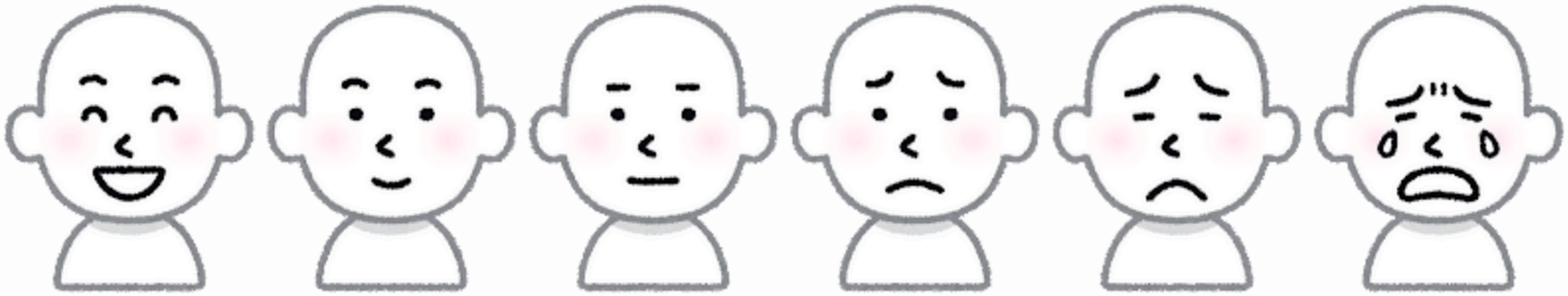


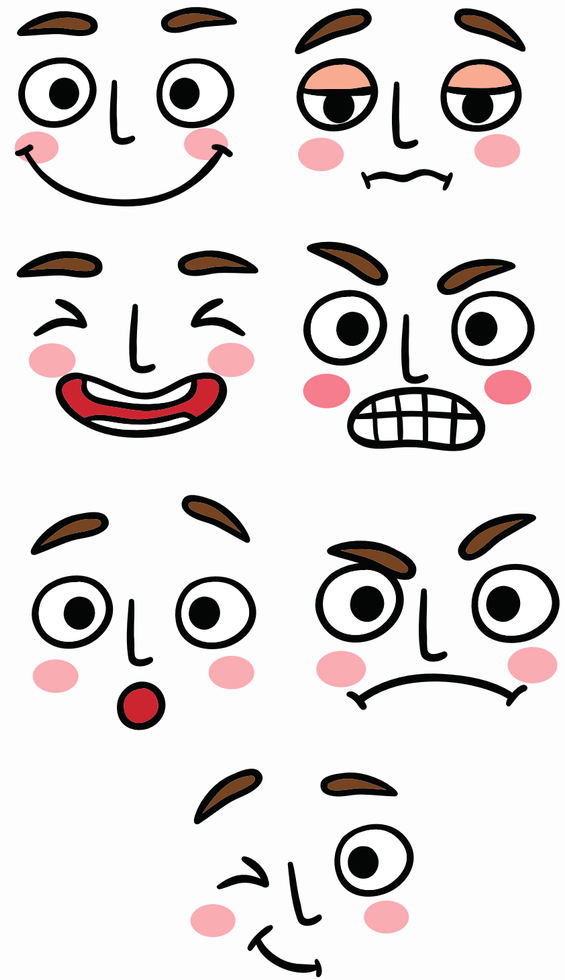
FACIAL EMOTION RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK (CNN)

DEEP LEARNING AND
ARTIFICIAL INTELLIGENCE

NATHALIA SILVA

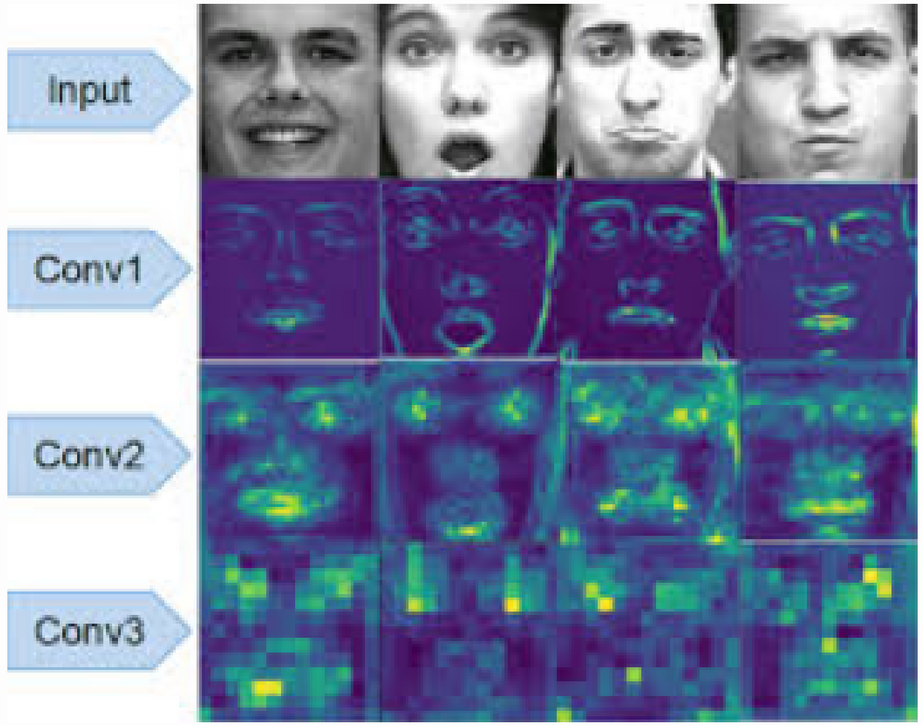
PROBLEM DEFINITION & MOTIVATION

- Facial expressions play a key role in human communication
- Automatic emotion recognition supports:
  - Human–computer interaction
  - Mental health monitoring
  - Smart devices and robotics
- Goal: Classify facial images into 7 emotion categories using a CNN



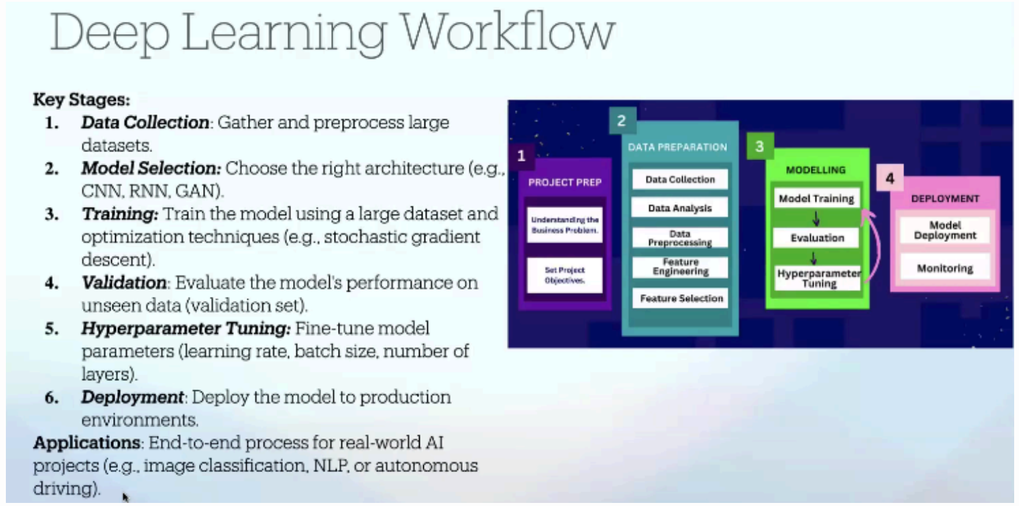
PROJECT OBJECTIVES

- Build a CNN capable of classifying emotions from 48×48 grayscale facial images
- Use image-based FER dataset with train/validation/test splits
- Apply data augmentation to improve generalization
- Evaluate using accuracy, loss curves, and confusion matrix
- Follow professor’s Deep Learning workflow



DEEP LEARNING WORKFLOW

There are six steps in the workflow: Project Prep, Data Preparation, Modeling, Training, Validation, Deployment/Results



Data Preparation and Augmentation

Dataset: FER2013 (Images Version) https://www.kaggle.com/datasets/deadskull7/fer2013/code?datasetId=28577&sortBy=dateRun&tab=bookmarked&excludeNonAccessedDatasources=false

- ~35,000 grayscale images (48×48)
- 7 emotion categories: Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral
- Dataset organized into:
- Train folder
- Validation split (20% of train)
- Test folder

```
TensorFlow version: 2.19.0
Train directory: /content/fer2013_images/train
Test directory : /content/fer2013_images/test
Train subdirectories: ['sad', 'surprise', 'disgust', 'angry', 'neutral', 'happy', 'fear']
Found 22968 images belonging to 7 classes.
Found 5741 images belonging to 7 classes.
Found 7178 images belonging to 7 classes.
Class indices: {'angry': 0, 'disgust': 1, 'fear': 2, 'happy': 3, 'neutral': 4, 'sad': 5, 'surprise': 6}
/usr/local/lib/python3.12/dist-packages/keras/src/layers/convolutional/base_conv.py:113: UserWarning: Do
```

Techniques applied to training images:

- Random rotation (±10°)
- Width/height shift (10%)
- Zoom (10%)
- Horizontal flip

DEEP LEARNING WORKFLOW

Modeling - CNN Architecture

- 3 convolutional blocks
- Each block:
- Conv2D → ReLU → Conv2D → ReLU → MaxPooling → Dropout
- Fully connected layer (256 units + dropout)
- Output: Softmax with 7 classes

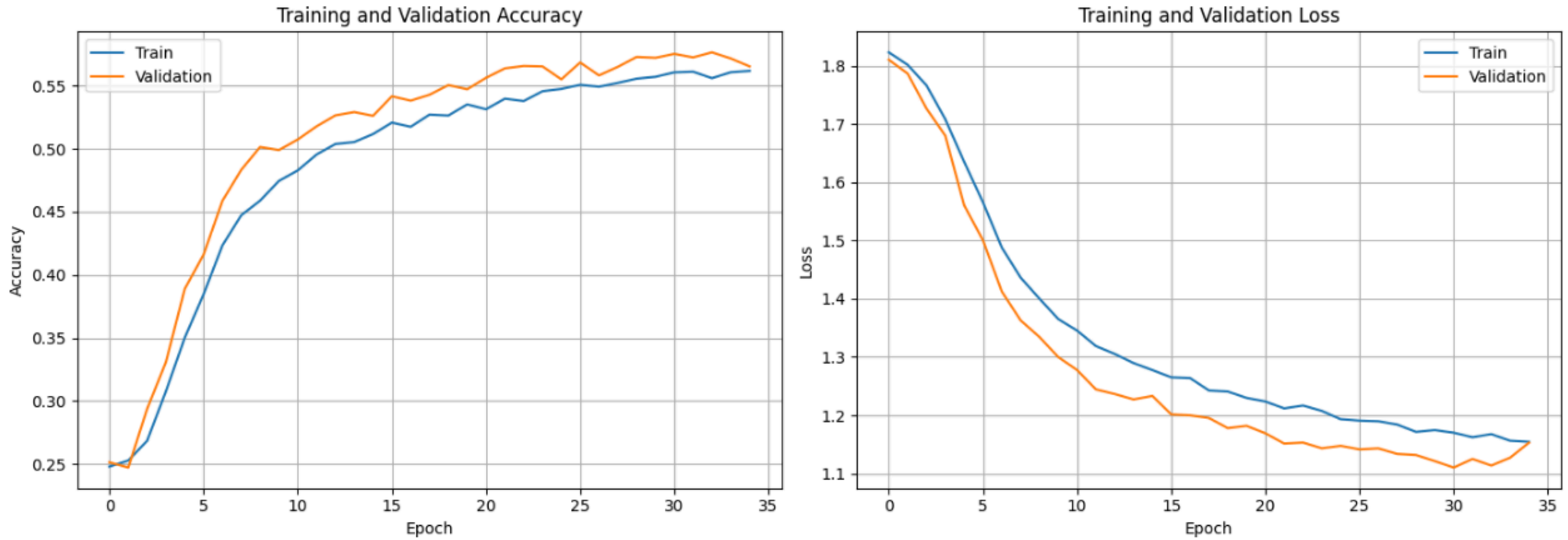
Training - Training Setup

- Optimizer: Adam (LR = 0.001)
- Loss: categorical crossentropy
- Batch size: 64
- Epochs: 35
- Callbacks: EarlyStopping, ReduceLROnPlateau, ModelCheckpoint

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 48, 48, 32)	320
conv2d_7 (Conv2D)	(None, 48, 48, 32)	9,248
max_pooling2d_3 (MaxPooling2D)	(None, 24, 24, 32)	0
dropout_4 (Dropout)	(None, 24, 24, 32)	0
conv2d_8 (Conv2D)	(None, 24, 24, 64)	18,496
conv2d_9 (Conv2D)	(None, 24, 24, 64)	36,928
max_pooling2d_4 (MaxPooling2D)	(None, 12, 12, 64)	0
dropout_5 (Dropout)	(None, 12, 12, 64)	0
conv2d_10 (Conv2D)	(None, 12, 12, 128)	73,856
conv2d_11 (Conv2D)	(None, 12, 12, 128)	147,584
max_pooling2d_5 (MaxPooling2D)	(None, 6, 6, 128)	0
dropout_6 (Dropout)	(None, 6, 6, 128)	0
flatten_1 (Flatten)	(None, 4608)	0
dense_2 (Dense)	(None, 256)	1,179,904
dropout_7 (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 7)	1,799

Total params: 1,468,135 (5.60 MB)
Trainable params: 1,468,135 (5.60 MB)
Non-trainable params: 0 (0.00 B)

Training Curves - Accuracy & Loss Curves




```
l13/113 ━━━━━━━━━━━━━━━━━━━━ 3s 26ms/step - accuracy: 0.5354 - loss: 1.1813

Test Accuracy: 59.18%
l13/113 ━━━━━━━━━━━━━━━━━━━━ 3s 24ms/step
```

Test Results & Evaluation

- Test Accuracy: ~59.53%
- Strong accuracy for “happy” and “surprise”
- Lower performance for “fear” and “disgust”

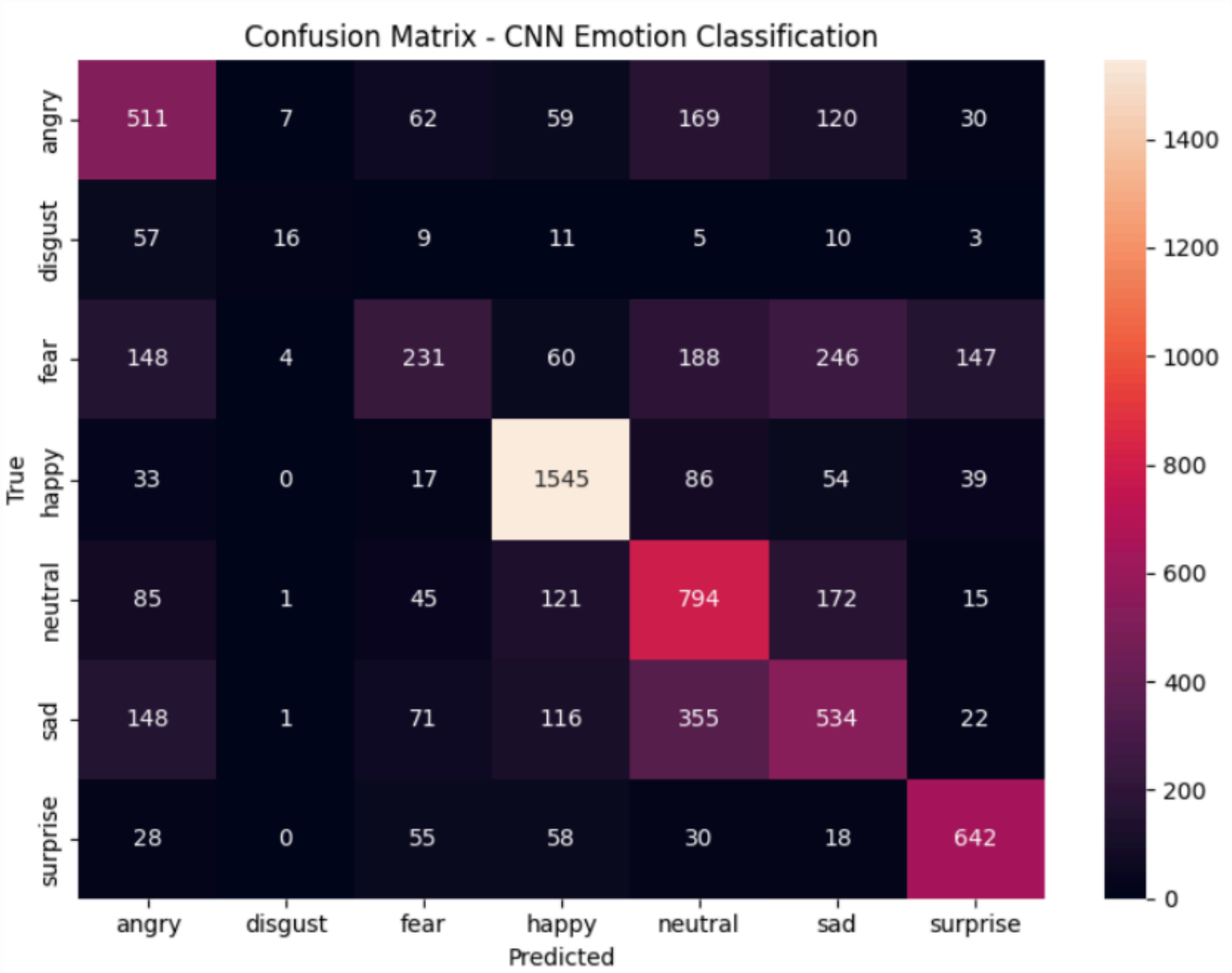
Classification Report

- Highest F1-score: Happy (~0.83)
- High accuracy: Surprise (~0.74), Neutral (~0.56)
- Difficult classes: Fear (~0.31), Disgust (~0.23)
-

```
113/113 ━━━━━━━━━━━━━━━━━━━━ 3s 23ms/step - accuracy: 0.5388 - loss: 1.1570

Test Accuracy: 59.53%
113/113 ━━━━━━━━━━━━━━━━━━━━ 3s 21ms/step

Classification Report:
              precision    recall  f1-score   support

       angry       0.51      0.53      0.52       958
     disgust       0.55      0.14      0.23       111
        fear       0.47      0.23      0.31      1024
       happy       0.78      0.87      0.83      1774
     neutral       0.49      0.64      0.56      1233
         sad       0.46      0.43      0.44      1247
    surprise       0.71      0.77      0.74       831

    accuracy                           0.60      7178
   macro avg       0.57      0.52      0.52      7178
weighted avg       0.58      0.60      0.58      7178
```

Confusion Matrix

- Happy and Surprise: very high recall
- Neutral: moderate accuracy
- Fear and Disgust: hardest classes
- Confusions mostly occur between visually similar expressions

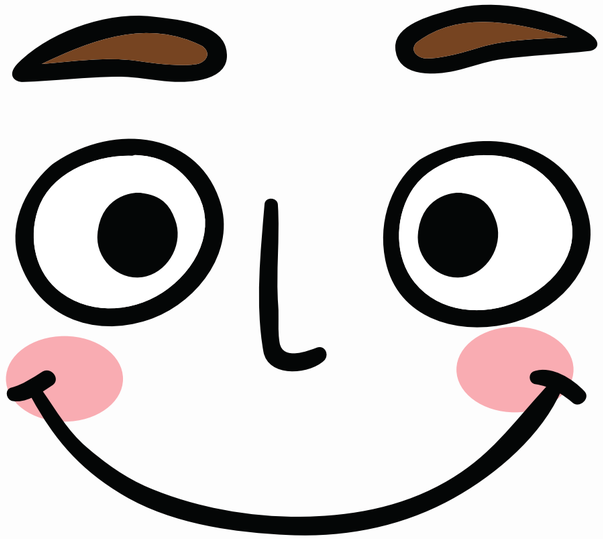


DISCUSSION& POSSIBLE IMPROVEMENTS

- Good performance for expressive faces (happy, surprise)
- Model struggles with subtle emotions
- Resolution (48×48) limits performance
- Dataset imbalance affects results

Improvements:

- More training data
- Class balancing
- Use deeper models (ResNet, VGG)
- Transfer learning
- Try Vision Transformers (ViT)
- Build real-time emotion recognition with webcam



THANK YOU!