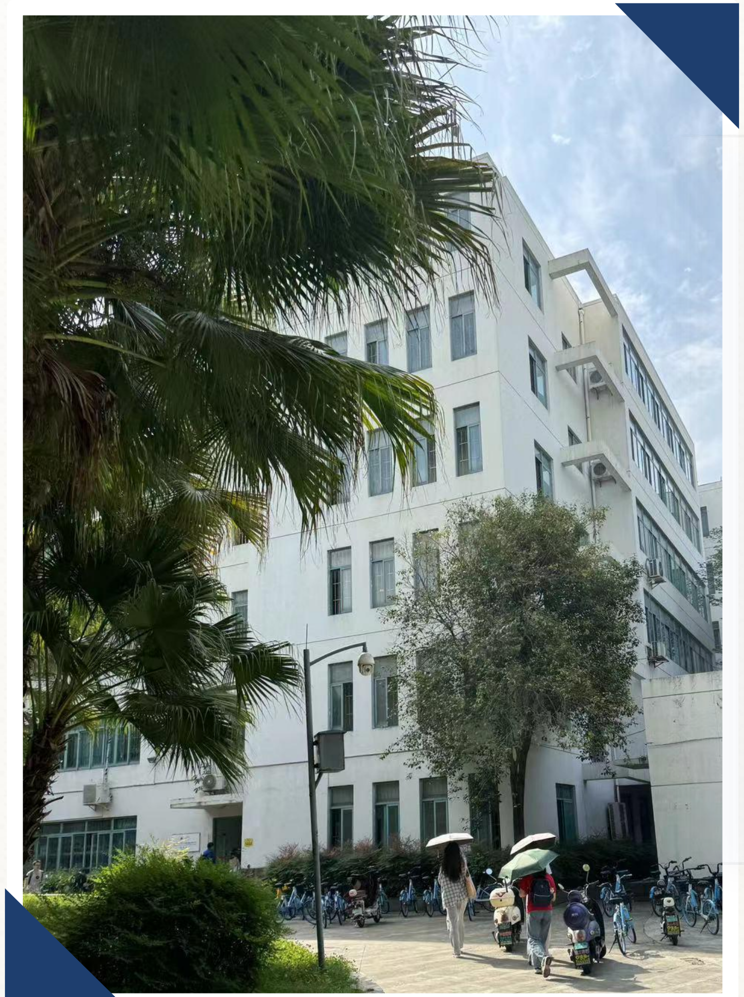
It provides strong support for applications in:

- Security
- Human-computer interaction
- Social media analysis
- Medical Judgment Assistance

However, despite these achievements, challenges remain in terms of:

- Data diversity
- Model complexity
- Interpretability
- Domain adaptability

Future work



Enhancing dataset diversity and coverage:

Explore the use of Generative Adversarial Networks (GANs) to create more varied training data, improving model performance in rare and extreme cases.

Model optimization and lightweighting:

Investigate more efficient model architectures and parameter compression techniques to reduce computational demands and enhance deployability on resource-constrained devices.

Reference

[1] ‘Gender Classification Dataset’. Accessed: Dec. 14, 2023. [Online]. Available: https://www.kaggle.com/datasets/gpiosenka/gender-classification-from-an-image

[2] ‘PlantVillage Dataset’. Accessed: Apr. 07, 2024. [Online]. Available: https://www.kaggle.com/datasets/abdallahalidev/plantvillage-dataset

[3] ‘Skin Cancer: Malignant vs. Benign’. Accessed: Apr. 07, 2024. [Online]. Available: https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign/data

The End

Thank you for Listening!!

Q & A