Homework 2

Stage 1:

You will be provided with several extracted visual features. You are required to choose one of them and build an event classification model using the specified training set. The learned model should be applied to the testing set as specified to get the prediction scores. The detailed steps of this assignment can be summarized as follows:

1) Find the extracted visual features of the experimental videos with the following command under the directory of /data/MM2/MED/labelsets:

perl select_files.pl filelists/sift_spbof.list vids_MED10_dev_evl.idx> YOUR OWN WORKING DIRECTORY/MED10.sift_spbof.list

Note that here SIFT spatial Bag-of-Words is used as an example. You can use other 4 features as provided, namely csift_spbof.list, stip_spbof.list, mosift_spbof.list and tch_spbof.list per your choice.

2) The MED10.sift_spbof.list includes the features of the positive examples of all 3 events and all the negative examples (NULL videos). You need the features of a number of positive examples of this event and the features of a number of negative examples for building a classifier. In this homework, you are required to select 10 positive examples for each event and 100 negative examples as your training set. The remaining positive examples and another 100 negative examples form your testing set. The following step is for such purpose.

3) In the following file:

/data/MM2/MED/labelsets/old/csv/MED10TRN_20101215_JudgementMD.csv

You find the ID of all the videos and if a video is a positive example of event 001, 002 or 003, or a negative example (null video). Take event 001 as an example. You should parse the file to generate an index file with the ID of the first 10 positive videos for this event. Meanwhile, you are required to choose the first 100 negative examples from MED10TRN_20101215_JudgementMD.csv and incorporate their ID into your own index file of training videos. Similarly, to generate your own index file of testing videos, parse the ID of all the remaining positive videos for this event and the ID for the second 100 negative examples from MED10TRN_20101215_JudgementMD.csv. The same procedure can be used to generate the index file for event 002 and 003. You should also generate the label file for the training and testing setsin this step (1 indicates a positive example whereas 0 indicates a negative example).

Update: For your convenience, we parsed the ID list for you. Please find them at: /data/MM22/xiaojun/Zhigang

There are two files, namely id.idx and label.file. The first one is the index of videos whereas the second one is the corresponding label file. 1 indicates that the corresponding video is a positive example for event P001; 2 indicates that the corresponding video is a positive example for event P002; 3 indicates that the corresponding video is a positive example for event P003; 0 indicates that the corresponding video is a negative example.

Now you can easily use these two files to generate your index files for the training and testing sets.

4) Now you can use your own index file to select features from MED10.sift_spbof.list to form the training set and the testing set.
5) Once the features and labels of the training set are ready, you can use libsvm to train a SVM classifier. Note that it is best to use Chi-square kernel for video analysis. Thus, you need to calculate the Kernels for both training set and between training and testing sets and use them as the input of libsvm.
The tool can be found at:
/data/MM22/xiaojun/Zhigang/libsvm-3.18
Copy the folder to your own directory first.
Note that svm-train is the script for training the classifier whereas svm-predict is the script for prediction on the testing data. You are strongly suggested to read the README file before you start your training and testing.
6) Apply your model to the testing set and calculated the average precision (AP).
The code for AP calculation is located at:
/data/MM22/xiaojun/Zhigang/mAP
Copy the folder to your own directory first.
ap is the script for calculating the average precision. The input should be the ground truth label and the prediction label from your model, both in a column vector format saved as a text file. There are two samples, namely label_file and prediction_file. For instance, you can use the following command to calculate the average precision:
./ap label_file prediction_file

Stage 2:
You will be provided with the raw videos. One is required to extract the SIFT features from the raw videos. After that, one is required to perform clustering, BoW generation and pooling to get the final representation of a video. Further, an event classification model should be trained and applied on the testing data. The detailed steps of this assignment can be summarized as follows:
1) Find the raw experimental videos with the following command under the directory of /data/MM2/MED/labelsets:
perl select_files.pl ../raw_videos/MED14.fps vids_MED10_dev_evl.idx > YOUR OWN WORKING DIRECTORY/MED10.video.list
2) The MED10.video.list includes the raw videos of the positive examples of all 3 events and all the negative examples (NULL videos). Similarly to Stage 1, use the following step to form a subset from these videos.
3) In the following file:
/data/MM2/MED/labelsets/old/csv/MED10TRN_20101215_JudgementMD.csv
You find the ID of all the videos and if a video is a positive example of event 001, 002 or 003, or a negative example (null video). You should parse the file to generate an index file with the ID of all the positive videos for all 3 events. Meanwhile, you are required to choose the first 200 negative examples from MED10TRN_20101215_JudgementMD.csv and incorporate their ID into your own index file.
4) Use the index file to select the raw videos. Now use ffmpeg to down sample these videos with the following setting:

Select the first 30 seconds of the videos and downsize them to 160×120 and 15 frames per second.
ffmpeg can be called with the following command:
/data/MM22/xiaojun/Zhigang/sfep/resize_code/rocks/ffmpeg
Note that you should execute the following command before using ffmpeg:
Export
LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/data/MM22/xiaojun/Zhigang/sfep/resize_code/rocks
5) Select the key frames.
6) Use OpenCV to extract SIFT features from the key frames. OpenCV can be found at:
/data/MM22/xiaojun/Zhigang/OpenCV-2.3.1
Copy the folder to your own directory first. Note that you should compile the OpenCV before using it.
7) Perform clustering and generate BoW model.
8) Perform pooling to get the final representations.
9) Normalize the representations.
10) Similarly to Stage 1, use your index file to generate the training and testing sets for each event with your own feature representations and the labels created in Stage 1.
11) Use libsvm to train a SVM classifier.
The tool can be found at:
/data/MM22/xiaojun/Zhigang/libsvm-3.18
Copy the folder to your own directory first.
12) Apply your model to the testing data and calculated the average precision (AP).
The code for AP calculation is located at:
/data/MM22/xiaojun/Zhigang/mAP
Copy the folder to your own directory first.