**Assignment 3 Report**

Kaiqiang Song

[Experiment Environment]

Nvidia GTX 1080 ti (11G), CUDA 8.0, CuDNN v5.1, Keras with Tensor Flow version 1.2

SkLearn Toolkit for SVM Implementation

[Feature Extraction]

Using features of the first fully connected layer of VGG16 network model pre-trained on ImageNet. It is a higher semantic features with 4096 dimensions.

I use the Keras version of Implementation from:

<https://github.com/fchollet/deep-learning-models>

To extract features, and modify some of codes to let it suitable for my environment with Tensor Flow version 1.2

Basicly you can extract all the features for each frame in this dataset.

[Classifier Design]

I give one frame of label which is same to the label of the video where it comes from. And design a multi-class linear-SVM using one-vs-one strategy for frame classification and then apply a voting mechanism to decide which class it belongs to.

In particularly, the frames are label in ‘ACTION-VIEW’ form, since different view has different center among the data distribution. I prefer to first fine-grained classified the label of both action and angles of sight and then evaluated with only the action label.

[Leave one out cross validation]

Since select one out of each class as test is a kind of odd with different number of instances in different class. I chose to do leave one out on the entire dataset (leave 1 video as test set and 149 as training set).

[Evaluation]

Finally we can get a matrix of confusion:

Where the element at i-th row and j-th col is the number which true label is i and predicted as j.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Dive | Golf | Kick | Lift | Ride | Run | Skate | Swing | Walk |
| Dive | 14 |  |  |  |  |  |  |  |  |
| Golf |  | 16 | 1 |  |  |  |  |  | 1 |
| Kick |  |  | 15 |  |  | 5 |  |  |  |
| Lift |  |  |  | 6 |  |  |  |  |  |
| Ride |  |  |  |  | 11 | 1 |  |  |  |
| Run |  |  | 2 |  |  | 6 | 1 |  | 4 |
| Skate |  |  |  |  |  | 2 | 9 |  | 1 |
| Swing | 1 |  |  |  |  |  |  | 32 |  |
| Walk |  |  | 2 |  |  | 1 | 2 |  | 17 |

Totally the accuracy is 84%

Precision, recall and F1 Score for each label are shows below:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Dive | Golf | Kick | Lift | Ride | Run | Skate | Swing | Walk |
| P | 93.33 | 100.00 | 75.00 | 100.00 | 100.00 | 40.00 | 75.00 | 100.00 | 73.91 |
| R | 100.00 | 88.89 | 75.00 | 100.00 | 91.67 | 46.15 | 75.00 | 96.97 | 77.27 |
| F1 | 96.55 | 94.12 | 75.00 | 100.00 | 95.65 | 42.86 | 75.00 | 98.46 | 75.56 |