CV2 Assignment 2: Repetitive Action Count

Dai ZiJia 2022233158

April 24, 2023

1 Introduction

In the assignment 2, I will use the state-of-art algorithm on repetitive action count named "PoseRAC: Pose Saliency Transformer for Repetitive Action Counting". I will show the network frame work and give the obtained results.

Innovation: They proposed pose saliency representation mechanism using the two most salient poses to represent each action, which captures essential information while reducing computational overhead. The approach estimates the pose of each frame and utilizes such core information to classify actions frame-by-frame reduces calculation. Finally, it uses a lightweight action-trigger to obtain the final result.

2 The Network

For an input video, the repetitive count can be obtained through Pose Estimation, Transformer Encoder, Pose Mapping and Action-trigger, where only the Encoder and the Pose Mapping need to be trained. The network uses Triplet Margin Loss to train the Encoder while Binary Cross Entropy Loss to train both the Encoder and the Pose Mapping.

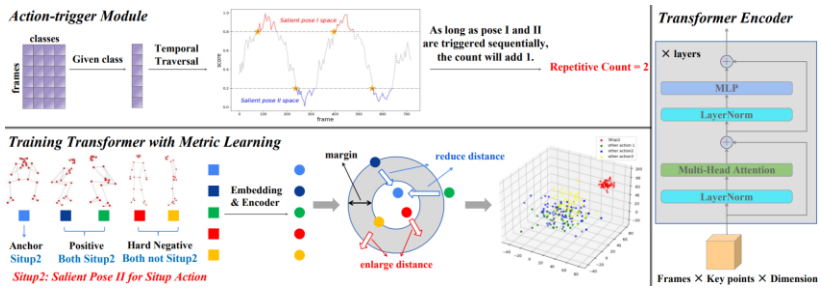
It's loss functions:

$$\mathcal{L}_{bce} = -\frac{1}{N}\sum_{i=1}^{N}(\frac{1}{C}\sum_{j=1}^{C} loss(i,j))$$

$$loss(i,j) = y_{ij}\log p_{ij} + (1-y_{ij})\log(1-p_{ij})$$

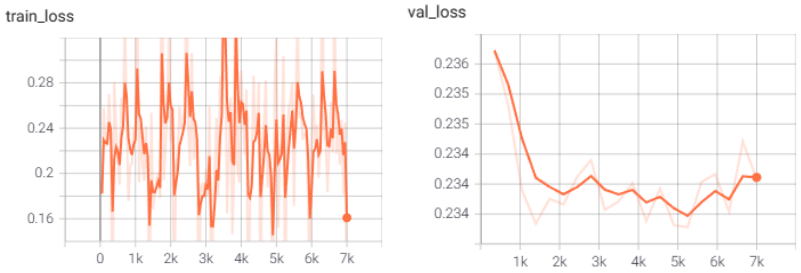
$$\mathcal{L}_{tri} = \max(\mathrm{CS}(a,p) - \mathrm{CS}(a,n) + \mathrm{margin}, 0)$$

It' s the modules:



3 Results

Training and Validation Loss:



MAE and OBE: The main evaluation metrics used in previous work are Off-By-One (OBO) count error and Mean Absolute Error (MAE). OBO measures the error rate of repetition count over the entire dataset, while MAE represents the normalized absolute error between the ground truth and the prediction. They can be defined as:

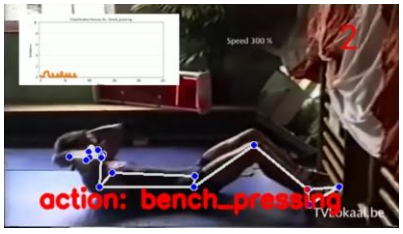
$$\mathbf{OBO} = \frac{1}{N}\sum_{i=1}^{N}[|\tilde{c}_i - c_i| \leq 1]$$

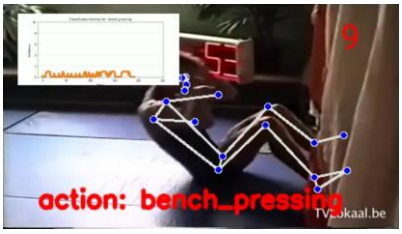
$$\mathbf{MAE} = \frac{1}{N}\sum_{i=1}^{N}\frac{|\tilde{c}_i - c_i|}{\tilde{c}_i}$$

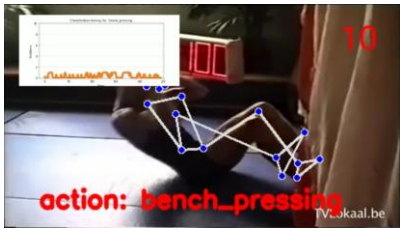
```
video input path ./RepCount_pose/video/test/stu12_32.mp4

video input path ./RepCount_pose/video/test/stu12_33.mp4
allation, or you can use:

video input path ./RepCount_pose/video/test/stu12_34.mp4
MAE:0.2356079854110582,OBO:0.5592105263157895
```

Visualization:


Speed 300 %
1
action: bench_pressing


Speed 300 %
2
action: bench_pressing


9
action: bench_pressing


10
action: bench_pressing