From Pixels to Pulse: Enhancing Remote PhotoPlethysmography Accuracy in Compressed Videos

Farhan Ansari

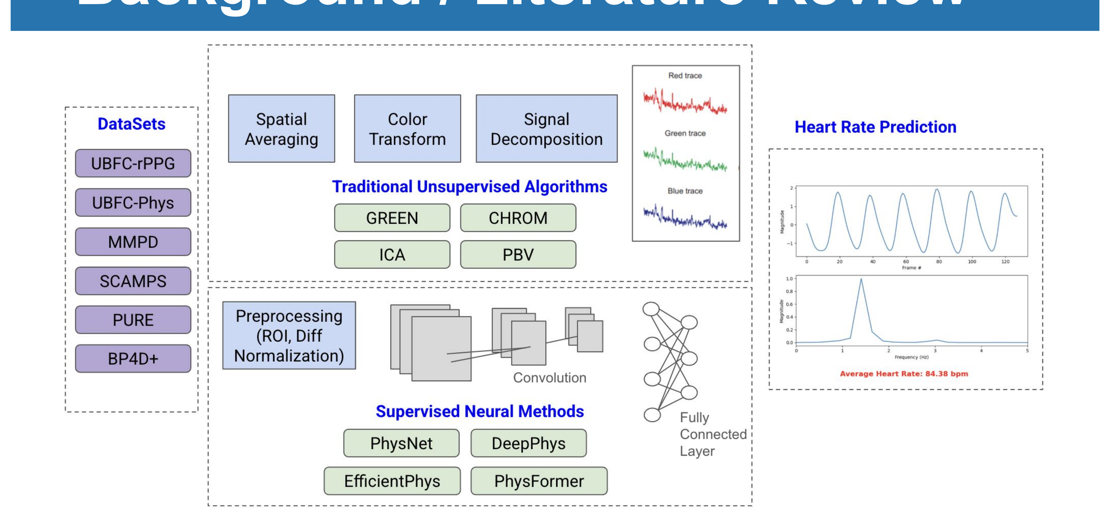
Introduction

- Remote photoplethysmography (rPPG) is a camera-based contactless method to measure physiological signals such as heart rate by analyzing subtle changes in the skin color related to blood flow.
- It has potential for use in telehealth, which would allow for non-invasive, low cost monitoring of vital signs, especially in remote/developing areas
- Nearly all the existing rPPG methods are designed based on uncompressed video data. However most video data transmitted over the internet is in compressed format. Therefore in order to scale this technology to real world use cases, it needs to work well on compressed data.
- Lossy compression can add artifacts such as noise, blur, blockiness, etc. which can degrade the signal-to-noise ratio of the extracted rPPG signal.

Engineering Goal

- To improve the accuracy of remote heart rate estimation (rPPG) from compressed video data, addressing the degradation of signal quality caused by compression for applications in telehealth or remote patient monitoring.

Background / Literature Review



- Conventional rPPG techniques involve localizing the face and identifying regions of interest (ROI), then isolating the pulse signal by removing the green channel from the red and stabilizing it by subtracting the average over a long fluctuation.
- Deep Learning approaches have become popular in recent years. Neural models such as PhysNet utilize a 3D CNN (a "spatio-temporal network") to identify the pulse over time.

Methods and Procedures

Dataset: UBFC-rPPG

- Consists of 42 videos at 30fps with a resolution of 640x480 in uncompressed 8-bit RGB format. Widely used in many research papers. Bitrate: 220 Mbps.
- The ground truth data is recorded from a pulse oximeter.

Model: PhysNet

- Uses 3D Convolutional Neural Network to extract rPPG features in the spatial and temporal domains simultaneously.
- Widely used in research studies.

Platform: rPPG-ToolBox

- Open source platform for developing new models for rPPG.
- Supports common rPPG datasets, models and evaluation methods.

Software and Tools

- GitHub for version control
- Uv for package management
- Google colab, Python, Numpy, Pandas
- Pytorch for model training
- OpenCV - library for image processing

Generating Baseline data

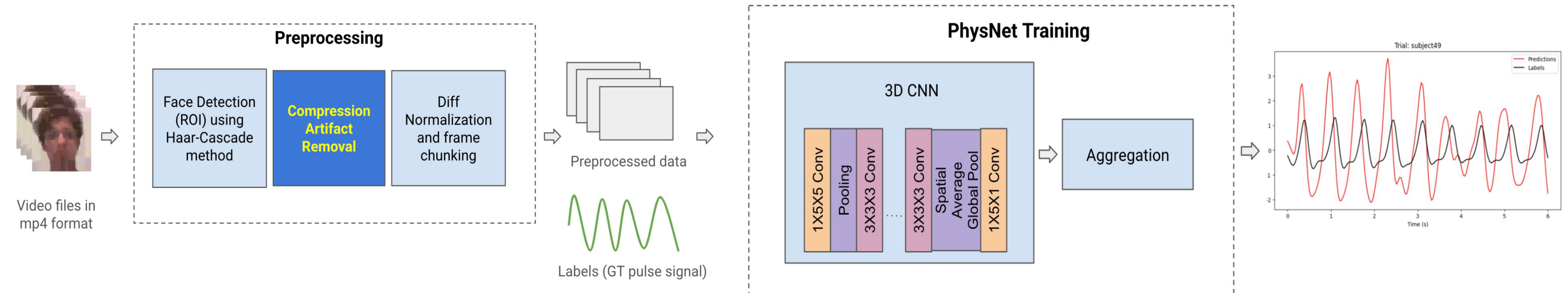
- Obtained UBFC-rPPG dataset
- The UBFC-rPPG dataset was compressed at several different ratios (70:1, 100:1 and 150:1) using the H.264 codec, and was then split 80/20 for training and testing.
- Obtained baseline numbers for uncompressed dataset with PhysNet.
- Trained PhysNet model on the compressed data to obtain the baseline results for each compression ratio.

Compression Artifact Removal Module

- Implemented a compression artifact removal module in the preprocessing step.
- Considered various denoising, deblurring, deblocking techniques such as gaussian blur, bilateral, non-local means denoising, etc.
- Considered deep learning based video restoration models: Video Restoration Transformer (VRT), MdVRNet (Deep Video Restoration under multiple distortions).

Testing and Evaluation

- UBFC-rPPG dataset was split 80:20 for training and testing.
- Used evaluation methods provided in the rPPG-Toolbox.
- Metrics used for evaluation:
  - Mean Absolute Error (MAE)
  - Root Mean Squared Error (RMSE)
  - Mean Absolute Percentage Error (MAPE)
  - Pearson's Correlation



Data

Compression Ratio	Bit Rate	Model	MAE (HR) ↓	RMSE (HR) ↓	MAPE (HR) ↓	Pearson's Correlation ↑
70:1	3,159,771	Base Model (PhysNet)	3.52	6.09	2.74	0.98
		PhysNet with artifact removal	**3.08**	**4.54**	**3.17**	**0.98**
100:1	2,211,840	Base Model (PhysNet)	5.86	7.32	6.07	0.99
		PhysNet with artifact removal	**3.73**	**5.03**	**3.79**	**0.98**
150:1	1,474,560	Base Model (PhysNet)	9	14.68	7.97	0.80
		PhysNet with artifact removal	**3.95**	**9.27**	**3.54**	**0.92**

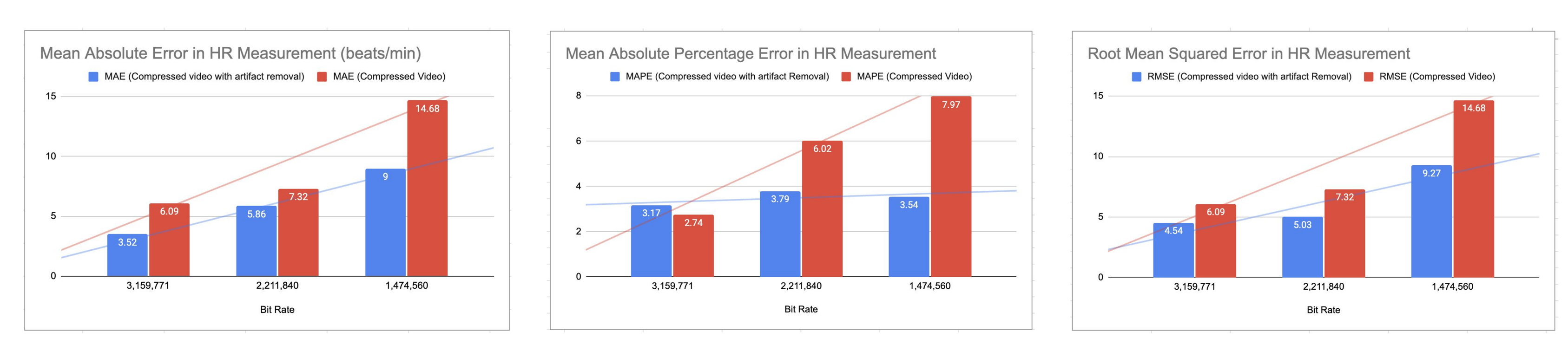
MAE = Mean absolute Error (beats/min), RMSE = Root Mean Square Error, MAPE = Mean Absolute Percentage Error

Evaluation Metrics for compressed data

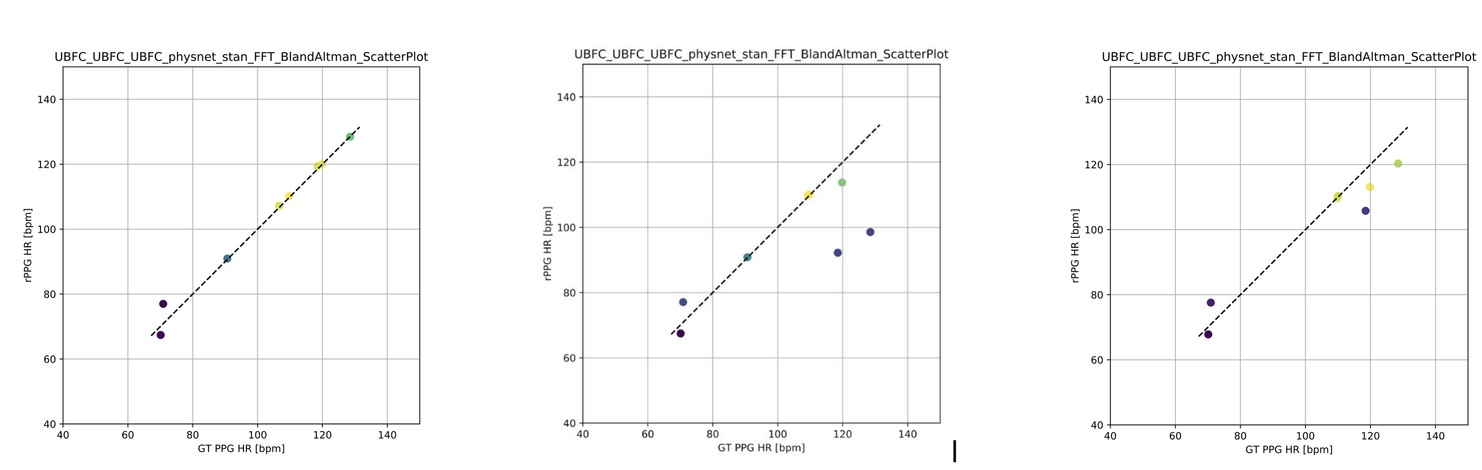
Compression Ratio	Bit Rate	Model	MAE (HR) ↓	RMSE (HR) ↓	MAPE (HR) ↓	Pearson's Correlation ↑
None	220 Mbps	PhysNet	**1.09**	**2.56**	**1.56**	**0.99**

MAE = Mean absolute Error (beats/min), RMSE = Root Mean Square Error, MAPE = Mean Absolute Percentage Error

Evaluation Metrics for original uncompressed data



Comparison of compression artifact removal with base model



Uncompressed

150:1 compression

With compression artifact removal

Bland-Altman Plots

Results / Findings

- As compression ratio was increased, MAE, MAPE and RMSE increased. For very high compression ratios (> 150:1), the signal-to-noise ratio decreased significantly.
- Addition of compression artifact removal module in the preprocessing step yielded lower MAE, MAPE, and RMSE for all levels of compression.
- Among all the different artifact removal techniques tried, the non local means denoising technique showed best results (as shown in the charts).
- The higher the compression ratio, the more effective artifact removal was.
- Neural models for video restoration such as VRT and MdVRNet increased preprocessing time significantly.

Further Research

- Explore deep learning based artifact removal techniques specifically for compression artifacts and develop an end-to-end model.
- Evaluate model performance on other datasets containing compressed videos, e.g. MMPD and COHFACE. Perform both intra-dataset and inter-dataset evaluations to assess how well the model generalizes.

Selected References

- Yu, Z., Li, X., & Zhao, G. (2019, May 7). Remote Photoplethysmograph Signal Measurement from Facial Videos Using Spatio-Temporal Networks. arXiv.org. https://arxiv.org/abs/1905.02419
- Liu, X., Narayanswamy, G., Paruchuri, A., Zhang, X., Tang, J., Zhang, Y., Sengupta, S., Patel, S., Wang, Y., & McDuff, D. (2022, October 3). RPPG-Toolbox: Deep Remote PPG toolbox. arXiv.org. https://arxiv.org/abs/2210.00716
- Y. Benezeth - UBFCRPPG. (n.d.). https://sites.google.com/view/ybenezeth/ubfcrppg
- A. Buades, Non-Local Means Denoising, Image Processing On Line, 1 (2011), pp. 208–212.https://doi.org/10.5201/ipol.2011.bcm_nlm