
Accuracy comparison of video classification using SVM and Convolutional Neural Netowrk

Smit Anand, Shama yazdani, Meet Kumar pandya
North Carolina State University
sanand2@ncsu.edu, syazdan@ncsu.edu, mhpandya@ncsu.edu

Abstract

1 Process of video analysis is a very interesting and current topic in machine learning.
2 Work in this field is being done by many top universities and companies like
3 Stanford and Google. Several tools are also available to carry out the process. We
4 will following a process of beaking video into frames and training out model with
5 the frames associated with a specific class then feeding that trained model with
6 a testing video, divided into frames, and printing first some classes with aximum
7 confidence.

8 1 Introduction

9 Current approaches to object recognition make essential use of machine learning methods. To
10 improve their performance, we can collect larger datasets, learn more powerful models, and use
11 better techniques for preventing overfitting [1]. This method sometimes helps us to achieve our result.
12 In this project we will be trying to analyze a video dataset to label the videos in predefined classes.
13 Experiments, involving different machine learning algorithms for object recognition has been carried
14 out to check the accuracy of algorithm and the used dataset.
15 There are several tools or libraries which can be used to achieve our target like OpenCV, Matlab,
16 Scikit learn, Tensorflow, Caffe. In this project we will be using Scikit learn and Tensorflow to train
17 and analyze the result.

18 Working idea of the model

19 We will be using, first a part and then whole, video data of Youtube provided by Google. This dataset
20 contains videos classified in 4700 classes. We need to train our learning algorithm to analyze the
21 video of a specified class and classify if if given as input.
22 To achieve this result we will be using scikit learn (a python machin learning library) to train SVM or
23 S3VM with the dataset and we will also be using Tensorflow to train a CNN on this dataset. Our final
24 result will be a table of comparison among those two learning algorithms.
25 We have first decided to break a video from the dataset into frames and then feed it to our learning
26 model to analyze a set of frames associated to a video which in tr=urn is associated to a class. Doing
27 so will train our algorithm to classify a frame to a class. We will then follow the same approach for
28 another algorithm. Then we will feed a testing video to our algorithm and give the first 3 or 4 classes
29 having maximum confidence.
30

31 References

32 [1] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional
33 neural networks." Advances in neural information processing systems. 2012.