Facial Expression Recognition

Group - 1

Krutik Parmar
201801199

Ayushi Patel
201801203

Shivangi Zala
201801457

Outline

01 Problem Statement

02 Dataset

03 Methodology

04 Results

Problem Statement

Video-based facial expression recognition aims to classify a video into one of seven basic emotions, i.e anger, contempt, disgust, fear, happiness, sadness, and surprise.

- Input
  - Video clip
- Output
  - One of seven expression classes i.e anger, contempt, disgust, fear, happiness, sadness, and surprise.

- Image source : https://unsplash.com/photos/ufK-deiLqY8

Dataset

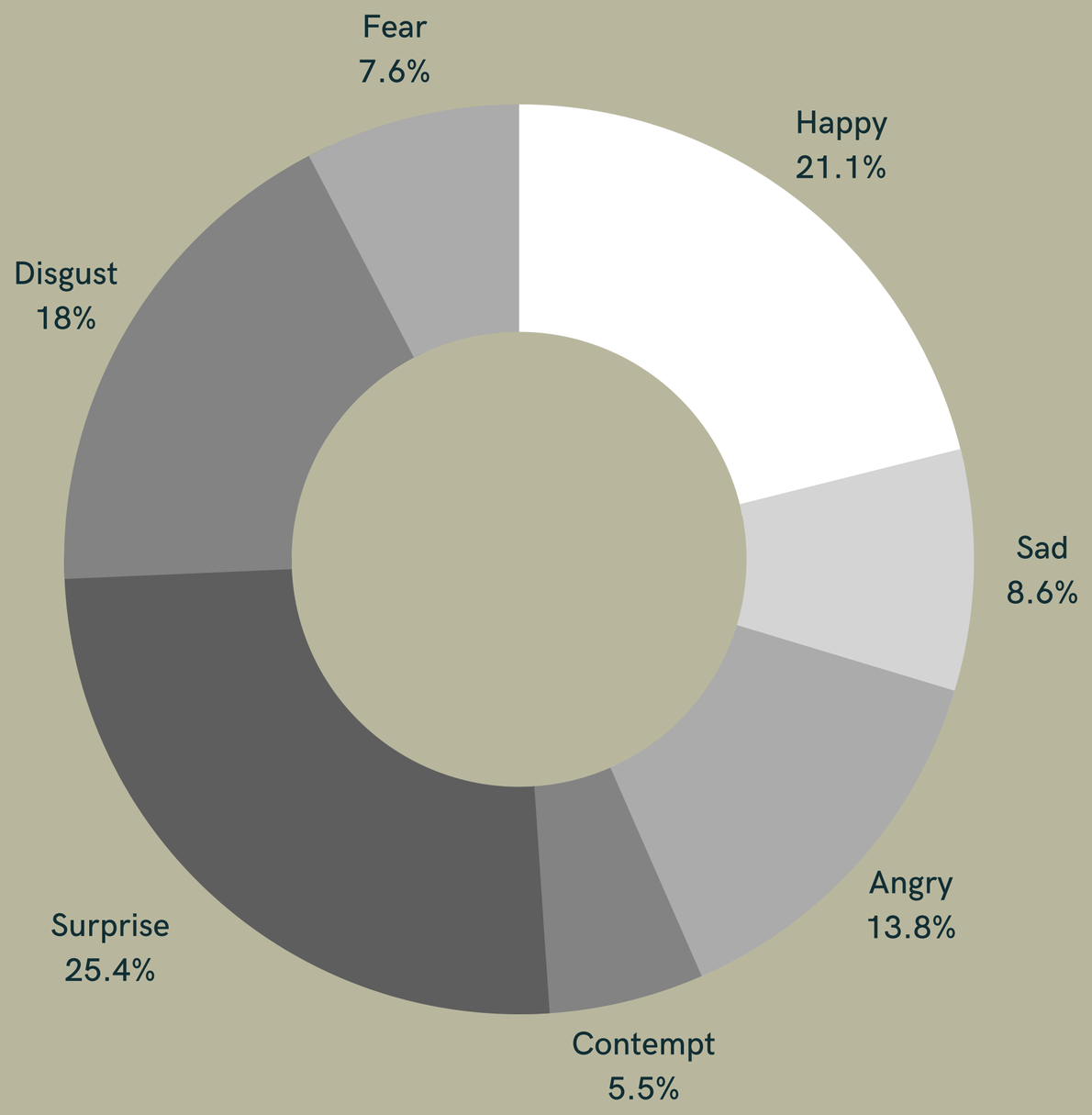
Extended Cohn-Kanade (CK+) dataset

Video Examples

- 593 video sequence
- 123 different subjects

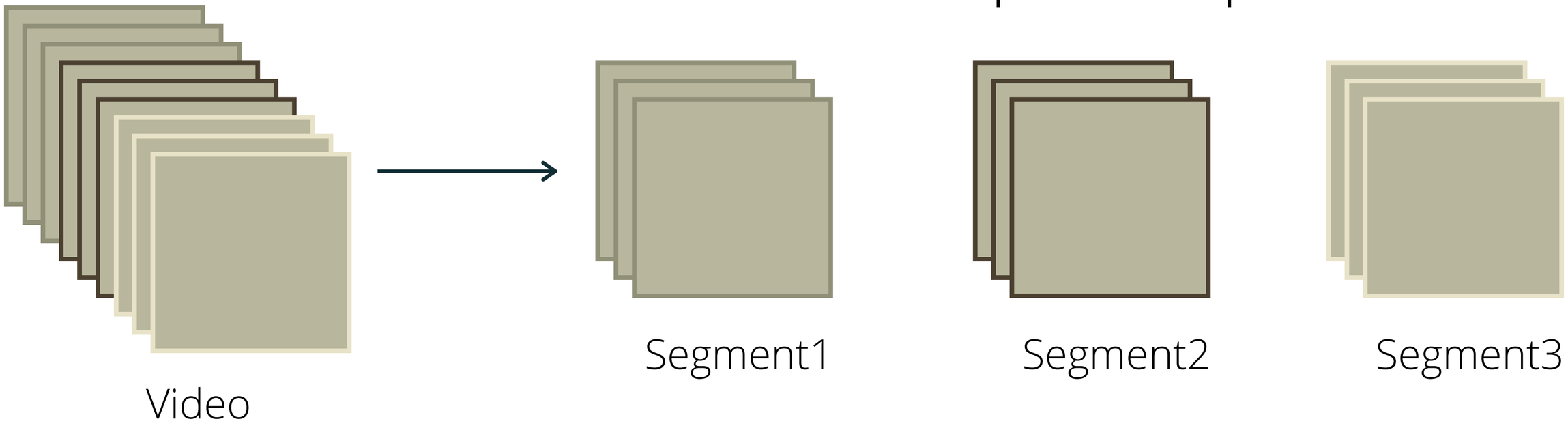
Frame

- 30 FPS
- 640x490 pixels


Fear
7.6%
Happy
21.1%
Disgust
18%
Sad
8.6%
Angry
13.8%
Surprise
25.4%
Contempt
5.5%

Data Preparation

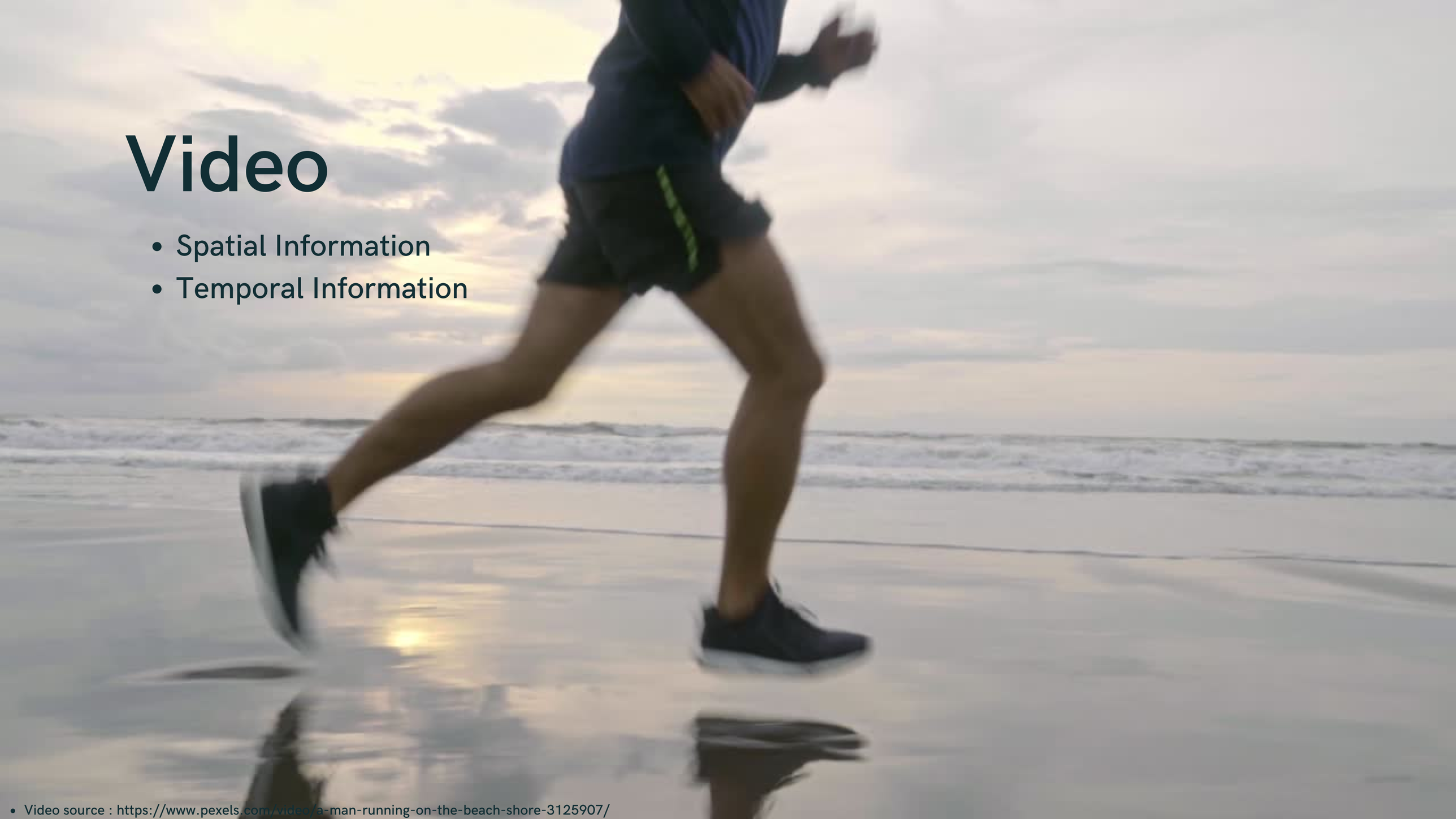
- Cropped faces from frames
- We have divided each video into 3 segments
- Randomly select one frame from each segment
- Create a stack of those selected frames
- Resize frames into 224x224 size and pass it as input



Methodology

Video

- Spatial Information
- Temporal Information



- Video source : https://www.pexels.com/video/a-man-running-on-the-beach-shore-3125907/

The Modules

01

deep CNN

We have used ResNet-18 as a deep CNN module.

02

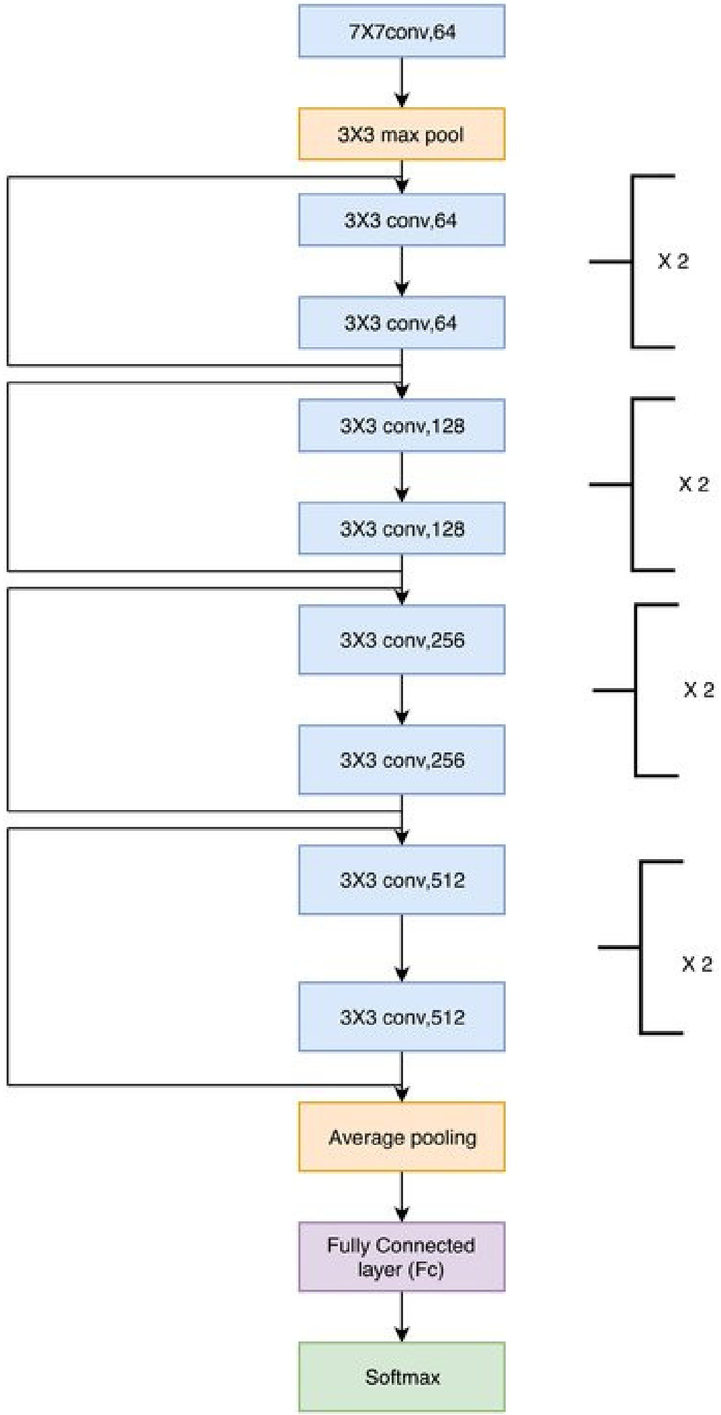
PAN

Persistent Appearance Network is used to capture temporal information.

03

Frame-Attention

To assign attention weights to frame of the video.



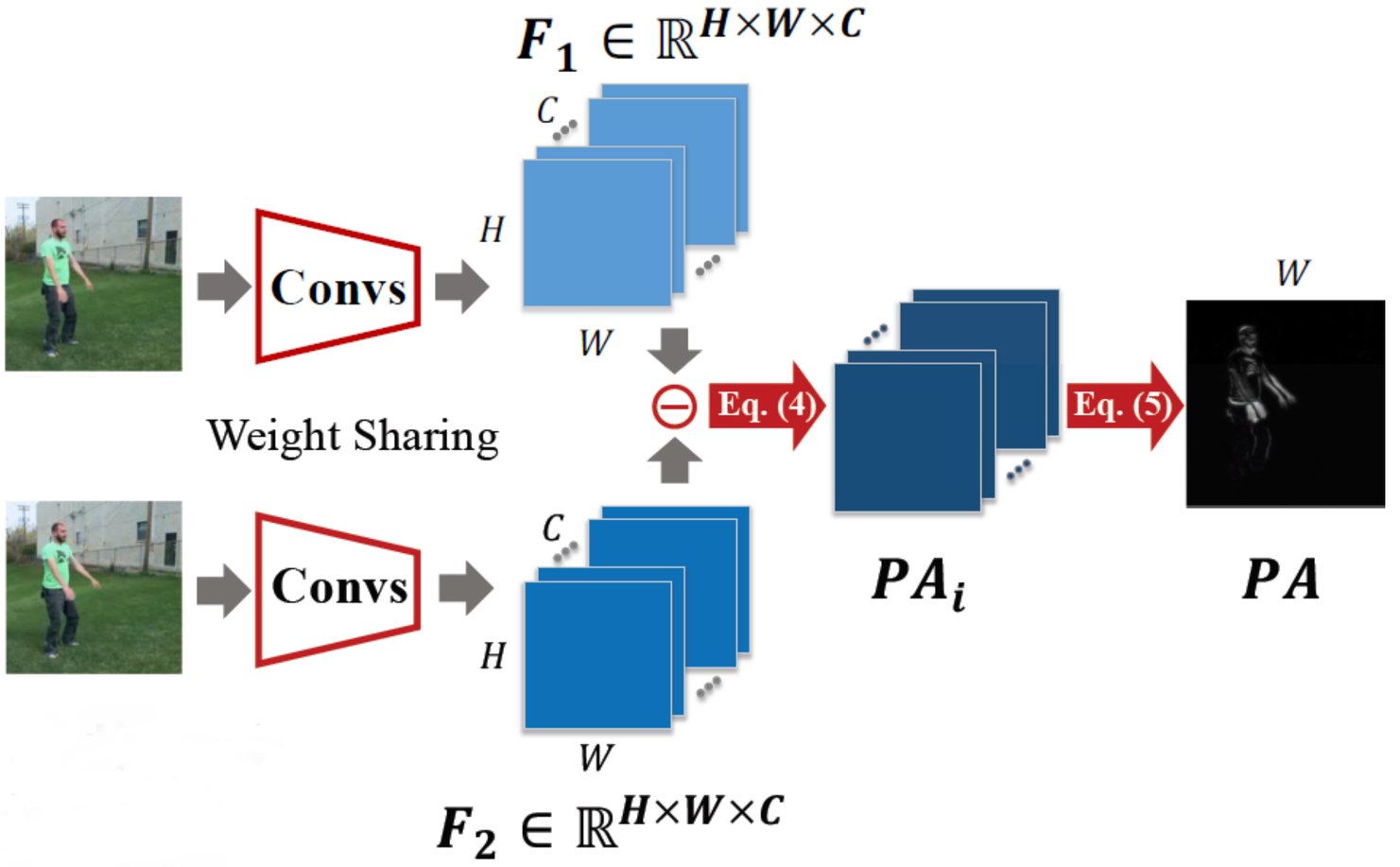
ResNet-18

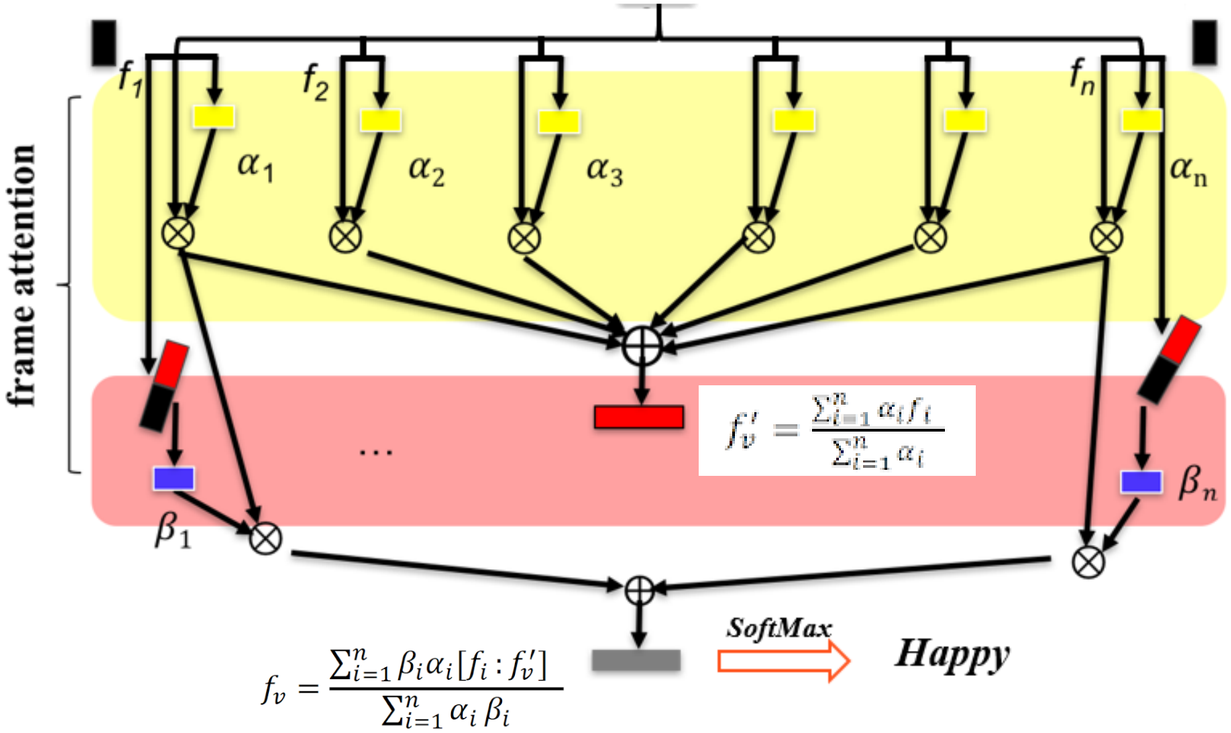
- Deep
- High Accuracy
- Advantage of residual block

PAN

Persistent Appearance Network

- PAN can model small displacements between frames.
- In FER, small displacements are mouth, forehead and eye's displacement.

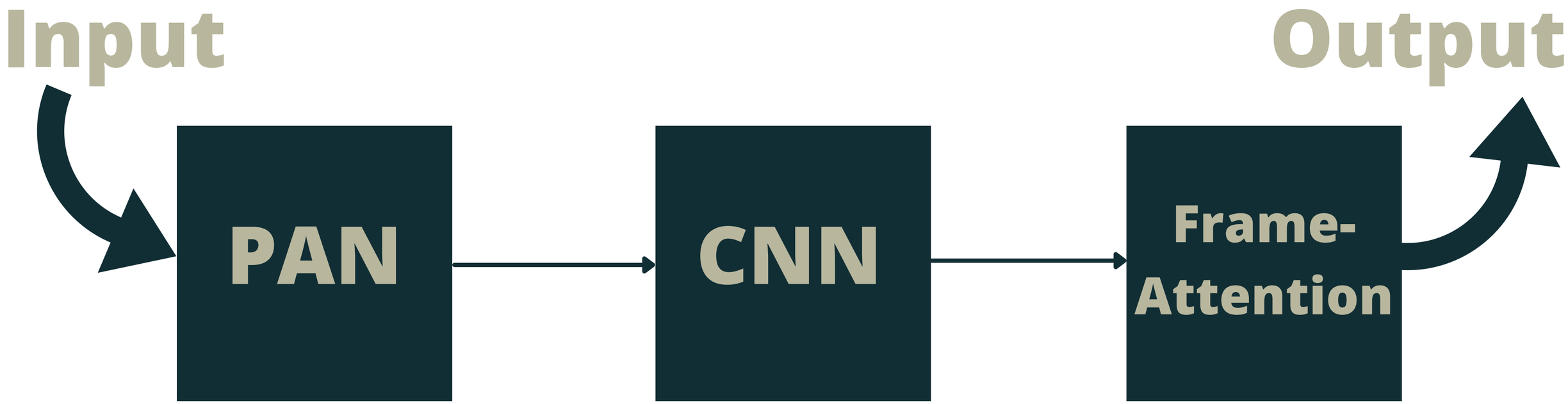




Frame-Attention

- Assign attention weights to the frames
- Generate a global feature vector f'_v
- Creates a final feature vector f_v

Final Model



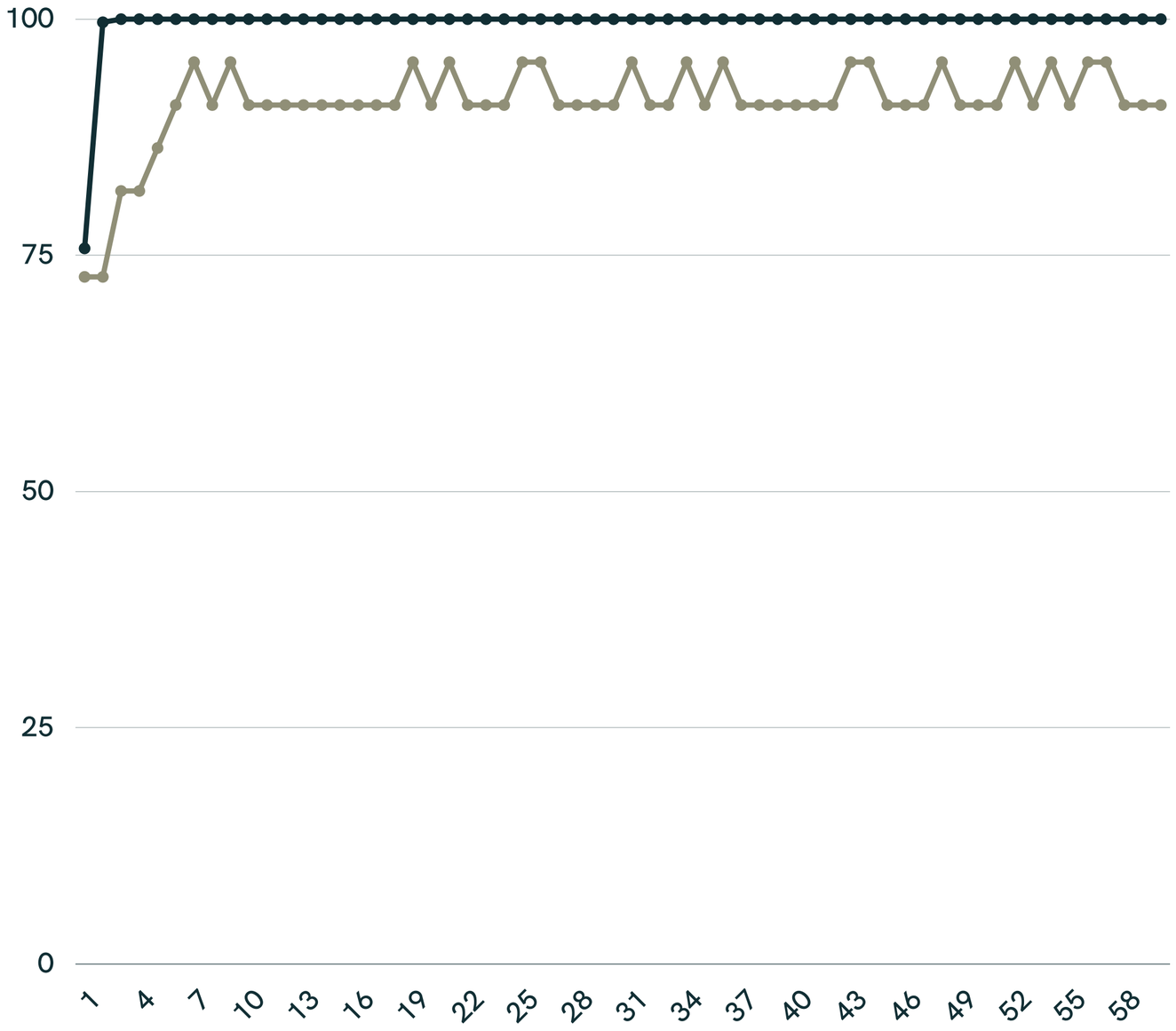
Trasnfer Learning

- We have used a ResNet-18 model which was pre-trained on the FER+ dataset.
- In our model, we trained Frame-attention network and PAN from scratch.

Optimizer and Hyperparameters

- SGD optimizer with momentum
- Learning rate - 0.02
- Epochs - 60
- Batch size - 48

Results

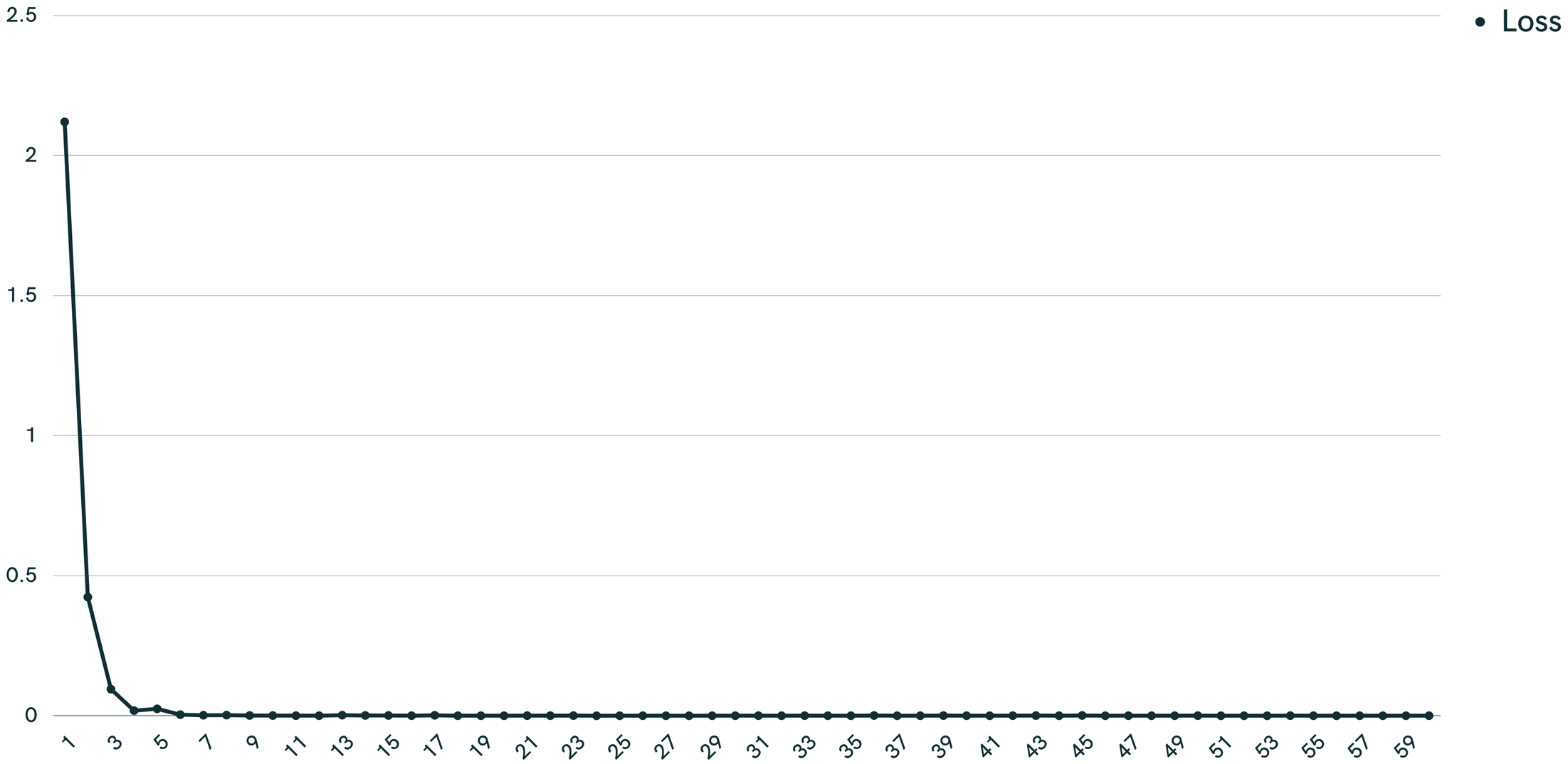


- Traning accuracy
- Validation accuracy

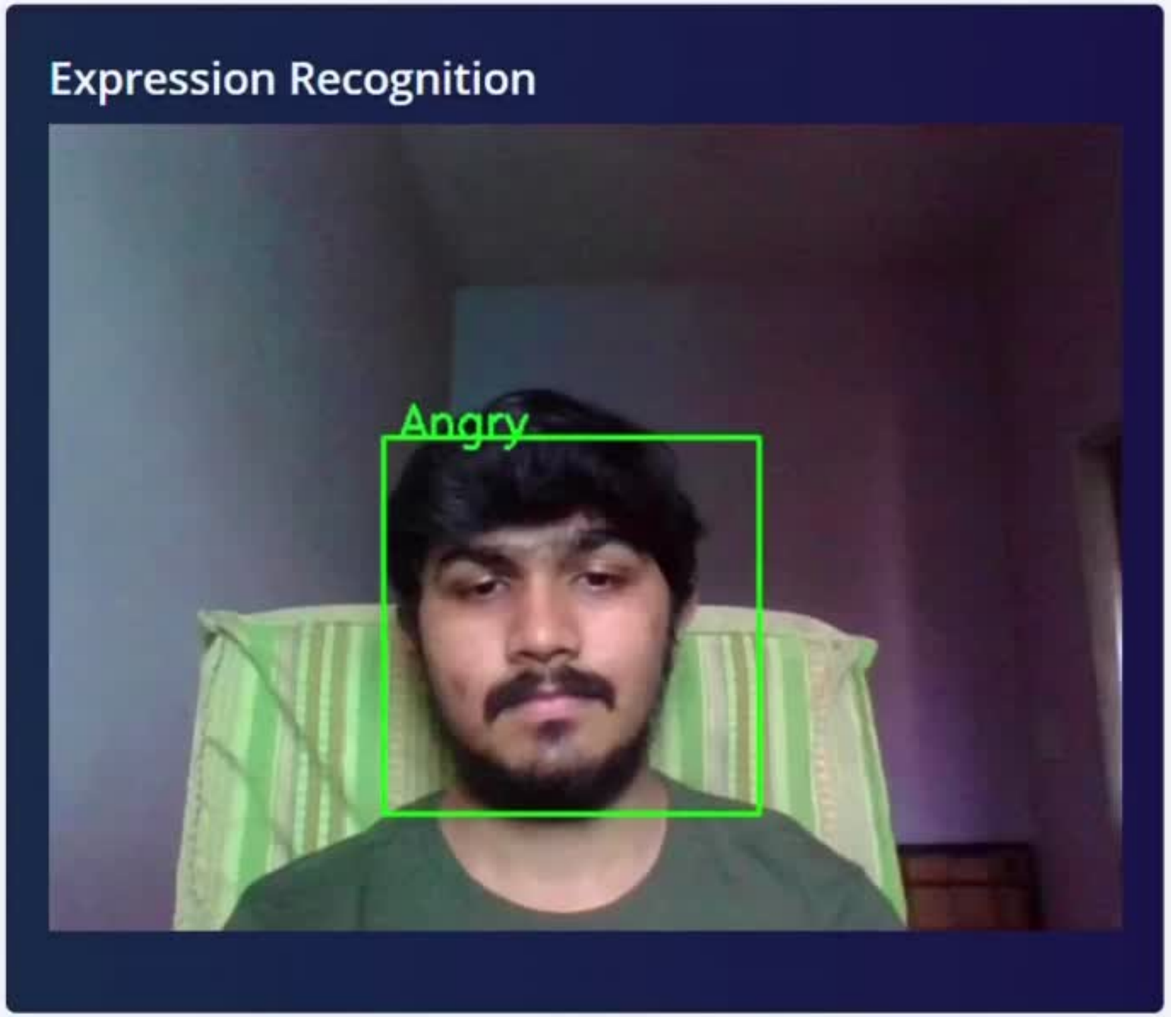
Accuracy

- We have achieved accuracy of 95.455

Cross Entropy Loss



Project : Facial Expression Recognition



Final Result

References

- Huang, Y.; Chen, F.; Lv, S.; Wang, X. Facial Expression Recognition: A Survey. Symmetry 2019, 11, 1189. https://doi.org/10.3390/sym11101189
- Meng, Debin & Peng, Xiaojiang & Wang, Kai & Qiao, Yu. (2019). Frame Attention Networks for Facial Expression Recognition in Videos. 3866-3870. 10.1109/ICIP.2019.8803603.
- Can Zhang, Yuexian Zou, Guang Chen, and Lei Gan. 2019. PAN: Persistent Appearance Network with an Efficient Motion Cue for Fast Action Recognition. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19). Association for Computing Machinery, New York, NY, USA, 500–509. DOI:https://doi.org/10.1145/3343031.3350876
- Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, 46-53.
- Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I.(2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.

Thank You