CARDIOVISION

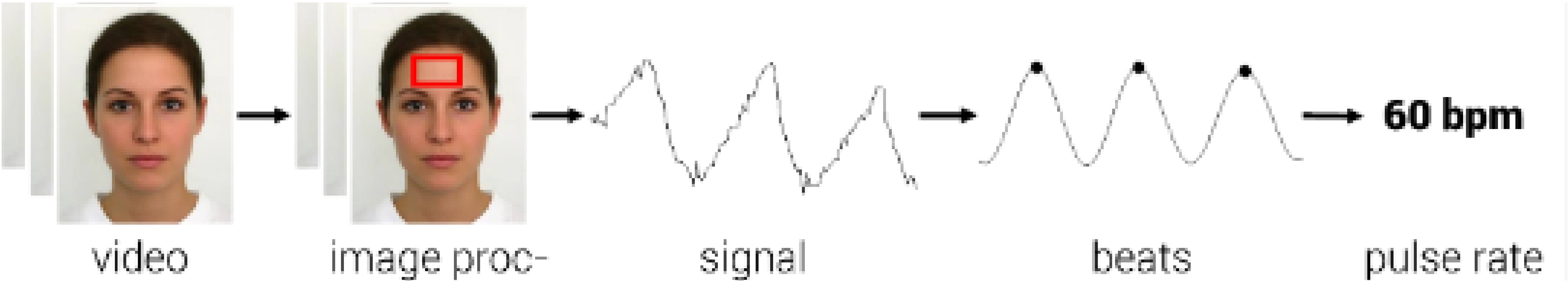
REAL-TIME NON-INVASIVE CONTACT-LESS HEART RATE DETECTION UNDER ADVERSARIAL CONDITIONS

Team PANDAVAS Members

- Allan Robey (22111007)
- Avnish Tripathi(22111014)
- Divyesh Tripathi (22111020)
- Kush Shah (22111033)
- Pulkit Sharma (22111048)

Problem Statement

- The problem statement is to develop a deep learning-based solution for contact-less heart rate detection from real-time video in adverse conditions.
- Most of the traditional methods for measuring heart rate require physical contact with the user, which can be uncomfortable and inconvenient.
- Therefore, the system aims to provide a non-invasive and user-friendly solution for heart rate detection.
- The system should be able to accurately detect heart rates from real-time video data, which can be captured using various devices such as smartphones, webcams, etc.



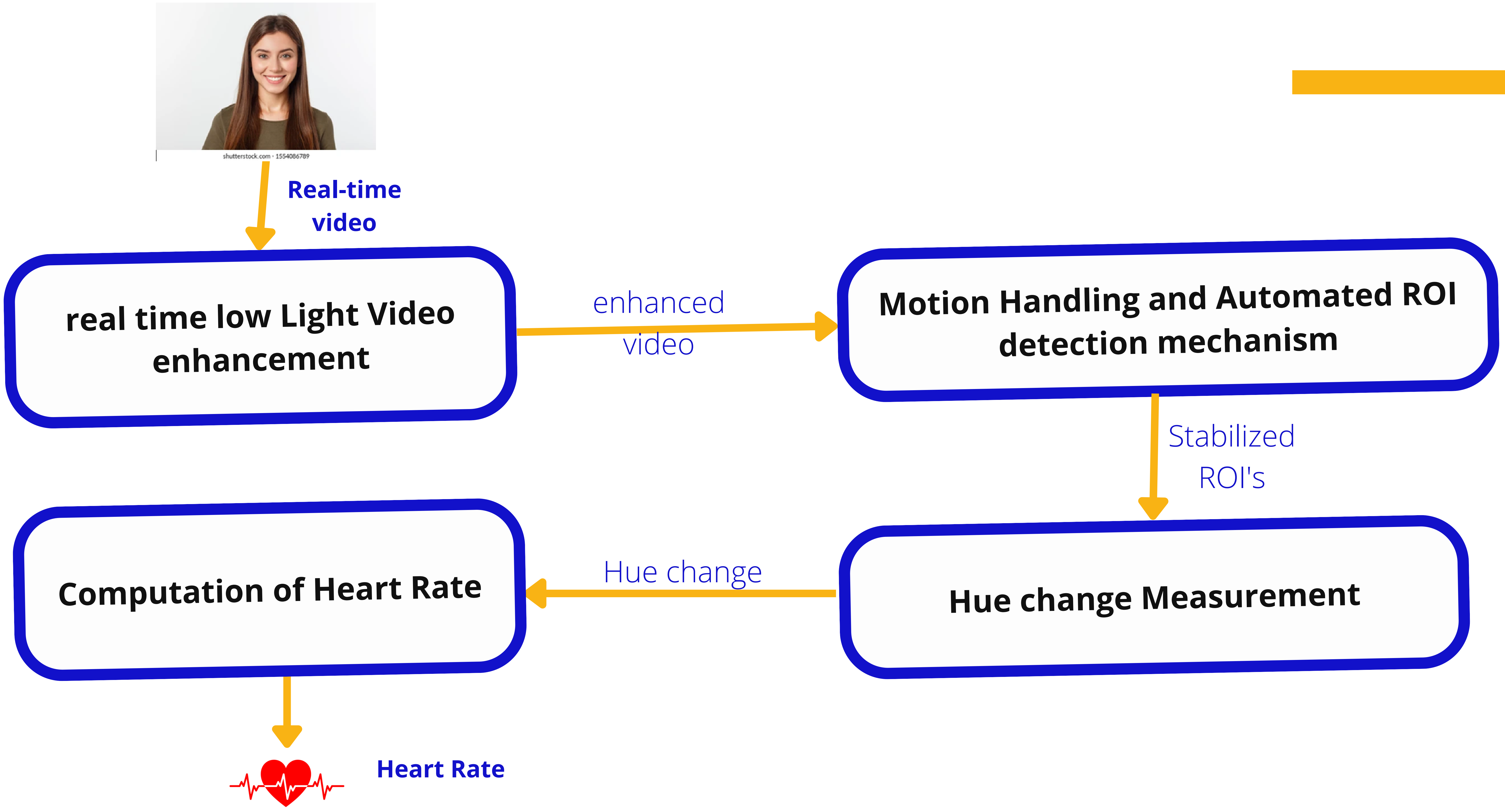
Problem Statement

- However, detecting heart rates from video data presents several challenges.
- Firstly, the video data may have low-light conditions that can affect the accuracy of the heart rate detection.
- Additionally, the video data may contain motion artifacts that can make it difficult to detect the heart rate accurately.
- Therefore, the system needs to be able to handle such adverse conditions and provide accurate heart rate measurements.

State of the Art Table

Approach	Effect of low light	Effect of motion	Real time	Comments
combines motion and color information from an RGB camera to estimate heart rate	✗	✗	✓	issue of generalizability due to relatively small dataset of 25 subjects.
signal processing techniques to estimate the heart and respiration rates in real-time by analyzing changes in the hue channel of the HSV color space.	✗	✗	✗	did not investigate the impact of different skin colors on the accuracy
a convolutional neural network (CNN) to analyze time-series color variation data from videos	✓	✗	✓	limited generalizability of the method to populations with different ages, health conditions, and skin tones
heart rate detection method using face detection and object tracking in video streams.	✗	✗	✓	achieves high accuracies on several public datasets, but faces challenges in extreme cases such as overtly bright or flickering lighting and large head and body movements (e.g., during exercising)
unsupervised rPPG analysis method	✗	✓	✓	achieves high accuracies on public datasets, but faces challenges in extreme cases
Our Model: Zero DCE ,VGG16-Automated ROI Detection(using our created data set) capable of handling effect of motion	✓	✓	✓	**a computationally efficient method for real-time heart rate detection in videos, capable of handling adverse conditions such as low-light and motion artifacts.**

Solution Approach



Real-Time Low Light Video enahancement

- The current solutions for low-light video enhancement, have shown promise but they are computationally expensive and they cause issues with inference time at real-time heart rate detection.
- To address these limitations, we have used Zero-DCE approach reduces the computational complexity while maintaining the accuracy of real-time heart rate detection.
- This approach had effectively resolved the challenges of low-light videos and enabled real-time monitoring of heart rate in low-light settings.

low light image

openCV

Mirnet

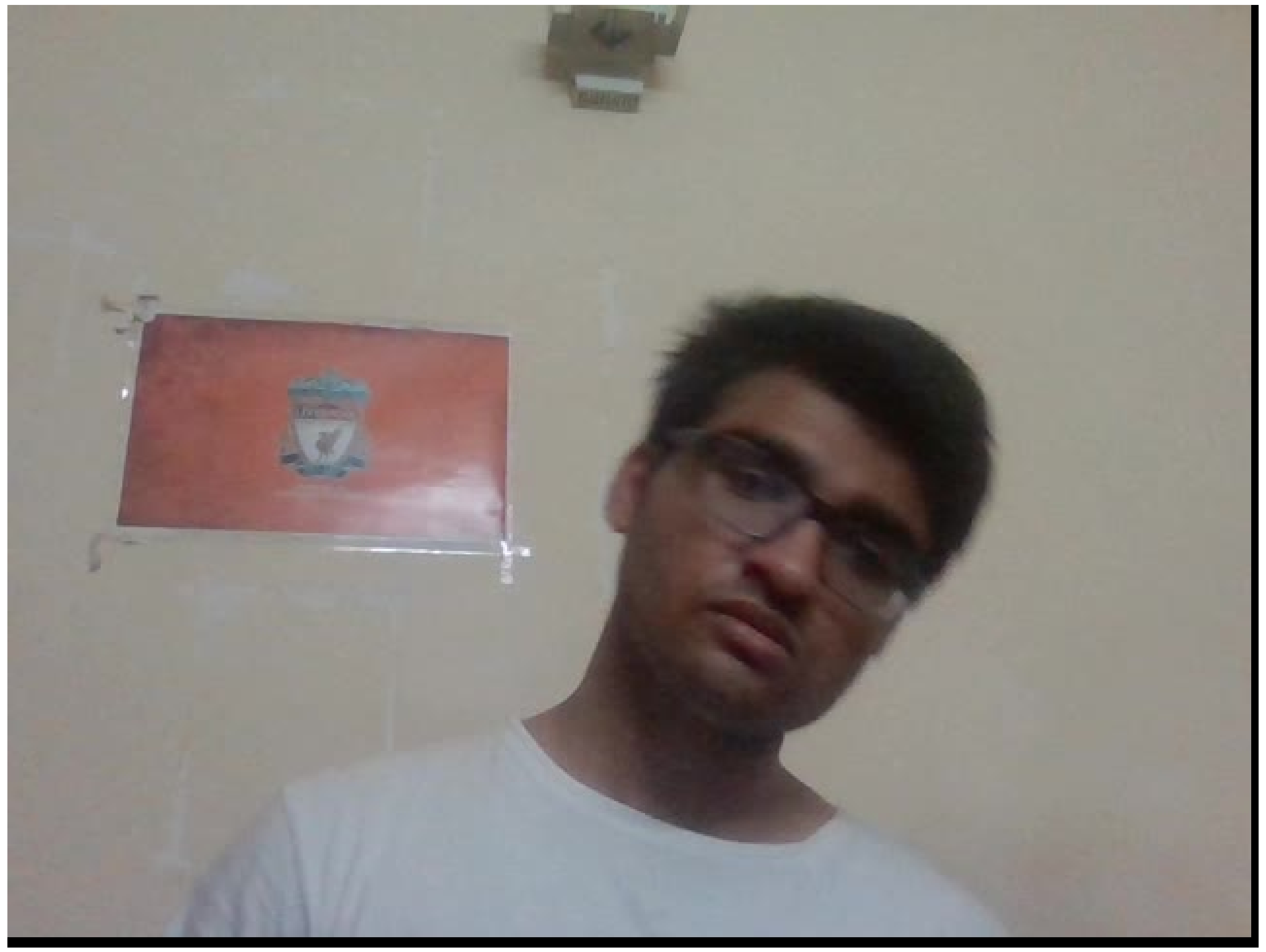
Zero-Dce

Solution Approach

Real-Time Video Motion Handling and Automated ROI Detection Mechanism

- To address the issues of motion artifacts and instability in video, we have developed a solution that leverages deep learning techniques for real-time heart rate measurement.
- Current solutions such as CNNs and recurrent neural networks (RNNs) to predict and correct camera motion, resulting in more stable and smoother video recordings are computationally expensive, making it difficult to handle motion artifacts and instability in real-time heart measurement.
- Our model will address these challenges with efficient computation and thereby, enabling more accurate and reliable heart rate measurement in challenging scenarios.
- A VGG16 model has been trained on a large and diverse dataset of forehead images for more robust and accurate detection of foreheads under the effect of motion artifacts such as head movements, etc thereby eliminating the need of video stabilization.

Real-Time Video Motion Handling and Automated ROI Detection Mechanism



ROI Detection using openCV

ROI Detection using our model

Real-Time Video Motion Handling and Automated ROI Detection Mechanism

effect of motion using openCV

effect of motion using our model

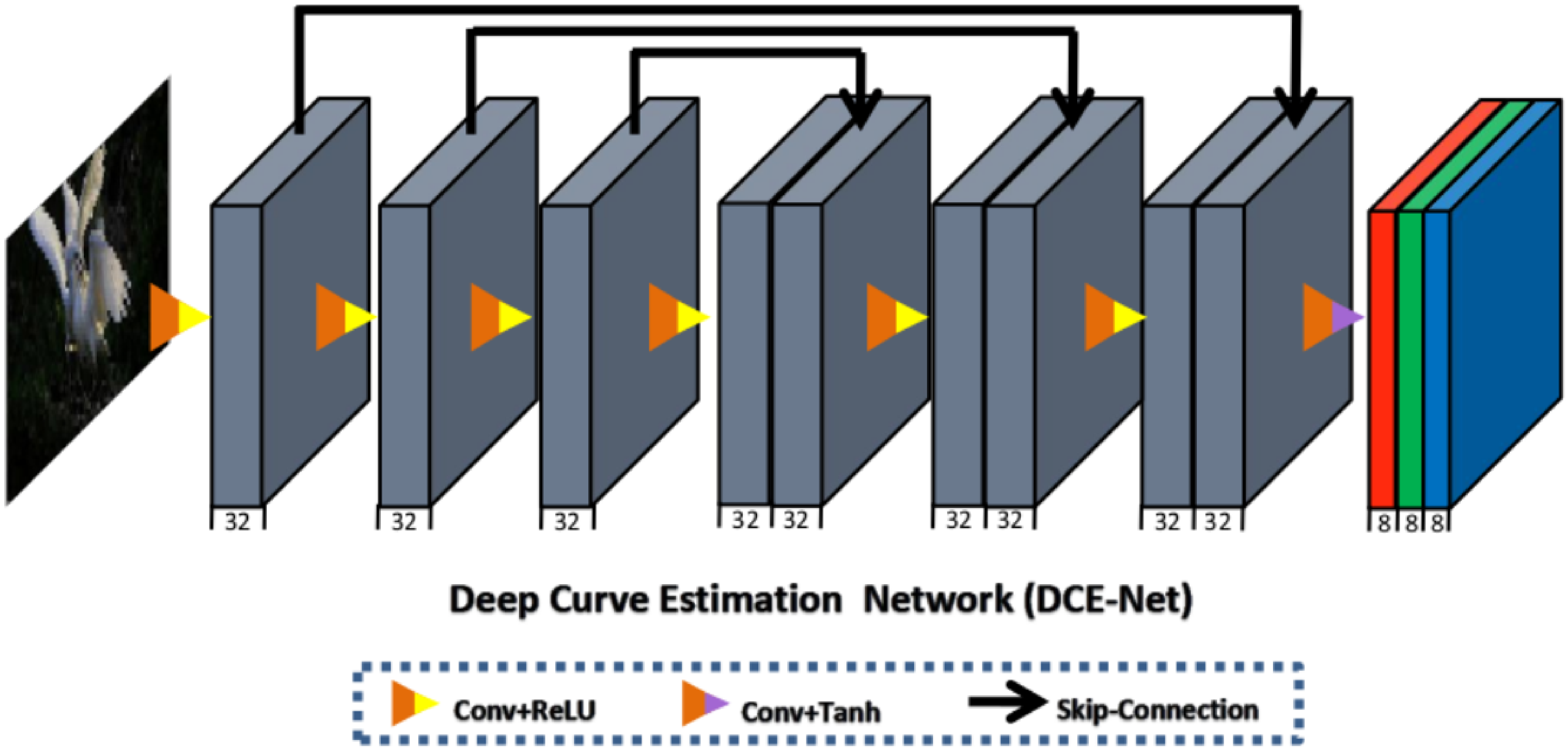
Methodology

Real-Time Low Light Video enahancement

- In addition to the 485 training and 15 testing images of the LoL dataset, we have also included 50 images of human subjects to enhance the performance of Zero-DCE in low-light image enhancement.
- Zero-DCE formulates low-light image enhancement as the task of estimating an image-specific tonal curve with a deep neural network. We train a lightweight deep network, DCE-Net, to estimate pixel-wise and high-order tonal curves for dynamic range adjustment of a given image.
- Zero-DCE takes a low-light image as input and produces high-order tonal curves as its output.
- DCE-Net is a lightweight CNN that learns the mapping between low-light input image and corresponding higher-order curve parameter maps. It has 7 convolutional layers, each with 32 convolutional kernels of size 3x3 and ReLU activation function, and produces 24 parameter maps for 8 iterations with Tanh activation function in the last layer.
- The loss functions used in this model is Color constancy loss, Exposure loss Illumination smoothness loss, Spatial consistency loss

Methodology

Zero DCE Architecture



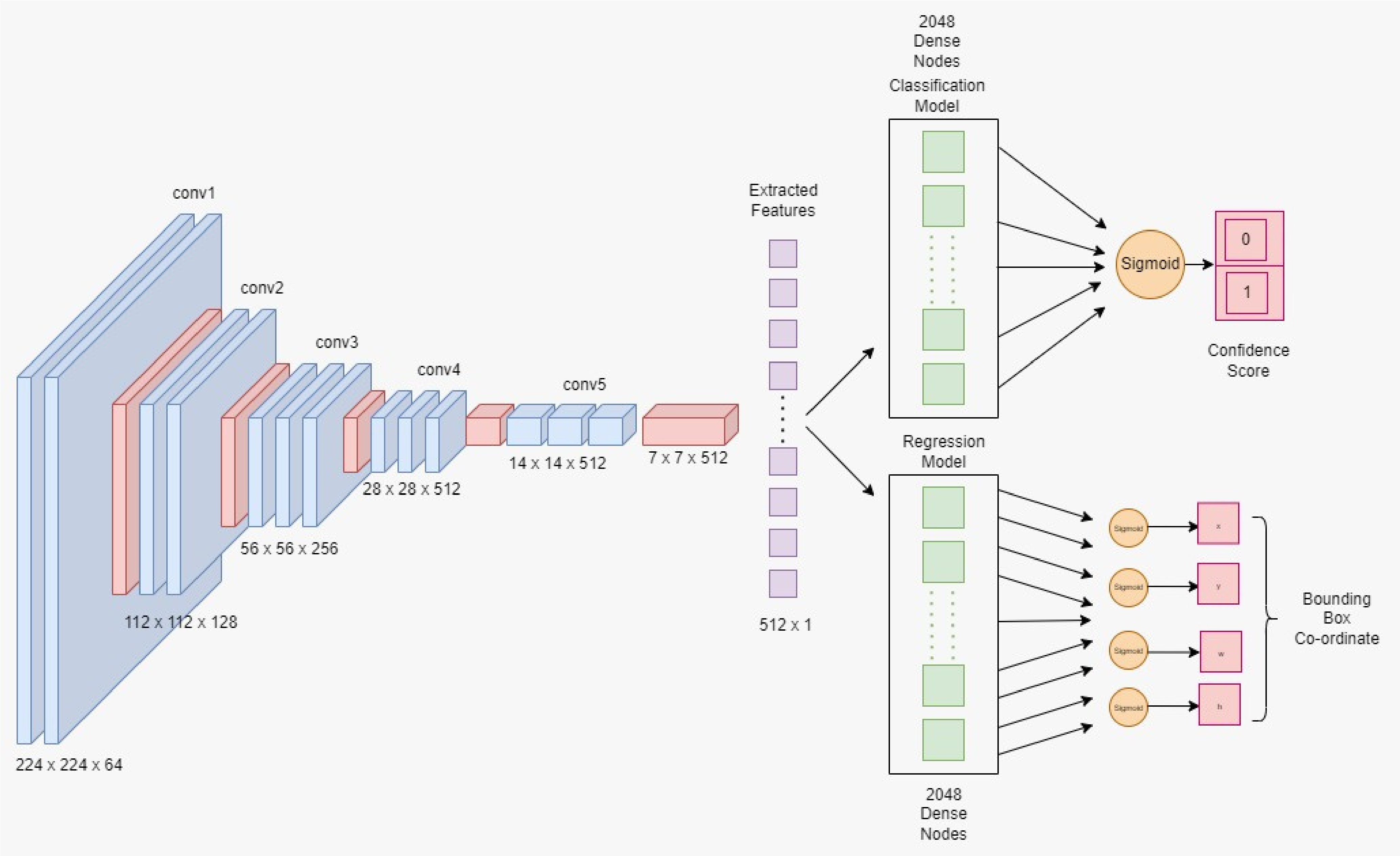
Methodology

Real Time Motion Handling and Automated ROI detection mechanism

- For our experimental setup, we created a dataset for training and testing our model. The dataset consisted of a total of 2455 images, with 2335 images for training, 120 images for testing, and 60 images for validation. We captured an additional 500 images of human subjects in different orientations and with varying motion artifacts using a Python script. The remaining 2000 images were taken from a publicly available human faces dataset. All images were manually labelled using LabelMe with bounding box annotations represented by two coordinates - the upper left and lower bottom points. The dataset was then augmented, resulting in a final dataset size of 23350 images for training, 1200 images for testing, and 600 images for validation.
- Train Test Split-Train = 2335, Test = 120, Validation = 60
- Augmented Train = 23350 , Test = 1200, Validation = 600
- Hyperpaarameters Batch Size = 8,Epoch = 50,Learning Rate=0.0001
- optimizer-Adam
- Activation layer-Relu ,softmax,sigmoid

Methodology

Real Time Motion Handling and Automated ROI detection mechanism



Real Time Motion Handling and Automated ROI detection mechanism

- The feature extraction process plays a crucial role in deep learning-based object detection techniques. In our approach, we utilized the VGG16 model as a feature extractor. Specifically, we dropped the fully connected layers of the VGG16 model and used the output of the last convolution layer, which has dimensions of None,None,None,512, as the input to the subsequent layers.
- Afterward, we employed a Global Max Pooling 2D layer to obtain a feature vector of dimensions [512,1]. These extracted features were then fed into two separate models: a classification model and a regression model.
- The classification model is composed of 2048 dense nodes, with a sigmoid activation function used to produce output in the form of 0 and 1 classes. On the other hand, the regression model also has 2048 dense nodes, and its output is in the form of four co-ordinates: the center co-ordinates of the bounding box, as well as its height and width.
- In summary, our approach uses the VGG16 model to extract features from input images, which are then passed through separate classification and regression models to produce the desired output. By utilizing these models, we are able to achieve our aim of detecting the ROI accurately under the effect of motion.

Trainable parameters

Model Summary:

```
Model: "vgg16"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_1 (InputLayer)        [(None, None, None, 3)]   0

 block1_conv1 (Conv2D)       (None, None, None, 64)    1792
 block1_conv2 (Conv2D)       (None, None, None, 64)    36928
 block1_pool (MaxPooling2D)  (None, None, None, 64)    0

 block2_conv1 (Conv2D)       (None, None, None, 128)   73856
 block2_conv2 (Conv2D)       (None, None, None, 128)   147584
 block2_pool (MaxPooling2D)  (None, None, None, 128)   0

 block3_conv1 (Conv2D)       (None, None, None, 256)   295168
 block3_conv2 (Conv2D)       (None, None, None, 256)   590080
 block3_conv3 (Conv2D)       (None, None, None, 256)   590080
 block3_pool (MaxPooling2D)  (None, None, None, 256)   0

 block4_conv1 (Conv2D)       (None, None, None, 512)   1180160
 block4_conv2 (Conv2D)       (None, None, None, 512)   2359808
 block4_conv3 (Conv2D)       (None, None, None, 512)   2359808
 block4_pool (MaxPooling2D)  (None, None, None, 512)   0

 block5_conv1 (Conv2D)       (None, None, None, 512)   2359808
 block5_conv2 (Conv2D)       (None, None, None, 512)   2359808
 block5_conv3 (Conv2D)       (None, None, None, 512)   2359808
 block5_pool (MaxPooling2D)  (None, None, None, 512)   0

=================================================================
Total params: 14,714,688
Trainable params: 14,714,688
Non-trainable params: 0
_________________________________________________________________
```

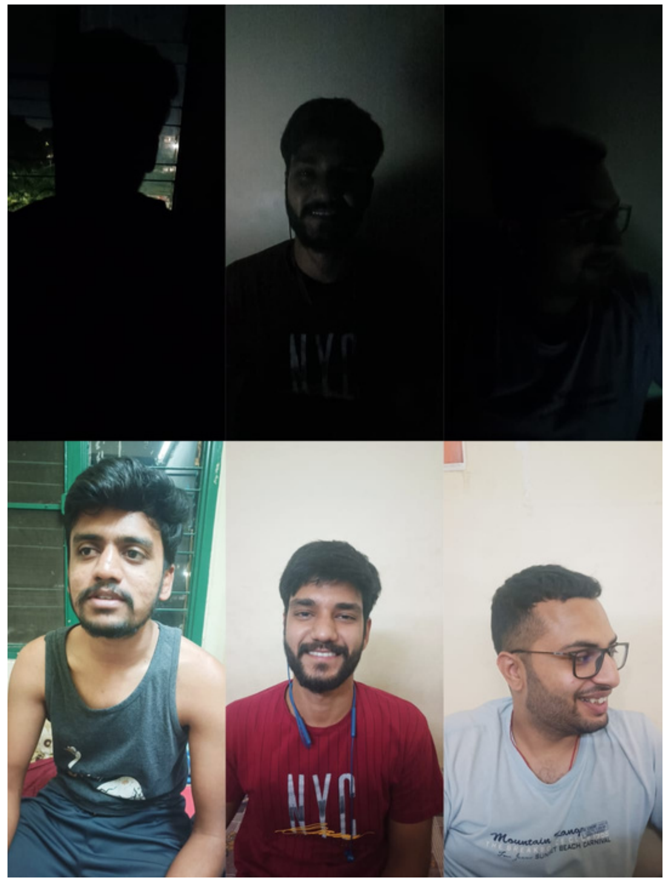
Methodology

Heart rate Measurement

- Once, the forehead region is extracted as the ROI. The forehead is chosen as it has a good blood flow and is less susceptible to motion artifacts.
- Hue Channel Extraction: The ROI is then converted from the RGB color space to the HSV color space. The hue channel is extracted from the ROI as it is less affected by lighting changes compared to other color channels.
- Signal Processing: The hue channel signal is pre-processed to remove noise and enhance the signal. This is done using a bandpass filter and a moving average filter.
- Peak Detection: The peaks in the pre-processed signal are detected using a peak detection algorithm. The peaks correspond to the heartbeats.
- Heart Rate Calculation: The heart rate is calculated by measuring the time difference between consecutive peaks and converting it to beats per minute (BPM

Experimental setup

Data set for motion handling and
Automated ROI detection



Data set low low light enhancement

Experimental setup

- For our experiment, we collected a dataset of 2335 images for training, 120 images for testing, and 60 images for validation. An additional 500 images were captured in various head orientations and motions, and manually labelled using LabelMe. After augmentation, the dataset was expanded to 23350 images for training, 1200 images for testing, and 600 images for validation.
- For the experimental setup, we used the LoL dataset, which comprises 485 low-light images for training and 15 for testing. Additionally, we incorporated 50 images of human subjects to improve the performance of Zero-DCE in low-light image enhancement.
- The heart rate detection experiments were conducted on an HP Pavilion laptop with a built-in webcam capable of capturing 30 frames per second at a resolution of 640*480 pixels.
- Python version 3.7.7 was used for all implementations, and an initial delay of 2 seconds was required for HR estimation due to the duration of each heartbeat, which ranges from 0.45 to 1.25 seconds.
- The average of all detected HR within the last second was considered after the initial estimation, and the stable values were obtained by averaging the ten most recent HR.
- The heart rate detection results were compared with those obtained from two contact-based devices: a Samsung Watch and a Boat Watch 6.

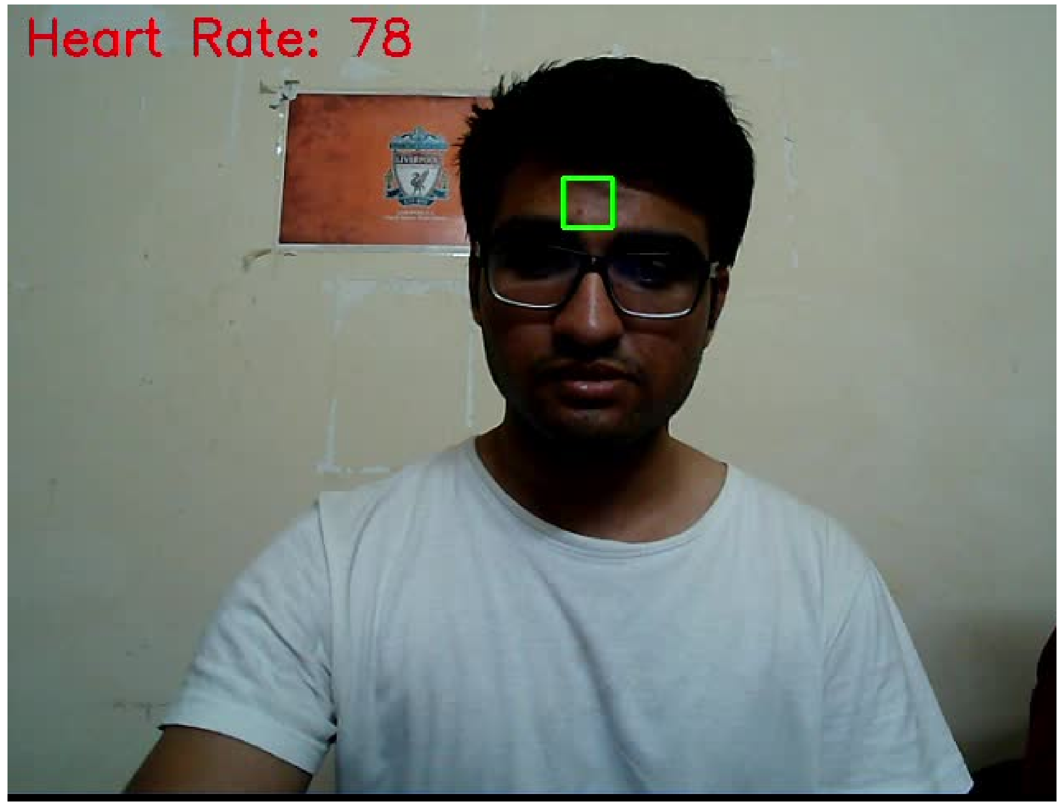
Results

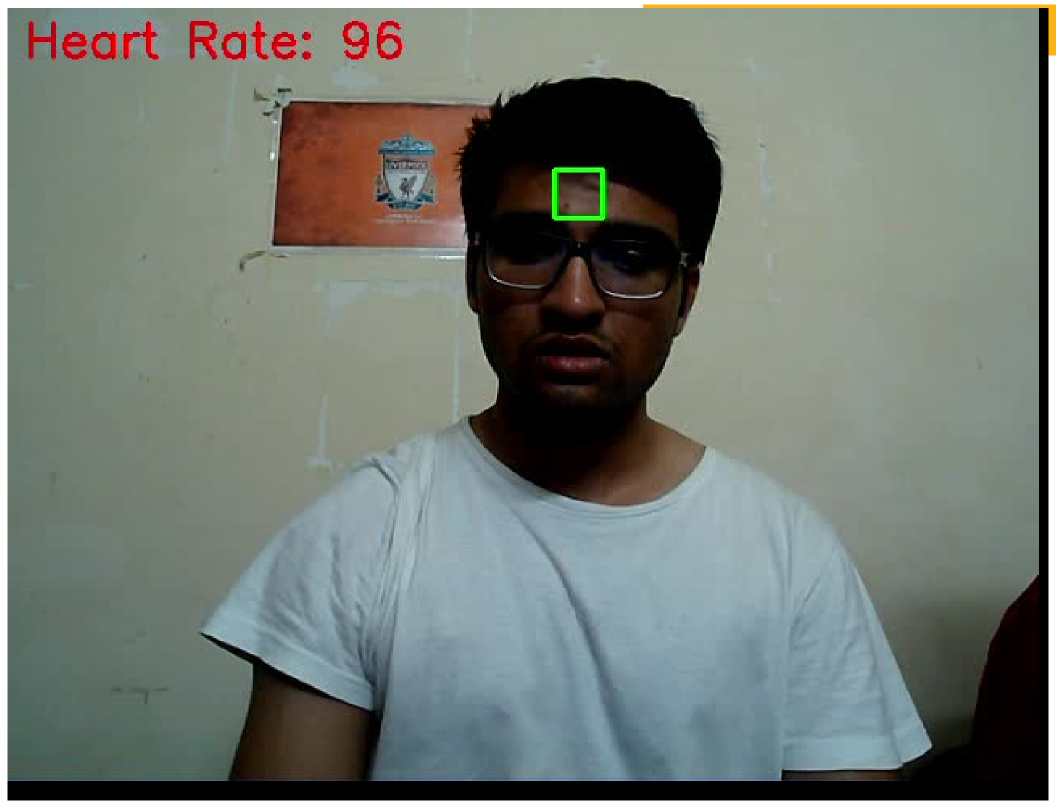
conditions	Boat Smart Watch	Samsung Smart Watch
Ideal conditions	4.02	5.69
low light condition	7.39	10.81
motion video	5.36	5.76

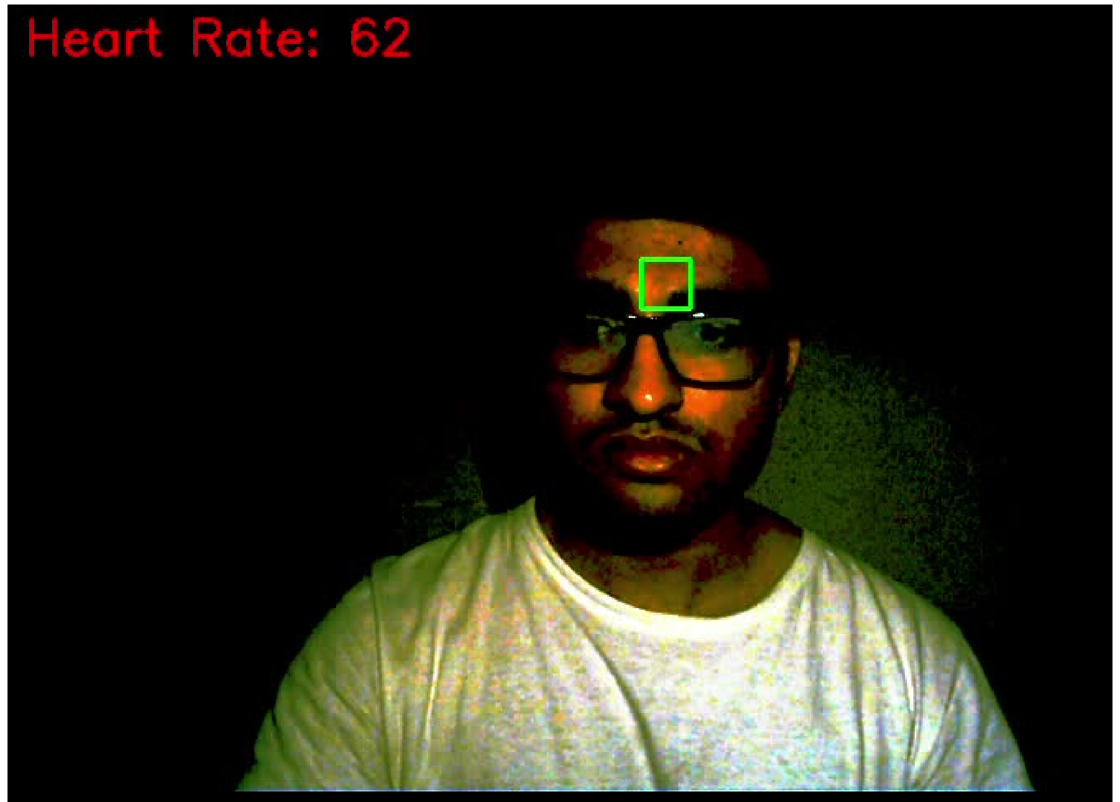
Discussion

In our heart rate detection models, we utilized the VGG16 model as a feature extractor, dropping its fully connected layers and using the output of the last convolution layer. We then employed a Global Max Pooling 2D layer to obtain a feature vector that was fed into separate classification and regression models. Our approach allowed for accurate detection of the ROI and measurement of heart rate in real-time through non-invasive means, making it a valuable tool in medical imaging and diagnosis. The efficient zero DCE model was also utilized in our heart rate detection model, further enhancing its accuracy and efficiency. In summary, our approach showcases the importance of feature extraction in deep learning-based techniques, and demonstrates the potential for these models in various fields such as medical imaging and diagnosis.

Project Demonstration







Contribution of each team member

1.	Allan Robey	20%
2.	Avnish Tripathi	20%
3.	Divyesh Tripathi	20%
4.	Kush Shah	20%
5.	Pulkit Sharma	20%

References

1) Heart Rate Measurement Using Face Detection in Video (Carmen Nadrag, Vlad Poenaru, and George S
Department, Beia Consult International, Bucharest, Romania{carmen.nadrag, vlad.poenaru, george}@beia.ro)
2) Real Time Video based Heart and Respiration Rate Monitoring (Jafar Pourbemany, Almabrok Essa, and Ye ZhuDep
of Electrical Engineering and Computer ScienceCleveland State University, Cleveland, OH, USA)
3) Real-Time Webcam Heart-Rate and Variability Estimation with Clean Ground Truth for Evaluation (Amogh Gudi
Bittner and Jan van Gemert)
4) Effects of Lighting and Window Length on Heart Rate Assessment through Video Magnification(Lee
Kassab,Andrew Law,Bruce Wallace,Julien Larivière-Chartier,Rafik Goubran,Frank Knoefel)
5) EnlightenGAN: Deep Light Enhancement without Paired Supervision (Yifan Jiang , Xinyu Gong, Ding Liu, Yu Che
Fang, Xiaohui Shen, Jianchao Yang,Pan Zhou , and Zhangyang Wang , Member, IEEE)
6)
7) Assessment of ROI Selection for Facial Video-Based rPPG (Dae-Yeol Kim, Kwangkee Lee, Chae-Bong Sohn)
8) Deep Online Video Stabilization With Multi-Grid Warping Transformation Learning (Miao Wang, Guo-Ye Yang, Jin
Song-Hai Zhang, Ariel Shamir, Shao-Ping Lu, Shi-Min Hu)
9) Heart Rate Measurement Using Facial Videos (Carmen Nadrag, Vlad Poenaru, George Suciu)
10) Contact-Less Heart Rate Detection in Low Light Videos (Tamal Chowdhury, Sukalpa Chanda, Saumik Bhattachar
Biswas & Umapada Pal)
11) Learning Video Stabilization Using Optical Flow" by Yanchao Yang, Deqing Sun, Huaizu Jiang, and Ming-Hsuan Ya
2017)
12) Low-Light Video Enhancement Using Generative Adversarial Networks With Channel Attention" by
Shangwen Liang, and Shiqi Wang (IEEE Access 2019)
13) Heart Rate Measurement Combining Motion and Color Information (2022) (Jean-Pierre Lomaliza, Hanhoon Park
Seok Moon)
14) Heart rate prediction from facial video with masks using eye location and corrected by convolutional neural netw
15)Multi-level Attention Network for Low-light Image/Video Enhancement" by W. Ren et al