VIDEO FORGERY DETECTION

DEEP LEARNING-FALL 2024

PRESENTED BY: GROUP 6

- KHUSH SHAH
- VISHAL BAKSHI
- NENA BEECHAM
- BHARAT KHANDELWAL

GUIDED BY: PROF. AMIR JAFARI

PRESENTATION OUTLINE

- Introduction
- Dataset
- Architecture
- Models
- Implementation
- Future Scope
- Questions and Answers

INTRODUCTION

- Image forgery has been one of the rising issues with growth of social media.
- Every day about 70000 hours of video content is uploaded on a daily basis.
- Impact of forgery can be immense politically, and is done for the purpose of revenge.

To overcome the serious issue of forgery, we have implemented multiple computer vision models and techniques to detect the forgery in images/videos.

DATASET: FAKE IMAGES GENERATION

- GAN
- MASKING
- FACESWAP
- DEEPFACELAB
- FACEAPP

REAL VS FAKE DATA

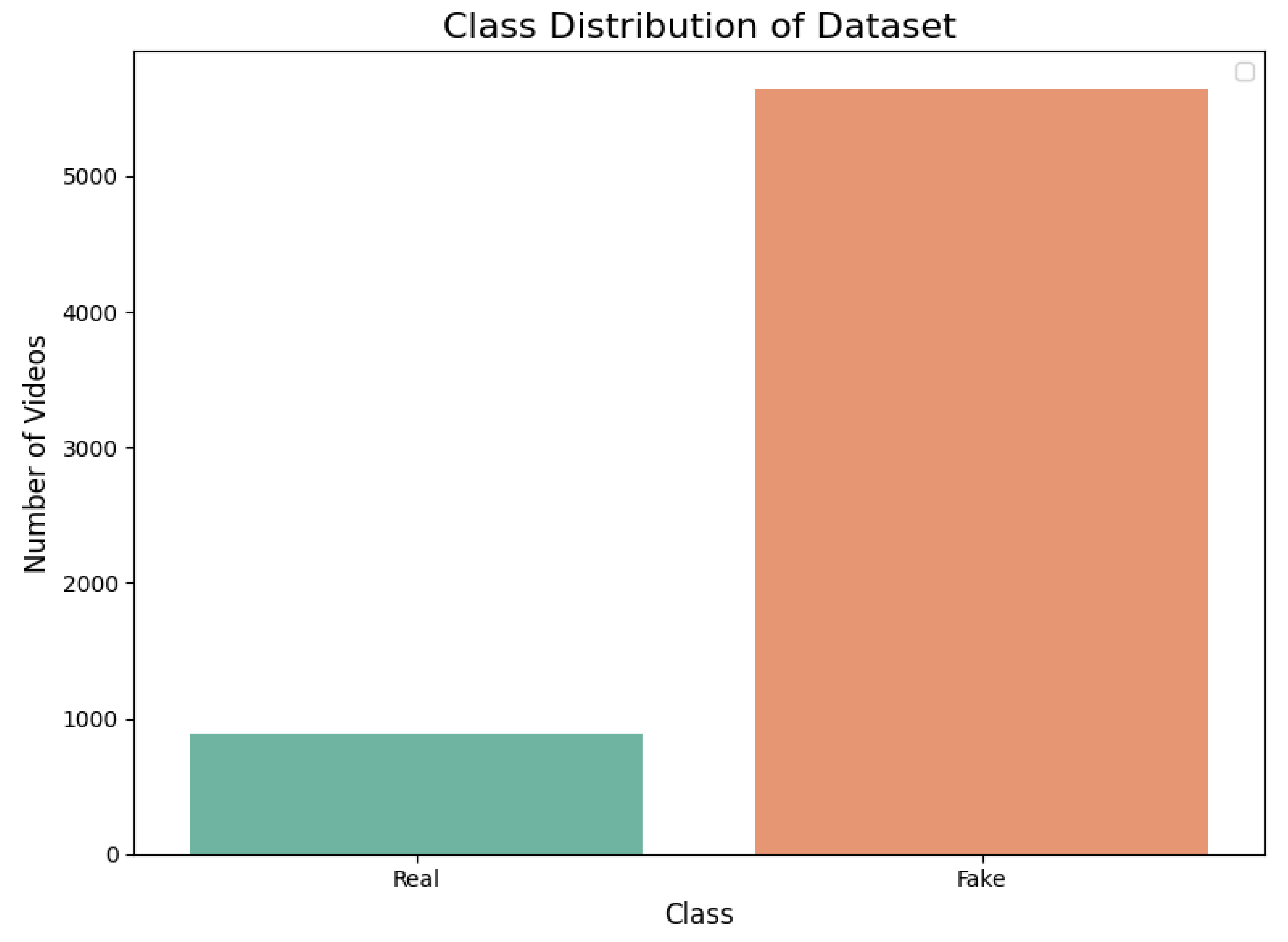
Source: Celeb DF v2

Dataset segmentation:

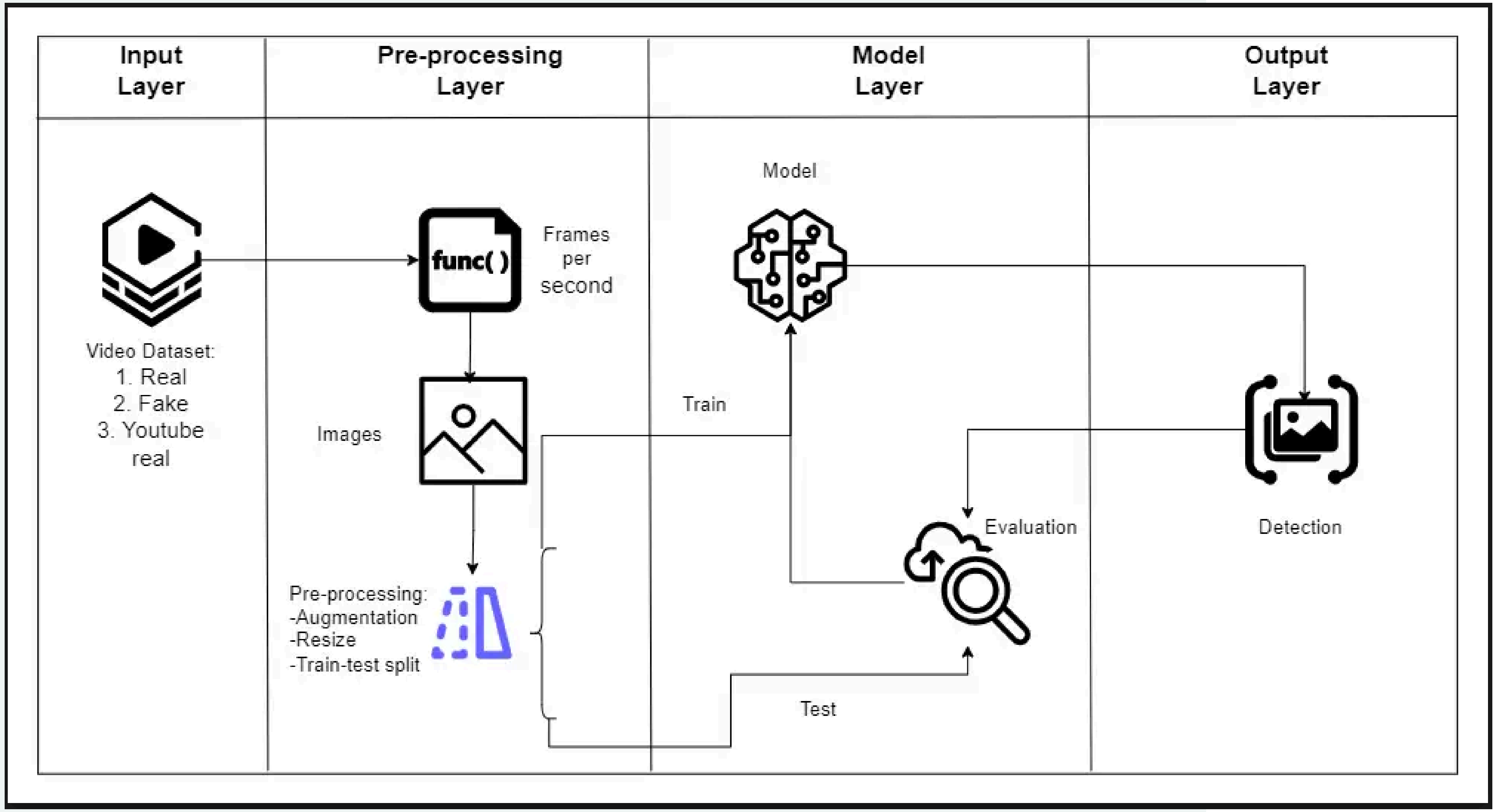
Real

Deepfake

CLASS DISTRIBUTION OF DATASET



DATA PIPELINE



Data Loaders

1. tf.data from generator
2. tf.flow from directory

MODEL DEVELOPMENT

1

CNN + GRU

2

CNN + DENSE

CNN + GRU

1. The model processes pre-extracted features from videos obtained using an InceptionV3 CNN.
2. It employs three stacked GRU layers to capture temporal dependencies in the sequence, with normalization layers following each GRU layer to stabilize training.
3. The output of the GRU layers is passed through fully connected dense layers to reduce the temporal representation for binary classification (e.g., real vs. fake).
4. The final layer is a single neuron with a sigmoid activation, outputting the probability for binary classification.
5. The focal loss function is used to address the significant class imbalance in the dataset.

Layer (type)	Output Shape	Param #	Connected to
frame_features (InputLayer)	(None, 16, 2048)	0	-
mask (InputLayer)	(None, 16)	0	-
gru (GRU)	(None, 16, 64)	405,888	frame_features[0… mask[0][0]
batch_normalizatio… (BatchNormalizatio…	(None, 16, 64)	256	gru[0][0]
gru_1 (GRU)	(None, 16, 32)	9,408	batch_normalizat…
batch_normalizatio… (BatchNormalizatio…	(None, 16, 32)	128	gru_1[0][0]
gru_2 (GRU)	(None, 16)	2,400	batch_normalizat…
dense (Dense)	(None, 32)	544	gru_2[0][0]
dropout (Dropout)	(None, 32)	0	dense[0][0]
dense_1 (Dense)	(None, 1)	33	dropout[0][0]

Total params: 418,657 (1.60 MB)
Trainable params: 418,465 (1.60 MB)
Non-trainable params: 192 (768.00 B)

REUSLTS

MODEL TEST RESULTS

	precision	recall	f1-score
0	0.6	0.9	0.72
1	0.99	0.94	0.96
accuracy			0.93
macro avg	0.8	0.92	0.84
weighted avg	0.95	0.93	0.94
F1-Macro score	0.84		

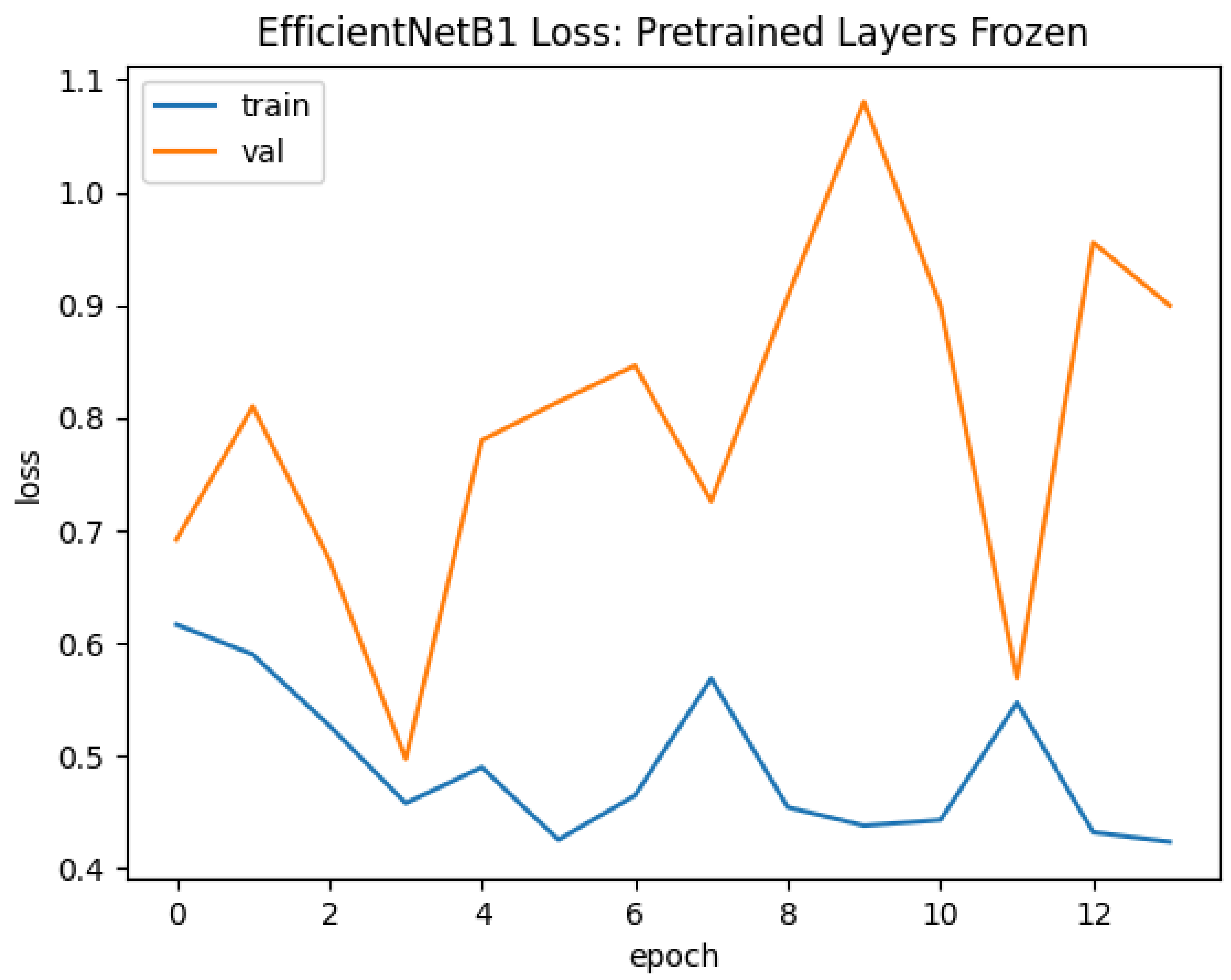
RESULT ON FAKE VIDEO OTHER THAN THE DATASET

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 154ms/step
Prediction: Fake (0.6670)
The video is classified as Fake with confidence 0.6670
```

	Precision	Recall	f1-score
Real	0.45	0.04	0.07
Fake	0.87	0.99	0.93
Accuracy			0.87
macro avg	0.66	0.51	0.5
weighted avg	0.82	0.87	0.81
F1 Macro Score	0.4969		

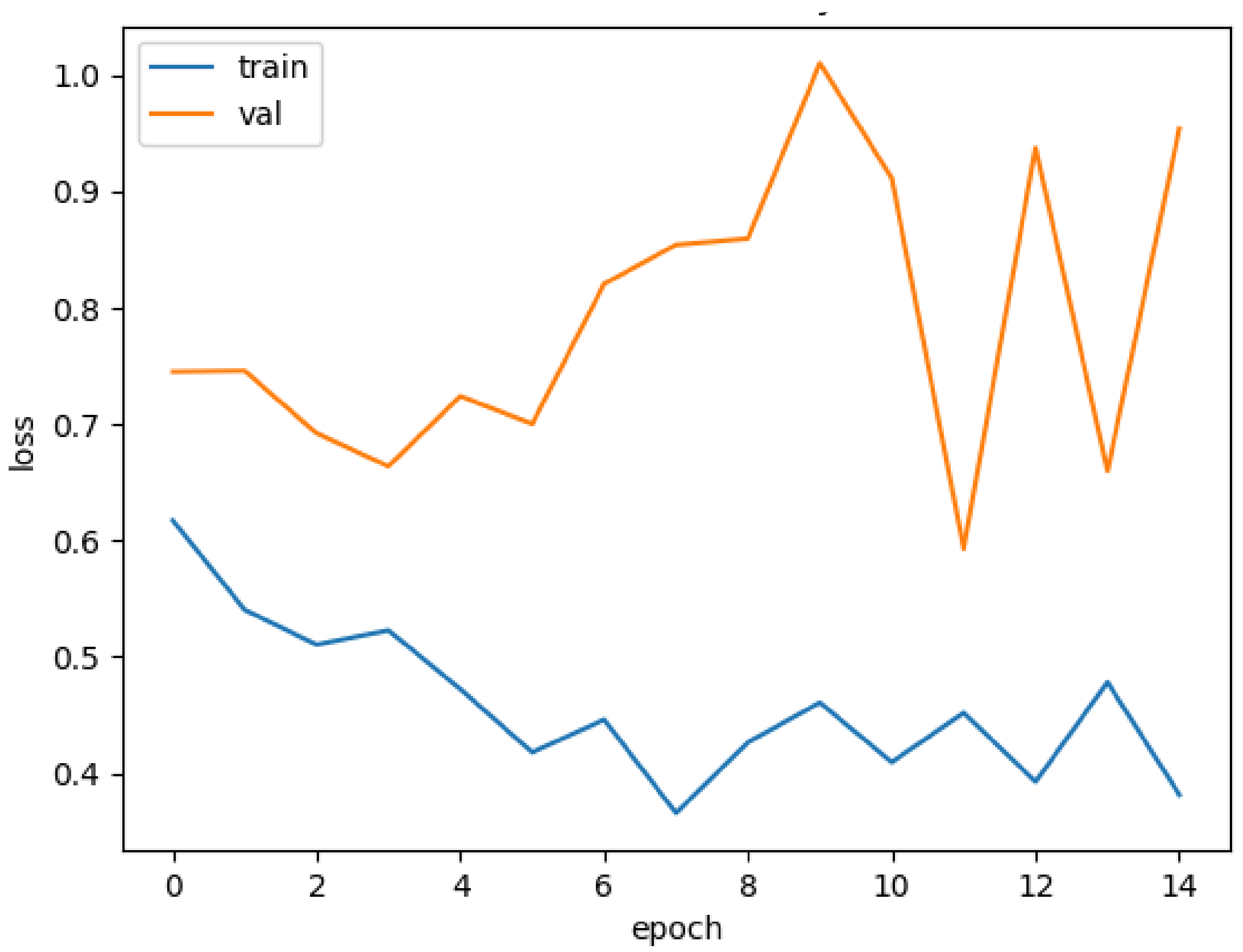
CNN + DENSE

1. Classification Head
    a. Add a GlobalAveragePooling and several dense layers together, taking base model's output as input
    b. Final Dense Layer using sigmoid transfer function for binary classification
2. Feature Extraction
    a. Tried ResNet50, VGG16, InceptionV3, and EfficientNetB1
    b. VGG16 and EfficientNetB1 produced best metrics
3. Finetuning of Base Model
    a. Experiment with unfreezing base layers
    b. 5-30 layers depending on the pretrained model

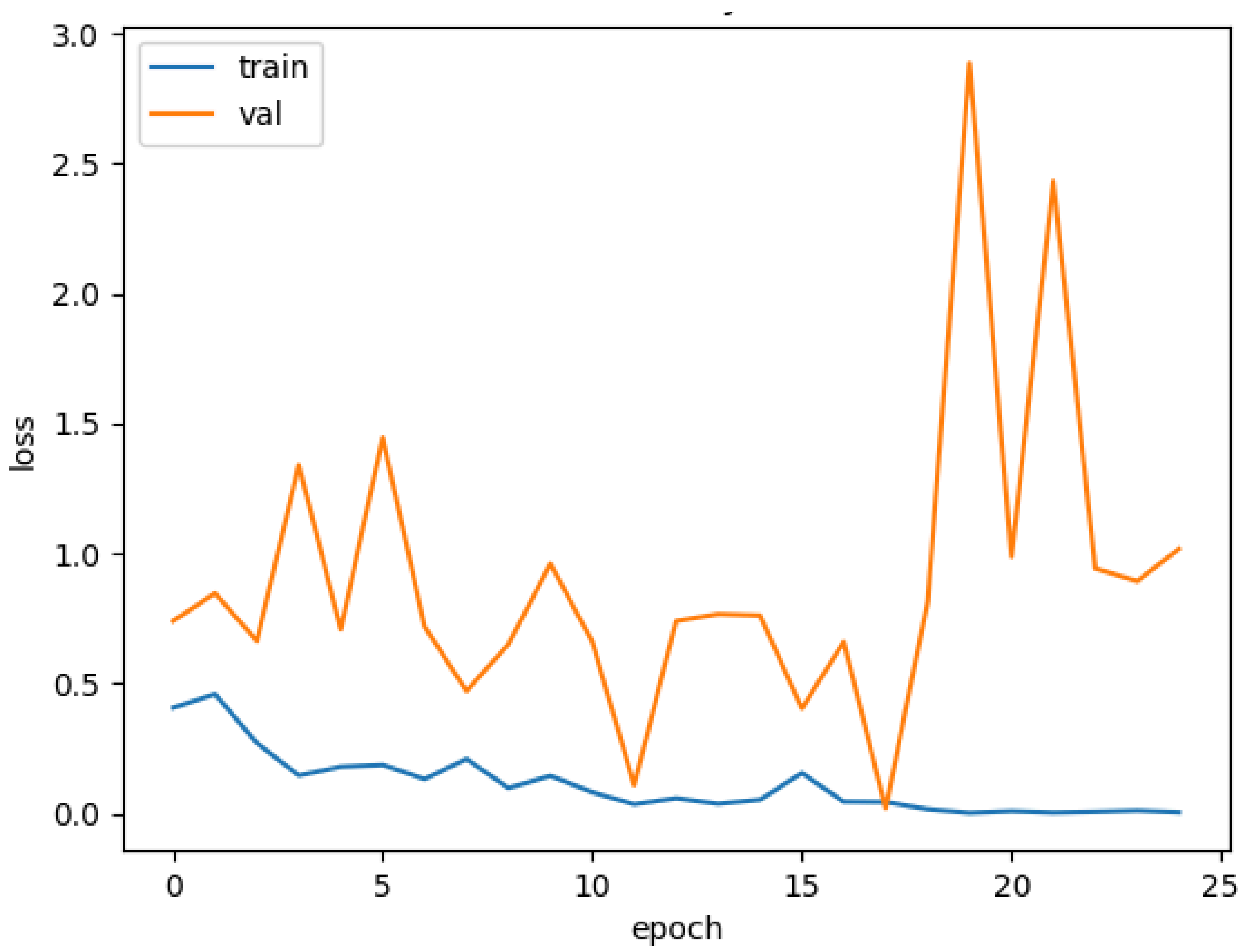


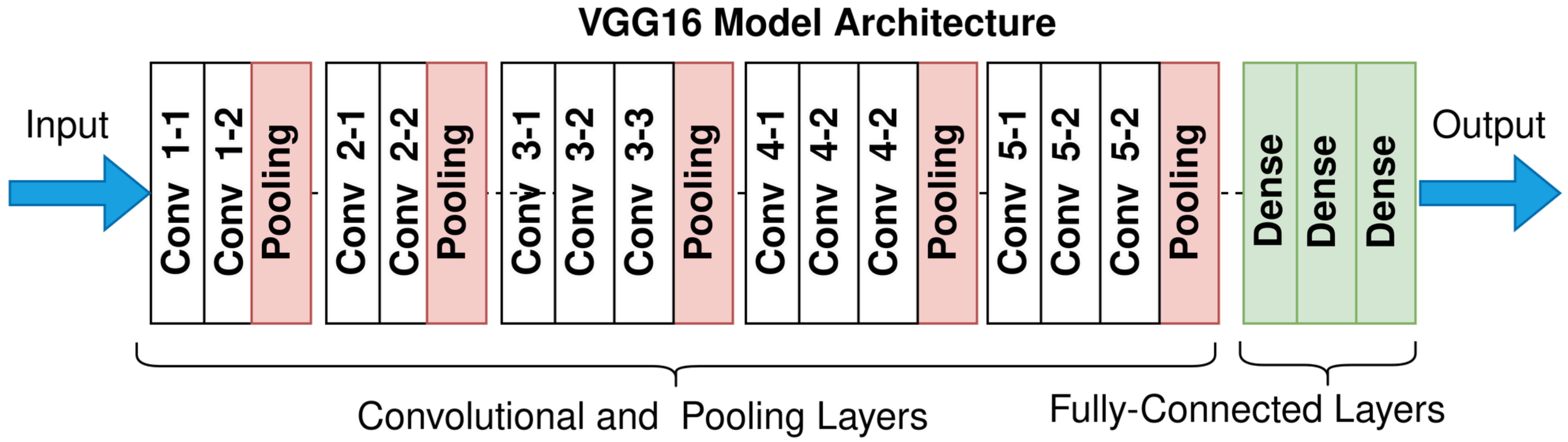
VGG 16 TRAINING AND ARCHITECTURE

VGG16 LOSS: PRETRAINED LAYERS FROZEN



VGG16 LOSS: ALL LAYERS UNFROZEN





NENA

VGG16 ON TEST CELEB DATA

Test F1 Macro Score: 75.52%
Test Accuracy: 83.66%

	precision	recall	f1-score	support
0	0.59	0.64	0.61	1180
1	0.91	0.89	0.90	4660
accuracy			0.84	5840
macro avg	0.75	0.76	0.76	5840
weighted avg	0.84	0.84	0.84	5840

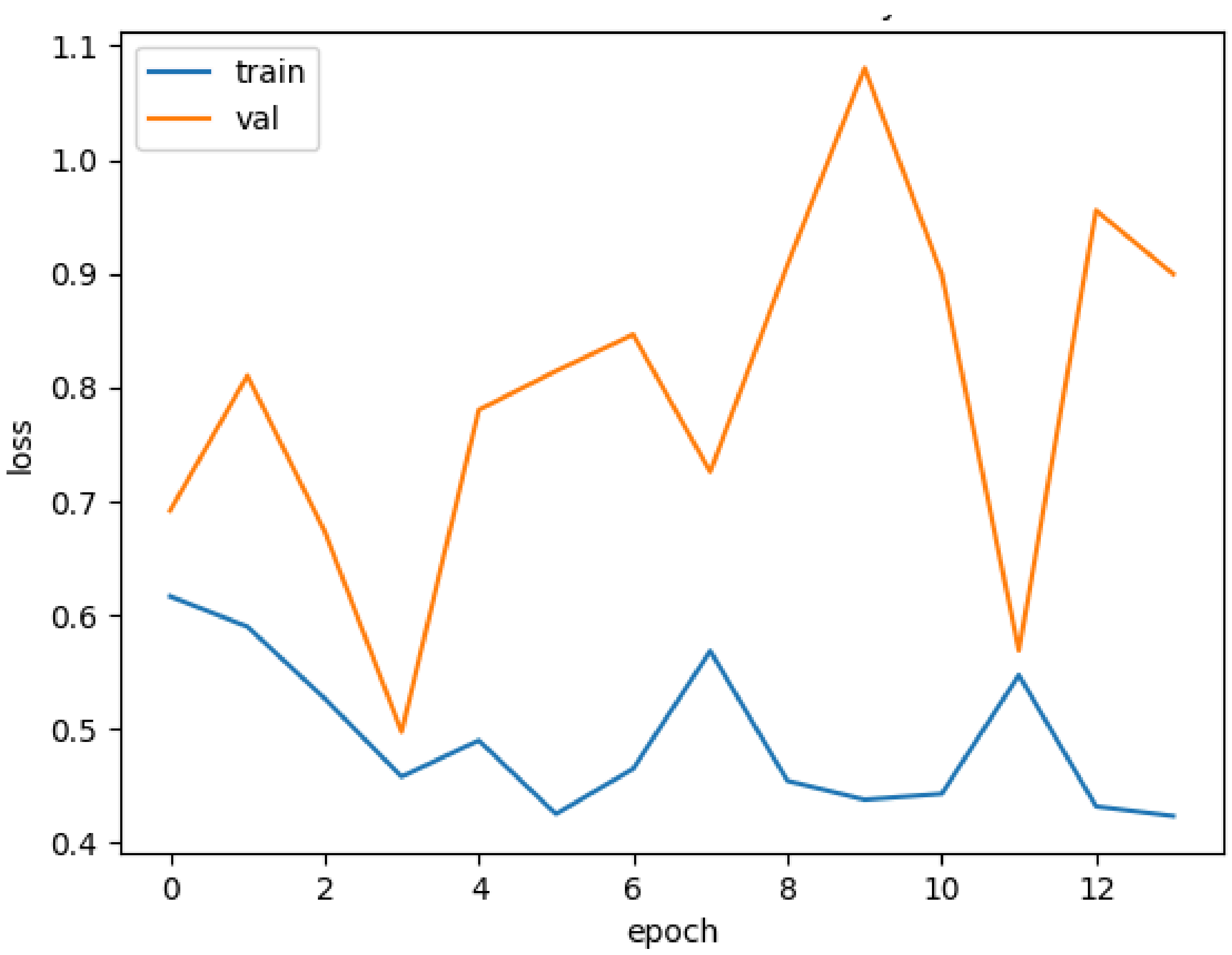
VGG16 ON YOUTUBE DATA

Test F1 Macro Score: 46.05%
Test Accuracy: 46.47%

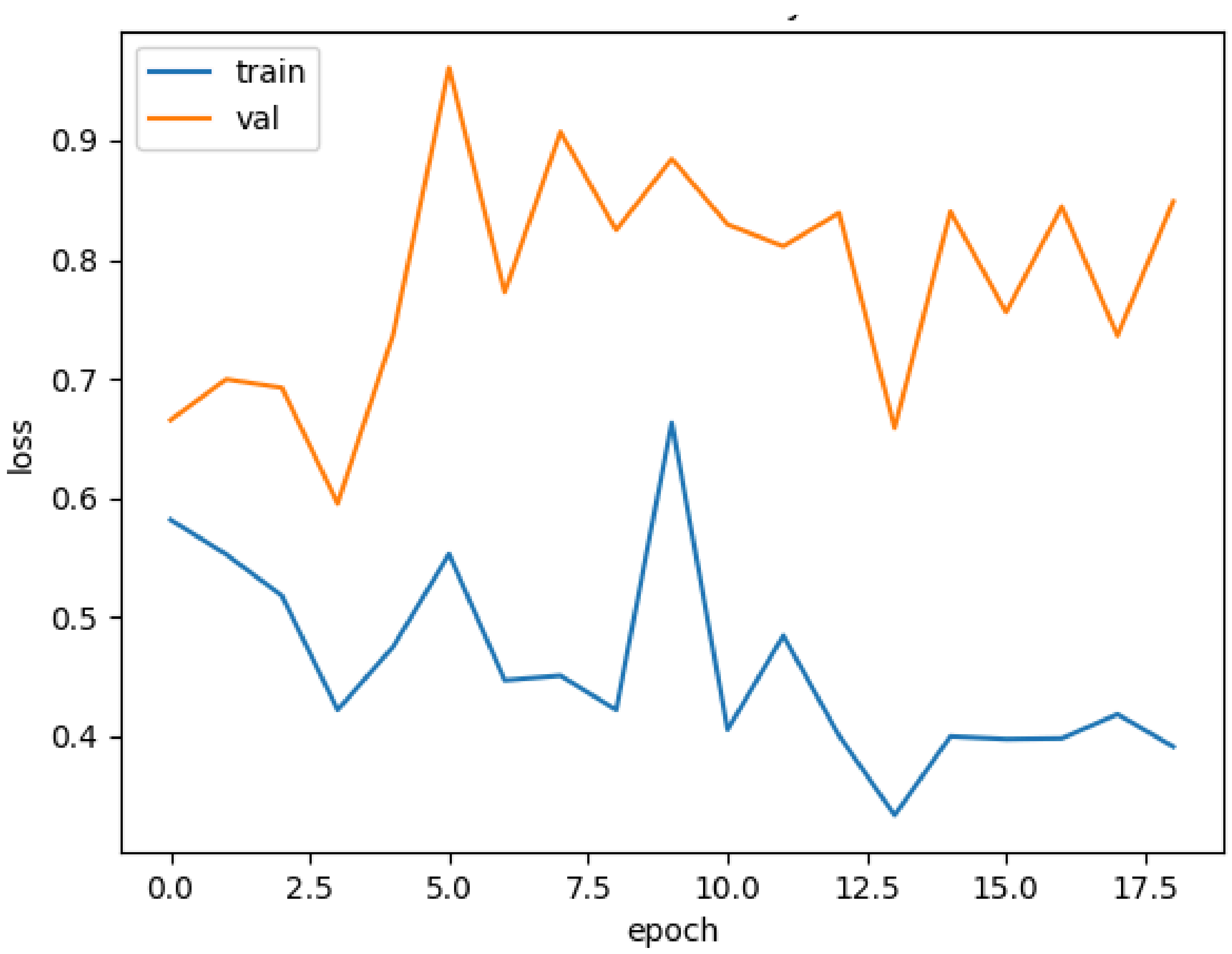
	precision	recall	f1-score	support
0	0.45	0.59	0.51	80
1	0.49	0.36	0.41	90
accuracy			0.46	170
macro avg	0.47	0.47	0.46	170
weighted avg	0.47	0.46	0.46	170

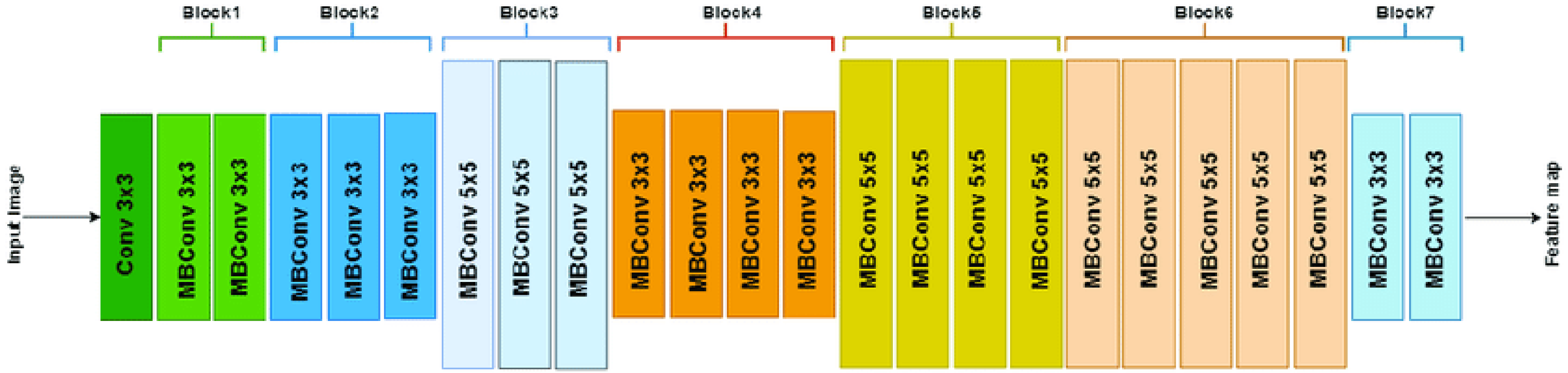
EFFICIENTNET B1 TRAINING AND ARCHITECTURE

EFFICIENTNETB1 LOSS: ALL LAYERS FROZEN



EFFICIENTNETB1 LOSS: LAST 80 LAYERS UNFROZEN





EFFICIENTNETB1
ON TEST CELEB DATA

```
Test F1 Macro Score: 48.62%
Test Accuracy: 60.41%
```

	precision	recall	f1-score	support
0	0.20	0.31	0.24	1180
1	0.80	0.68	0.73	4660
accuracy			0.60	5840
macro avg	0.50	0.49	0.49	5840
weighted avg	0.67	0.60	0.63	5840

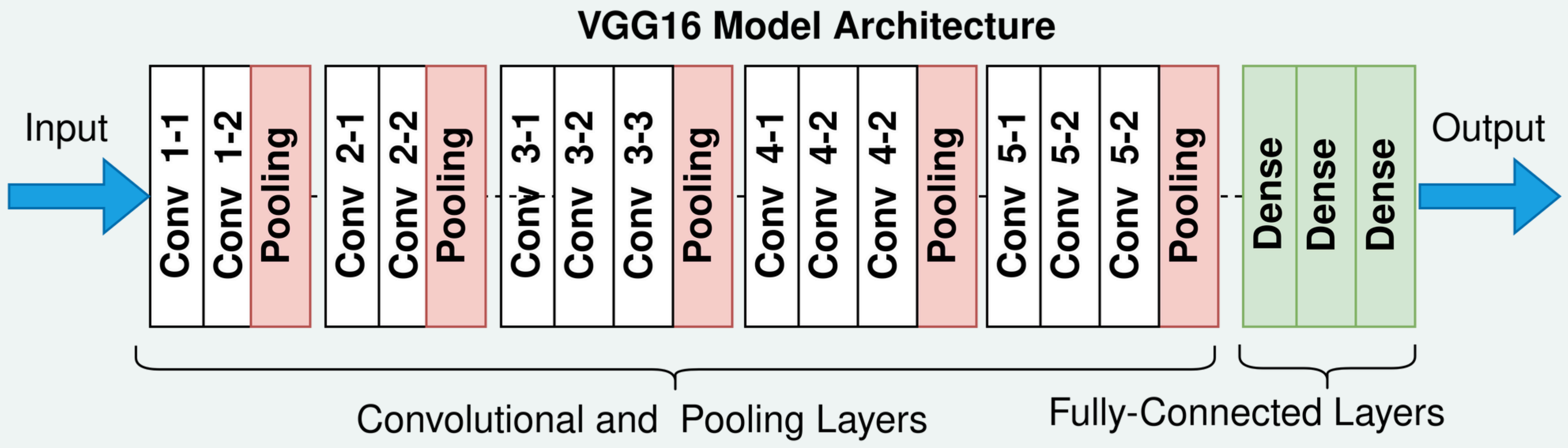
EFFICIENTNETB1
ON YOUTUBE DATA

```
Test F1 Macro Score: 53.86%
Test Accuracy: 54.71%
```

	precision	recall	f1-score	support
0	0.51	0.72	0.60	80
1	0.61	0.39	0.48	90
accuracy			0.55	170
macro avg	0.56	0.56	0.54	170
weighted avg	0.57	0.55	0.53	170

CNN + DENSE CONCLUSION

1. VGG16 outperformed EfficientNetB1 on the Celeb test dataset
   a. Higher precision for Celeb deepfake videos
   b. 50% increase: From 50% to 75% precision score
2. On the YouTube test dataset, EfficientNetB1 outperformed VGG16
   a. Higher precision for fake YouTube videos
   b. 20% increase: From 47% to 56% precision score



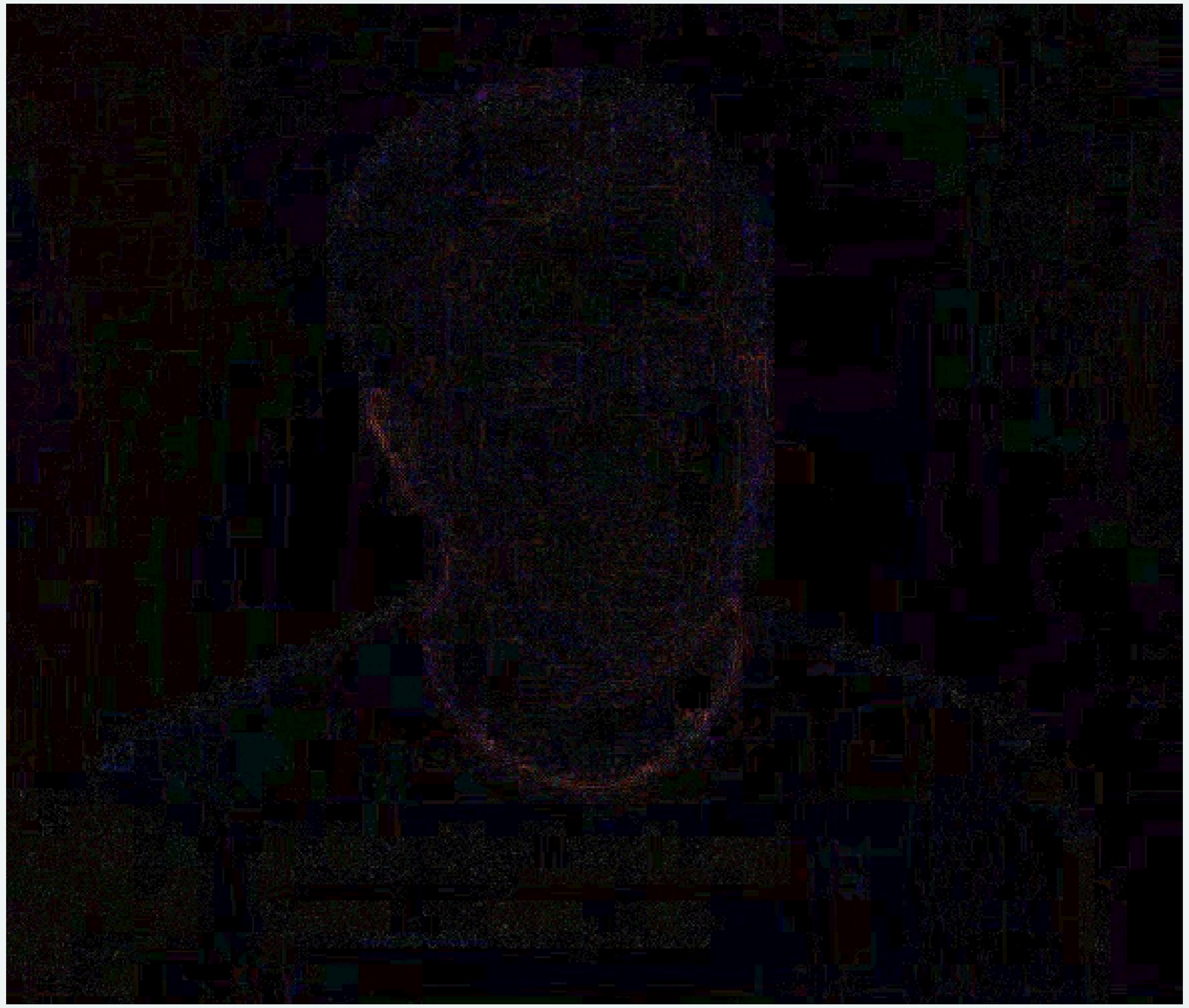
FUTURE SCOPE

- Multi-modal analysis for deepfake.
- Try performing masking techniques and add it as auxiliary data.
- Train the model on ELA (Error-Level Analysis) image.
- Implement encoder-decoder models.

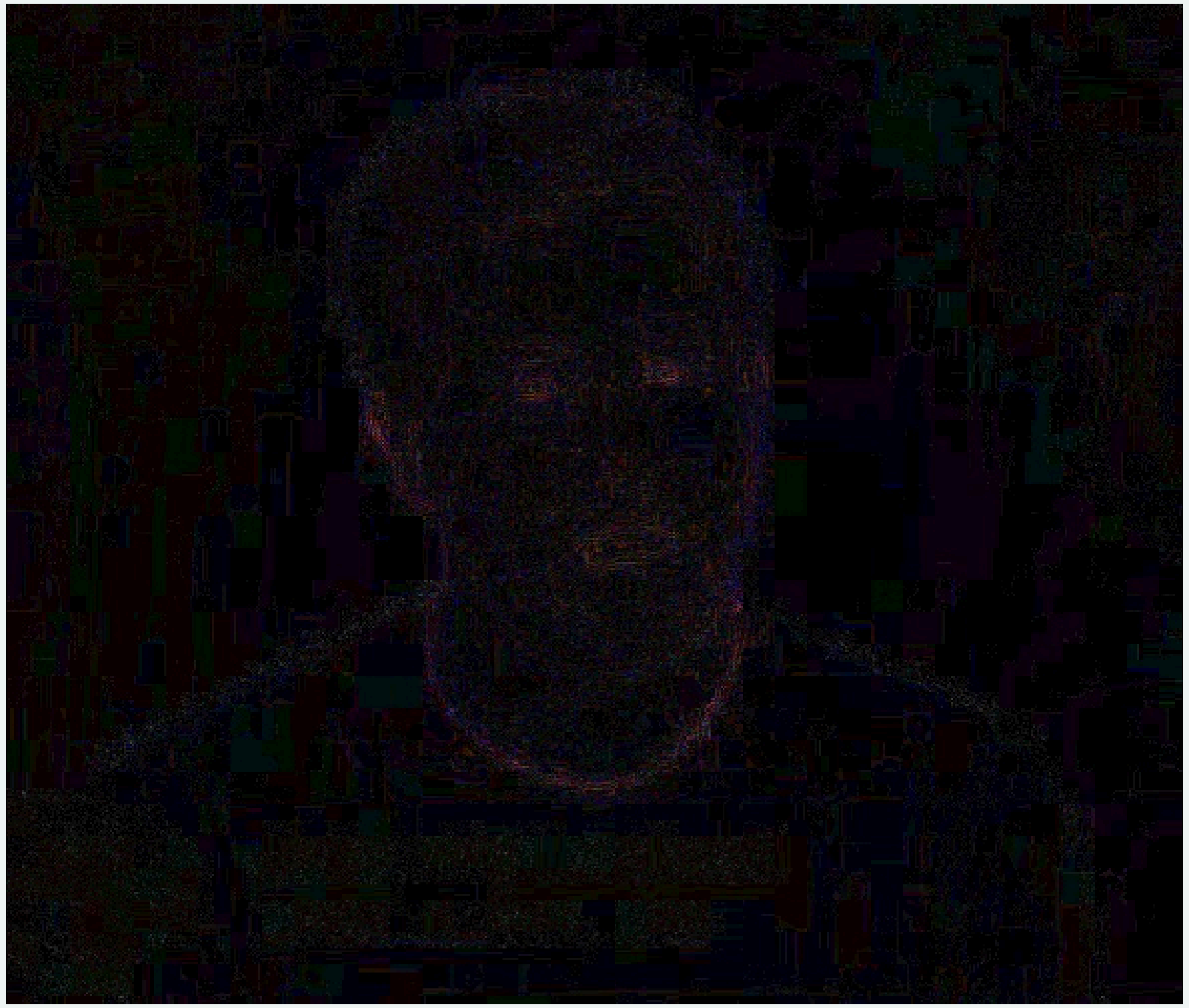
BASIC FUTURE SCOPE IMPLEMENTATION

Error Level Analysis (ELA) detects image tampering by comparing original and recompressed images, highlighting inconsistencies in compression artifacts. It visually exposes manipulated regions, aiding forensic analysis and deepfake detection.

REAL



FAKE



ELA RESULTS

**EFFICIENTNETB1
ON YOUTUBE DATA**

Test F1 Macro Score: 44.74%
Test Accuracy: 45.29%

	precision	recall	f1-score	support
0	0.44	0.59	0.50	80
1	0.48	0.33	0.39	90
accuracy			0.45	170
macro avg	0.46	0.46	0.45	170
weighted avg	0.46	0.45	0.44	170

REFERENCES

- https://openaccess.thecvf.com/content_CVPR_2020/papers/Li_Celeb-DF_A_Large-Scale_Challenging_Dataset_for_DeepFake_Forensics_CVPR_2020_paper.pdf
- https://openaccess.thecvf.com/content_CVPR_2020/papers/Li_Celeb-DF_A_Large-Scale_Challenging_Dataset_for_DeepFake_Forensics_CVPR_2020_paper.pdf

LEARNINGS

- Reading research papers and saving model for each epoch is the most crucial part that I have learned
- Use Focal loss correctly and optimize data loading

QUESTIONS AND ANSWERS