EMOTION DETECTION

Computer Vision (2023-2024) Project
Weidong Cai 1836167
Elena Jiang 1846716

Paper Ispiration

“Pre-Trained Multi-Modal Transformer for Pet Emotion Detection”

- The pet economy has been growing rapidly, especially during the Covid-19 pandemic, with increased pet adoptions.
- Animals show **emotional responses** similar to humans, and their facial expressions can be used to infer their current mood.
- There is a **lack of studies on animal emotions**, and most studies rely on human emotion recognition data and methods.

Several studies in the paper

- AI Advancements in Emotion Recognition
- Human Emotion Studies Across Disciplines
- Multimodal Data Analysis

This paper proposes a **pre-trained multimodal transformer emotion detection system**

- trained on a human emotion detection dataset
- applied to classify animal voice and expression data
- model pet-based emotion recognition focuses on cats and dogs

Our project

We concentrate on the **computer vision** → **Emotion Detection**

We begin by differentiating between human faces and pet faces.

Our dataset contains more than 2,000 RGB images with dimensions approximately around 1000x1000 pixels.

Human Emotion Detection

Pet Emotion Detection

Distinguishing between human faces and pet faces.

The Training dataset has 1421 images.
The Validation dataset has 365 images.
The Testing dataset has 244 images.

The training dataset has:
Class **'human'** has 725 images
Class **'pet'** has 696 images

We designed to set up and configure image data **generators** for training, validation, and testing datasets.

→ **ImageDataGenerator from Keras**

An example of plotting generator images

Distinguishing between human faces and pet faces.

MACHINE LEARNING

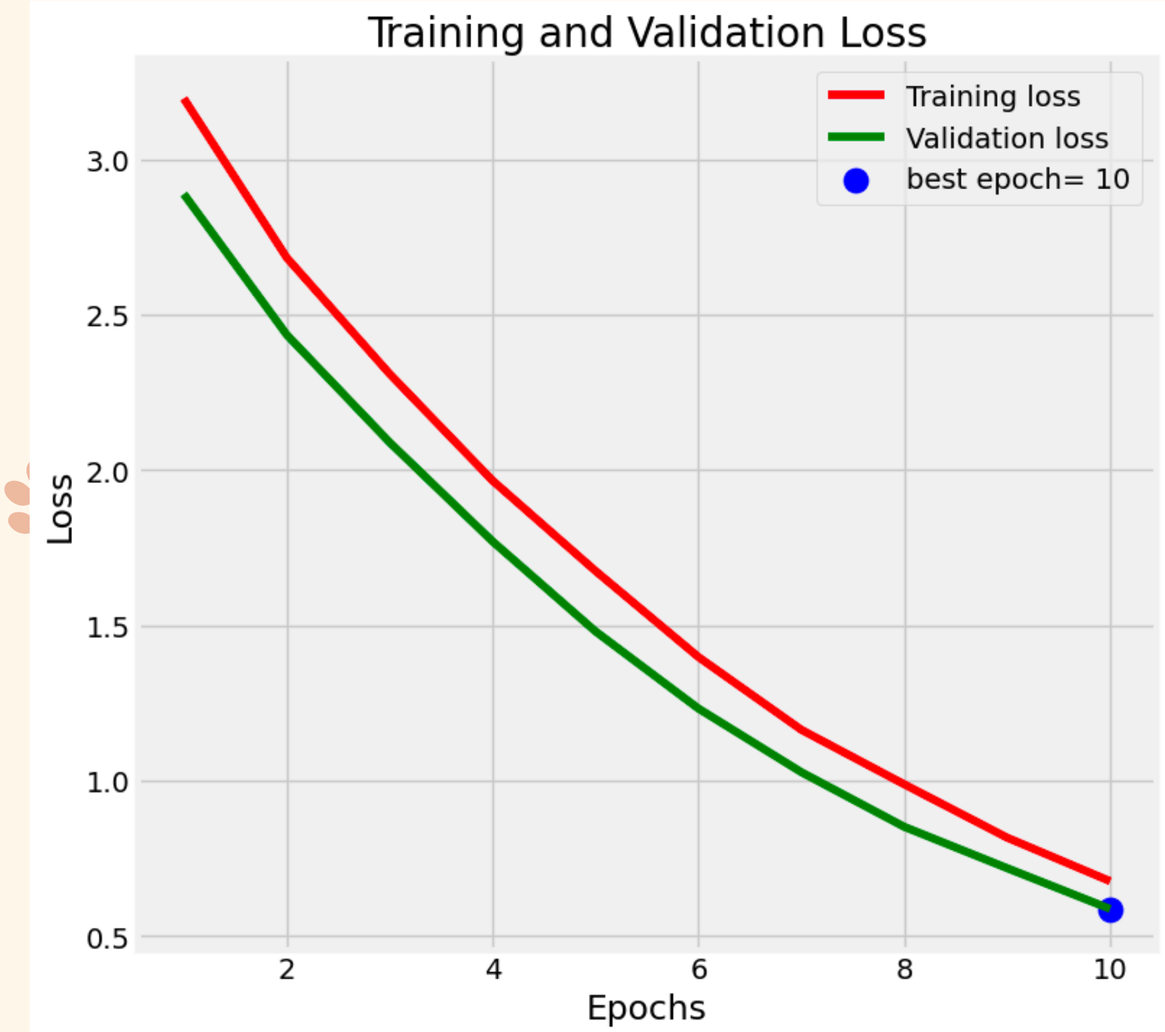
We use a **CNN model** based on the **EfficientNetB5** architecture for image classification. The model is built using **TensorFlow** and **Keras.**

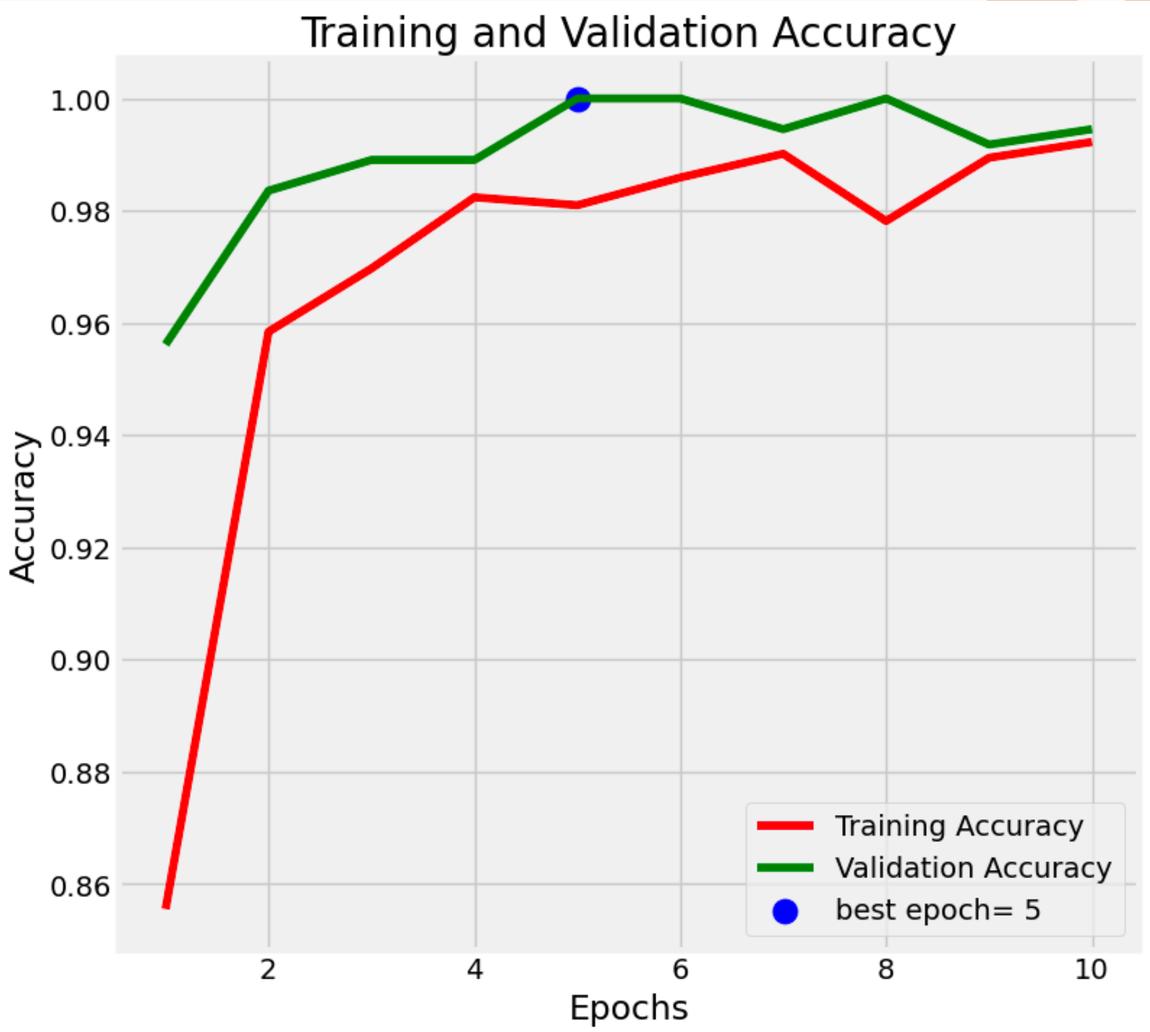
Sequential model: batch normalization, two dense layers with ReLU activations, a dropout layer for regularization and a final dense layer with softmax activation for multi-class classification.

For training the model we set **batch size** as **16** and we run **10 epochs**

TRAINING ACCURACY **100%**

Training history plot





Differentiation Model Evaluation

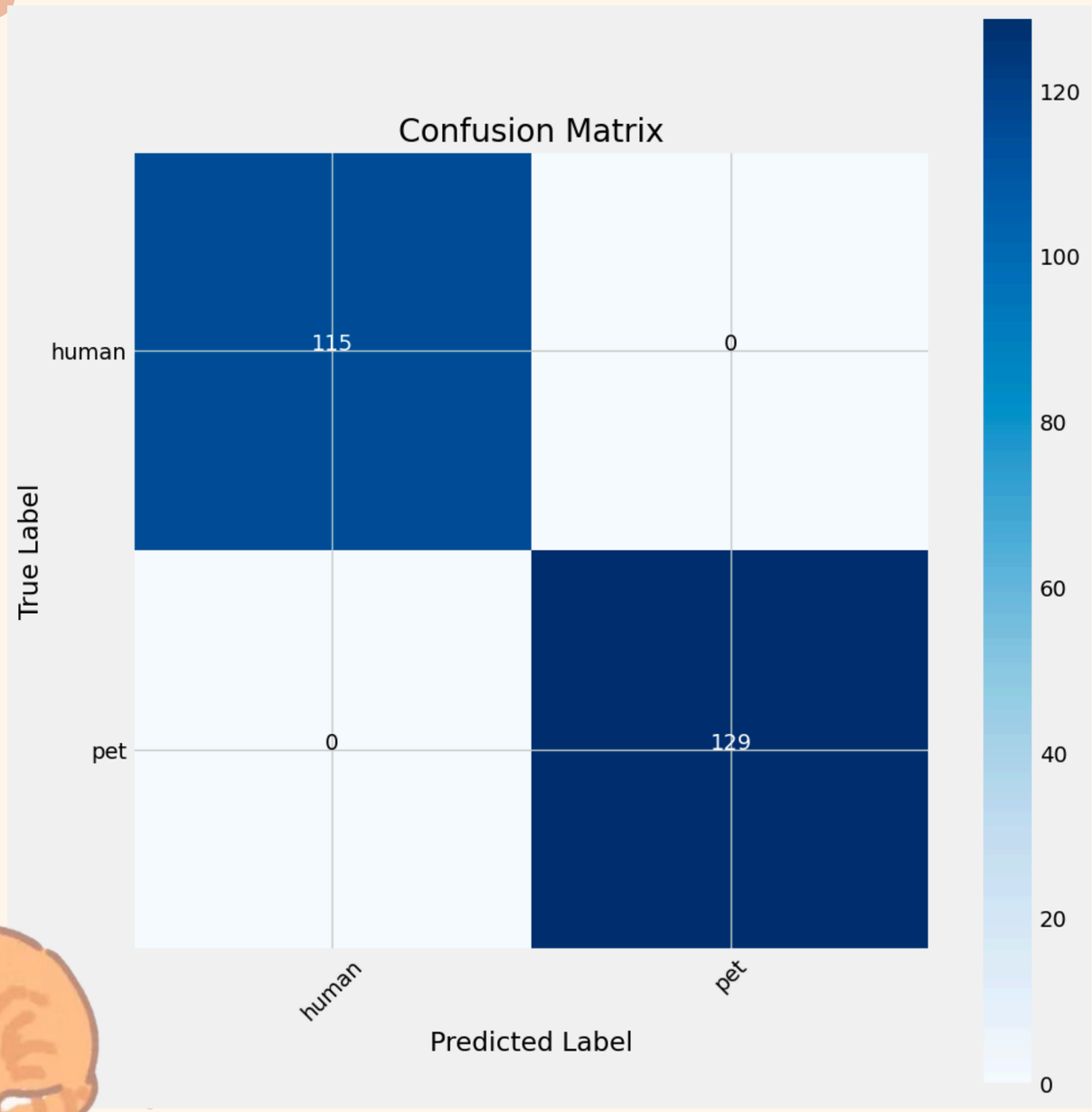
Train Loss: 0.57

Train Accuracy: 1.0

Test Loss: 0.58

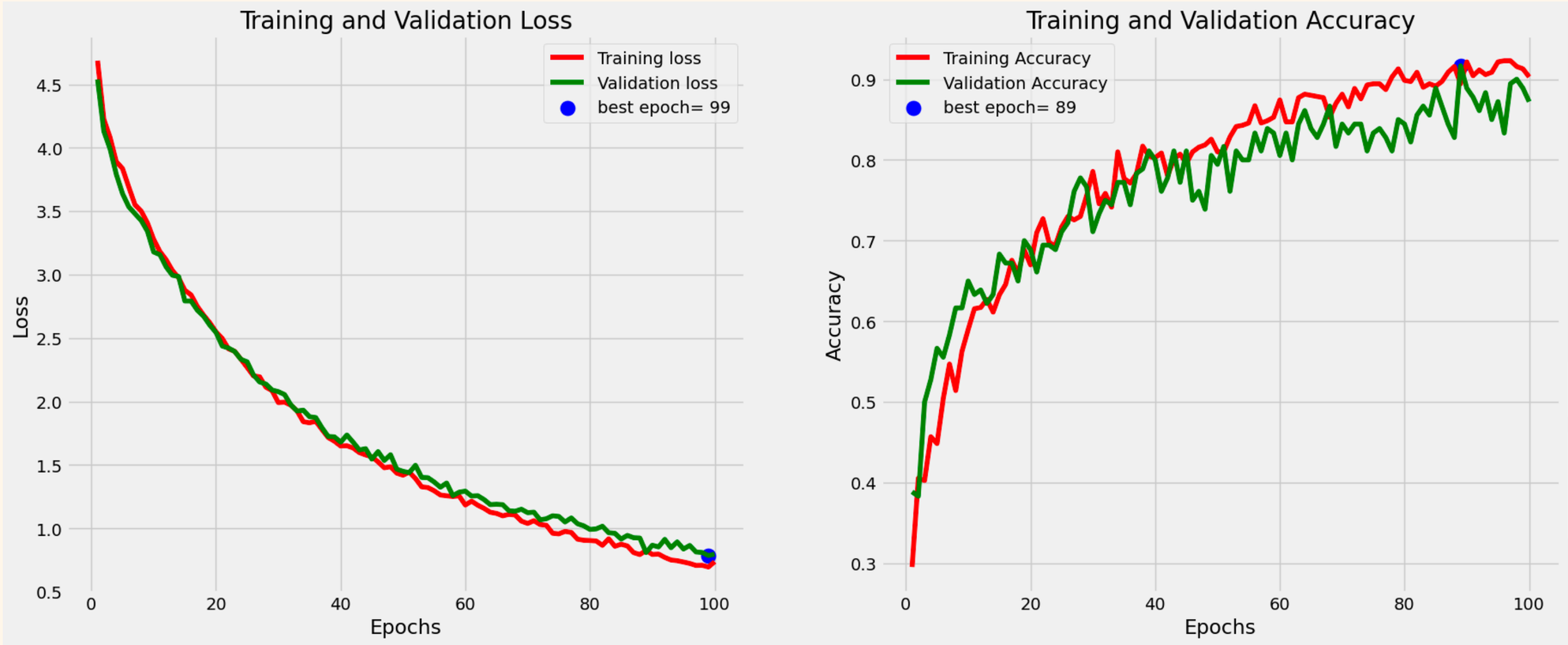
Test Accuracy: 1.0

	precision	recall	f1-score	support
human	1.00	1.00	1.00	115
pet	1.00	1.00	1.00	129



Pet Emotion Detection

- The Pet Emotion dataset has 1000 images divided in 4 classes
- CNN model based on the EfficientNetB5 architecture for image classification
- For training the model we set batch size as 16 and we run **100 epochs**



Pet Emotion Detection Model Evaluation

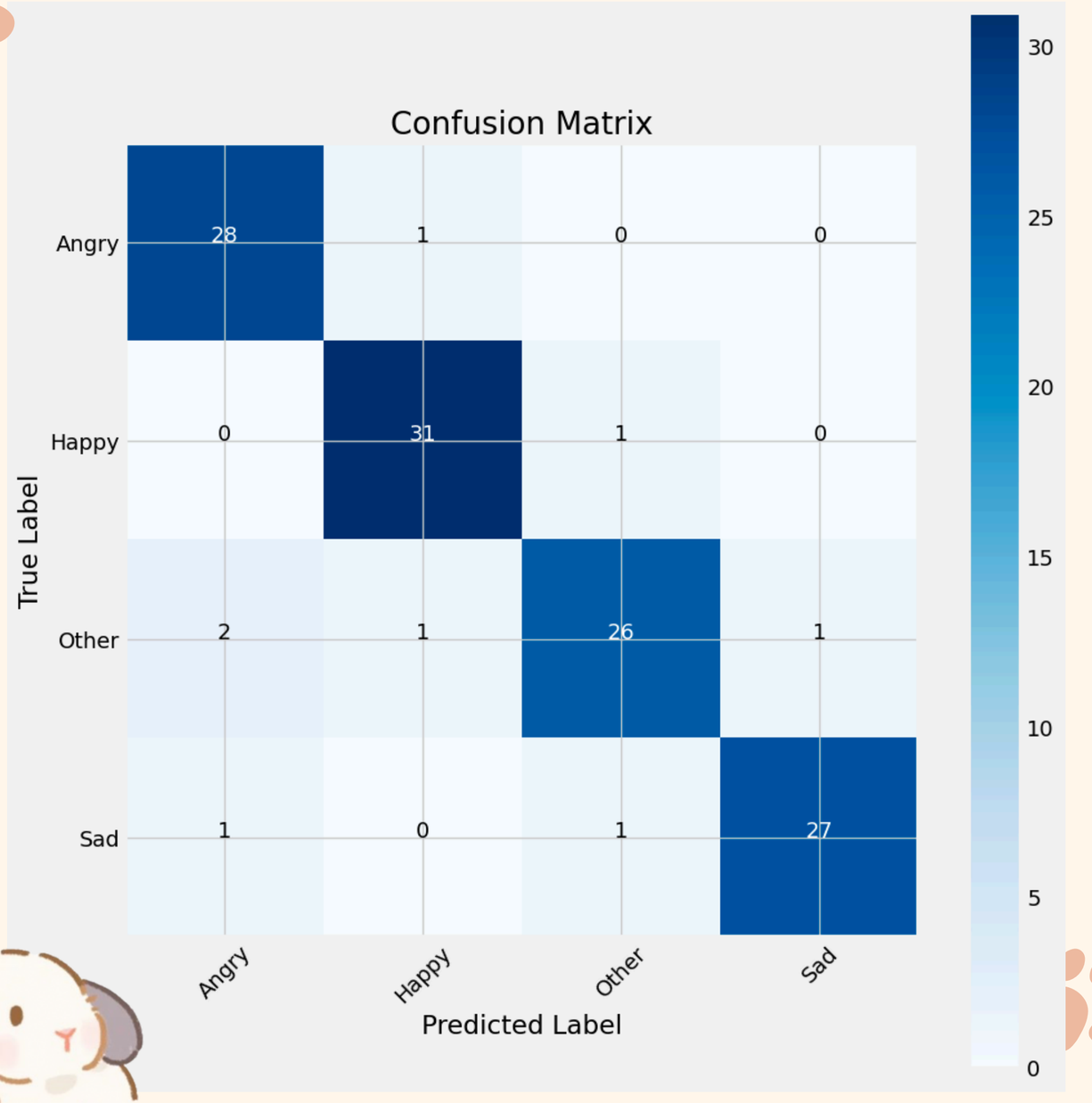
Train Loss: 0.55

Train Accuracy: 0.96

Test Loss: 0.76

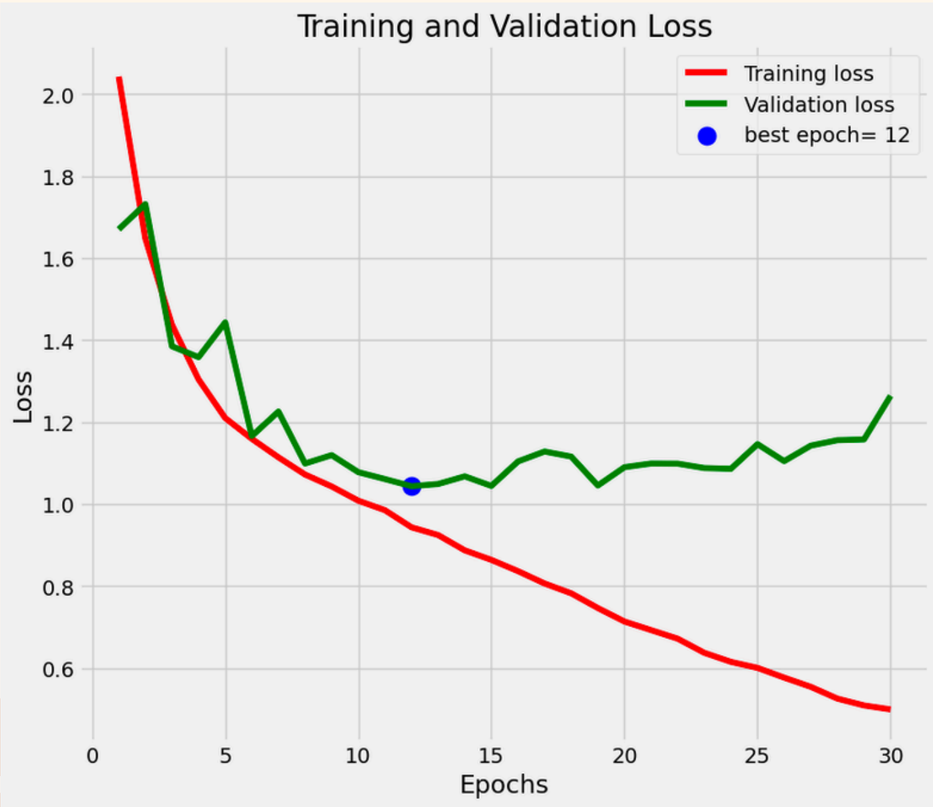
Test Accuracy: 0.88

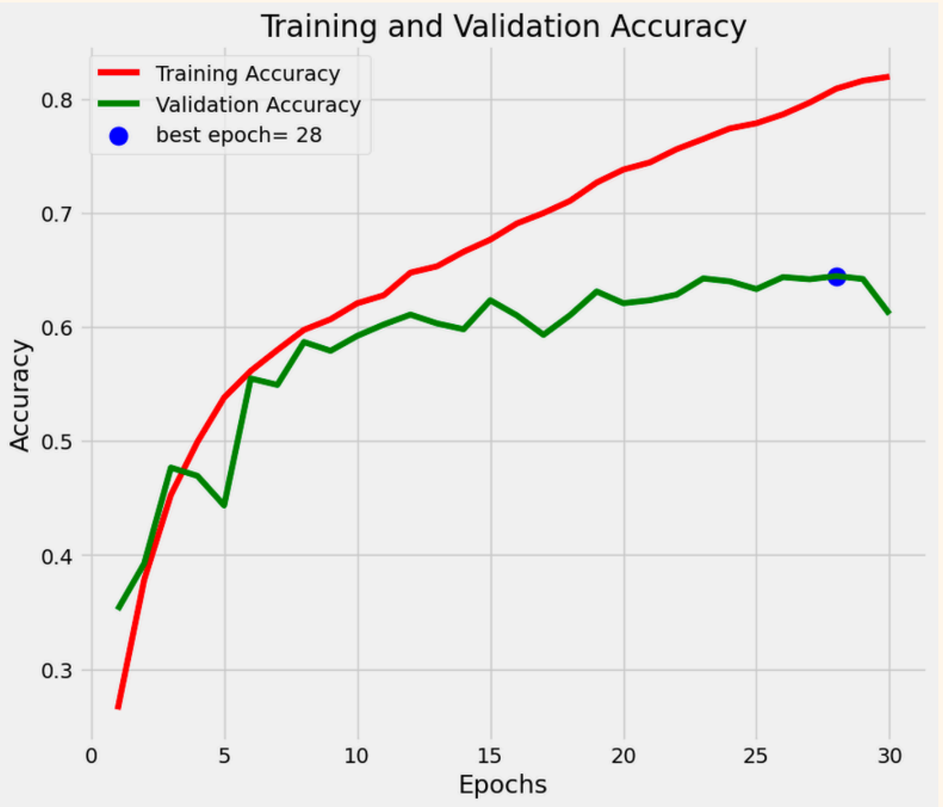
	precision	recall	f1-score	support
Angry	0.90	0.97	0.93	29
Happy	0.94	0.97	0.95	32
Other	0.93	0.87	0.90	30
Sad	0.96	0.93	0.95	29



Human Emotion Detection

- We use **FER-2013** dataset which has 28709 images divided in 7 classes
- 48x48 pixel **grayscale** images of faces
- We define a CNN model using **Sequential**
- We trained the model for **30 epochs**





Human Emotion Detection Model Evaluation

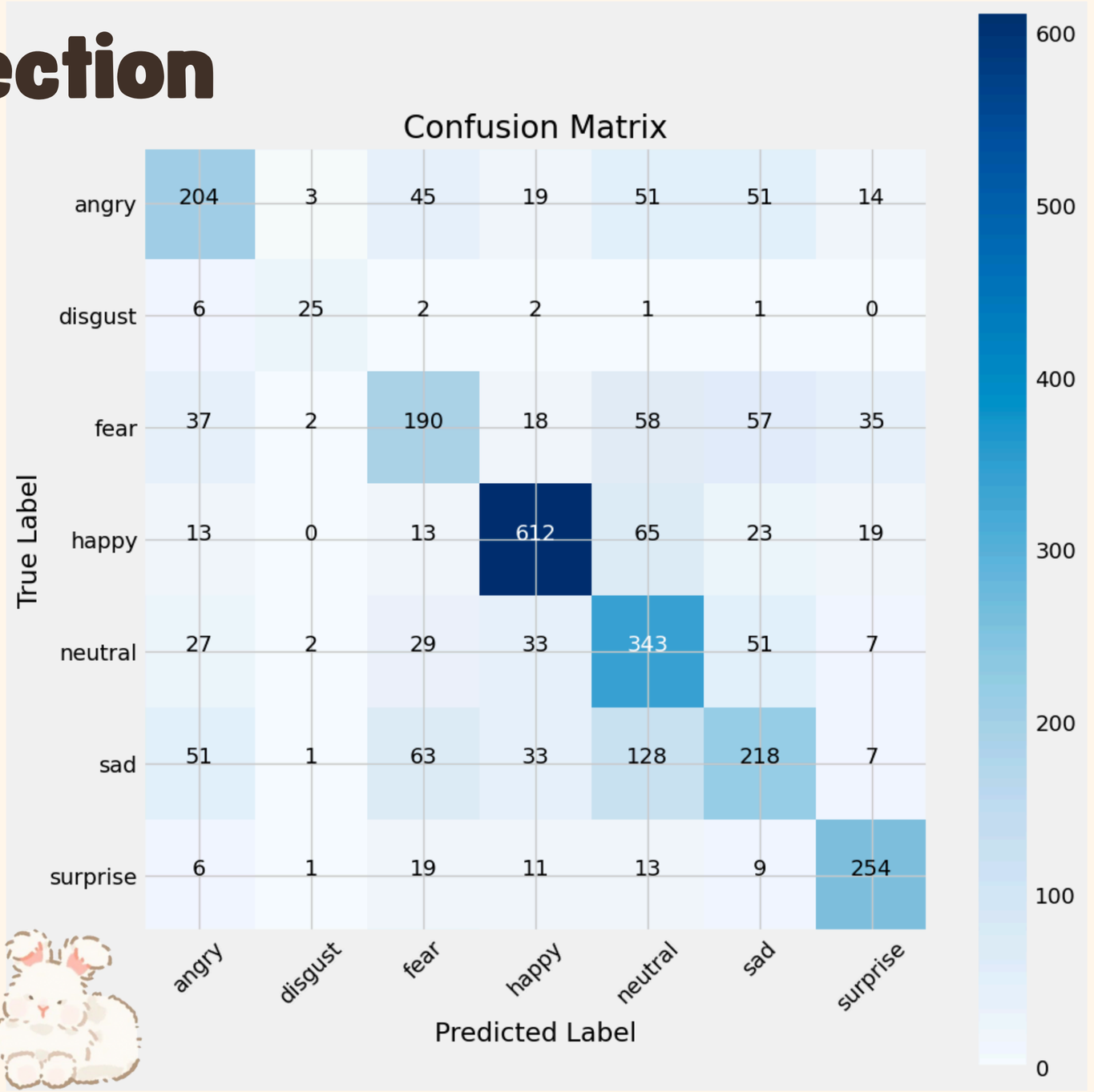
Train Loss: 0.34

Train Accuracy: 0.88

Test Loss: 1.11

Test Accuracy: 0.63

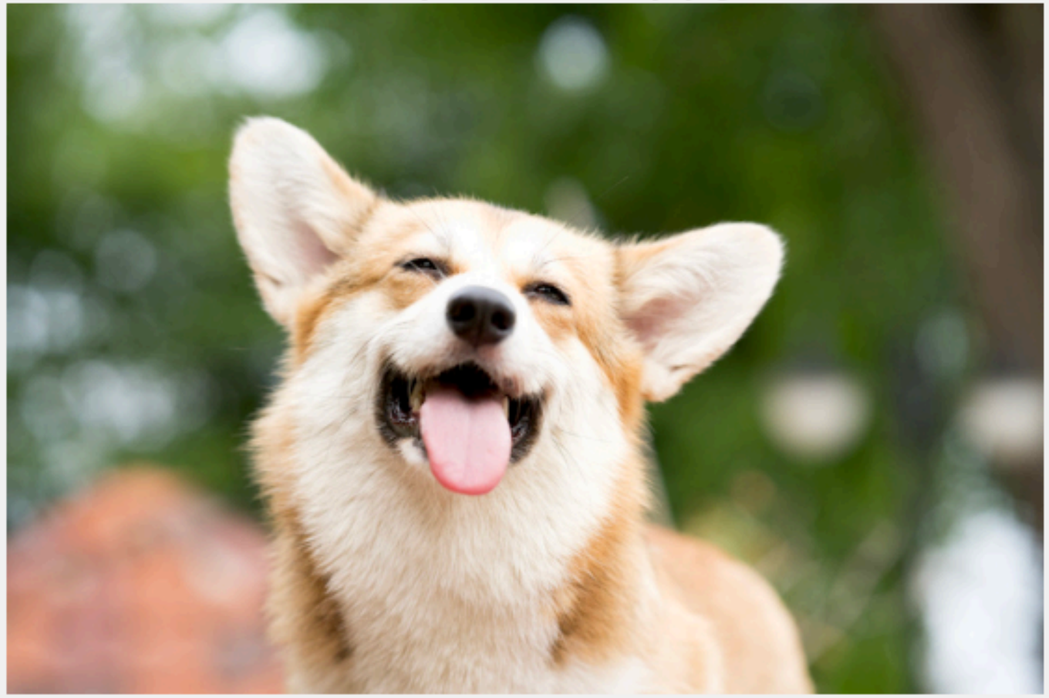
	precision	recall	f1-score	support
angry	0.59	0.53	0.56	387
disgust	0.74	0.68	0.70	37
fear	0.53	0.48	0.50	397
happy	0.84	0.82	0.83	745
neutral	0.52	0.70	0.60	492
sad	0.53	0.44	0.48	501
surprise	0.76	0.81	0.78	313



Demo

Examples of some results

The pet is Happy



The human is sad

Streamlit

Pet or Human Emotion Detection

Choose an image...

Drag and drop file here

Limit 200MB per file • PNG, JPG, JPEG

Browse files

Submit