CSE 4206

Enhanced Deepfake Detection and Localization via 3D CNN Batch Processing of Time Series Frames with Specialized Feature Mapping.

Presented By

Srabonti Deb

Roll: 1803163

Dept. of Computer Science & Engineering

Rajshahi University of Engineering & technology

Supervised By

Tasmia Jannat

Lecturer

Dept. of Computer Science & Engineering

Rajshahi University of Engineering & technology

20th January, 2024

Table of Content

Introduction

DeepFake

Real image

Manipulated image

Synthetic media -

Manipulated images-

Altered videos or audio -

Generated by AI -

Introduction

Issues in DeepFake

Real Image

Mask

Fake Image

Traditional Deepfaking Approach

- DeepFake
- Face2Face
- FaceSwap
- Neural Textures

DeepFake Detection

Misinformation Spread:
-Spread misinformation
and false narratives.

Legal and Ethical Complications:
-Render unreliable evidence
-Compromise the fairness
of legal proceedings.

Figure-1: Deepfake problems

Cybersecurity Threats:
-Underscores the risks
associated with cybercrime,
-Identity theft
-Fraud.

Social and Emotional Impact:
-Cause emotional
distress and
-Societal harm.

Deep fake detection and classification using error-level analysis and deep learning (Scientific Reports-2023) [1]

Contribution

- Error-level analysis utilization.
- Hyper-parameter optimization.
- Have low computational cost.

Limitations

- Encounter inconsistencies in handling time series data.
- Sensitivity to data noise.
- Struggle to generalize deepfakes created with different techniques or in unfamiliar contexts.

Contrastive learning-based general Deepfake detection with multi-scale RGB frequency clues. (Journal of King Saud University-2023) [2]

Contribution

- Employs reverse engineering for high-accuracy deepfake detection.
- Performs effectively with noisy images.
- The addition of MSE, CMA, and SCL models enhanced the model's capabilities.

Limitations

- Needs extensive training time and significant processing power.
- Struggles to detect deepfakes not generated by GAN models.
- Potential overfitting issues.

Multi-attention-based approach for deepfake face and expression swap detection and localization.(EURASIP Journal on Image and Video Processing -2023) [3]

Contribution

- Generates localized mask.
- Integrates attention with spatial and spectral features.
- Adaptable model and can effectively handle diverse manipulation types.

Limitations

- Unable to track inconsistencies within time series data.
- Model's frequency-based approach limits effectiveness on noisy images.

Objectives of Our Study

To ensure the integrity and ethical standards of multimedia content.

To detect time series inconsistency.

To utilize batch parallel processing for time series data handling.

To generate localized heatmaps.

To train on in-dataset and test on cross-dataset

To design a time and memory efficient model.

Datasets

FaceForensics++

- 1,000 Original Videos
- Manipulated by Four Automated Techniques
- 509 GB Total Size
- Self-reenactment and Actor Swap Scenarios.

Celeb-DF

- 590 Real Videos
- 5639 Fake Videos
- Celebrities' Footage
- Variety of Backgrounds

DFDC

- 104,500 Videos
- 23.5 TB Total Size
- Diverse Set of 960 Actors
- Binary Classification Labels

Workflow of Proposed System

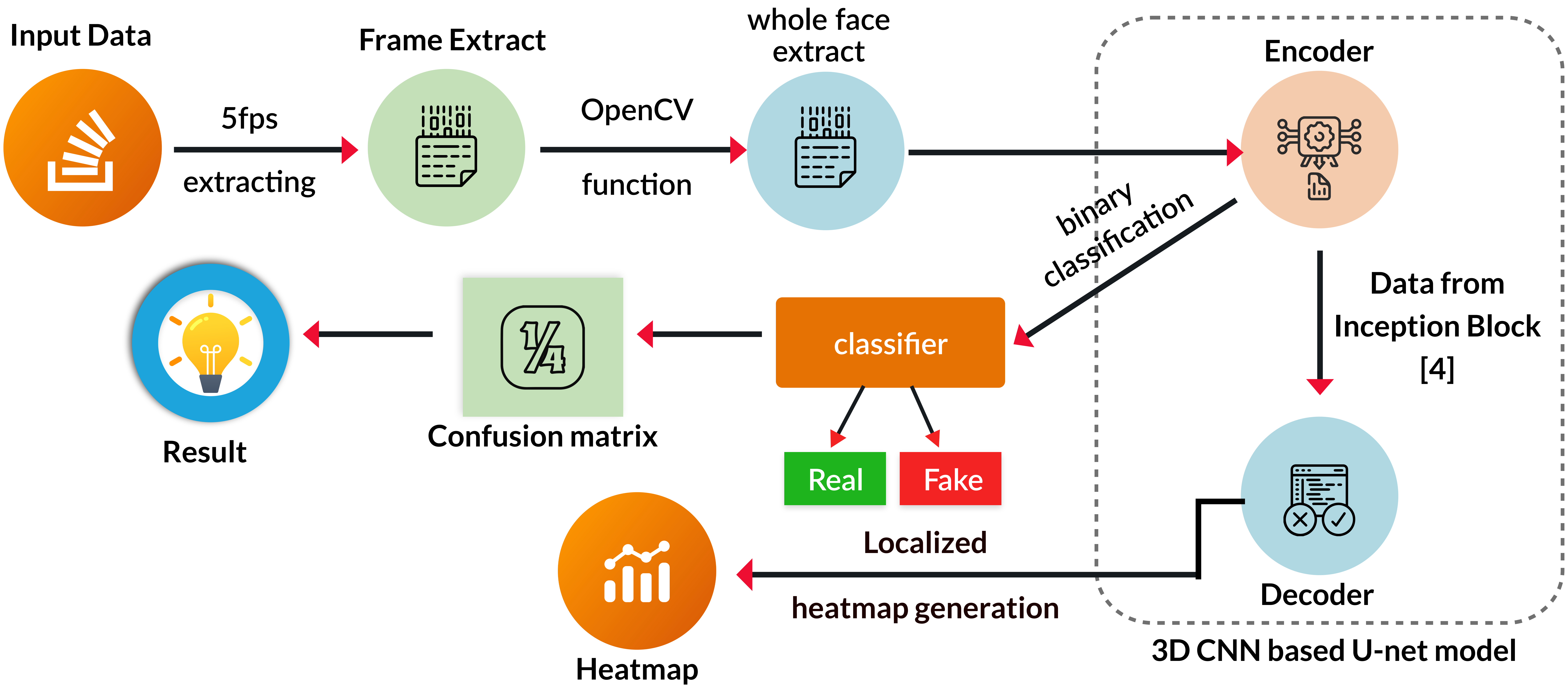


Figure-8: Flow diagram of 3D CNN based deepfake detection system

Proposed Method

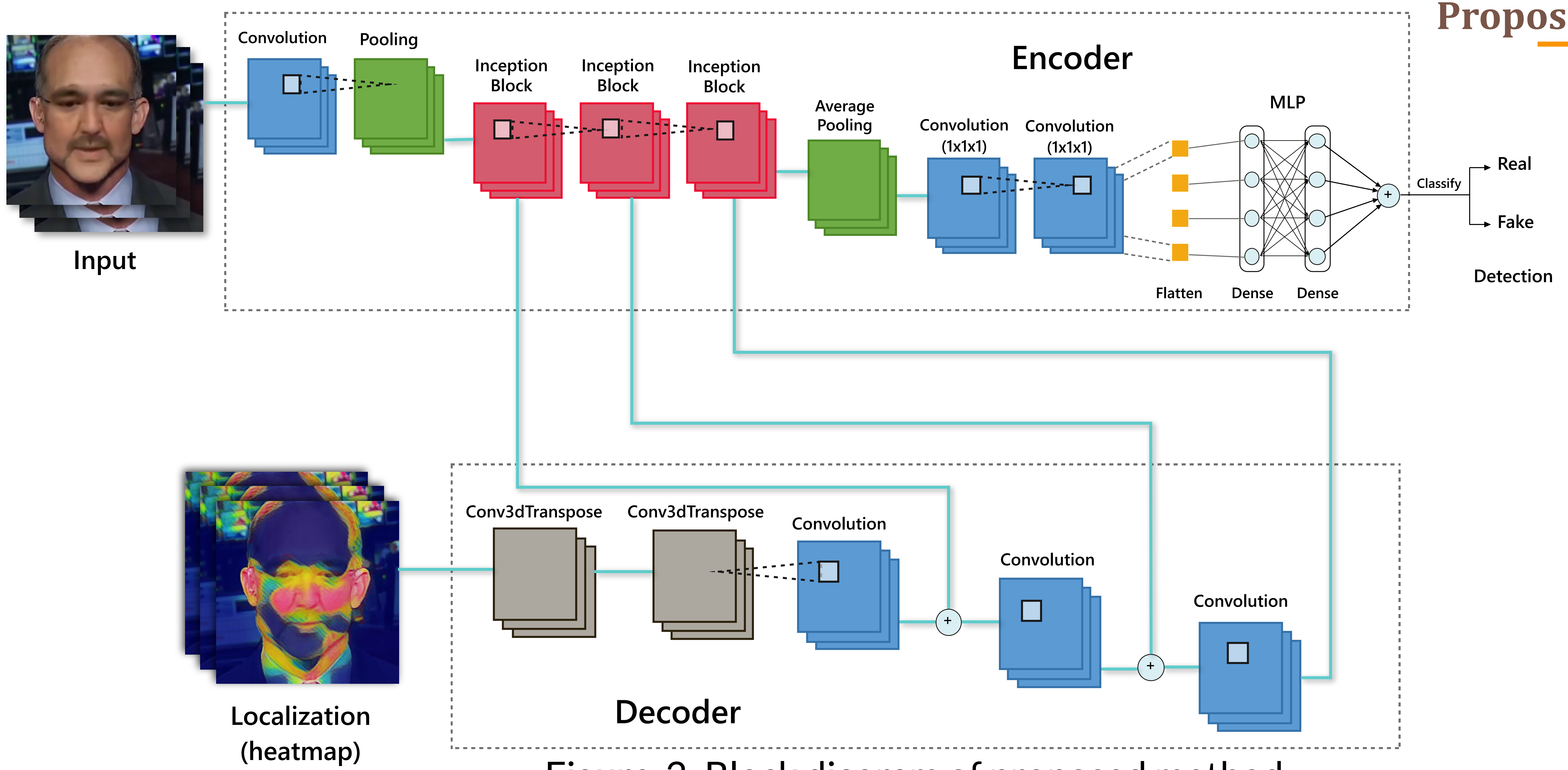


Figure-2: Block diagram of proposed method

Implementation

Data Preprocessing

Video data

Image frame extraction

OpenCV Library Function
(detectMultiScale)
+
Haar Cascade Classifier
[5]

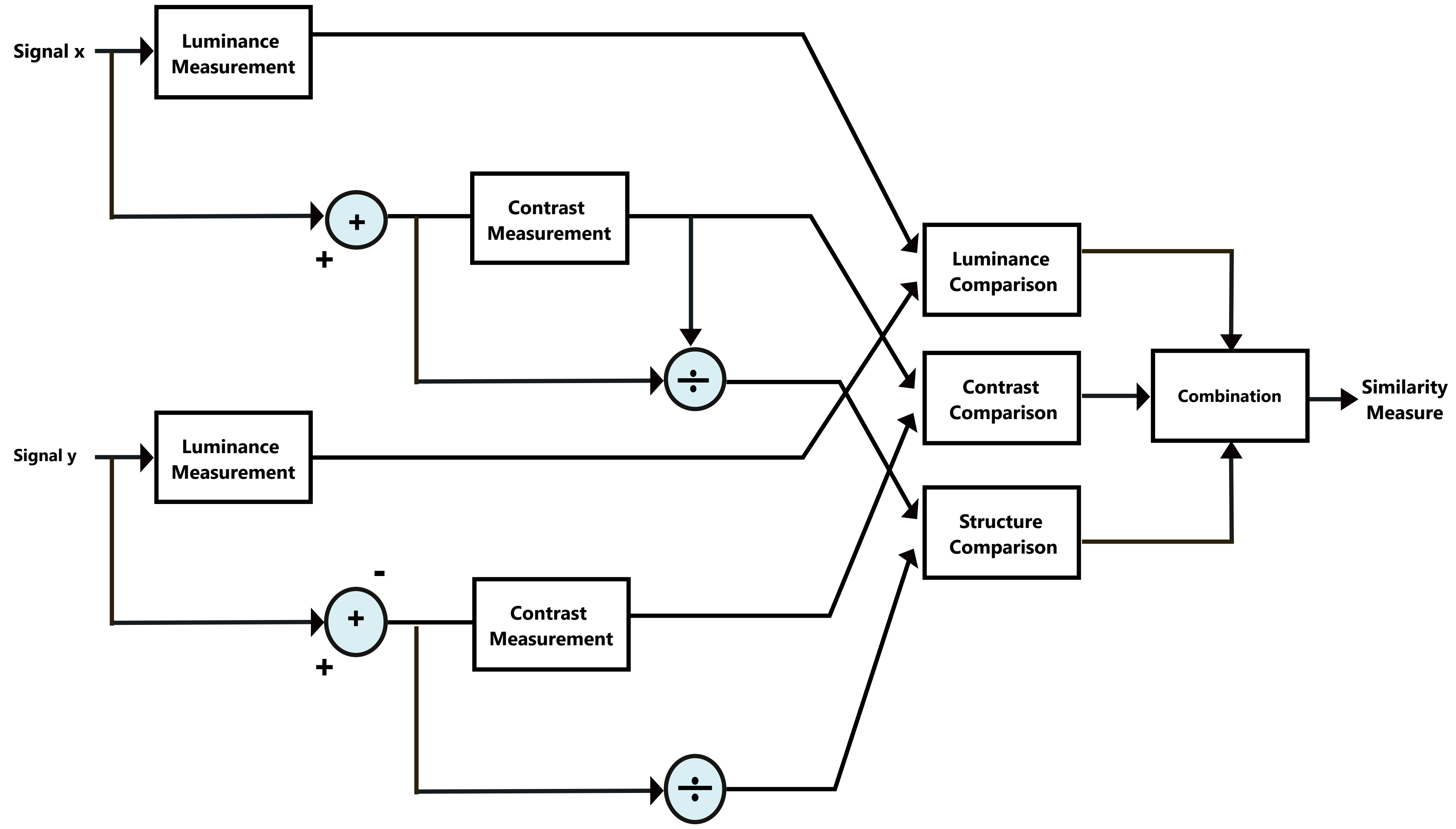
Facial landmark detection

Whole face extraction

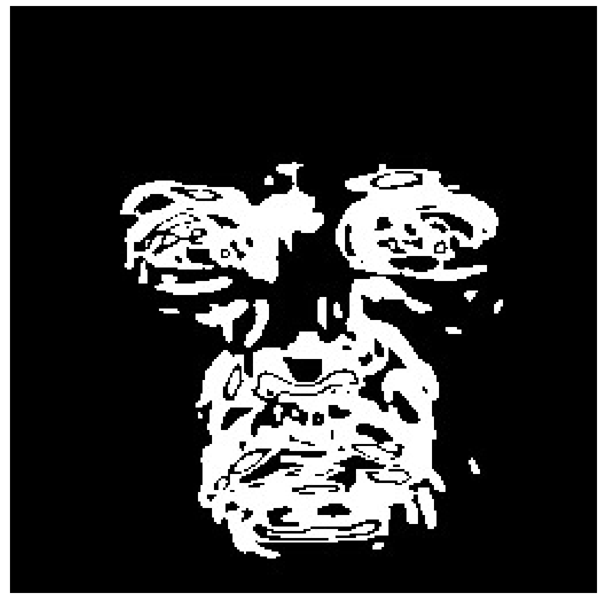
Data Preprocessing

Original Image

Manipulated Image



The Structural Similarity Index (SSIM) [6]



Mask

Calculating Dissimilarity

Experimental Setup

Python 3.8

Pytorch

NVDIA CUDA

1.80GHz i7

350MX with 2GB GDDR5

8GB DDR4

Challenges

Large Datasets Processing	High GPU Requirements
Software compatibility	Availability of Resources

Evaluations: Accuracy & Convolution Matrix

- Confusion Matrix of Faceforensics++ dataset with 60% training data.

Accuracy	**91.26%**

Confusion Matrix:

16(TP)	5(FN)
6(FP)	99(TN)

Conclusion

Work Done So Far

- Involved a comprehensive comparison of state-of-the-art models in the field.

- Developed model based on the analysis.

- Executed pre-processing steps.

Future Works

- Hyperparameter tunning

- Implementing model on entire dataset.

- Evaluating on in-dataset and cross dataset

References

[1] Rafique, R., Gantassi, R., Amin, R., Frnda, J., Mustapha, A., & Alshehri, A. H. (2023). Deep fake detection and classification using error-level analysis and deep learning. Scientific Reports, 13(1), 7422.

[2] Dong, F., Zou, X., Wang, J., & Liu, X. (2023). Contrastive learning-based general Deepfake detection with multi-scale RGB frequency clues. Journal of King Saud University-Computer and Information Sciences, 35(4), 90-99.

[3] Waseem, S., Abu-Bakar, S. A. R. S., Omar, Z., Ahmed, B. A., Baloch, S., & Hafeezallah, A. (2023). Multi-attention-based approach for deepfake face and expression swap detection and localization. EURASIP Journal on Image and Video Processing, 2023(1), 14

References

[4] Lu, C., Liu, B., Zhou, W., Chu, Q., & Yu, N. (2021, September). Deepfake video detection using 3D-attentional inception convolutional neural network. In 2021 IEEE International Conference on Image Processing (ICIP) (pp. 3572-3576). IEEE.

[5] OpenCV 3 Object Detection : Face Detection using Haar Cascade Classfiers - 2020. (n.d.). https://www.bogotobogo.com/python/OpenCV_Python/python_opencv3_Image_Object_Detection_Face_Detection_Haar_Cascade_Classifiers.php

[6] Kang-Long, Y., Zhao-kui, M., & Ming-jie, S. (2013, September). Image Quality Assessment: A Reduced Reference Algorithm for the Super-resolution Reconstruction Image. In 2013 Third International Conference on Instrumentation, Measurement, Computer, Communication and Control (pp. 171-175). IEEE.

THE END

THANK YOU

FOR CONSIDERATE AUDIENCE

Srabonti (1803163)