FACE EMOTIONS CLASSIFICATION CNN & VGG16

Presented By : Rishitha thoka(E0322026)
Afrin Banu R(E0322050)

OVERVIEW

- Introduction
- Problem Statement
- Objectives
- Data collection & Preprocessing
- Methodology
- Feature Extraction
- Model Training & Model evaluation
- Application
- Result
- Visualization
- Conclusion

INTRODUCTION

Facial expressions are a universal method of conveying emotions. With the advancement of artificial intelligence and computer vision, machines can now be trained to recognize human emotions using image processing techniques. This project explores the use of a deep learning model specifically the VGG16 Convolutional Neural Network (CNN) to accurately classify emotions based on facial features. The system aims to improve emotion detection systems by combining the power of transfer learning and facial image analysis.

PROBLEM STATEMENT

- Emotion detection using traditional methods is inaccurate and lacks adaptability.
- Facial expressions vary widely across individuals, lighting conditions, and angles.
- The project aims to develop a deep learning model for real-time facial emotion classification using Convolutional Neural Networks (CNN) and VGG16 architecture.
- The objective is to accurately identify and categorize human emotions (e.g., happy, sad, angry,) from facial expressions, enhancing human-computer interaction and emotional intelligence in AI systems.

OBJECTIVES

Objective 1

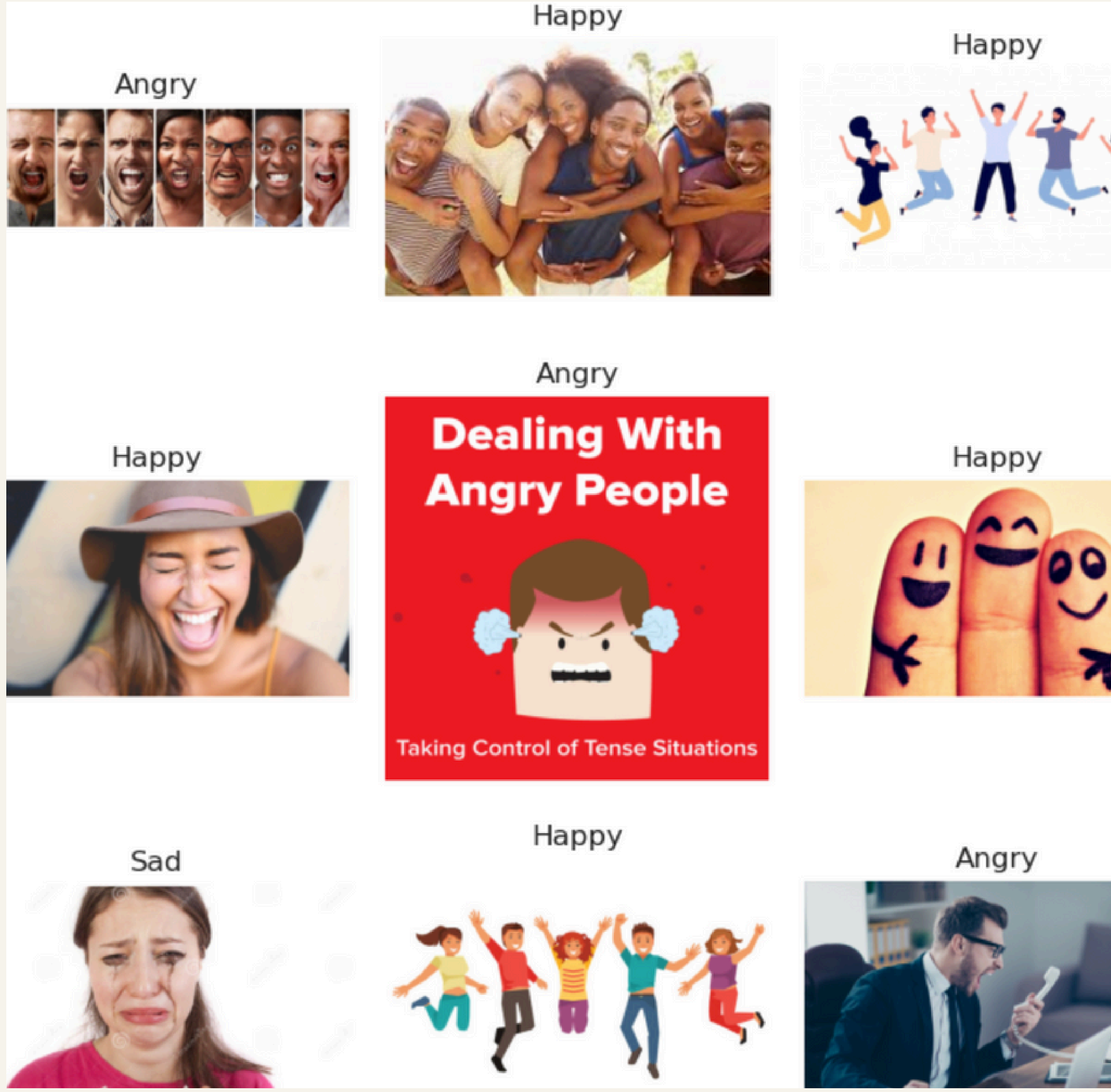
- The main goal of this project is to design and train a deep learning model that can accurately classify human facial expressions into three distinct emotional categories.
- To utilize transfer learning with VGG16 to improve learning efficiency.

Objective 2

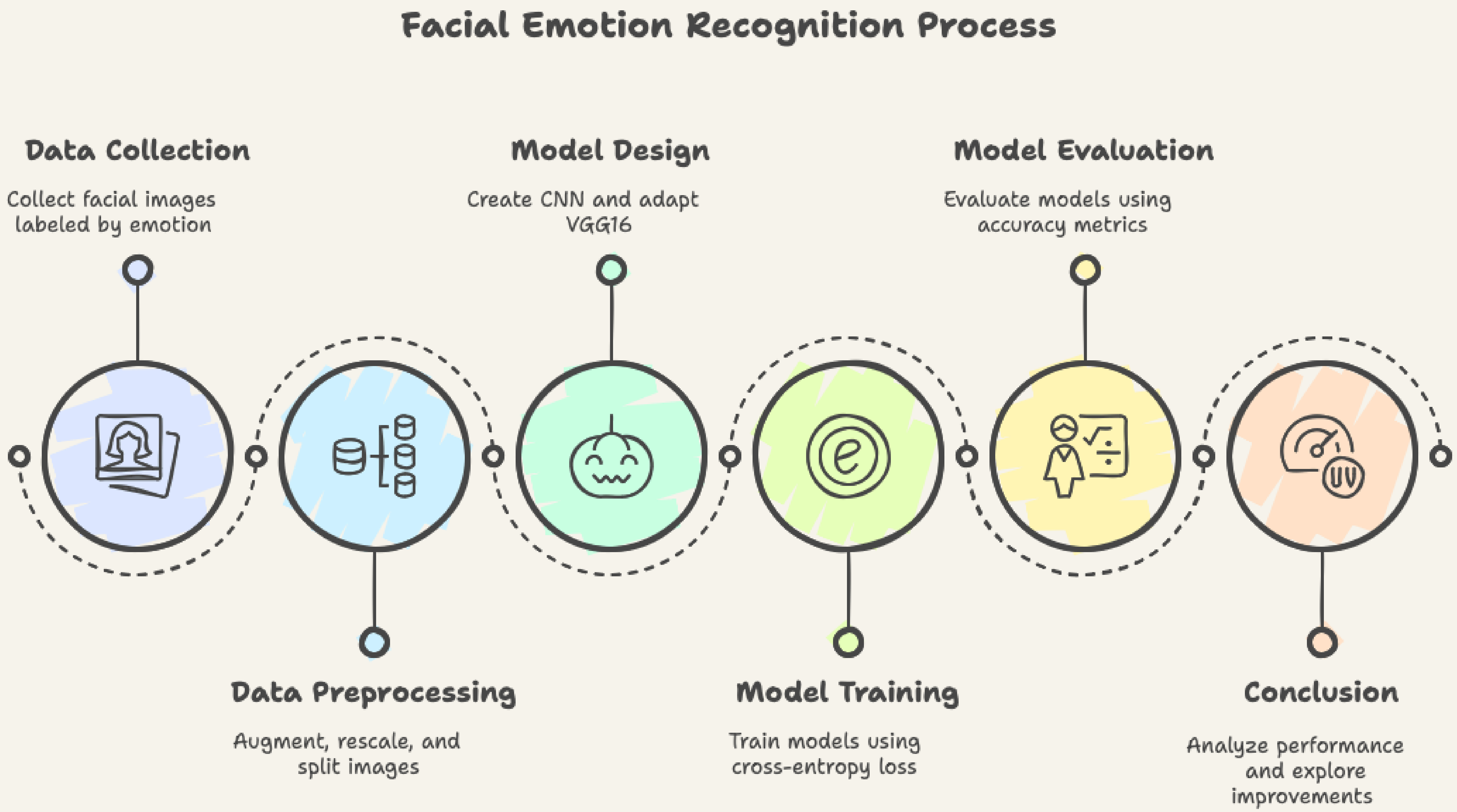
- To evaluate model performance with real-world facial expression datasets.
- The final model is expected to be efficient, scalable, and potentially integrable into real-world applications such as education, security, and mental health assessment.

DATA COLLECTION & PREPROCESSING

- The dataset consists of three folders: Happy, Sad, and Angry.
- Each folder contains 100 images representing the respective facial expression.
- The dataset is suitable for emotion classification using CNN and Computer Vision techniques.
- Preprocessing steps include:
- Rescaling pixel values to a 0–1 range for normalization.
- Applied data augmentation techniques like flipping and shifting to improve model generalization.



METHODOLOGY



FEATURE EXTRACTION AND MODEL ARCHITECTURE

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 222, 222, 10)	280
conv2d_1 (Conv2D)	(None, 220, 220, 10)	910
max_pooling2d (MaxPooling2D)	(None, 110, 110, 10)	0
conv2d_2 (Conv2D)	(None, 108, 108, 10)	910
conv2d_3 (Conv2D)	(None, 106, 106, 10)	910
max_pooling2d_1 (MaxPooling2D)	(None, 53, 53, 10)	0
flatten (Flatten)	(None, 28090)	0
dense (Dense)	(None, 128)	3,595,648
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 3)	387

Total params: 3,599,045 (13.73 MB)
Trainable params: 3,599,045 (13.73 MB)
Non-trainable params: 0 (0.00 B)

The project utilizes two CNN-based approaches:

- A custom-built CNN model
- A pre-trained VGG16 model using transfer learning

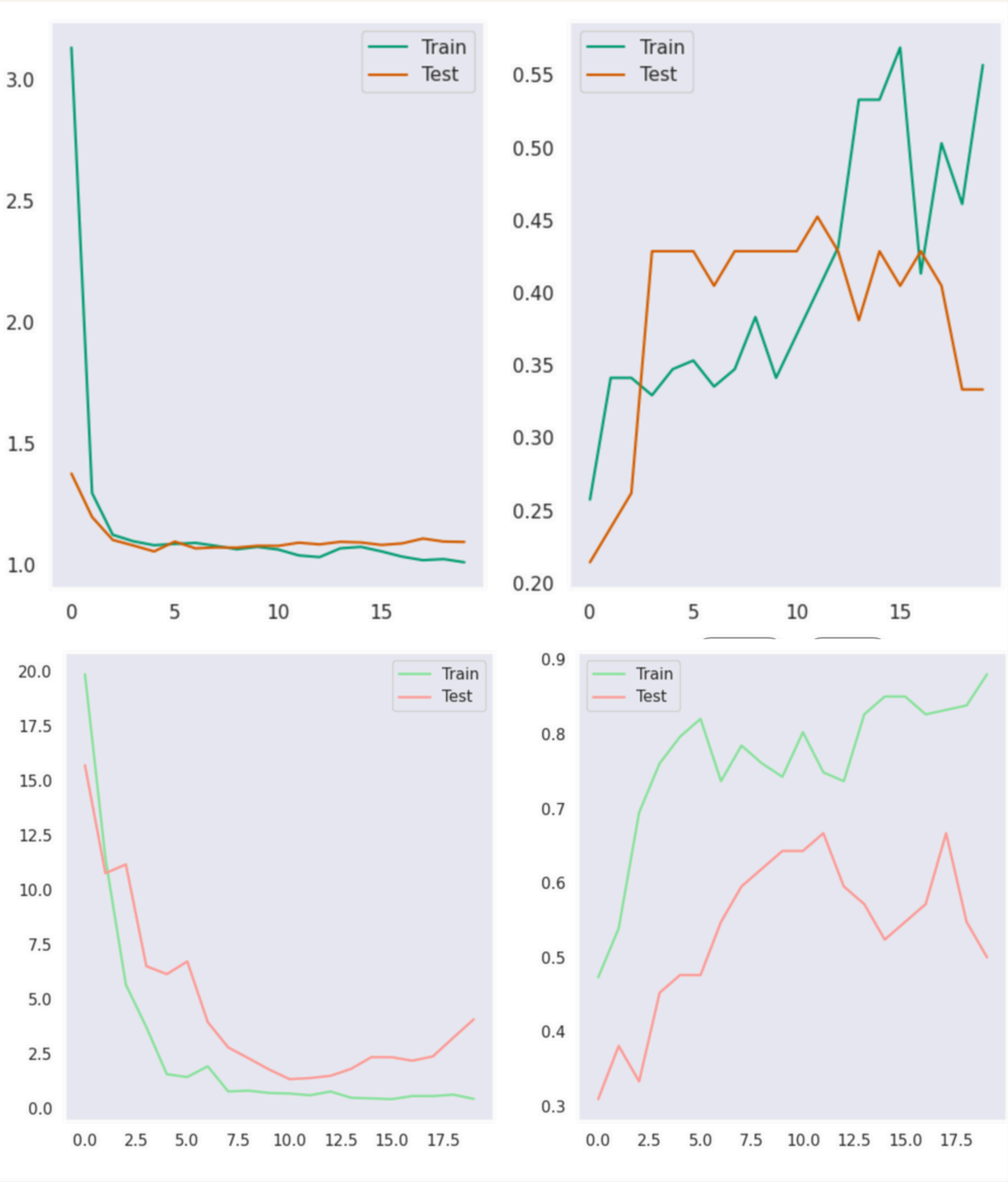
Custom CNN Model:

- Built from scratch using multiple convolutional layers followed by max pooling layers.
- Followed by Dense (fully connected) layers to learn complex representations.
- Dropout layers used to reduce overfitting.
- Final layer uses softmax activation for multi-class classification.

VGG16 Transfer Learning:

- VGG16 loaded with pre-trained ImageNet weights.
- Custom layers added: Flatten, Dense (ReLU activation), Dropout, and a softmax output layer.
- Helps leverage powerful feature extraction from a pre-trained model to improve accuracy.
- Feature extraction focuses on recognizing expression-specific patterns from Happy, Sad, and Angry face images.

MODEL TRAINING AND MODEL EVALUATION



- All models were trained using the Adam optimizer and categorical crossentropy loss function, suitable for multi-class classification (Happy, Sad, Angry).
- Training was conducted over multiple epochs, with early stopping employed to prevent overfitting by monitoring the validation loss.
- Performance metrics such as accuracy and loss were tracked for both the training and validation datasetsthroughout the training process.
- A custom CNN model served as the baseline architecture to compare against more advanced models like VGG16, ResNet50, and EfficientNetB0, used via transfer learning.
- For pre-trained models, the top classification layers were removed and replaced with custom dense layers, including Global Average Pooling, Dense (ReLU), Dropout, and a final Dense softmax layer for output.

APPLICATION

1)Human-Computer Interaction

- Enhances user experience by enabling systems to respond to emotional cues in real-time.

2)Mental Health Monitoring

- Assists in detecting emotional distress or mood changes in patients through automated observatioN.

3)Smart Surveillance Systems

- Identifies suspicious or abnormal behavior based on emotional expressions in public or restricted areas.

4)Gaming and Virtual Reality

- Adjusts gameplay and environment dynamically based on player emotions for a more immersive experience.

5)E-learning Platforms

- Monitors student engagement and emotional state to personalize learning experiences

RESULTS AND DISCUSSION

BEST PERFORMING MODEL:

- The VGG16 model outperformed other tested architectures, achieving the highest validation accuracy (~89.6%) in classifying facial expressions into Happy, Sad, and Angry categories.

PERFORMANCE METRICS:

- Model accuracy and loss were tracked throughout training on both training and validation sets.
- The plotted performance curves indicated that VGG16 maintained consistent performance with minimal signs of overfitting, unlike some other models.
- In comparison, custom CNN and simpler models showed moderate accuracy (~70–75%), highlighting the advantage of transfer learning.

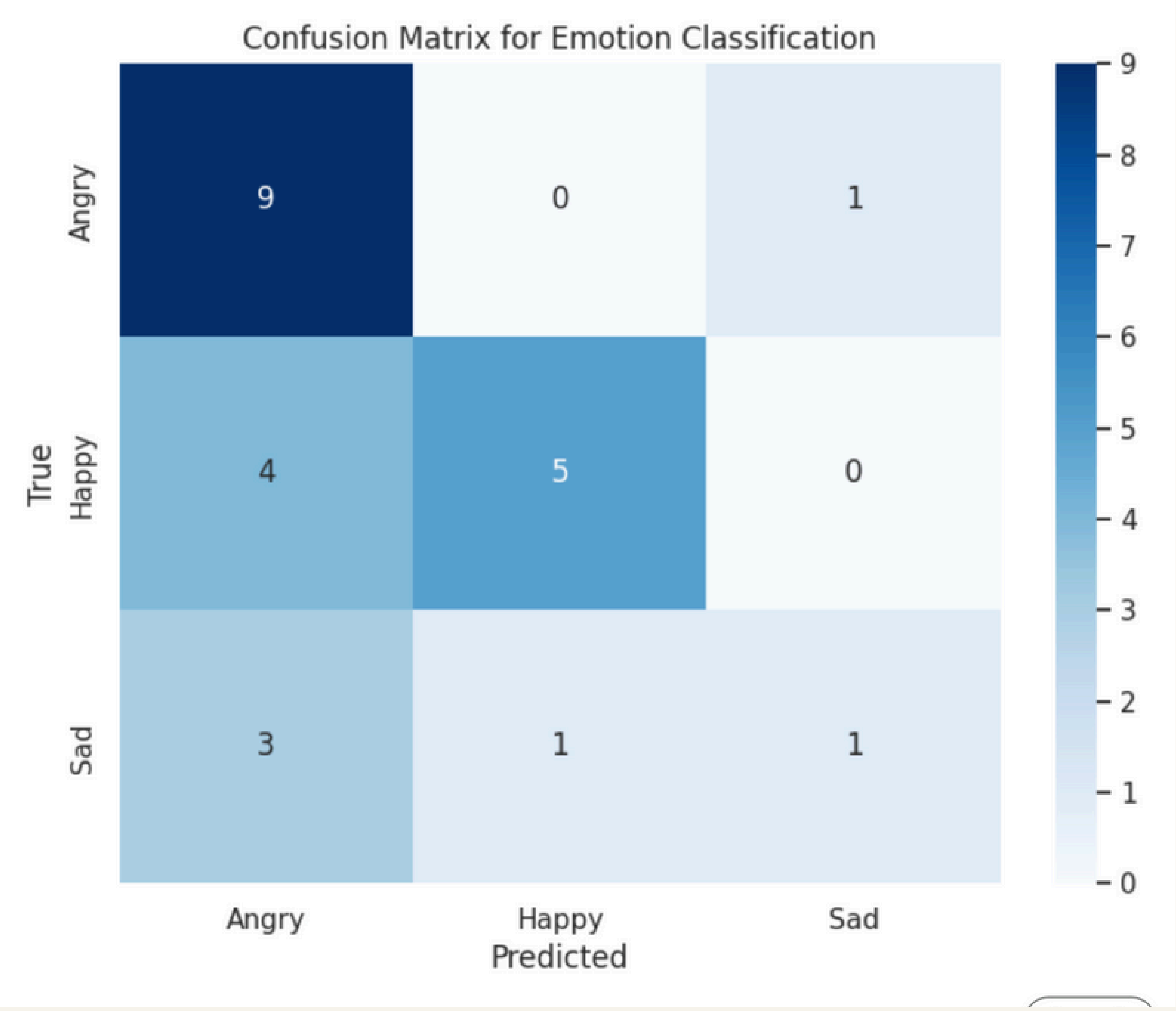
FEATURE LEARNING:

- Pre-trained models like VGG16 efficiently captured key facial features such as eyebrow shape, mouth curvature, and eye openness, which are essential for differentiating emotions.

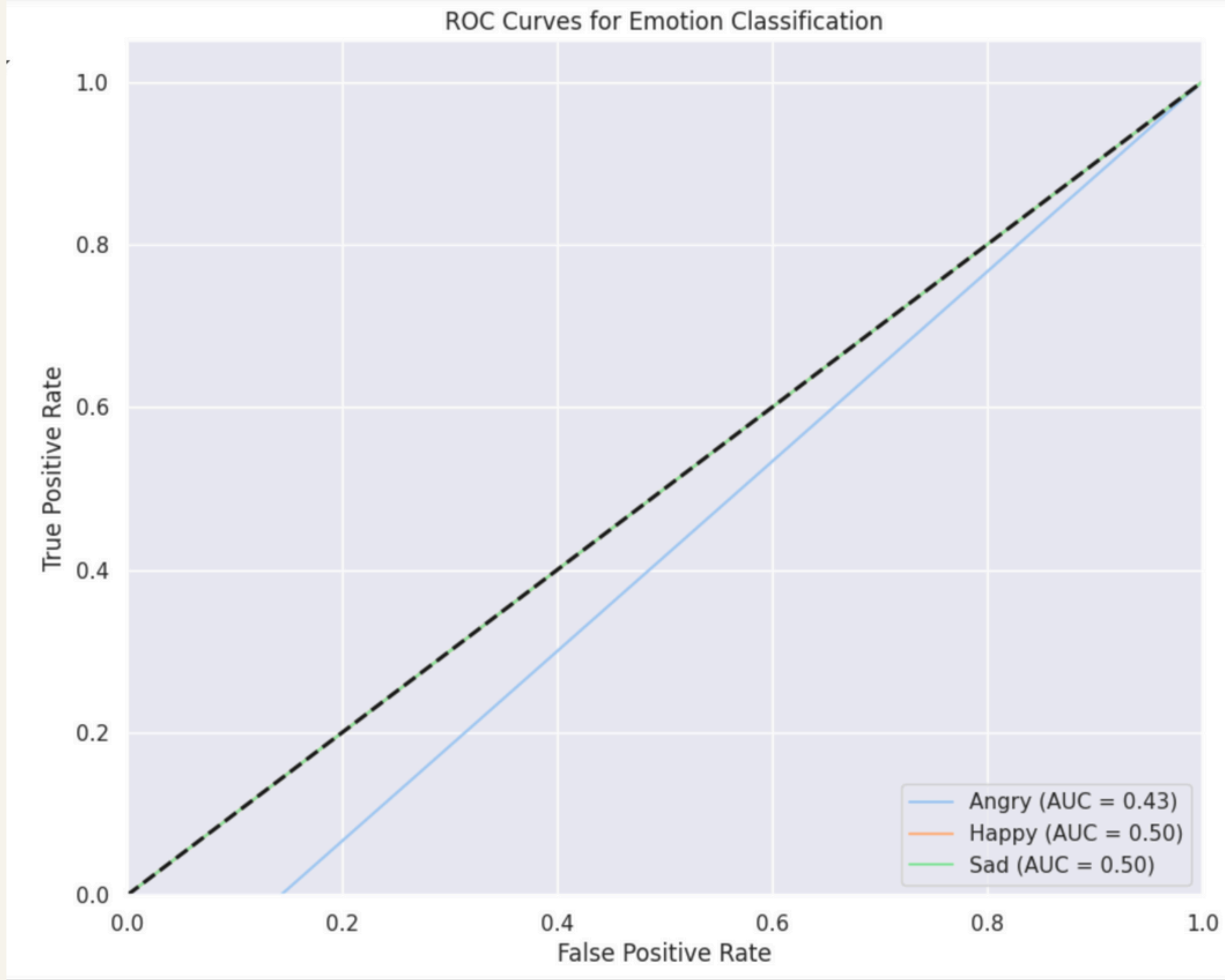
INSIGHTS:

- The model demonstrated a strong ability to distinguish between emotional states, even with subtle facial variations.
- This suggests that AI models like VGG16 can be reliably used in emotion-aware systems for applications like smart classrooms, healthcare monitoring, and user-adaptive interfaces.

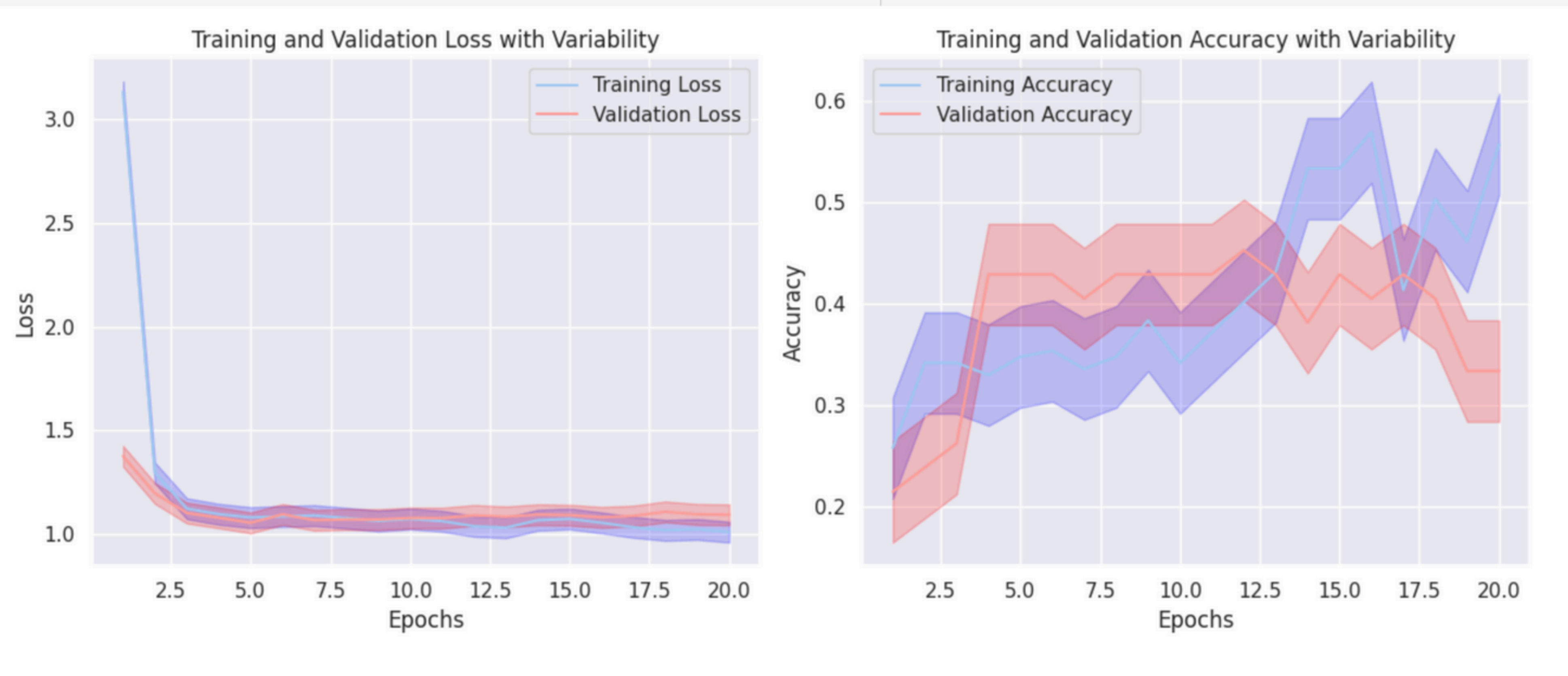
VISUALIZATION



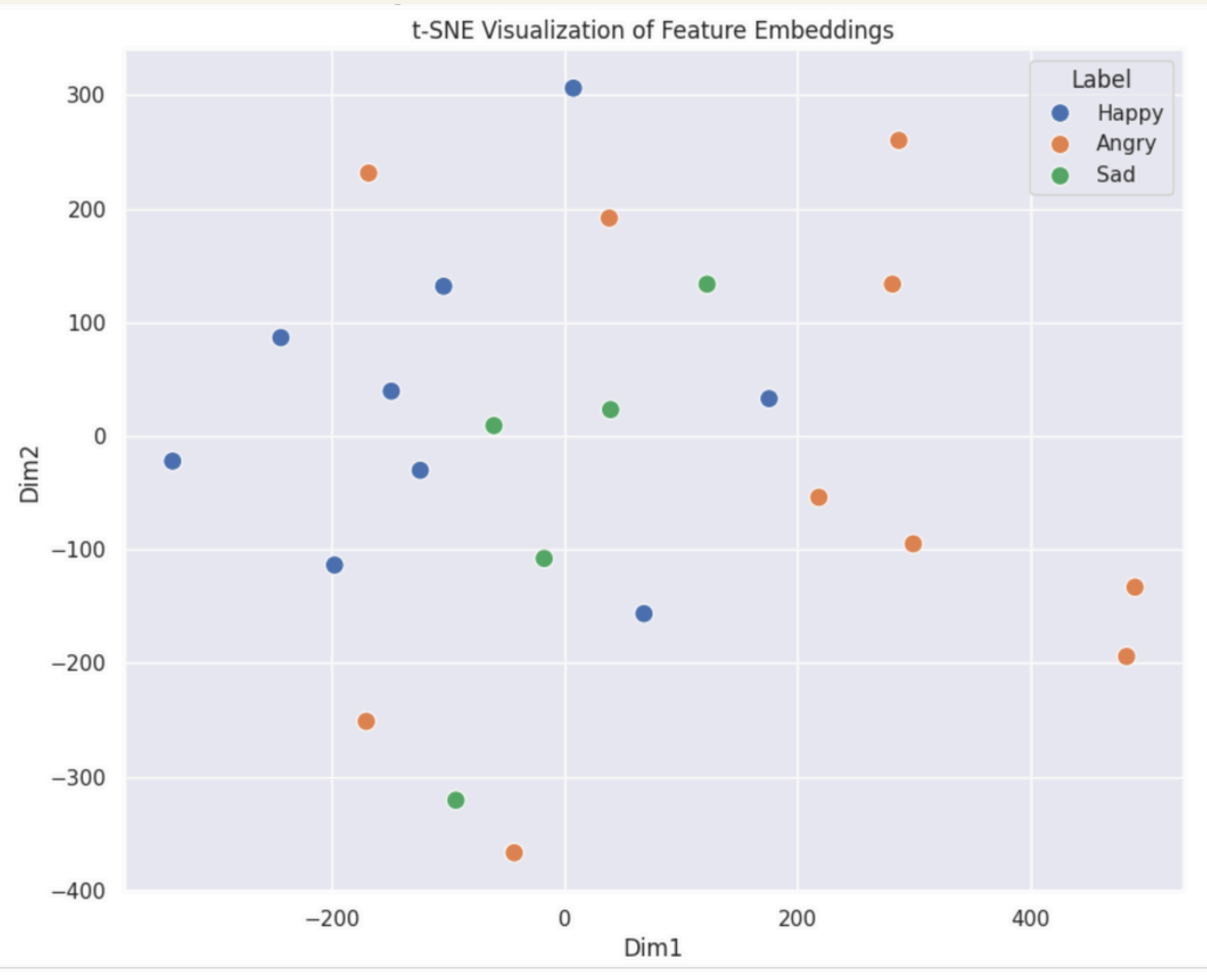
a)confusion Matrix for emotion classification



b)RDC Curves for emotion classification



c)Test and Train for epochs



d) t-SNE visualization

CONCLUSION

- This study demonstrates that deep learning models, particularly those based on transfer learning like VGG16, are highly effective in classifying facial emotions from images.
- Among the models evaluated, **VGG16** emerged as the best-performing architecture, achieving the highest validation accuracy and showing consistent generalization across Happy, Sad, and Angry classes. While the custom CNNand other baseline models performed moderately well, their results emphasized the superior feature extraction capability of pre-trained architectures.
- The analysis highlights the importance of selecting robust model architectures when applying deep learning to emotion recognition tasks.

Future work may include:

- Expanding classification to include additional emotions (e.g., surprise, fear, neutral).
- Deploying models on edge devices for real-time emotion detection in interactive applications.
- Enhancing performance through data augmentation, attention mechanisms, and temporal emotion tracking using video sequences.
- Integrating emotion classification with audio and gesture recognition for multi-modal human-computer interaction systems.