OXFORD BROOKES UNIVERSITY

Food Classification using Convolutional Neural Network

Department of Computing Science,
Chengdu University of Technology, Oxford Brookes University

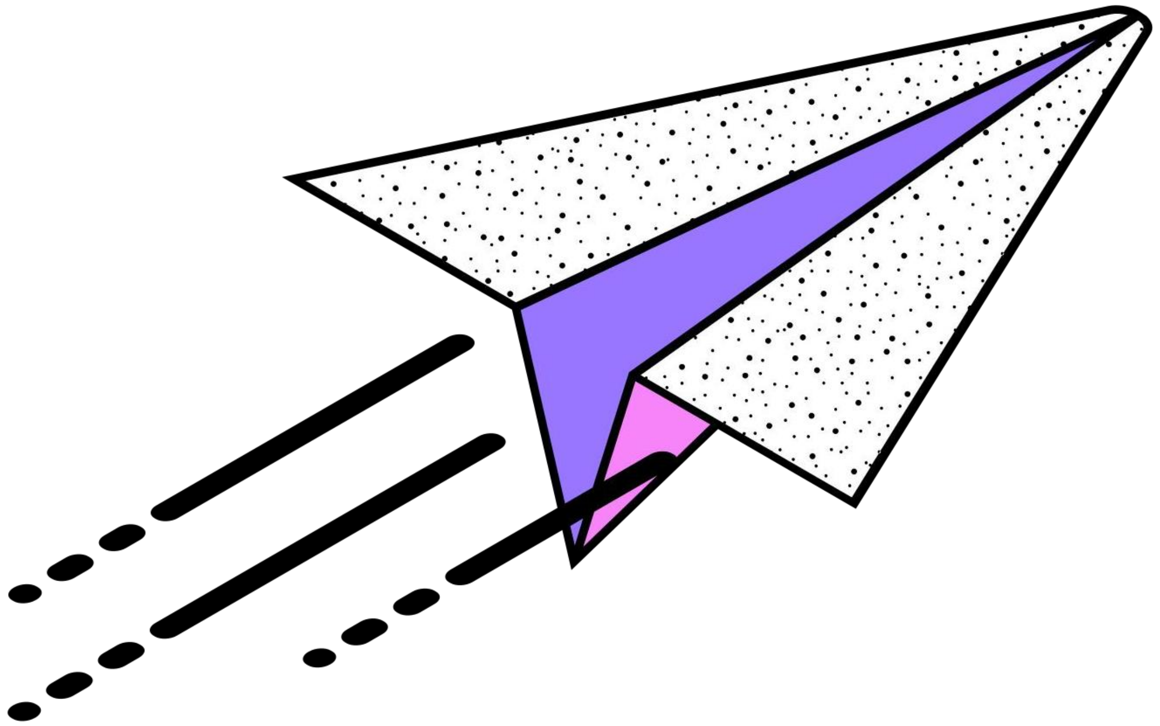
Emil
202018010421
Supervised by Dr. Happy Nkanta Monday

Date: June 4, 2024

OXFORD BROOKES UNIVERSITY

CONTENT

Food Classification using Convolutional Neural Network



OXFORD BROOKES UNIVERSITY

01 Background

Importances:

- Dietary Management
  - Tracking nutritional intake
  - Managing food-related health issues
- Food Safety
  - Ensures food quality and safety through non-destructive evaluation methods [1]

Challenges:

- Traditional food assessment techniques are time-consuming and laborious [1]

- High intra-class difference and inter-class similarity [2]

01 Background

Solution - Using CNN

- Advanced architectures: ResNet, MobileNetV2, and InceptionV3
- Techniques: transfer learning and ensemble learning, etc

Aim:

- Accurate food image recognition system
- A user-friendly GUI

OXFORD BROOKES UNIVERSITY

01 Background

Objectives:

- Evaluate the performance
- Combine multiple models
- Evaluate the model and demonstrate its interpretability
- Create a GUI for food image classification

Audiences:

- Individuals Seeking Healthier Eating Habits
- Organizations in the Agriculture and Food Supply Chain
- Scientific and Academic Communities
- Government and Food Regulatory Authorities

OXFORD BROOKES UNIVERSITY

02 Dataset

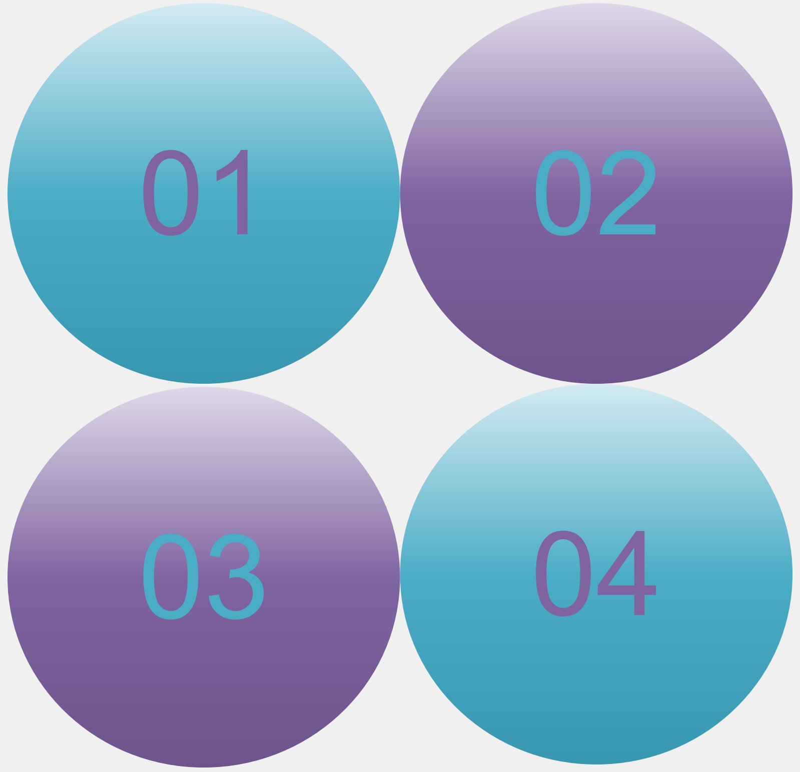
Dataset

- Food-101 Subset:
  - 5000 images across 5 categories
  - 1000 images per category

Figure 1: Visualize Images of The Dataset

Resizing & Normalization

- Training Input Size: 299x299x3
- Testing Input Size: Resize to 320, CenterCrop to 299x299
- Normalization:
  - Means: [0.485, 0.456, 0.406]
  - Stds: [0.229, 0.224, 0.225]



Data Splitting

- 80% for training, 20% for testing and validation

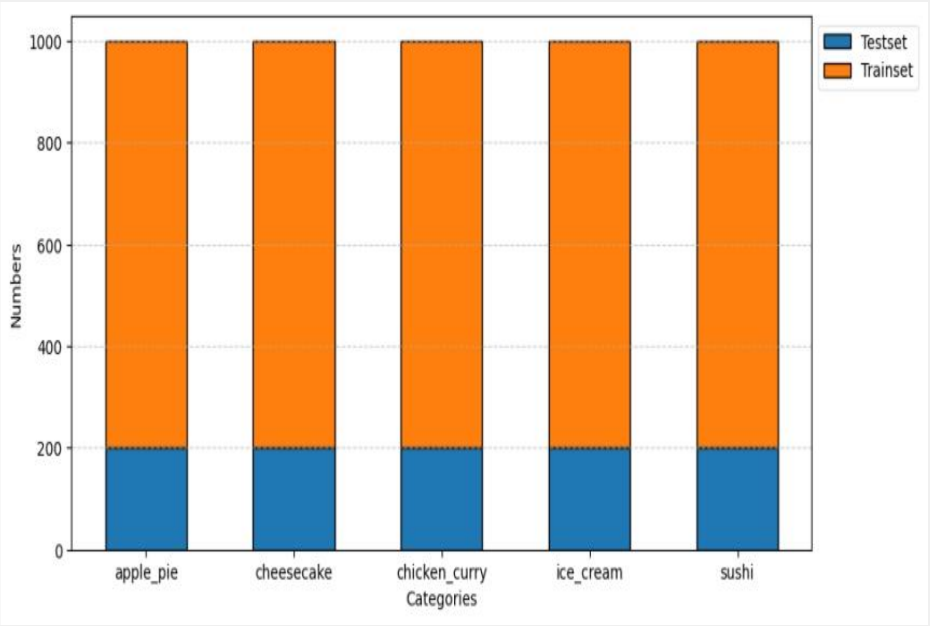


Figure 2: Number of Testsets and Trainsets in Each Category

Data Augmentation:

- RandomHorizontalFlip
- RandomRotation: ±30°
- ColorJitter

Figure 3: Preprocessed Images and Original Images

03 Methodology

Transfer Learning:

- Using Pre-trained Models
  - ResNet18, MobileNetV2, InceptionV3
  - Adapt models pre-trained to the Food-101 subset
  - Evaluate models

Ensemble Learing:

- Strategy - Weighted Averaging
  - Calculating weights for each model
  - Integrate models based on weights
  - Assess the combined performance of the ensemble model

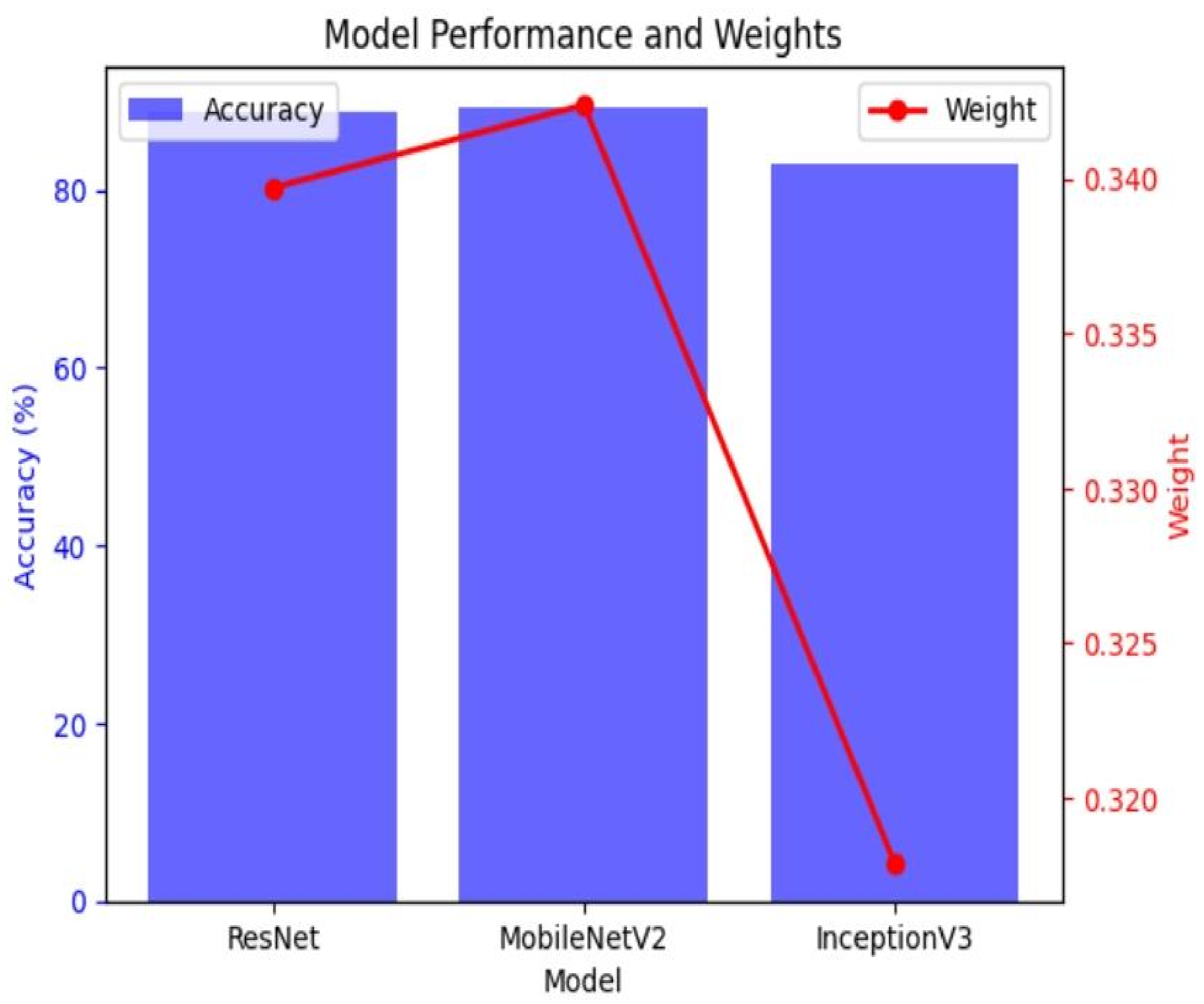


Figure 4: Model Performance and Weights

03 Methodology

Ensemble Model Details:

- Does not have a unified network hierarchy
- Combine the outputs of ResNet, MobileNetV2, and InceptionV3
- Weighted averaging based on the performance metrics of each individual model
- The integrated features are processed through the output layer

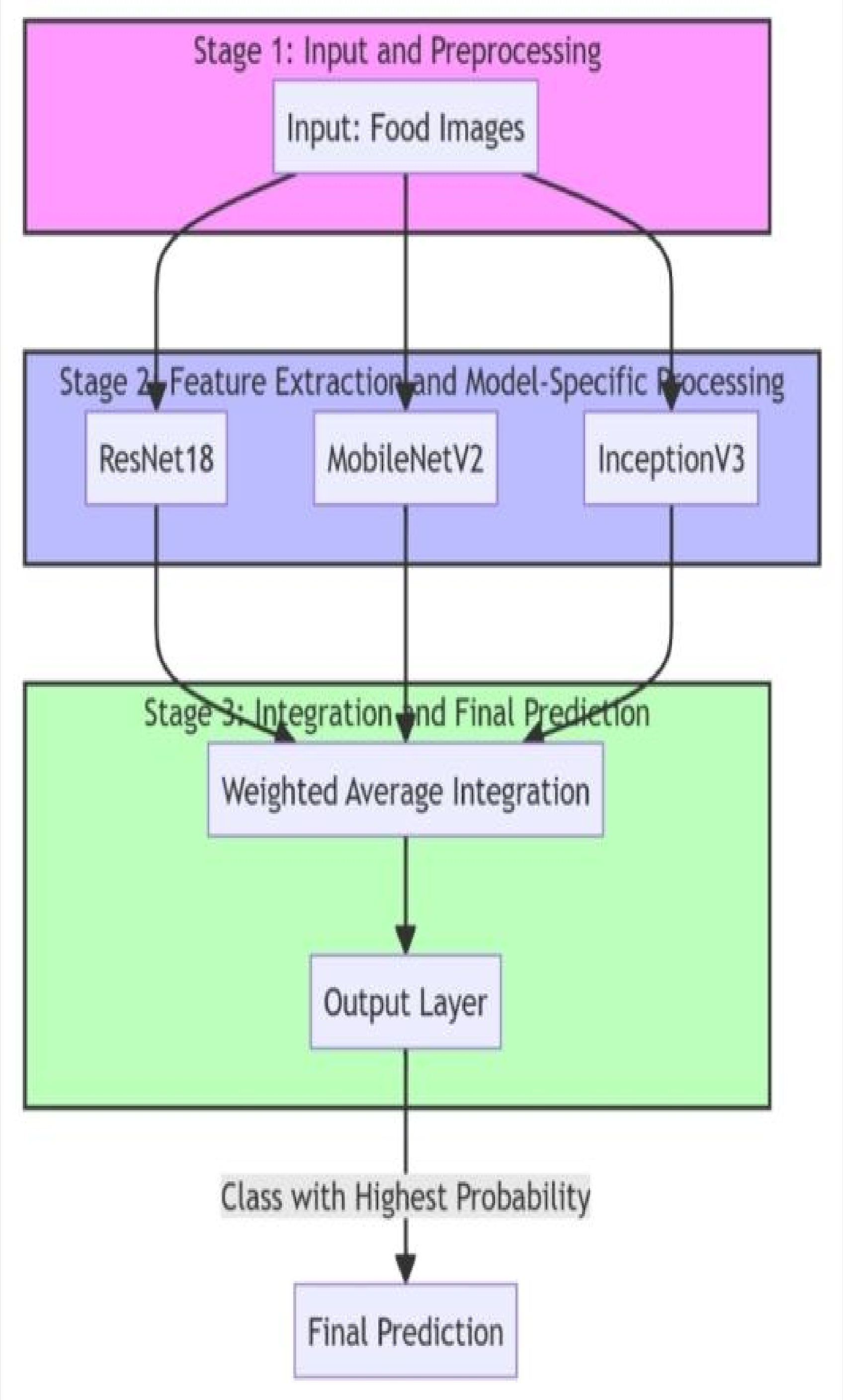


Figure 5: Architecture Structure

OXFORD BROOKES UNIVERSITY

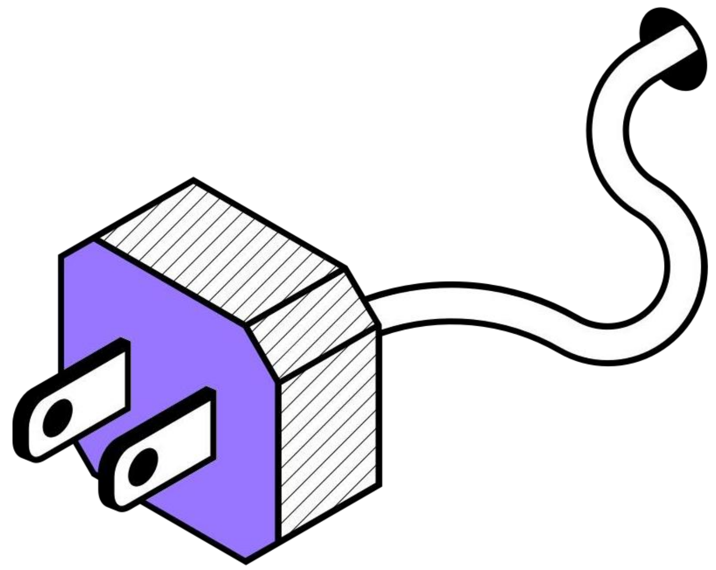
03 Methodology

Early Stopping:

- Monitor validation loss during training
- Stop training when validation loss does not improve
- Enhances generalization and reduces overfitting

Learning Rate Scheduler:

- Dynamically adjust the learning rate
- Improves convergence speed and training efficiency
- Improves overall model performance

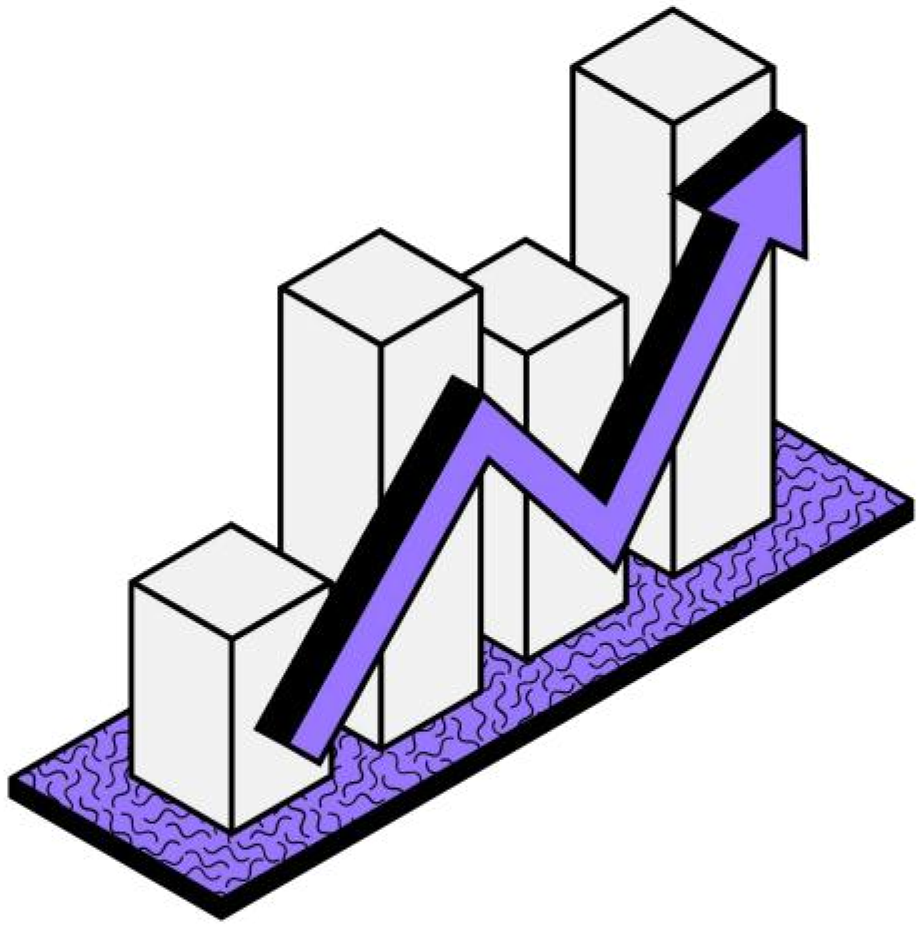


03 Methodology

Tools:

Category	Tool/Resource
Platform	Google Colab
Hardware	NVIDIA Tesla T4 GPU
	Intel(R) Xeon(R) CPU @ 2.20GHz
Software/Frameworks	Python (Programming Language)
	PyTorch (Deep Learning Framework) Version:2.2.1+cu121
	scikit-learn (Machine Learning Library)
Deep Learning Models	ResNet, MobileNetV2, InceptionV3
Data Augmentation	RandAugment (Data Augmentation Technique)
	ImgAug (Image Augmentation Library)
Datasets	Food-101
Preprocessing Tools	torchvision.transforms (preprocessing images)

Table 1: The Technologies of The Project



OXFORD BROOKES UNIVERSITY

04 Experiments & Results

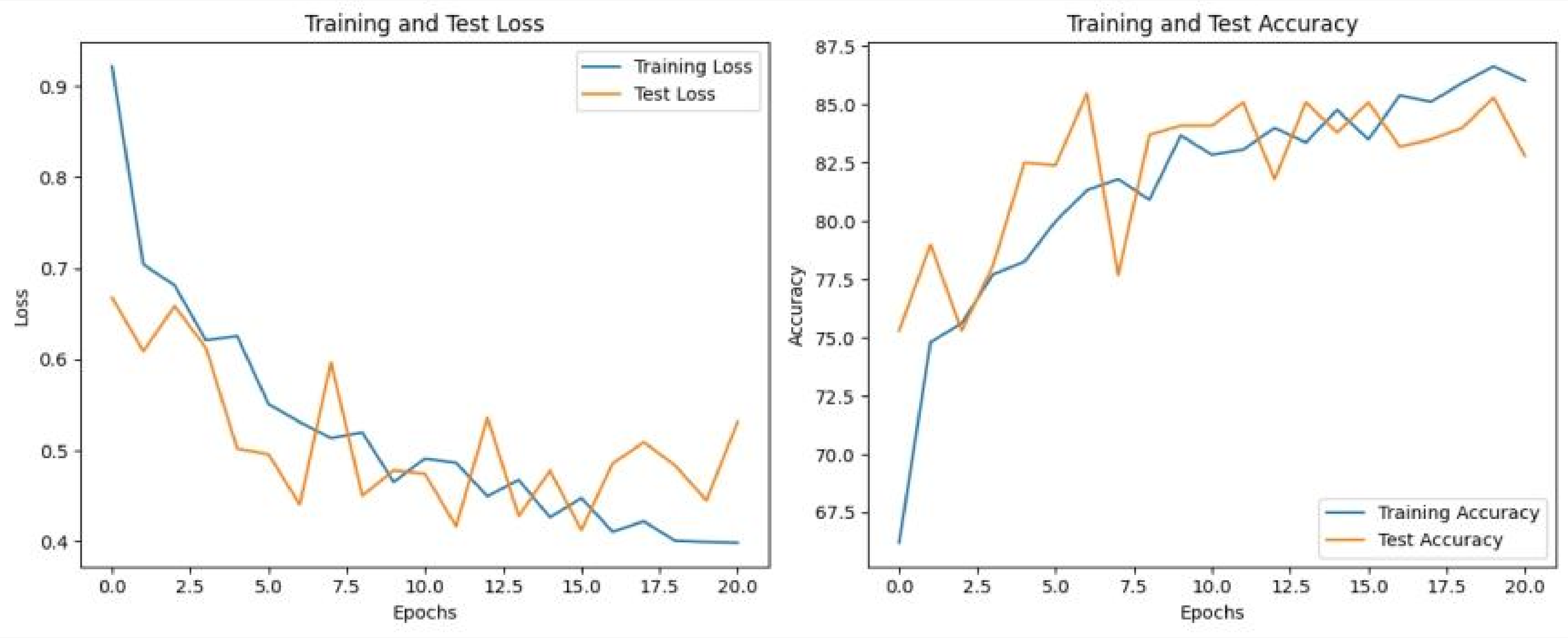


Figure 6: Loss and Accuracy of MobileNetV2

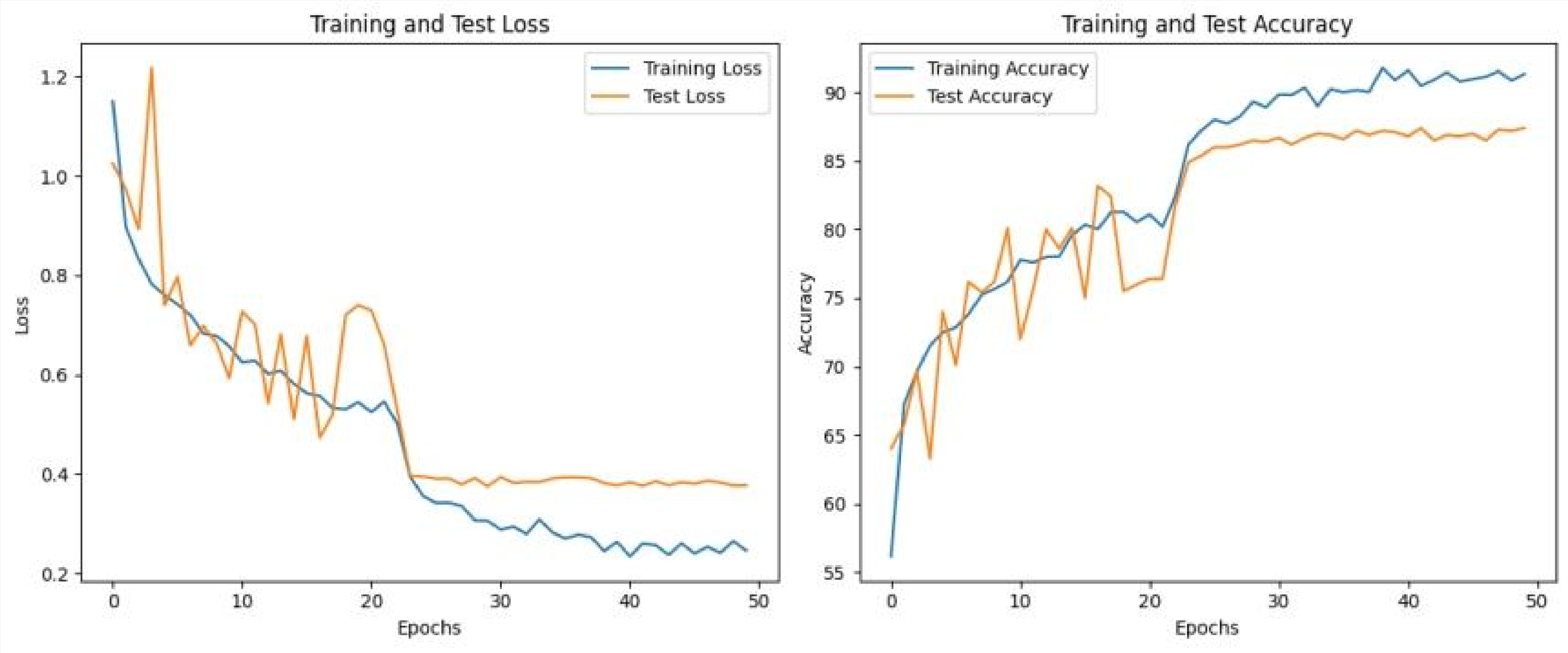


Figure 7: Loss and Accuracy of ResNet

Figure 8: Loss and Accuracy of InceptionV3

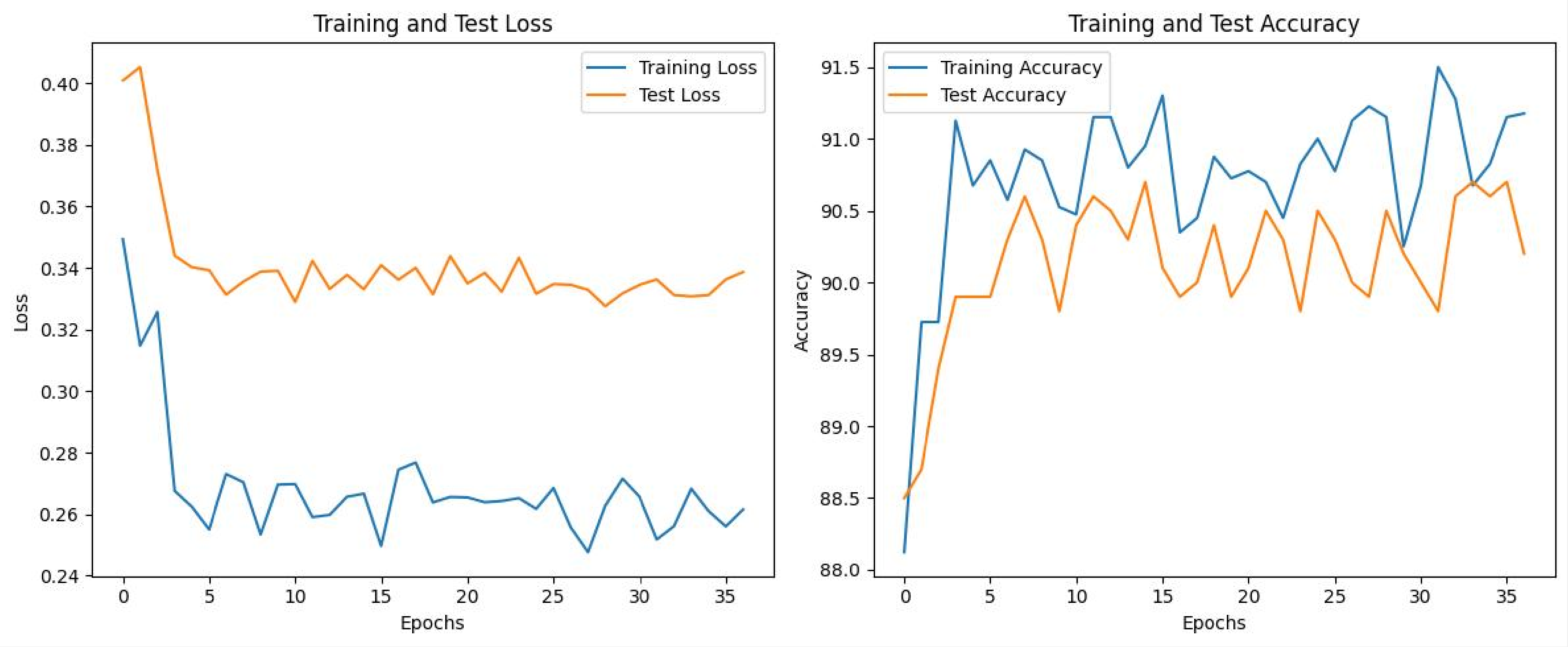


Figure 9: Loss and Accuracy of Ensemble Model

04 Experiments & Results

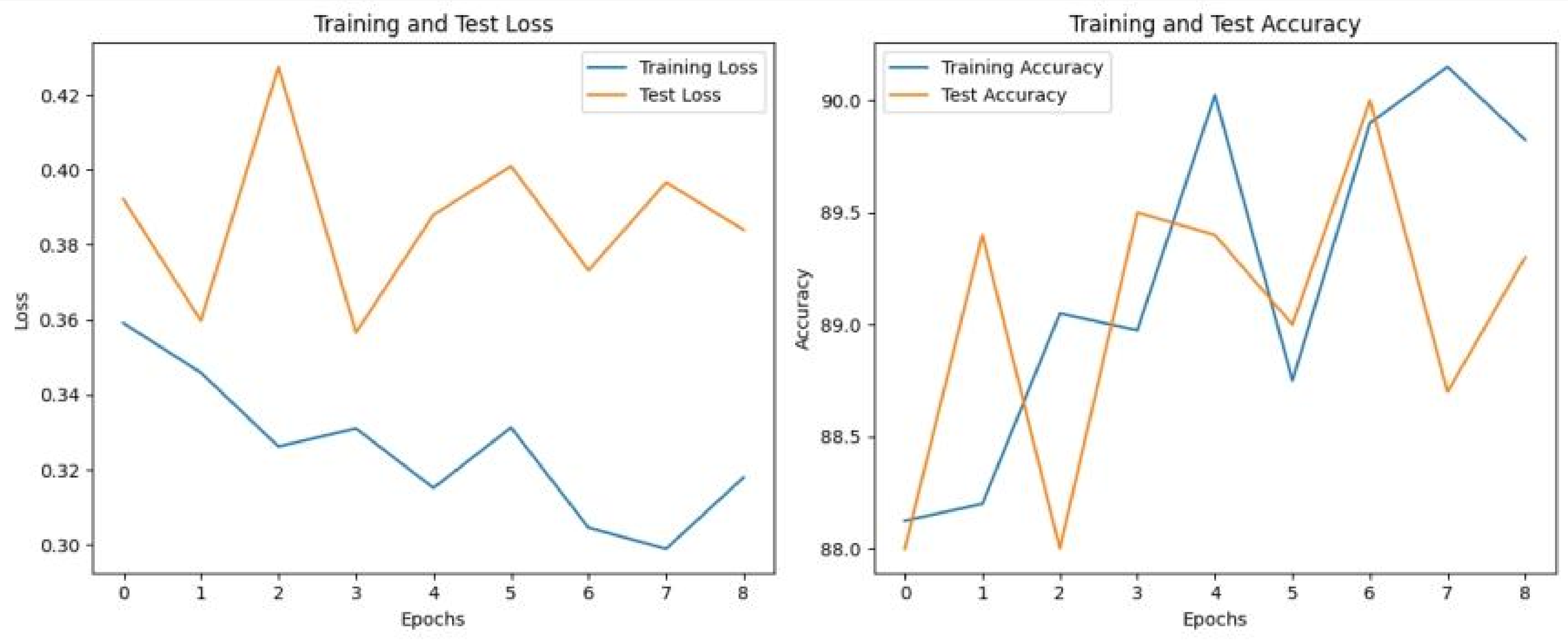


Figure 10: Loss and Accuracy of 0.01 Learning Rate

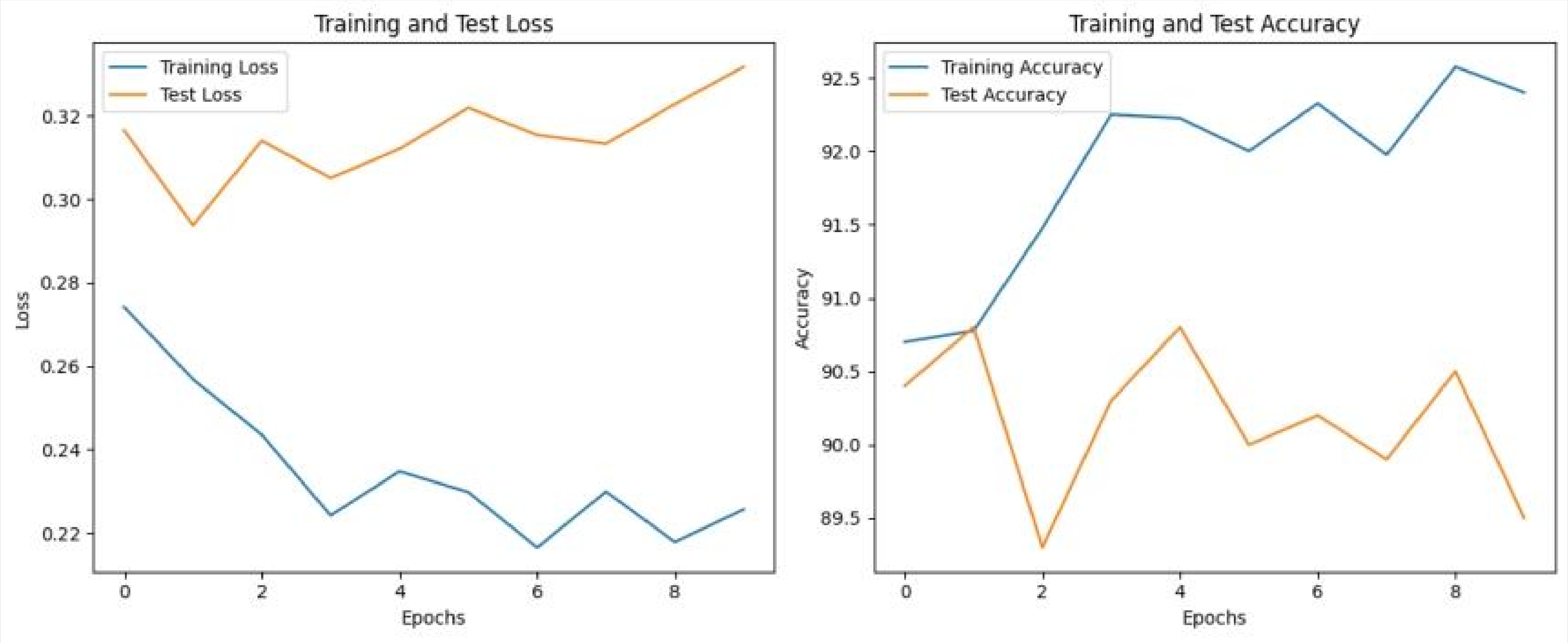


Figure 11: Loss and Accuracy of 0.001 Learning Rate

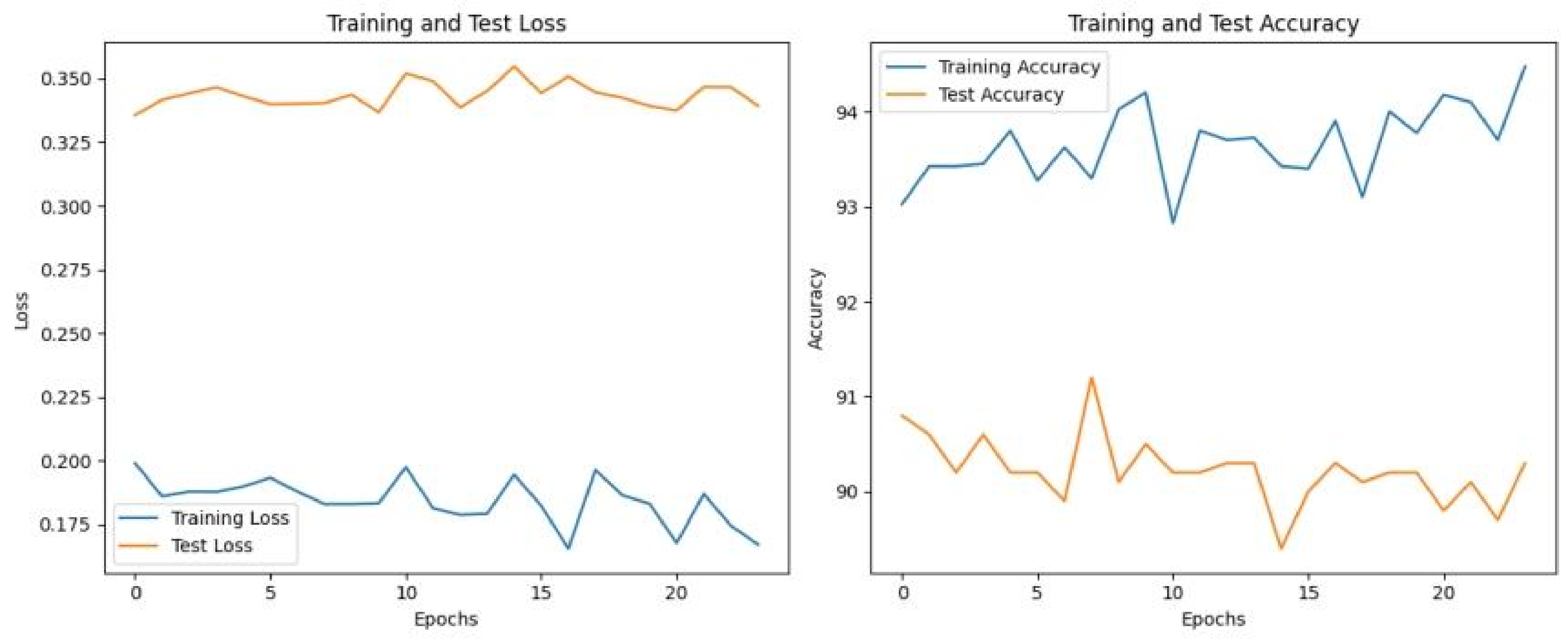


Figure 12: Loss and Accuracy of 0.0001 Learning Rate

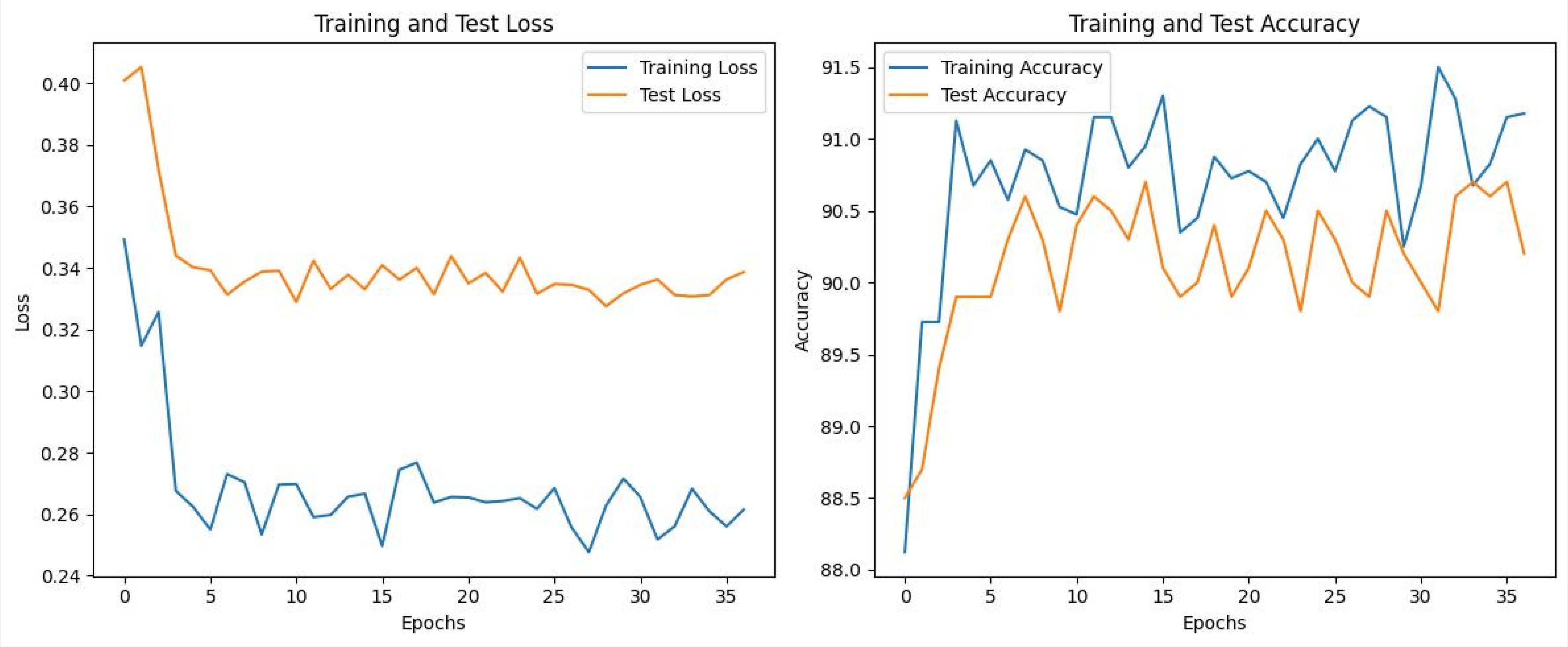


Figure 13: Loss and Accuracy of the Learning Rate Scheduler

04 Experiments & Results

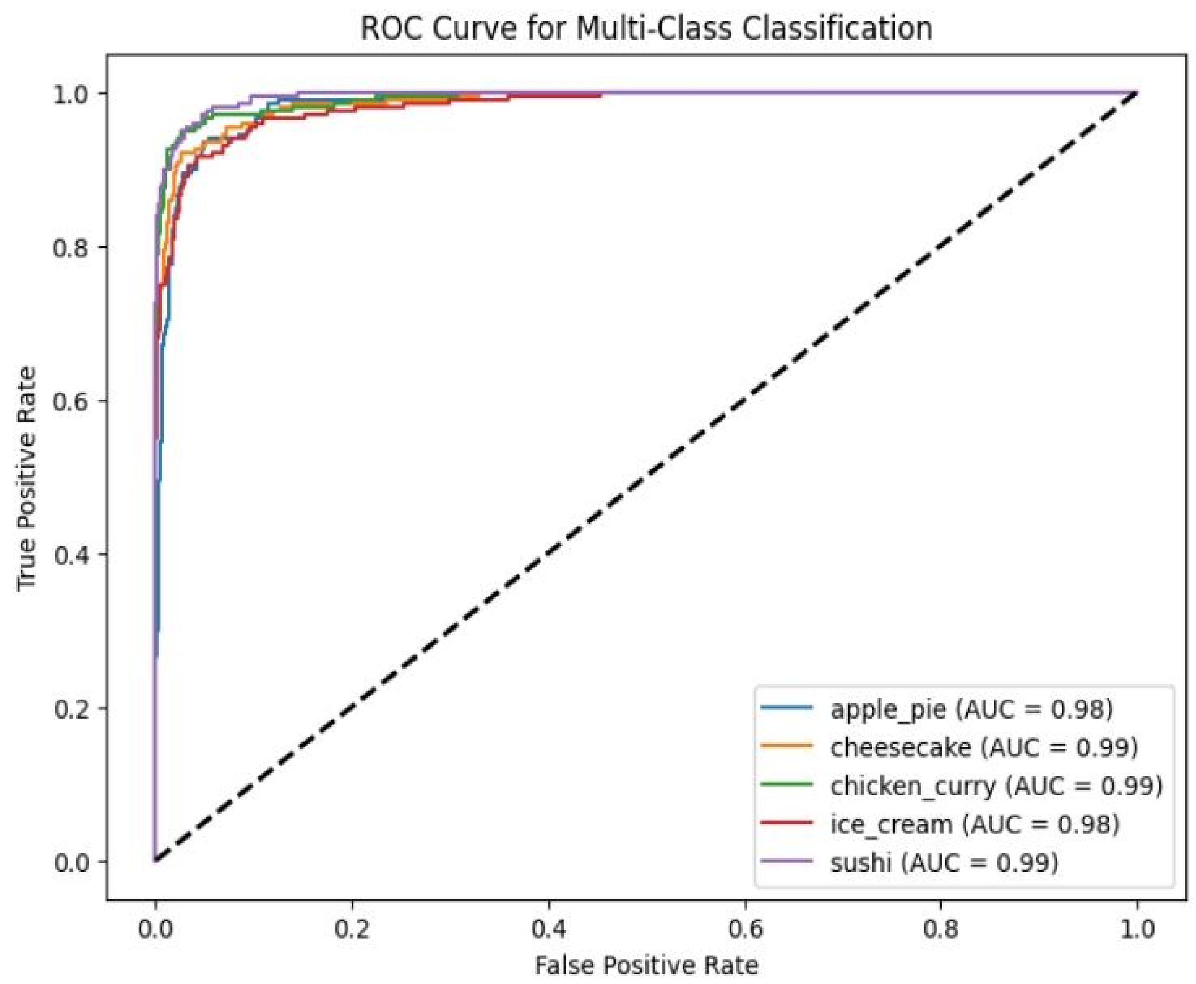


Figure 14: ROC Curve Diagram

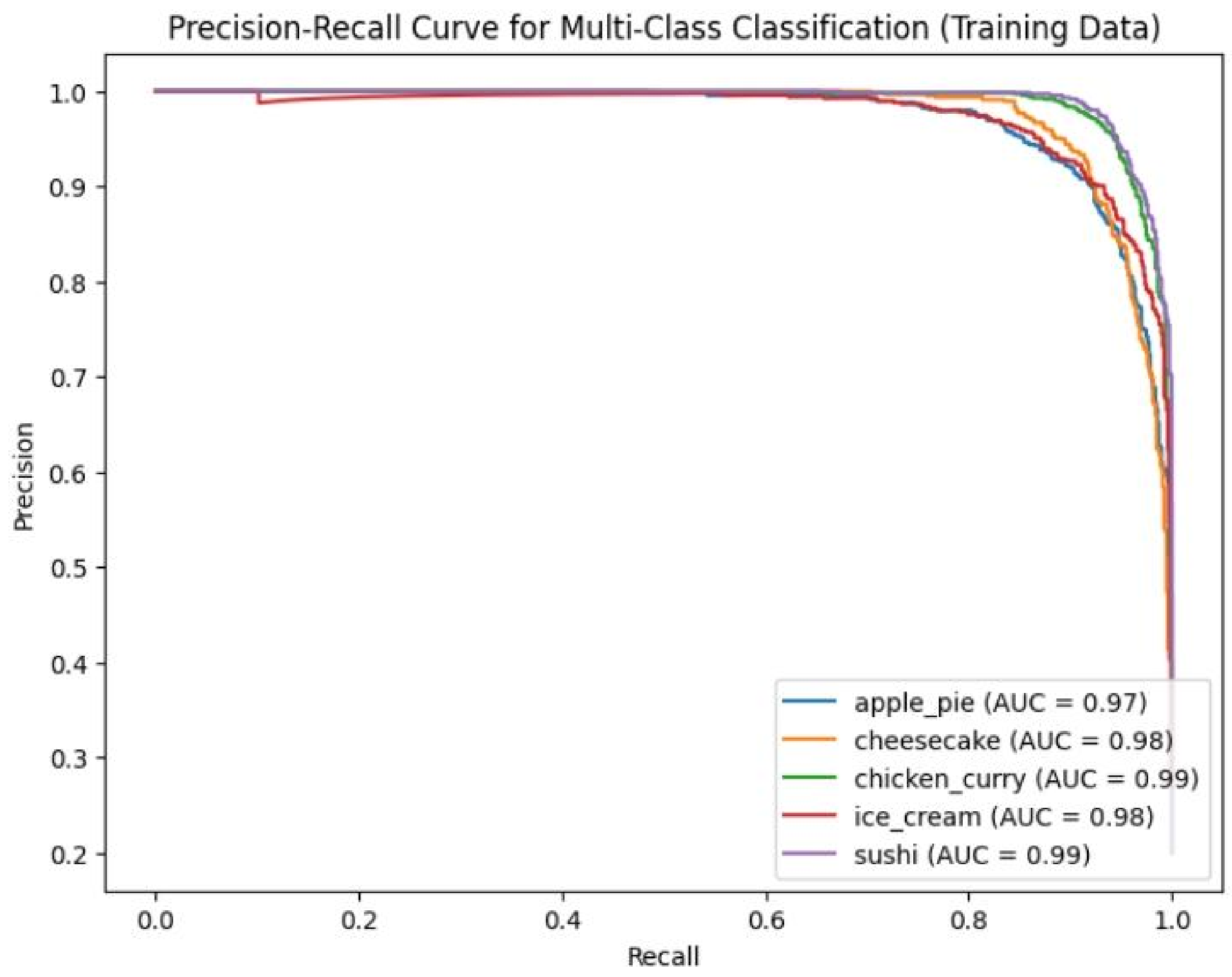


Figure 15: Precision-Recall Curve Diagram

- The AUC values of all categories in the ROC curve and Precision Recall curve diagram are close to 1.0
- The integrated model has excellent classification performance and strong predictive ability.

OXFORD BROOKES UNIVERSITY

04 Experiments & Results

Class	Acc	Pre	Sen	Spe	F1	mAP
apple_pie	0.9520	0.8585	0.9100	0.9625	0.8835	0.9423
cheesecake	0.9600	0.9124	0.8850	0.9788	0.8985	0.9623
chicken_curry	0.9740	0.9780	0.8900	0.9950	0.9319	0.9751
ice_cream	0.9530	0.9005	0.8600	0.9762	0.8798	0.9503
sushi	0.9730	0.9303	0.9350	0.9825	0.9327	0.9830

Table 2: Evaluation Indicators of Single Class

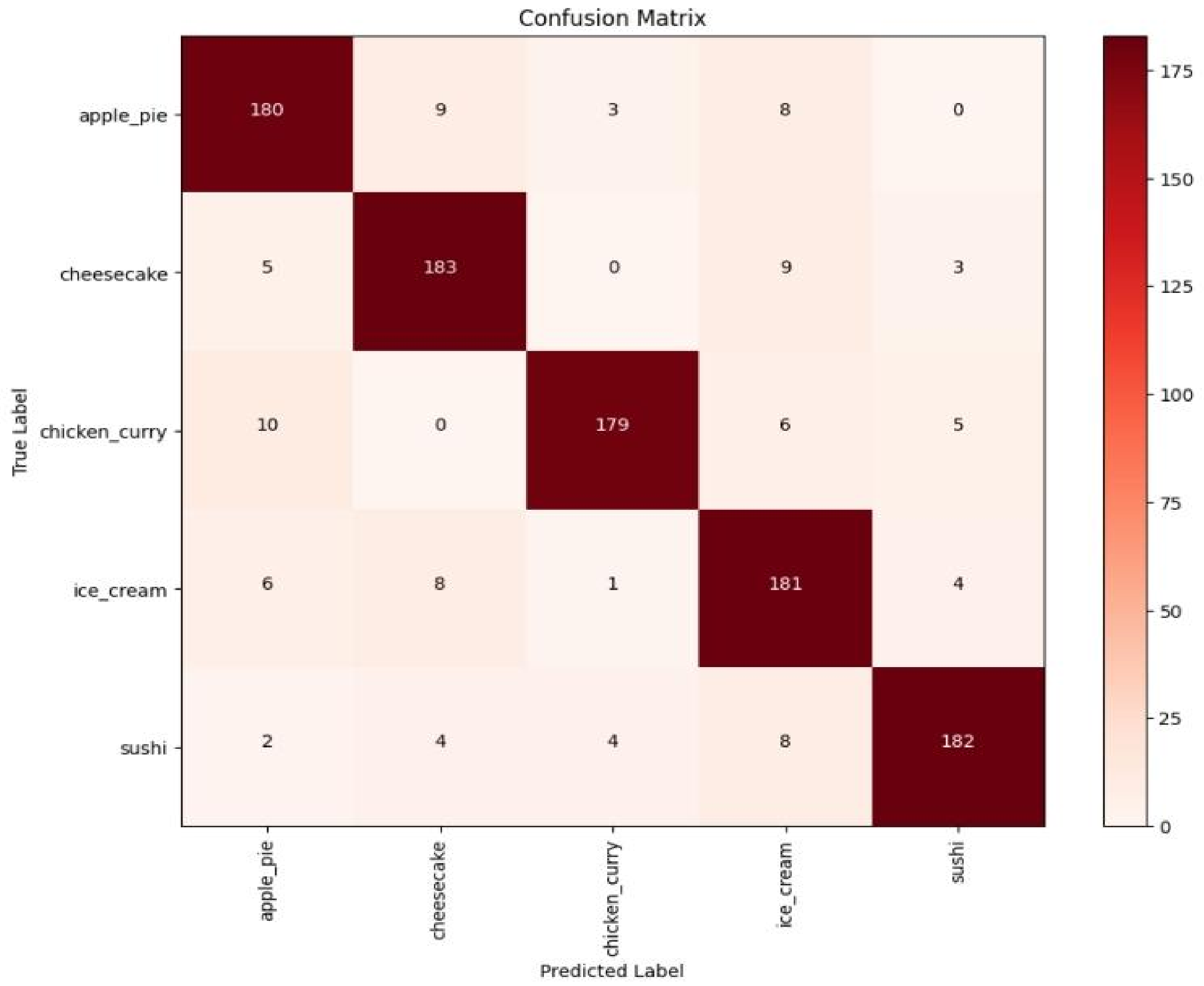


Figure 16: Confusion Matrix

- The confusion matrix shows a high degree of diagonal concentration
- The evaluation table showcases strong performance across all classes

04 Experiments & Results

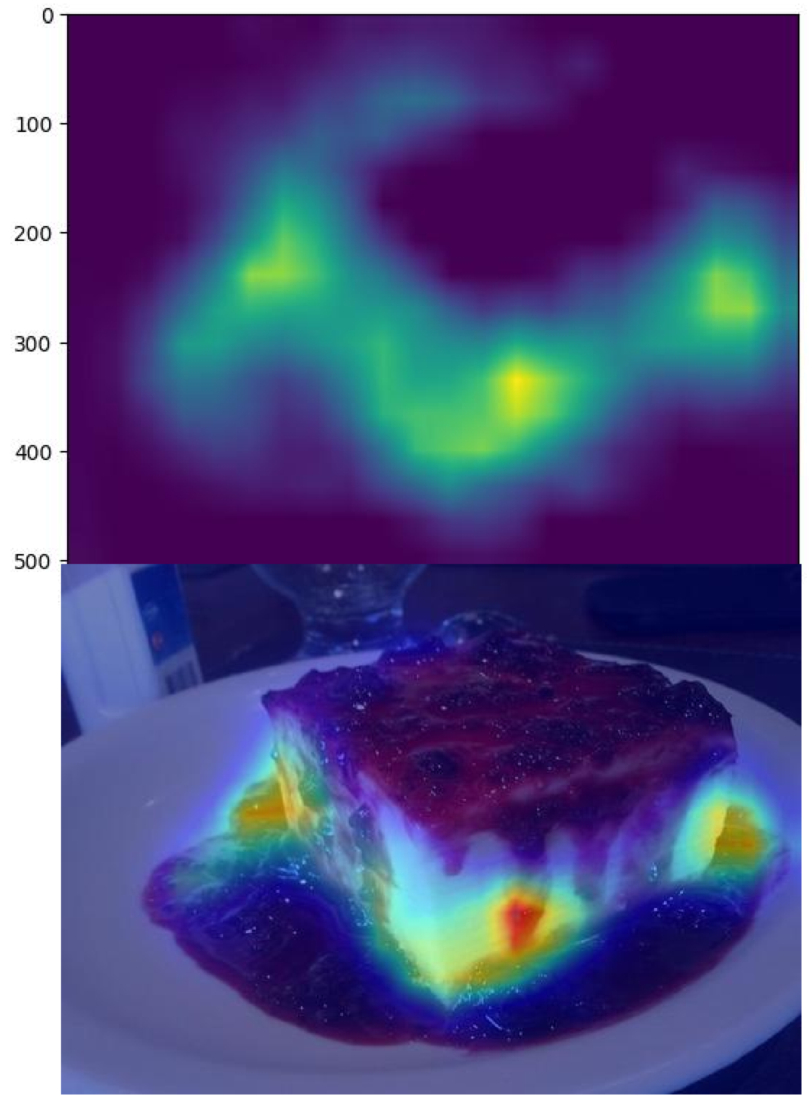


Figure 17: Grad-CAM Heatmap

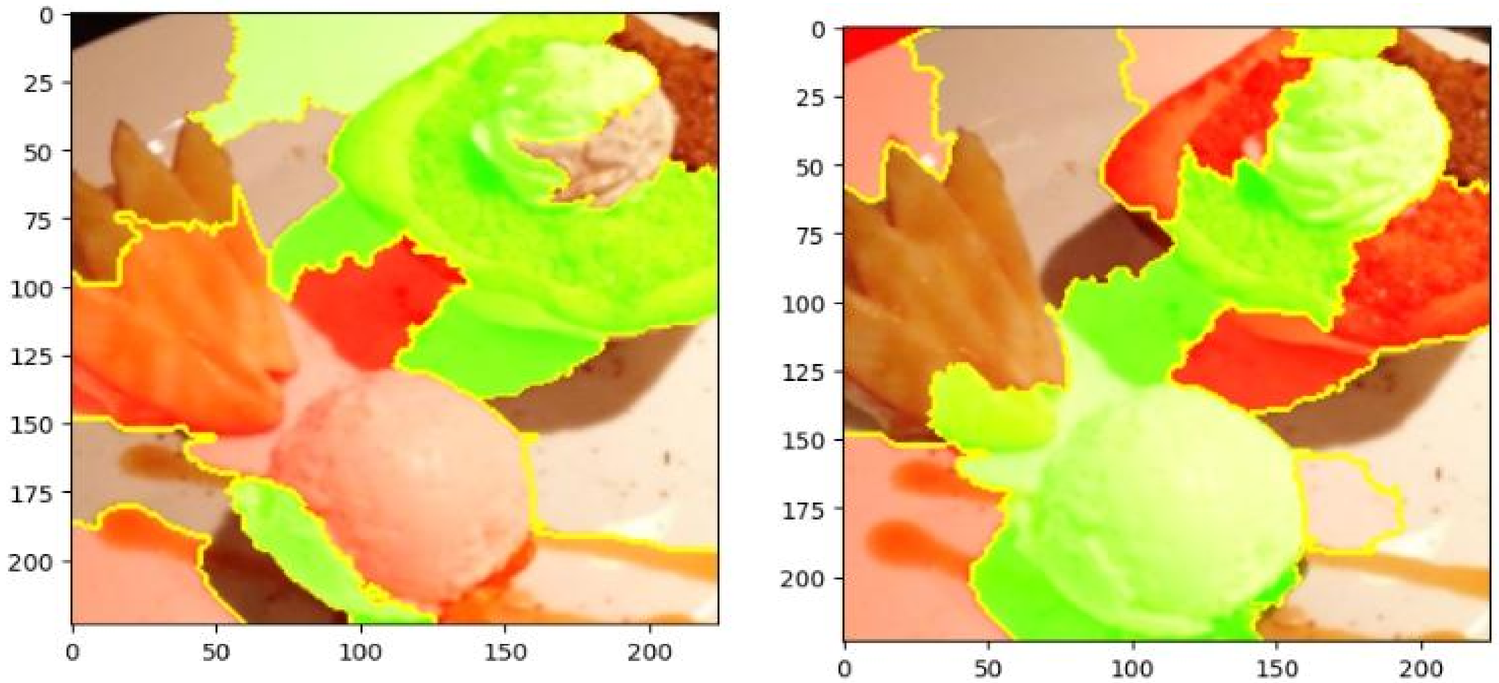


Figure 18: LIME Analysis Map

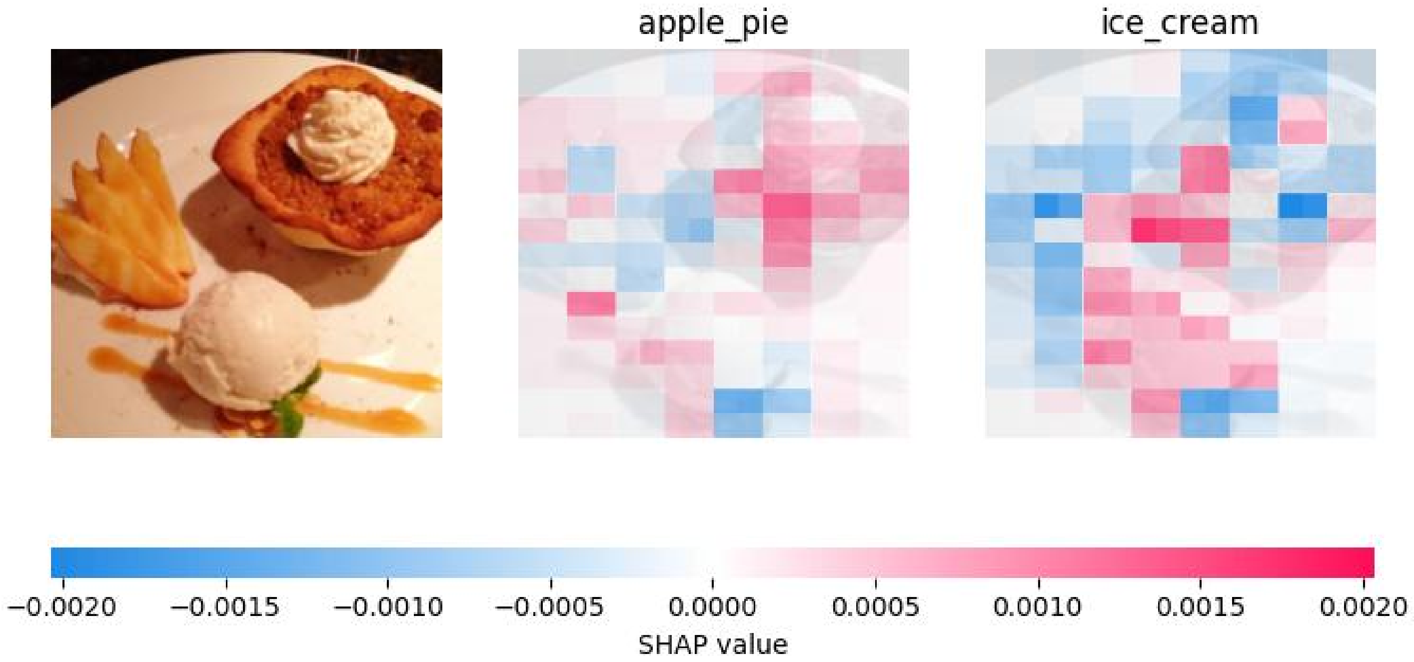
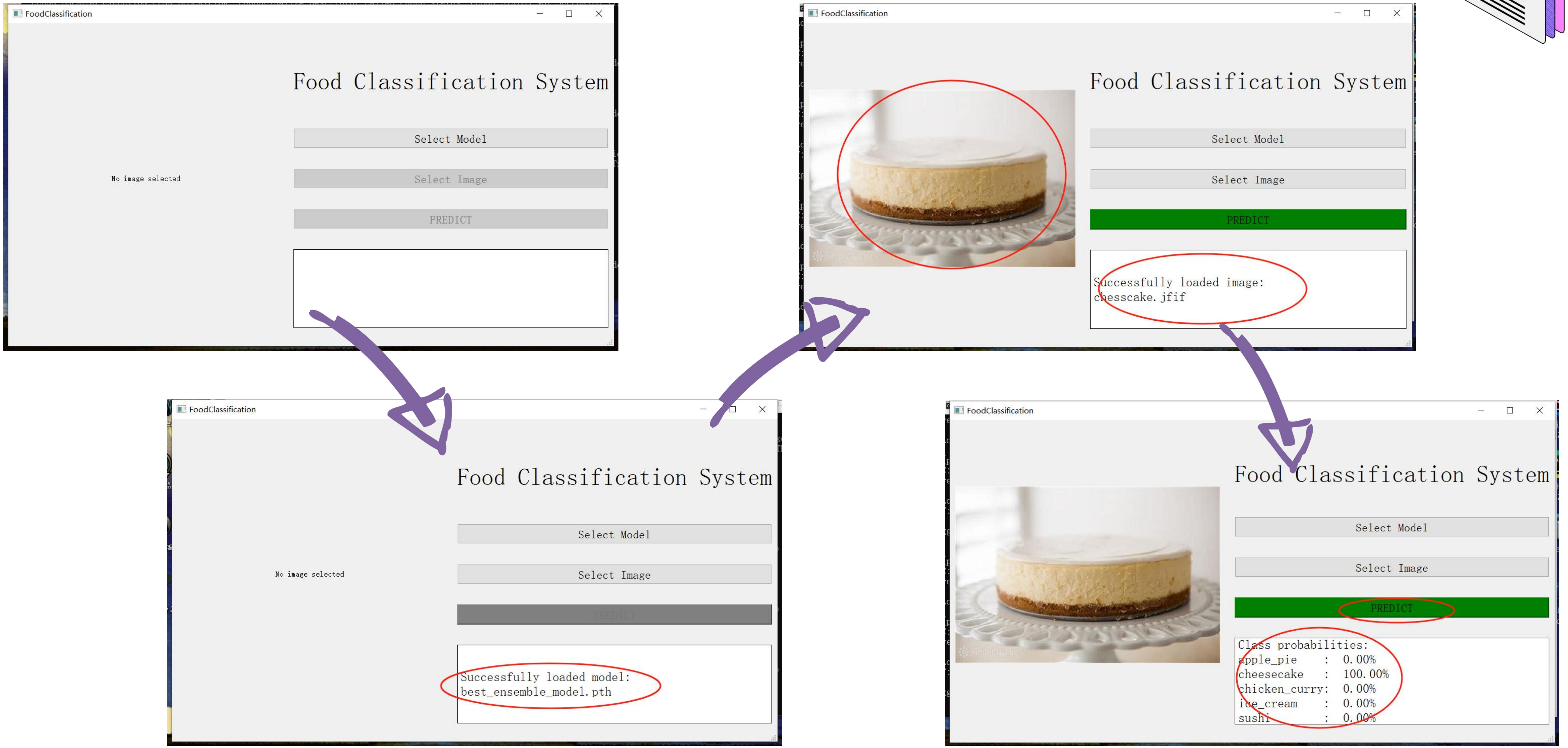


Figure 19: SHAP values

- Grad-CAM - Visualizes areas of significance in image-based models [2]
- LIME - Offers insights into model predictions by approximating the vicinity of data points [3]
- SHAP - Quantifies feature contributions to predictions [3]

04 Experiments & Results



05 Reflections & Conclusion

Key Findings

- High Accuracy
  - Achieved mainly by ensemble learning and learning rate scheduler
- Model Efficiency - MobileNetV2

Challenges and Limitations

- Small Dataset:
  - Limits generalization capability.
- Computational Resources:
  - High computational cost for ensemble models

Future Work

- Expanding Dataset:
  - Include more food categories
- Improved Models:
  - Explore lightweight CNN architectures

OXFORD BROOKES UNIVERSITY

Thank you for your listening!

OXFORD BROOKES UNIVERSITY

06 Q&A

Any questions?

07 References

1 Y. Liu, H. Pu, and D.-W. Sun, “Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices,” Trends Food Sci. Technol., vol. 113, pp. 193–204, Jul. 2021, doi: 10.1016/j.tifs.2021.04.042.

2 L. Xiao, T. Lan, D. Xu, W. Gao, and C. Li, “A Simplified CNNs Visual Perception Learning Network Algorithm for Foods Recognition,” Comput. Electr. Eng., vol. 92, p. 107152, Jun. 2021, doi: 10.1016/j.compeleceng.2021.107152.

3 K. R, S. B, and K. V, “Integrating Explainable AI with Infrared Imaging and Deep Learning for Breast Cancer Detection,” in 2023 OITS International Conference on Information Technology (OCIT), Dec. 2023, pp. 82–87. doi: 10.1109/OCIT59427.2023.10431160.

4 S. Gawde, S. Patil, S. Kumar, P. Kamat, K. Kotecha, and S. Alfarhood, “Explainable Predictive Maintenance of Rotating Machines Using LIME, SHAP, PDP, ICE,” IEEE Access, vol. 12, pp. 29345–29361, 2024, doi: 10.1109/ACCESS.2024.3367110.