Hand Gesture Tracking and Recognition in Videos

Suchibrata Bhowmik

Instructor : Dr. Sujoy Biswas

Machine Learning Systems

RKMVERI, Belur

Motivation

- Keyboard, Mouse have significant role in HCI.
- Now speech recognition and gesture recognition receive great attention in the field of HCI.

Fig. Indian sign language

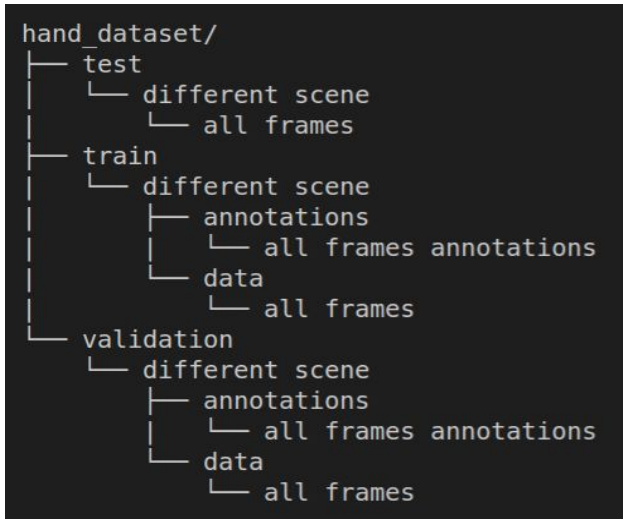
Fig. Robot control

Introduction

- Deep neural networks are a powerful tool for learning image representations in computer vision application.
- Object tracking is an important challenging task within the area in computer vision.
- I present deep learning methods to track the hand gesture and recognise it.
- Extract the feature and feature representation of the target object is compared to the search region.

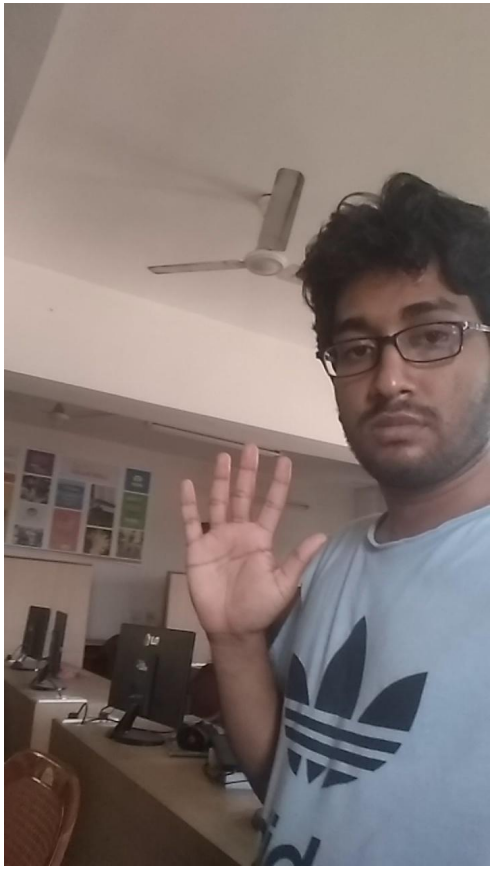
Dataset

- Collect 24 videos by laptop camera and phone camera in different scene with different gesture.
- Extract frames from all videos and save frames in following file structure

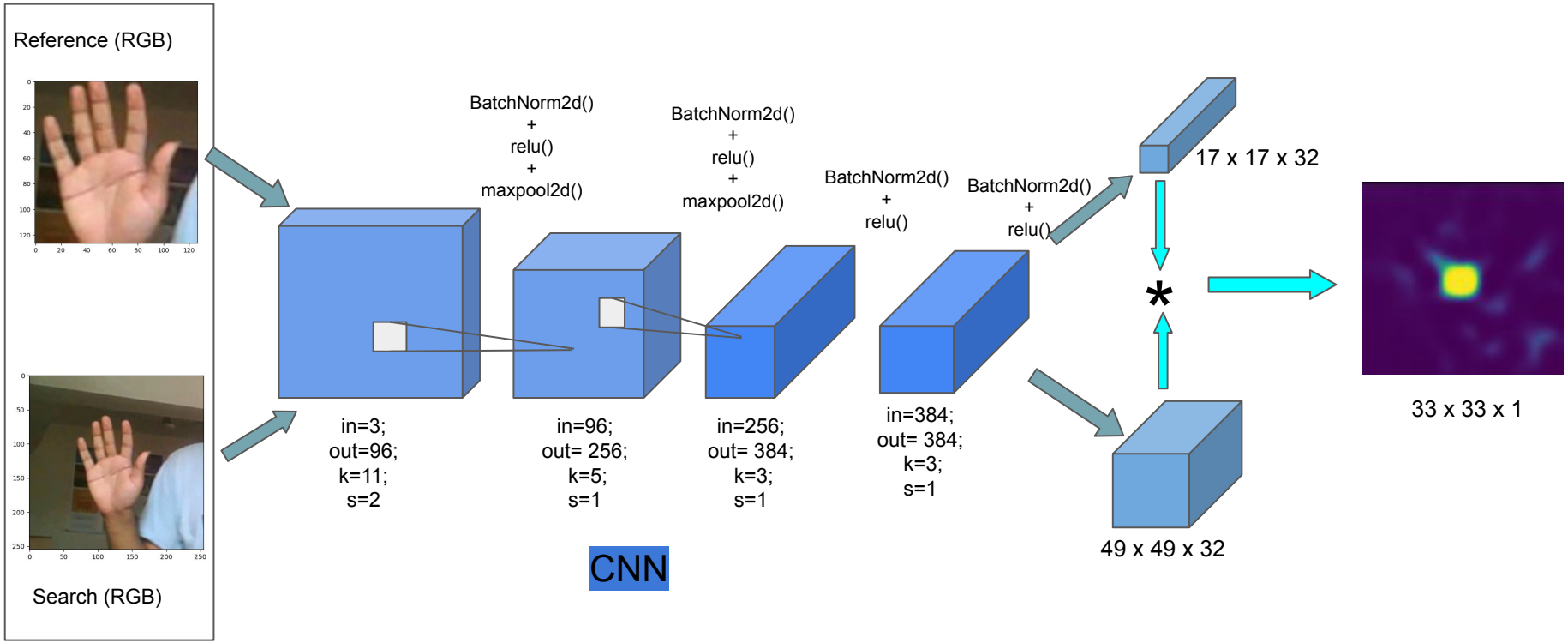


- Train set contain 16 videos, where total number of frame is 12339.
- Validation set contain 8 videos, where total number of frame is 3020

Data Sample



Fully-Convolutional Siamese Network



Dataset Preprocess

- Choose two frame one for reference image another for search image.
- First resize the object then crop from frames.

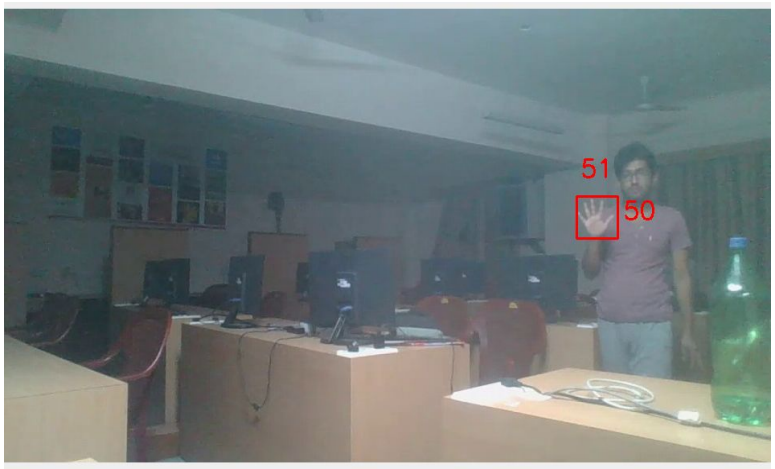


Fig. original frame

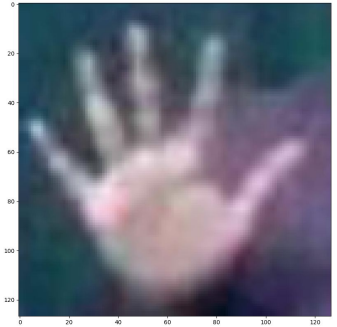


Fig. reference image

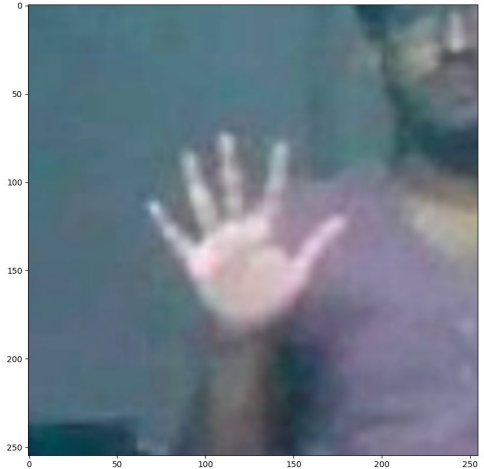


Fig. search image

- Another sample where add padding.

Fig. original image

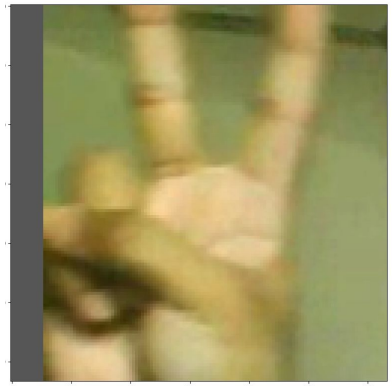


Fig. reference image

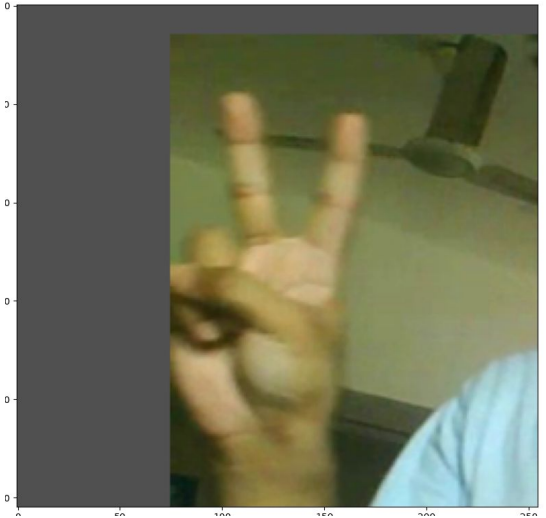


Fig. search image

Training

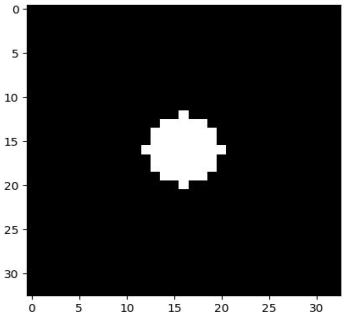


Fig. original label

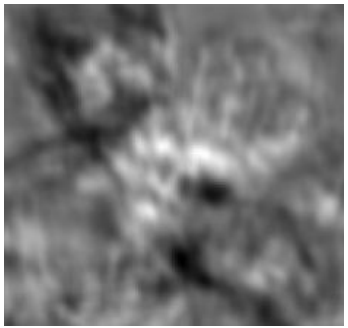


Fig. predicted label

Batch size	Epoch	Learning rate	Optimizer	Loss
64	15	1e-2	SGD	BCEWithLogitsLoss

Training Result

- Calculate the Area under the ROC Curve, Center error, Loss over validation dataset.


```
{
    "AUC": 0.9603479744922644,
    "center_error": 9.910661631253152,
    "loss": 0.07811034641539057,
    "epoch": 13.0
}
```

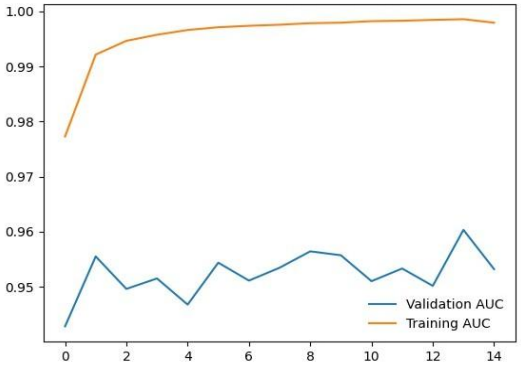



Fig. AUC plot

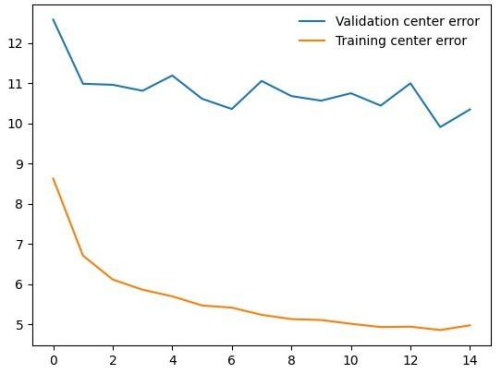


Fig. center error plot

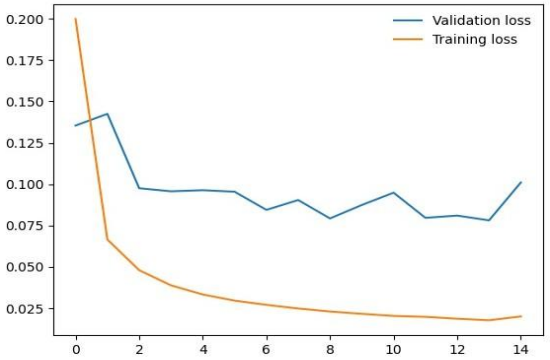
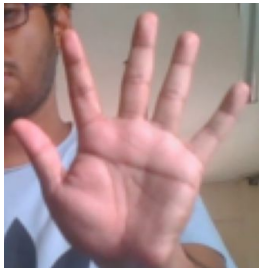
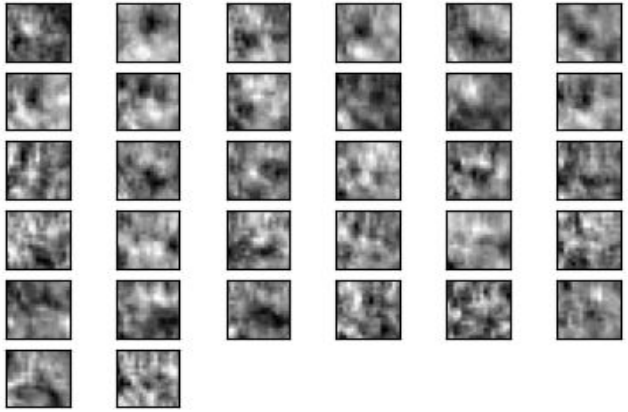


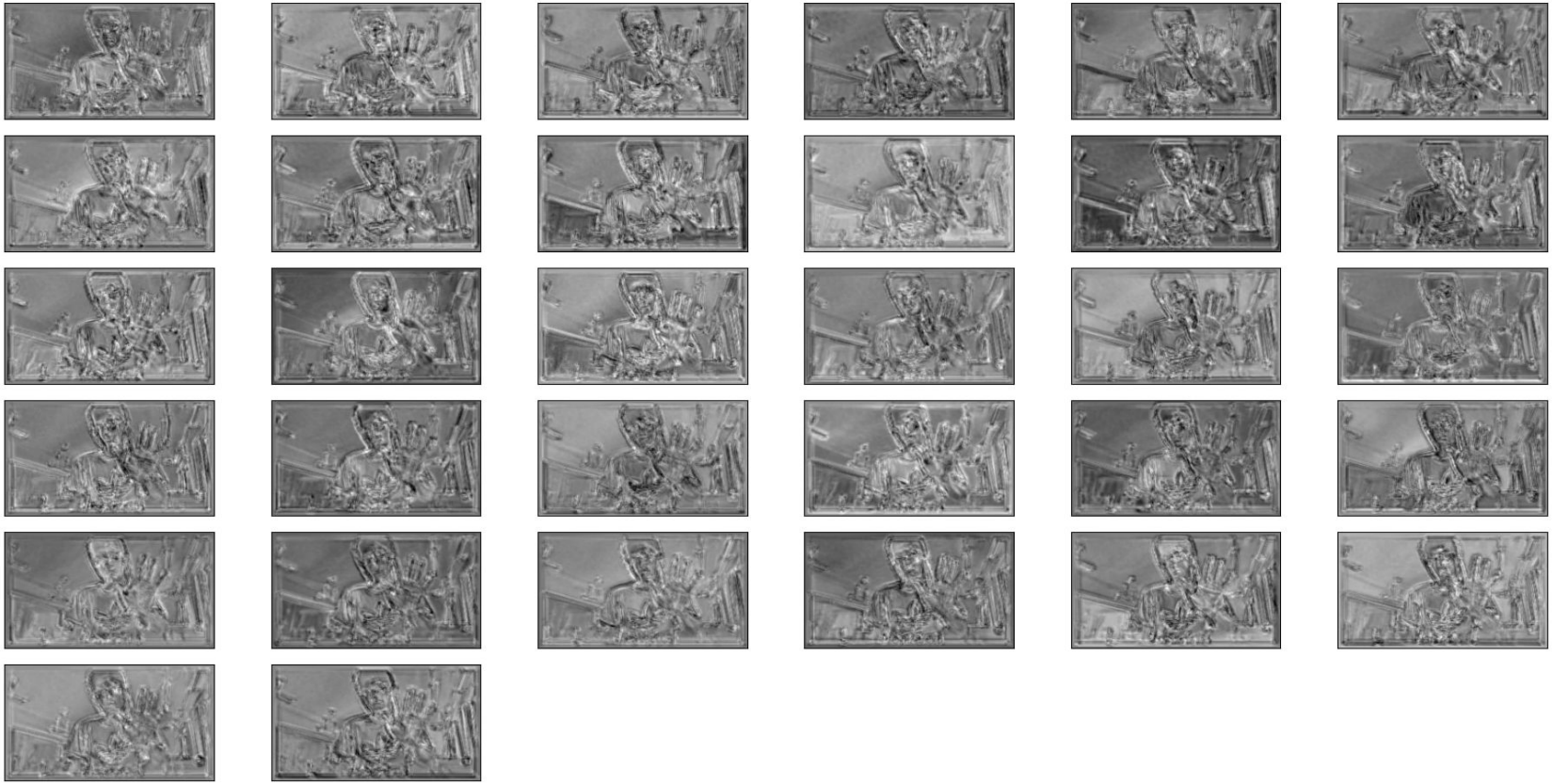
Fig. loss plot

Inference





- These are the outputs of Convolution Layers



- Get cross correlation then pass the map in sigmoid function

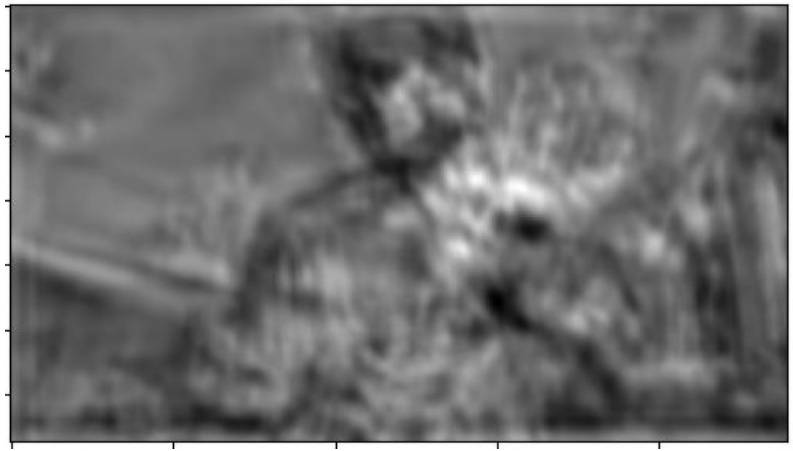


Fig. Cross Correlation

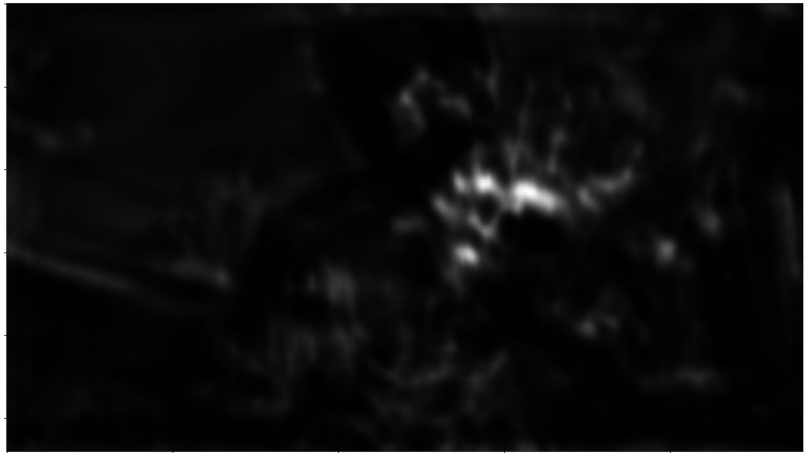
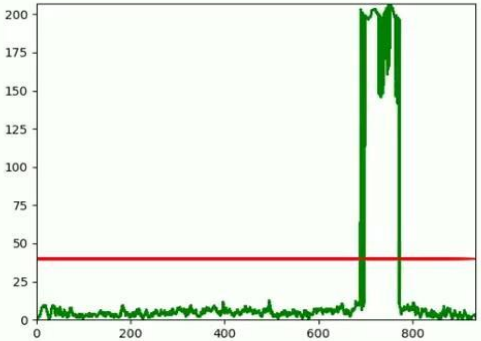


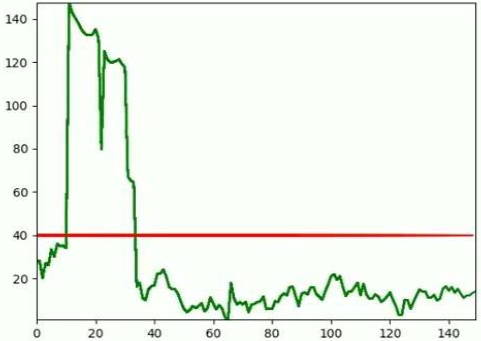
Fig. Score map

Evaluation

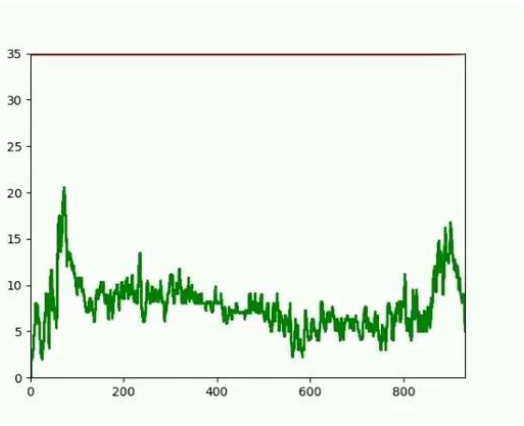
frm_no : 934
center_err : 1.414
score : 1.0

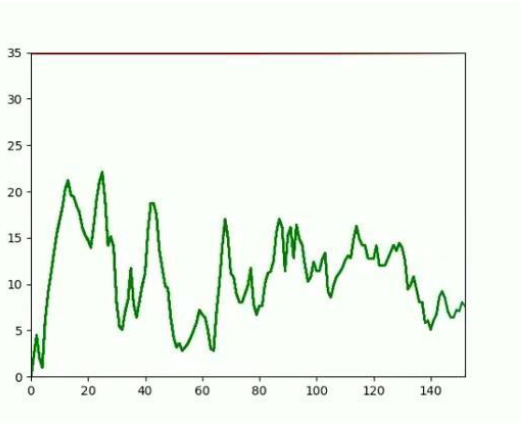

200
175
150
125
100
75
50
25
0
0
200
400
600
800

frm_no : 149
center_err : 14.036
score : 0.774
adidas


140
120
100
80
60
40
20
0
20
40
60
80
100
120
140

Optical Flow





Classification

Reference

- https://openaccess.thecvf.com/content_cvpr_2017/papers/Valmadre_End-To-End_Representation_Learning_CVPR_2017_paper.pdf
- https://github.com/rafellerc/Pytorch-SiamFC