THAI CANNABIS PLANTS

Tanishk N Shinde

TABLE OF CONTENTS

01 Data and information visualization

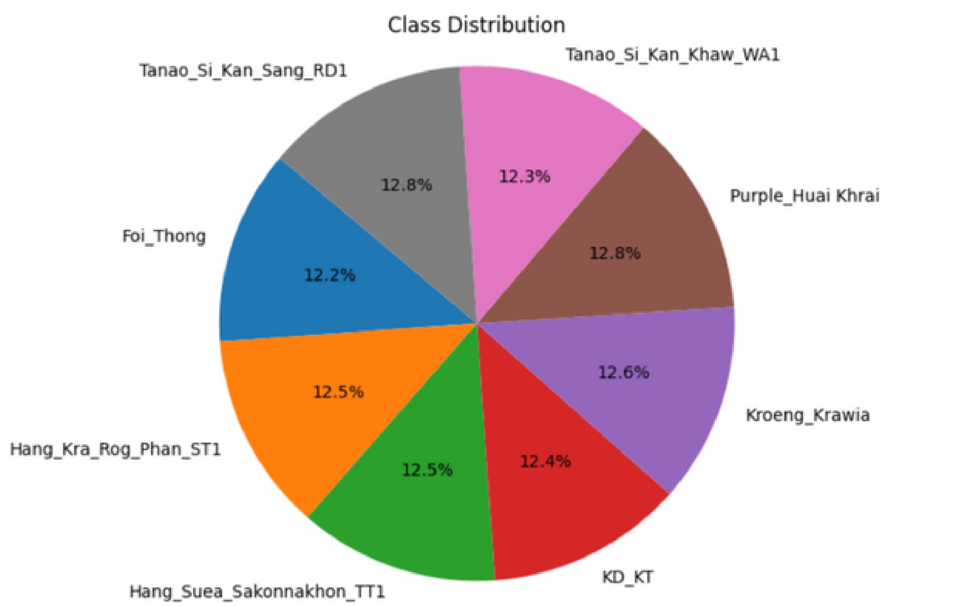
02 Comparison

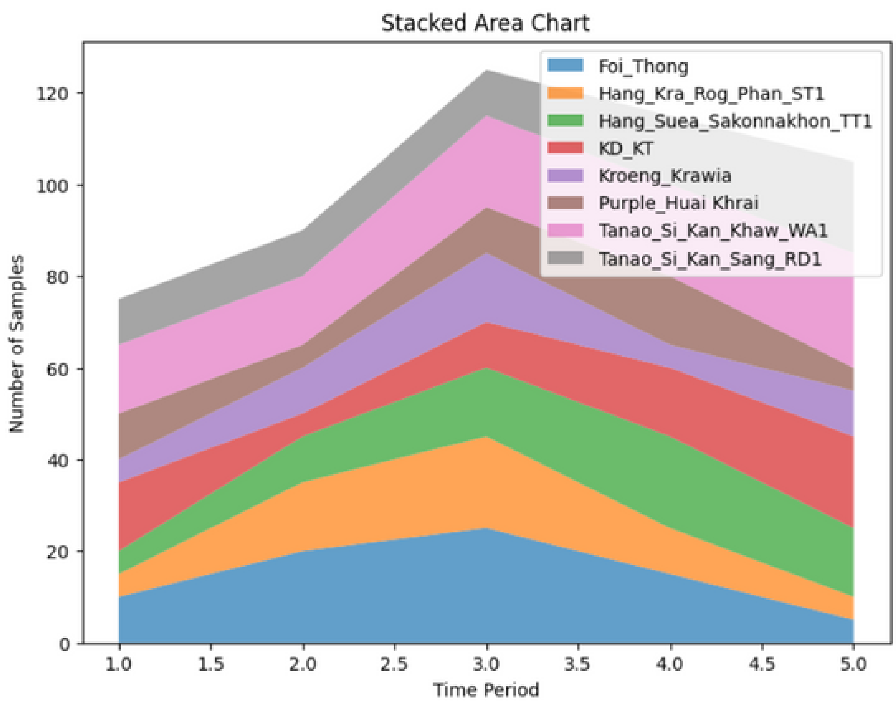
03 VGG-16

04 ResNet50

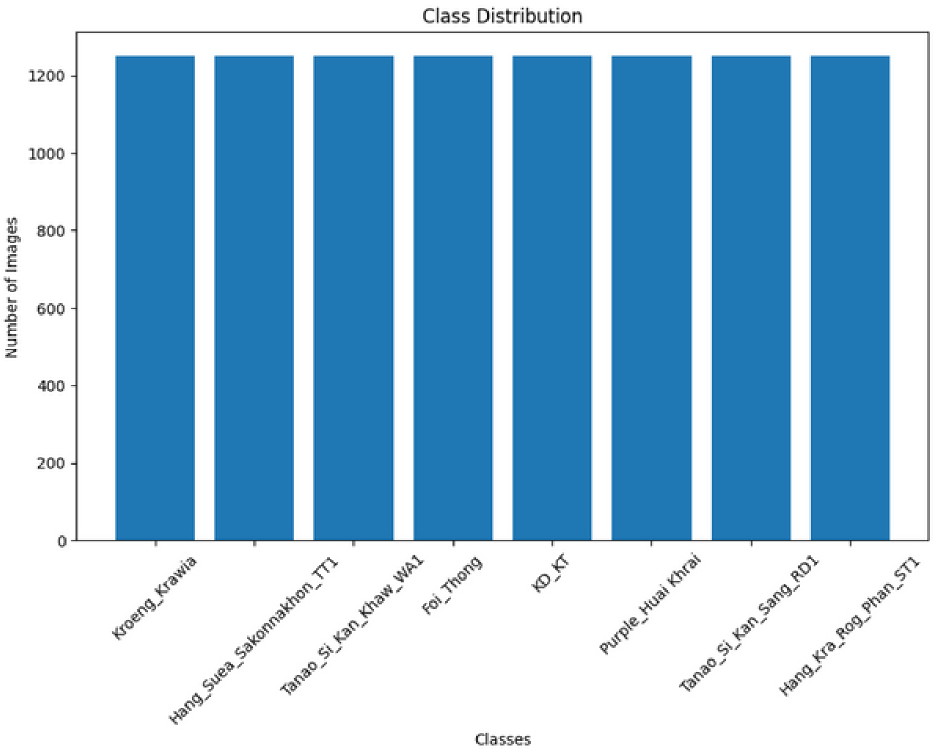
05 MobileNetV2

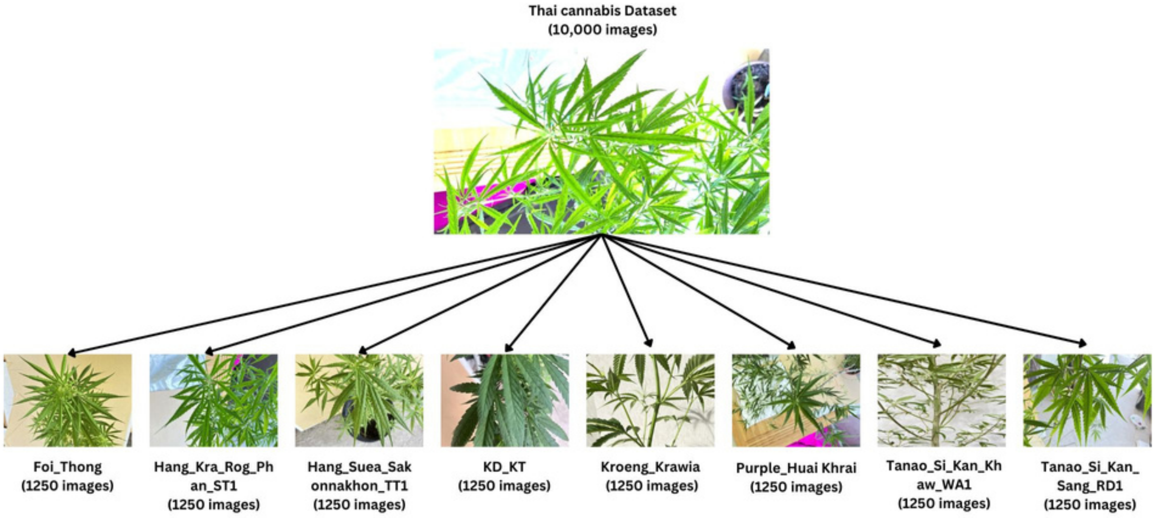
DATA AND INFORMATION VISUALIZATION






Kroeng_Krawia
Tanao_Si_Kan_Sang_RD1
Hang_Kra_Rog_Phan_ST1
Tanao_Si_Kan_Khaw_WA1
Foi_Thong
Hang_Suea_Sakonnakhon_TT1
KD_KT
Purple_Huai Khrai


Class Distribution
Number of Images
0
200
400
600
800
1000
1200
Kroeng_Krawia
Hang_Suea_Sakonnakhon_TT1
Tanao_Si_Kan_Khaw_WA1
Foi_Thong
KD_KT
Purple_Huai Khrai
Tanao_Si_Kan_Sang_RD1
Hang_Kra_Rog_Phan_ST1
Classes


Thai cannabis Dataset
(10,000 images)
Foi_Thong
(1250 images)
Hang_Kra_Rog_Ph
an_ST1
(1250 images)
Hang_Suea_Sak
onnakhon_TT1
(1250 images)
KD_KT
(1250 images)
Kroeng_Krawia
(1250 images)
Purple_Huai Khrai
(1250 images)
Tanao_Si_Kan_Kh
aw_WA1
(1250 images)
Tanao_Si_Kan_
Sang_RD1
(1250 images)

COMPARISON

Model	Accuracy	Val_Accuracy	Recall	F1-Score	ROC Curve	Loss	Val_Loss
VGG16	**0.9833**	**0.9382**	**0.9378**	**0.9394**	**0.9965**	**0.1452**	**0.5889**
ResNet50	0.9866	0.9791	0.9792	0.9791	0.9986	0.0609	0.1438
MobileNetV2	0.9645	0.9636	0.9628	0.9635	0.9977	0.1192	0.1415

VGG-16

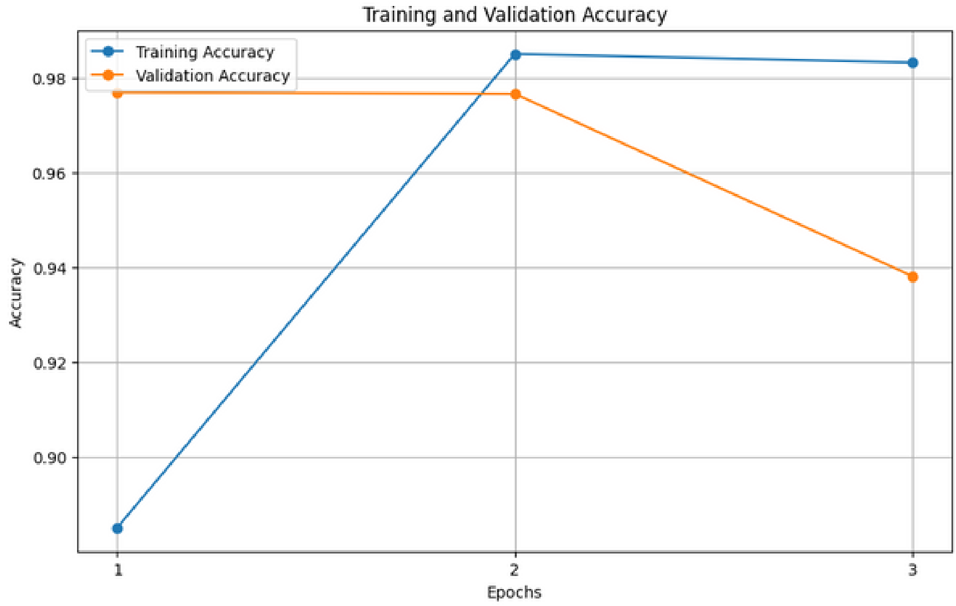
VGG-16, short for Visual Geometry Group 16, is a convolutional neural network (CNN) architecture widely recognized for its exceptional performance in image classification and object recognition tasks. Developed by the Visual Geometry Group at the University of Oxford, VGG-16 is a key milestone in the evolution of deep learning models for computer vision.

The VGG-16 architecture consists of 16 weight layers, which include 13 convolutional layers and 3 fully connected layers. The simplicity and uniformity of this architecture are notable, with small 3x3 convolutional filters used throughout the network. This design choice allows the model to learn intricate features and patterns in images, making it highly effective in recognizing objects of varying complexities.

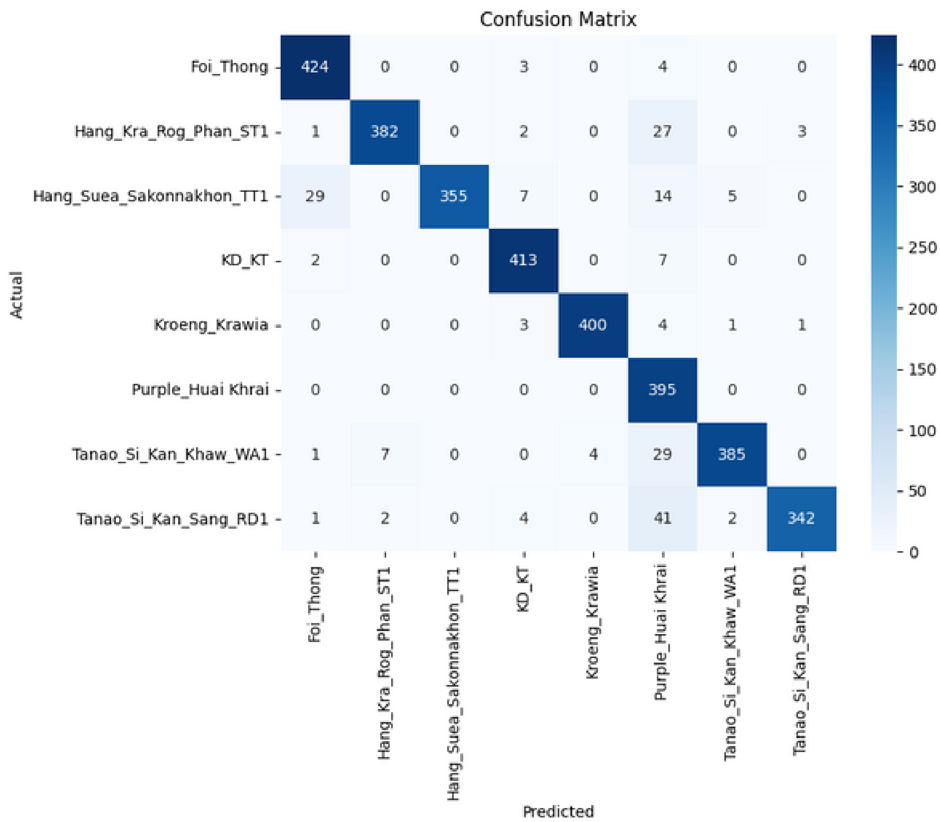
VGG-16 achieved its prominence by winning the 2014 ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition. It demonstrated remarkable accuracy in classifying a diverse range of objects in images, solidifying its reputation as a benchmark in the field of deep learning. Researchers and practitioners often use VGG-16 as a base architecture for transfer learning in various computer vision applications, thanks to its strong generalization capabilities.

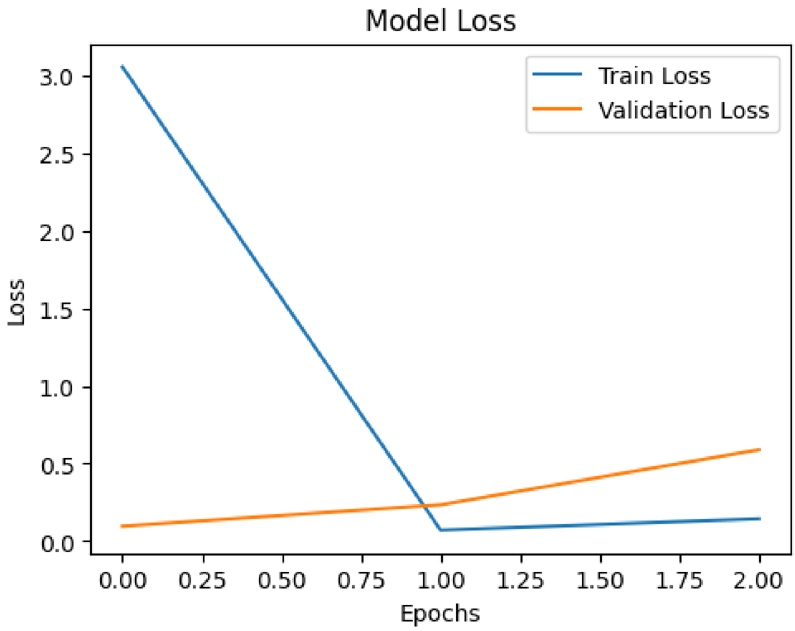
Despite its impressive performance, VGG-16 has certain drawbacks. Its architecture is relatively deep and computationally expensive, making it less suitable for real-time applications on resource-constrained devices. Also, VGG-16 does not incorporate recent advancements like residual connections or attention mechanisms, which are present in more recent models like ResNet and Transformer-based networks.

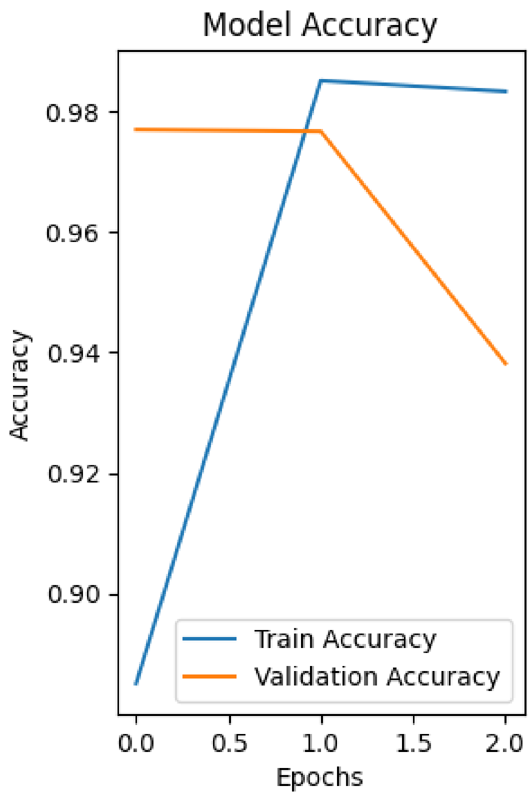
In summary, VGG-16 is a groundbreaking CNN architecture that set a standard for image classification. While it may not be the most efficient model in terms of computational resources, its simplicity and effectiveness continue to make it a valuable tool in the deep learning community, serving as a reference point for researchers and engineers working on computer vision tasks.

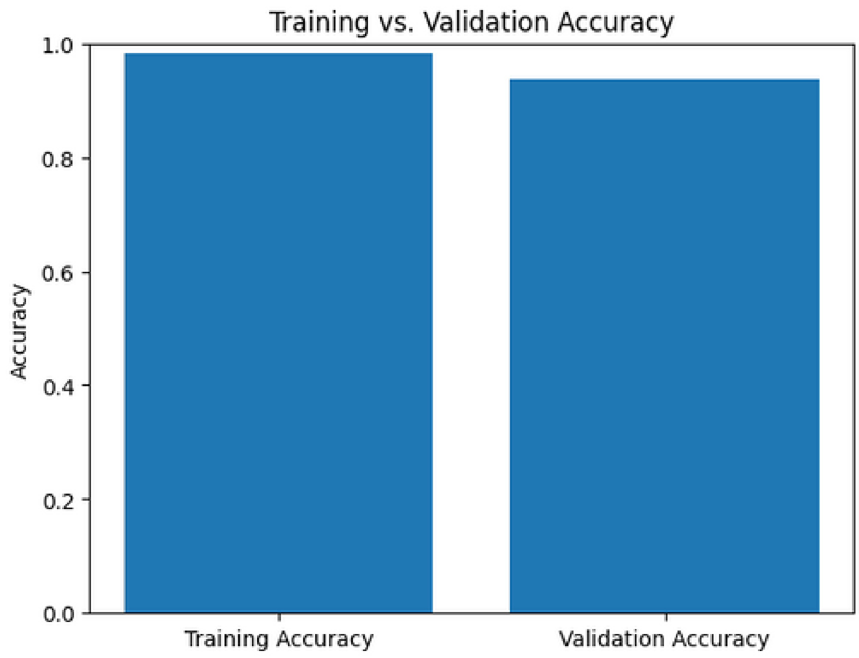



```
Final Training Accuracy: 0.9833
Final Validation Accuracy: 0.9382
```

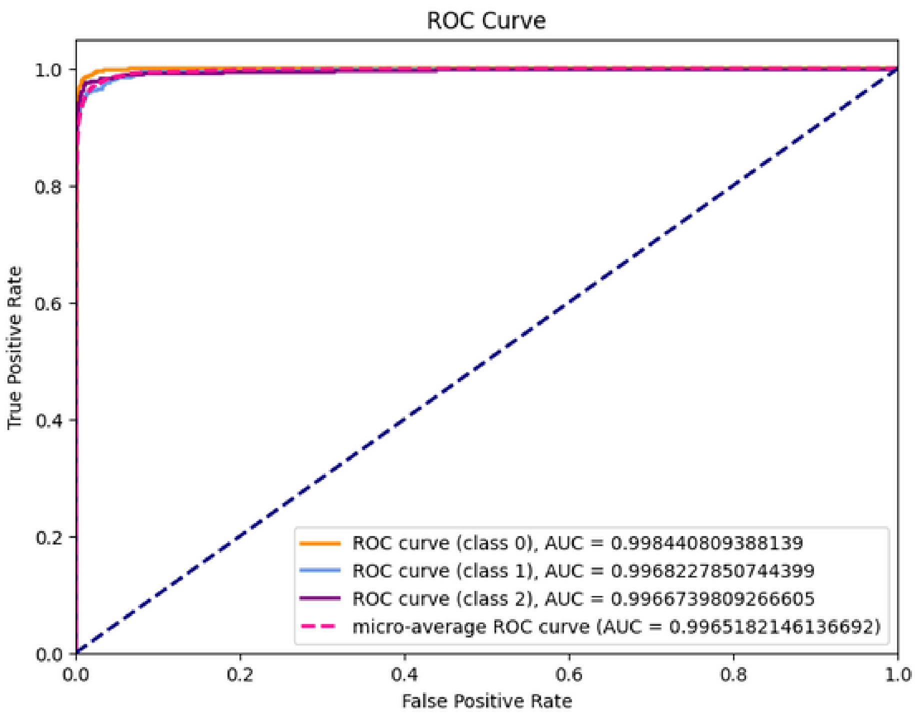



Model Loss
Train Loss
Validation Loss
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Loss
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
Epochs


Model Accuracy
0.98
0.96
0.94
0.92
0.90
Accuracy
Train Accuracy
Validation Accuracy
0.0
0.5
1.0
1.5
2.0
Epochs



Training vs. Validation Accuracy
Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Training Accuracy
Validation Accuracy



ROC Curve
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC curve (class 0), AUC = 0.998440809388139
ROC curve (class 1), AUC = 0.9968227850744399
ROC curve (class 2), AUC = 0.9966739809266605
micro-average ROC curve (AUC = 0.9965182146136692)

RESNET-50

ResNet-50, short for "Residual Network with 50 layers," is a deep convolutional neural network architecture designed for image classification tasks. It's a variant of the ResNet family, which was introduced by Kaiming He and his colleagues in 2015 and is known for its remarkable performance in various computer vision tasks.

ResNet-50's key innovation is the introduction of residual blocks, which allow for the training of extremely deep neural networks. Each residual block contains skip connections, or "shortcut connections," that skip one or more layers. This enables the network to learn residual functions, making it easier to train and mitigating the vanishing gradient problem.

ResNet-50 comprises 50 layers, including convolutional, batch normalization, activation, and fully connected layers. It is pre-trained on large-scale image datasets such as ImageNet, where it learns to recognize a wide range of objects and features.

The architecture of ResNet-50 can be summarized as follows:

1. Input Layer: Accepts images of a fixed size (typically 224x224 pixels with three color channels).

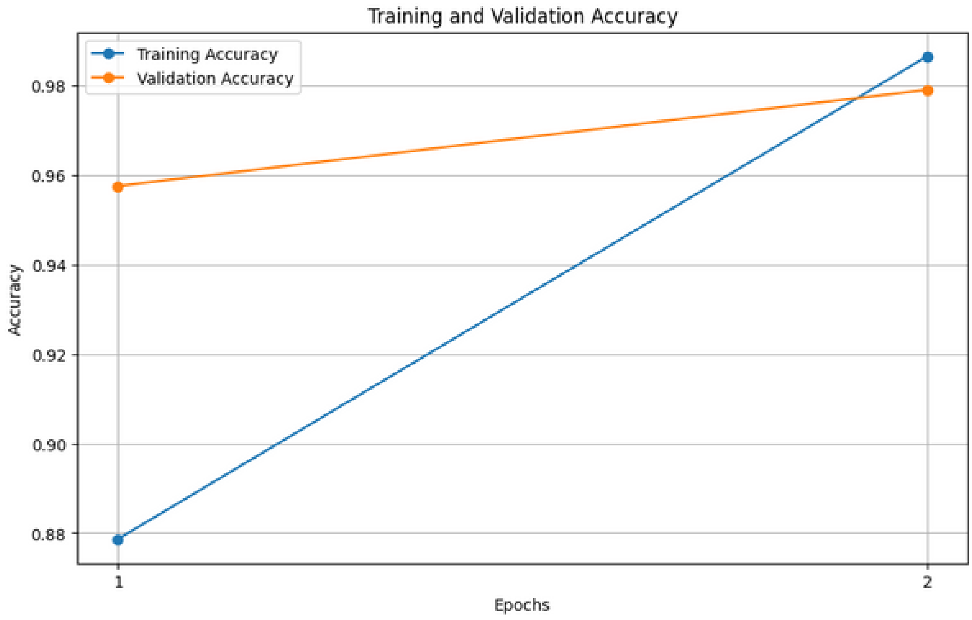
2. Convolutional Layers: Multiple layers of 3x3 convolution operations.

3. Residual Blocks: Comprising a series of convolutional and batch normalization layers, with skip connections.

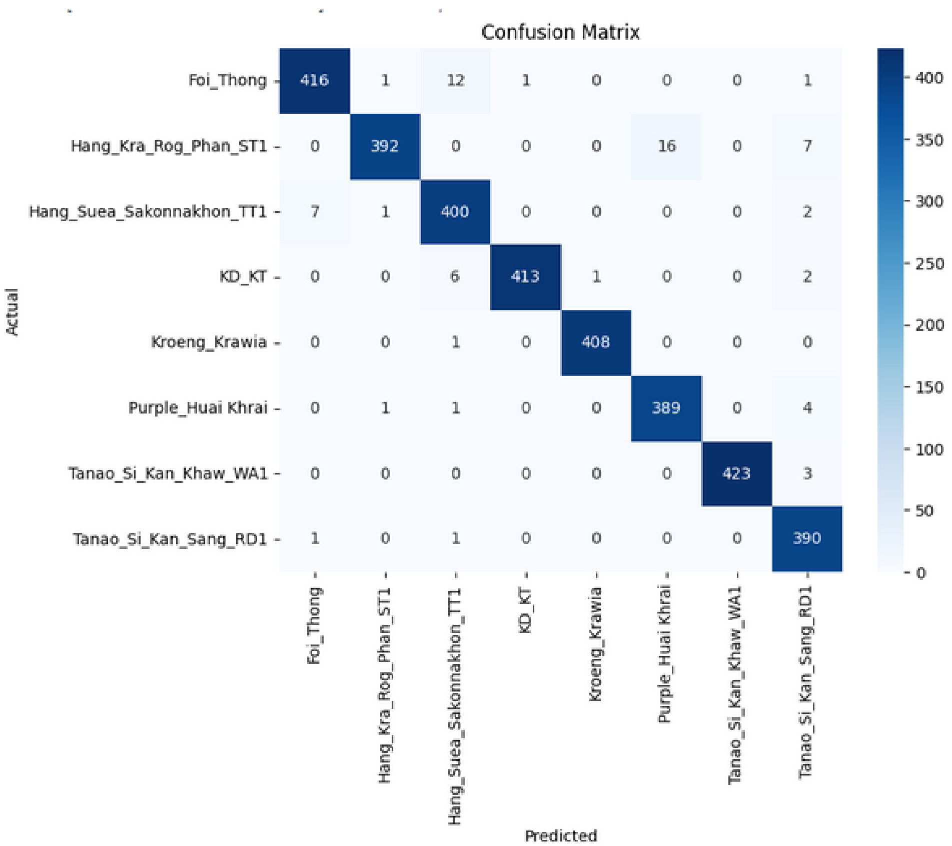
4. Global Average Pooling Layer: Reduces the spatial dimensions of the feature maps.

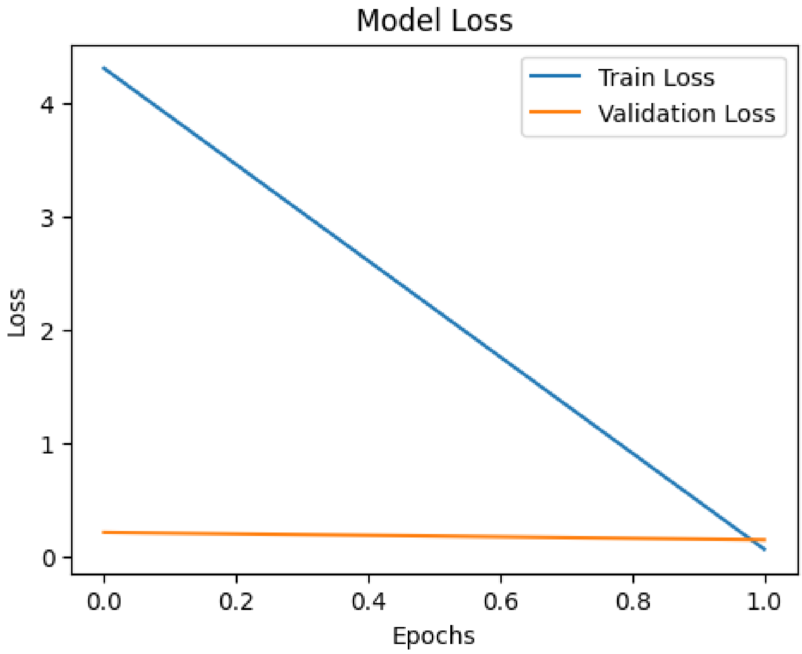
5. Fully Connected Layer: The output layer for classifying images into specific categories.

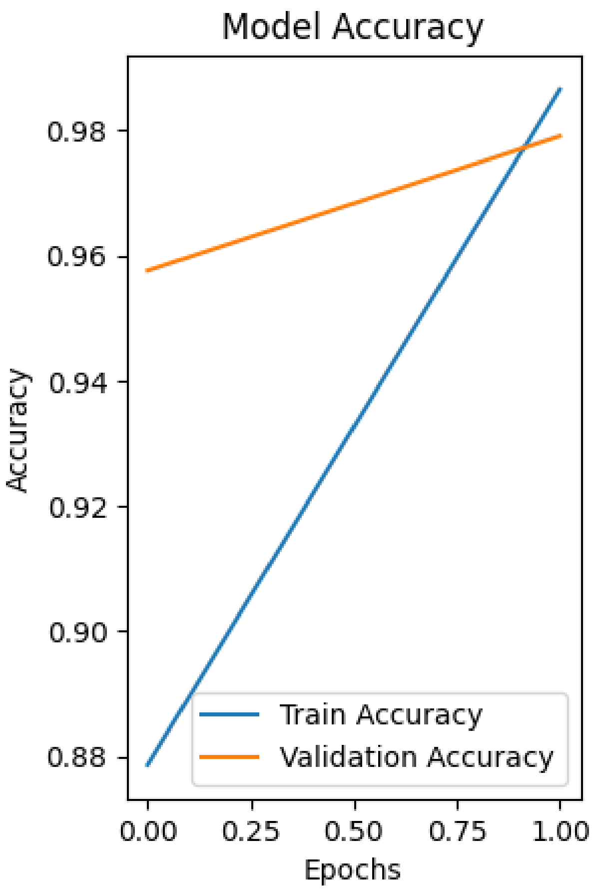
ResNet-50 has become a popular choice for various computer vision tasks due to its deep architecture, excellent accuracy, and the availability of pre-trained models. Researchers and practitioners often fine-tune it for specific tasks or use it as a feature extractor by removing the fully connected layers. Its ability to handle complex visual data has made it an essential tool in the field of deep learning and image recognition.

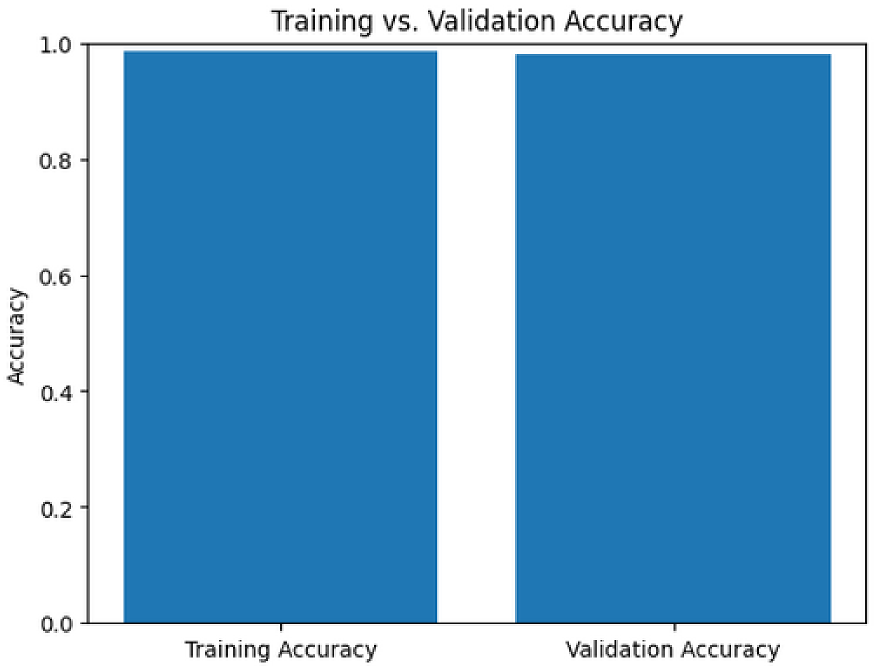


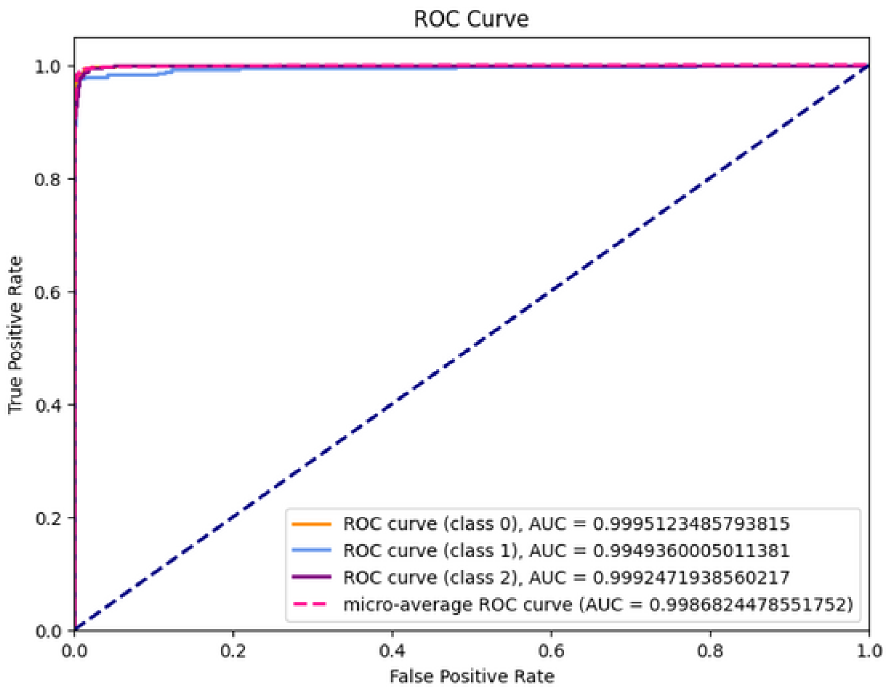

```
Final Training Accuracy: 0.9866
Final Validation Accuracy: 0.9791
```




Model Loss
Train Loss
Validation Loss
Loss
4
3
2
1
0
0.0
0.2
0.4
0.6
0.8
1.0
Epochs


Model Accuracy
Train Accuracy
Validation Accuracy
Accuracy
0.98
0.96
0.94
0.92
0.90
0.88
0.00
0.25
0.50
0.75
1.00
Epochs


Training vs. Validation Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
Training Accuracy
Validation Accuracy


ROC Curve
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
ROC curve (class 0), AUC = 0.9995123485793815
ROC curve (class 1), AUC = 0.9949360005011381
ROC curve (class 2), AUC = 0.9992471938560217
micro-average ROC curve (AUC = 0.9986824478551752)

MOBILENETV2

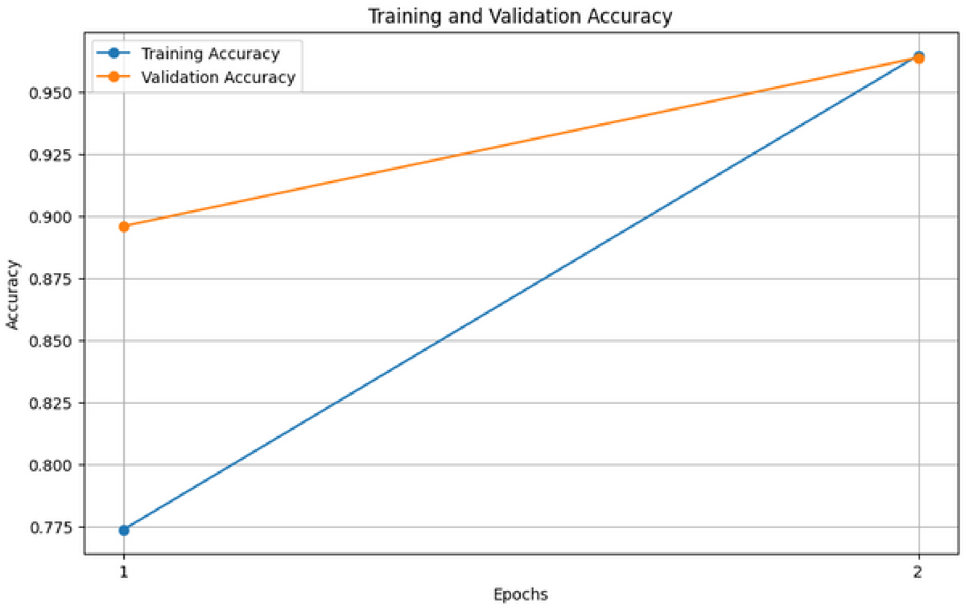
MobileNetV2 is a deep learning neural network architecture developed by Google as an evolution of the original MobileNet. It is designed for efficient on-device computer vision tasks, particularly on mobile and embedded devices with limited computational resources.

MobileNetV2 introduces several key features to improve its performance and efficiency. One of the most notable innovations is the use of inverted residual blocks, which consist of a lightweight bottleneck layer followed by an expansion layer. This structure reduces the overall computational cost while maintaining model accuracy. Additionally, MobileNetV2 utilizes depthwise separable convolutions, which separate the convolutional operation into depthwise and pointwise convolutions, further reducing computation and model size.

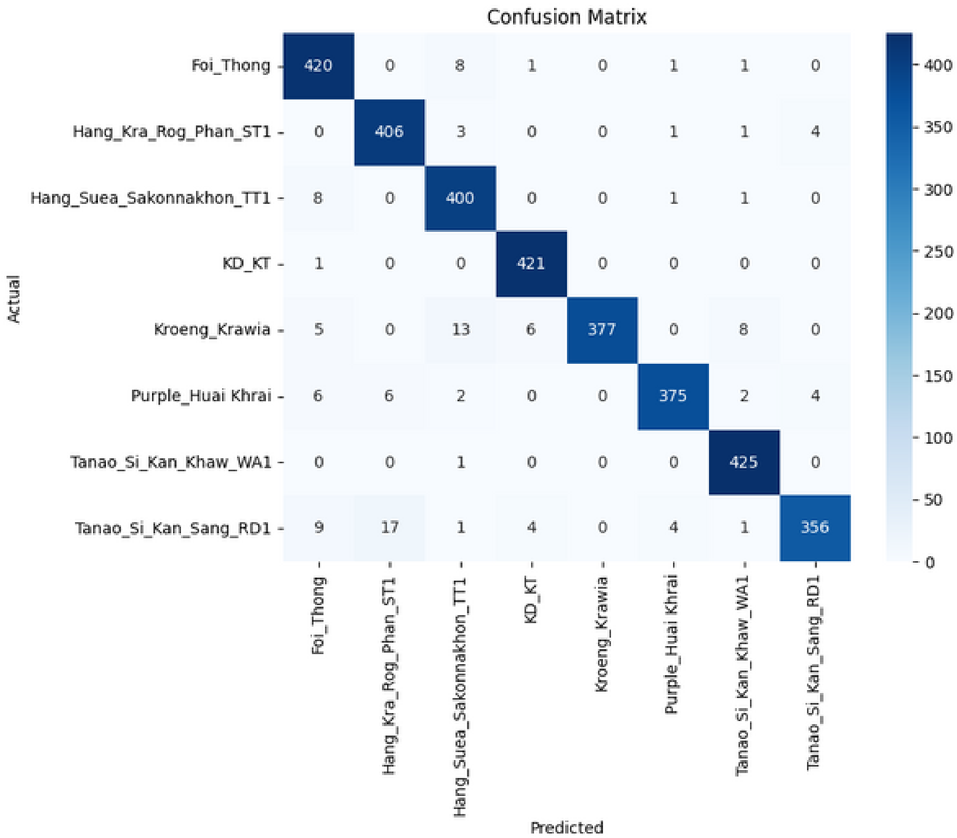
The architecture also incorporates a feature called linear bottlenecks, which prevents information loss during the transformations, and skip connections that enable better gradient flow during training. MobileNetV2 uses a set of hyperparameters to control model size and computational efficiency, allowing for customization to suit specific hardware and deployment constraints.

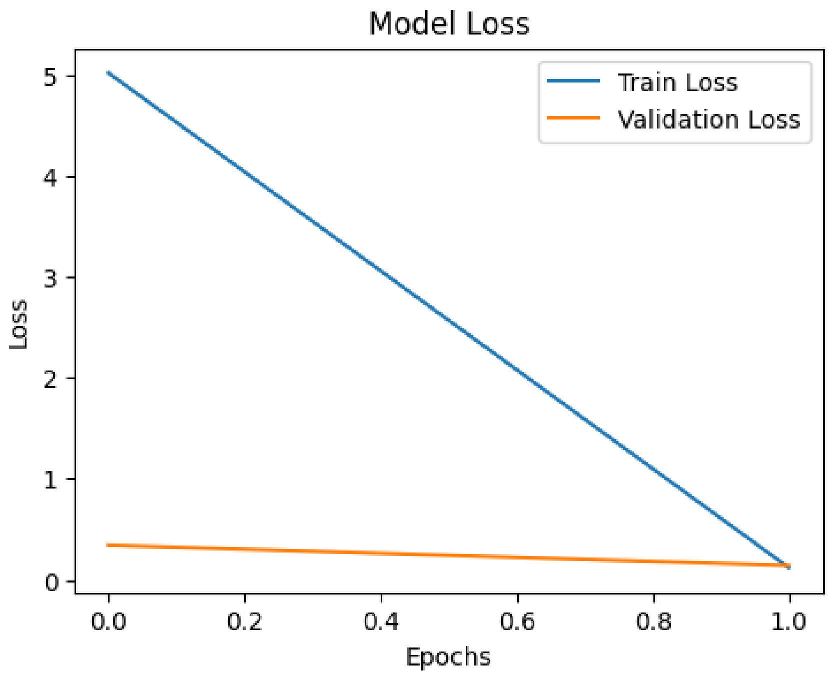
MobileNetV2 has proven to be highly efficient and effective for various computer vision tasks, including image classification, object detection, and semantic segmentation. Its lightweight design and compatibility with mobile and embedded devices make it a popular choice for real-time applications, such as mobile photography, augmented reality, and autonomous vehicles.

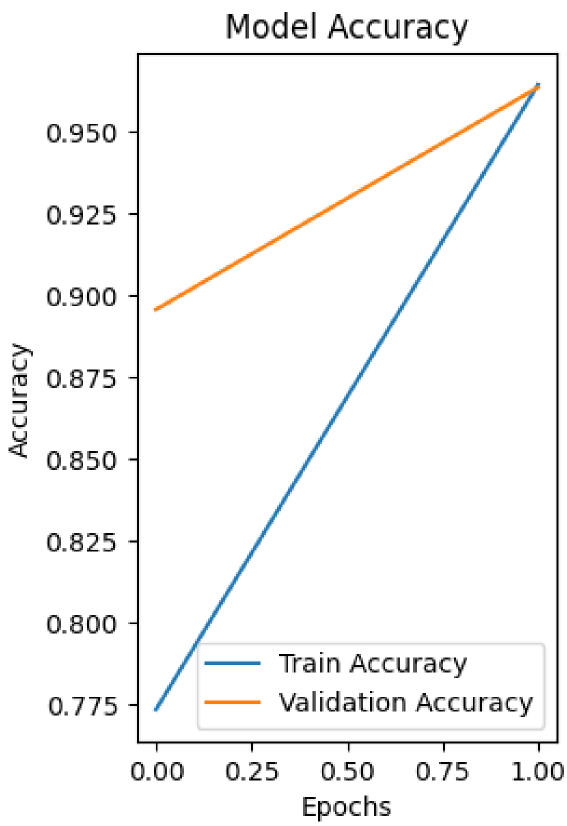
In summary, MobileNetV2 is a state-of-the-art neural network architecture optimized for mobile and embedded devices. Its innovative features, including inverted residual blocks and depthwise separable convolutions, enable efficient and accurate computer vision tasks, making it a valuable tool for applications requiring low-power and real-time processing.

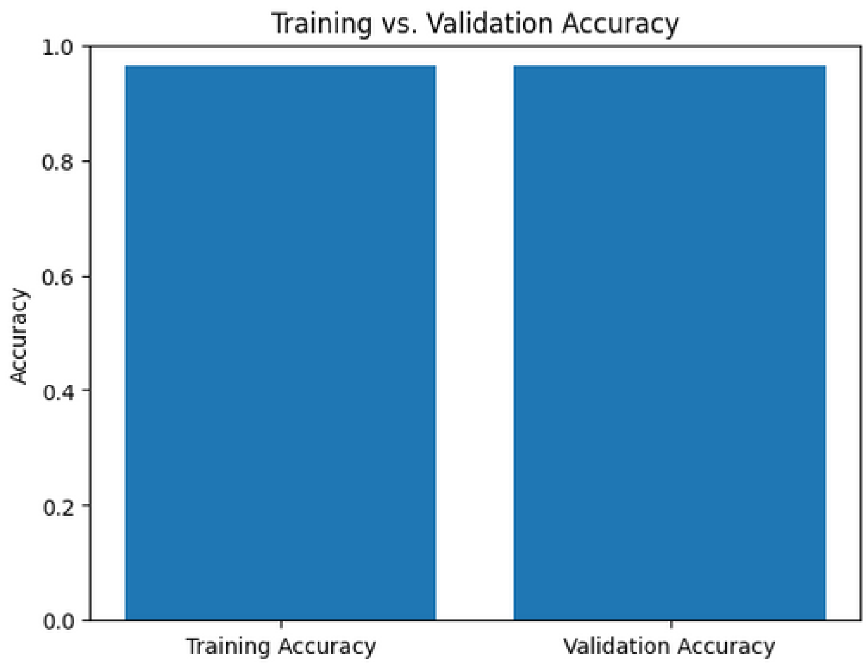


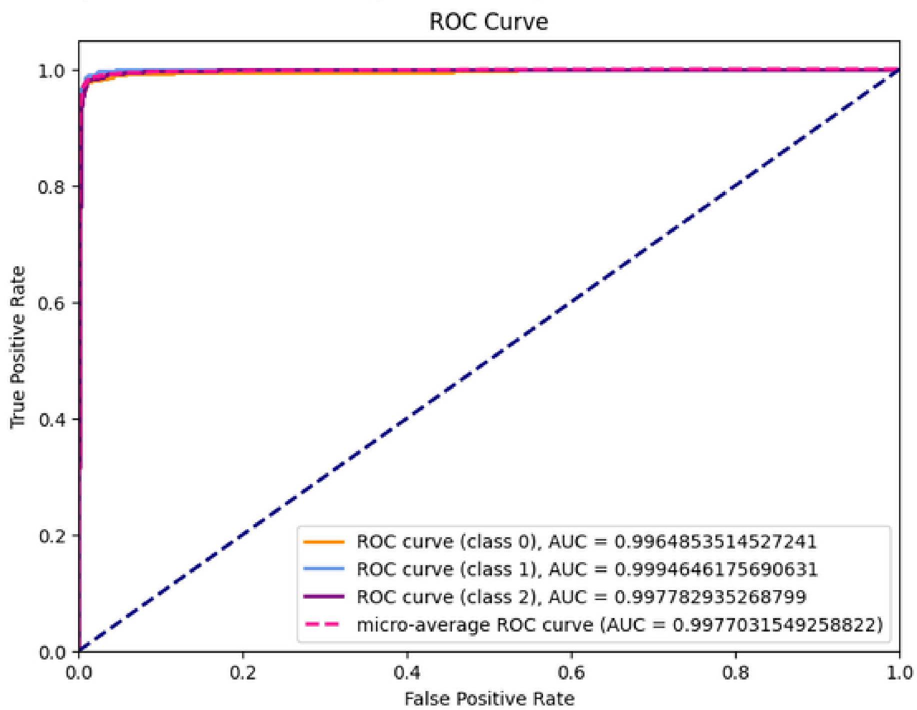

```
Final Training Accuracy: 0.9645
Final Validation Accuracy: 0.9636
```




Model Loss
Train Loss
Validation Loss
Loss
Epochs
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0


Model Accuracy
Train Accuracy
Validation Accuracy
Accuracy
Epochs
0.775
0.800
0.825
0.850
0.875
0.900
0.925
0.950
0.00
0.25
0.50
0.75
1.00


Training vs. Validation Accuracy
1.0
0.8
0.6
0.4
0.2
0.0
Accuracy
Training Accuracy
Validation Accuracy


ROC Curve
1.0
0.8
0.6
0.4
0.2
0.0
True Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate
ROC curve (class 0), AUC = 0.9964853514527241
ROC curve (class 1), AUC = 0.9994646175690631
ROC curve (class 2), AUC = 0.997782935268799
micro-average ROC curve (AUC = 0.9977031549258822)