Tool for detecting deepfakes based on biological factors

Student: Andrei Kulinkovich Supervisor: Ing. Milan Šalko.

MOTIVATION

- Deepfakes threaten media trust and privacy.
- Biological signals are hard to fake.
- Study the features of biological signals.
- Study breathing patterns for detection deepfakes.

TECH & DATASET

- PhysNet for extract PPG signal
- FaceForensics++ for training models
- Celeb-DF (V2) for testing models
- PyTorch
- MTCNN face detector for sequence-based method
- dlib face detector for image-based method

IMAGE-BASED

- Extract three face regions.
- Transform these face regions.
- Extract rPPG using the CHROM method.
- Generate PPG maps per region.
- Classify maps using a CNN.

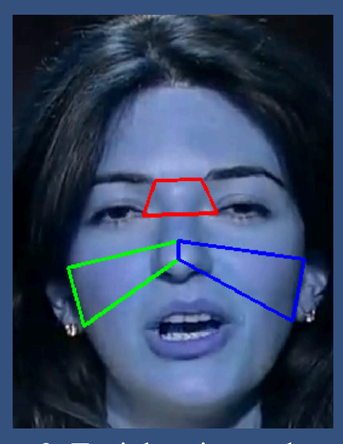

Figure 2: Facial regions selected for PPG map for image-based approach.

IMPLEMENTATION

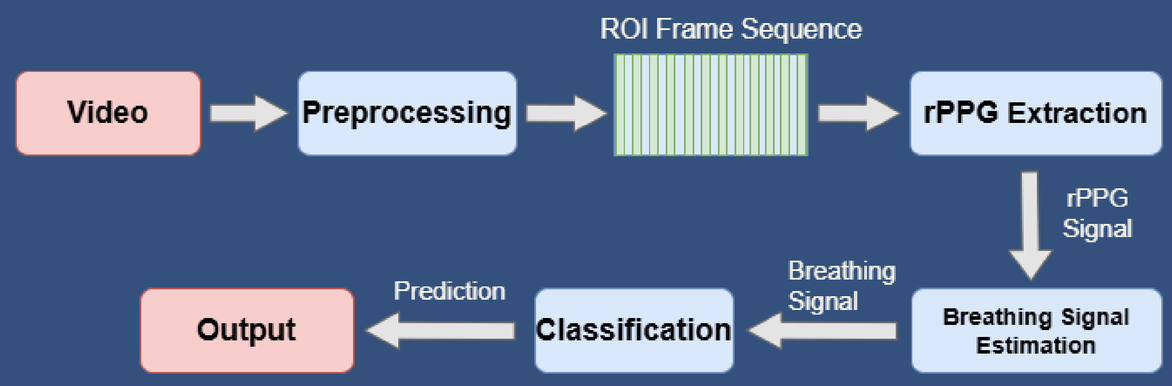

Figure 1: Breath rate estimation model architecture.

SEQUENCE-BASED

- Extract full face region.
- Model breathing pattern with LSTM/GRU.
- Predict labels based on rPPG signal dynamics.

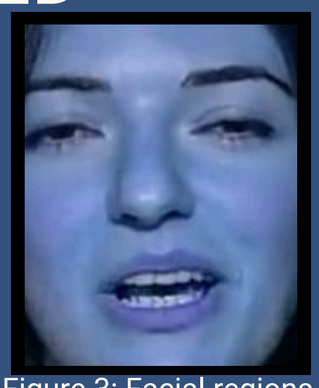

Figure 3: Facial regions selected for CNN approach.

- Extract face regions frame sequences from video.
- Process frame sequence.
- Extrac rPPG signal from facial regions.
- Filter signal to isolate breathing pattern.
- Classify using CNN or RNN models.

EXPERIMENT RESULTS

Model	FaceForensics++ (test)		Celeb-DF (v2)	
	Accuracy (%)	F1-score	Accuracy (%)	F1-score
Image-based (CNN)	60.7	0.58	52.3	0.51
Sequence-based (LSTM)	54.7	0.52	53.8	0.53
Sequence-based (GRU)	53.2	0.49	50.1	0.46

Table 1: Cross-dataset evaluation of model performance on FaceForensics++ and Celeb-DF (v2)

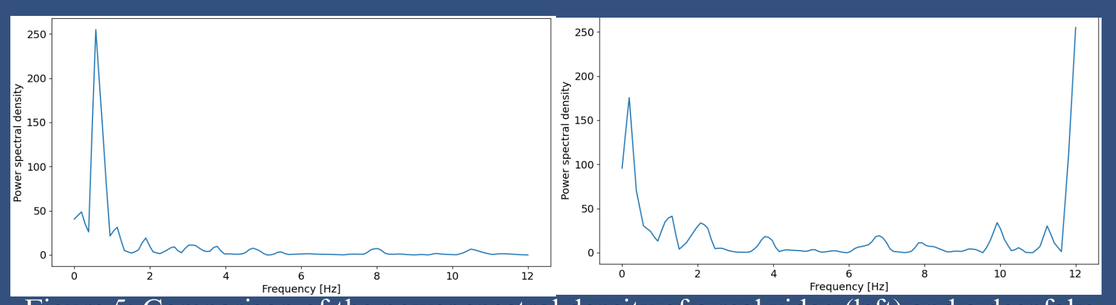

Figure 5: Comparison of the power spectral density of a real video (left) and a deepfake video (right).

Method	Accuracy (%)
Visual artefacts	98.7
DeepFakesON-Phys	98.5
CNN-RNN	97.0
FakeCatcher	93.5
Inconsistent head poses	89.0
DeepVision	87.5
DeepLie	72.7
Image-based (CNN)	60.7
Sequence-based (LSTM)	54.7
Sequence-based (GRU)	53.2

Table 2: Comparison of different deepfake detection methods. The names of our models are highlighted in bold.

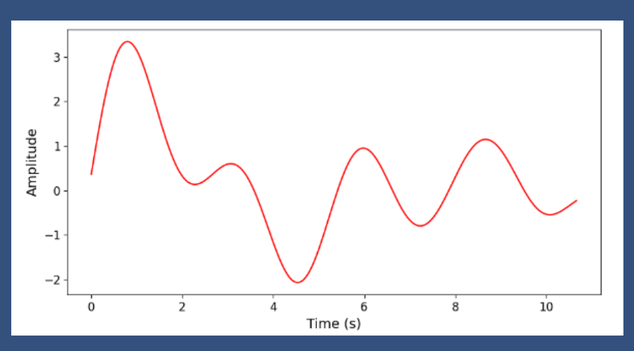

Figure 4: Output of the PhysNet model for a sequence of duplicated static frames.

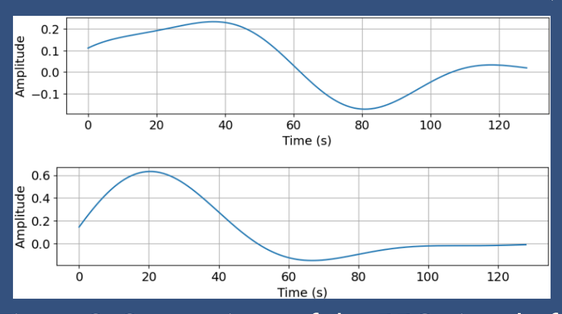

Figure 3: Comparison of the PPG signal of a real video (top) and a deepfake video (bottom)

Frequency band (Hz)	Accuracy (%)
0.1-0.4	60.7
0.7-2.5	80.2
0-12	83.5

Table 3: Comparison of different frequency bands for image-based method

BRNO UNIVERSITY OF TECHNOLOGY FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2025