Object Recognition using CALTECH 256 Dataset

Anubhab Majumdar
amajumd@ncsu.edu

Toshal Phene
tphene@ncsu.edu

Shubham Munot
samunot@ncsu.edu

1 Introduction

Object recognition is a quintessential machine learning problem of modern times. It is difficult for machines to recognize a large variety of objects in different conditions of lighting, occlusion and skew. In recent years, this field has seen huge leaps with machines becoming adept at object recognition. In this project, we explore different approaches of tackling this problem

We are using the CALTECH 256 dataset (link) for object recognition [1]. The dataset has 256 unique categories of images. The number of samples in each category varies from 80 to over 200. This dataset is quite a standard dataset in the field of object recognition and has been been used in important research papers like [2]. Our goal is to train a machine learning algorithm using 80% of the data and test using the remaining 20%. We experimented on a small subset (4-5 categories) of the dataset.

The experiments have been implemented using student version of MATLAB and python/tensorflow. Some of the less computationally intensive models are trained on personal computers, while the deep networks are trained on VCL servers.

2 Methodology

A typical classification system is shown in Fig.1. The paper focuses on the machine learning block shown in Fig.1. Different learning algorithms ranging from simple and lazy methods like KNN to deep convolution neural network is applied and results are compared for a comprehensive understanding of which method works better. The paper also tries to draw conclusion as to why certain methods work better than others.

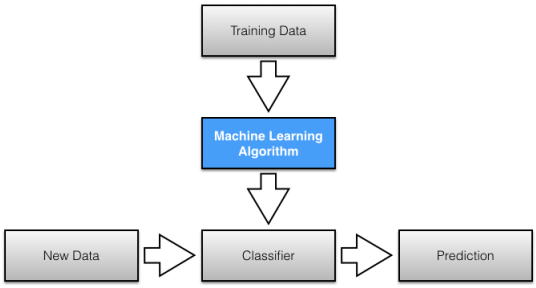

Figure 1: A typical prediction system

Before jumping into the algorithms applied, let's talk about the data itself. We have chosen the following 4 categories for our experimentation:

Figure 2: Few samples from our dataset

- Backpack (Sample size - 151)
- Binoculars (Sample size - 216)
- Eiffel Tower (Sample size - 83)
- Fried egg (Sample size - 90)

Sample of some images from the dataset are shown in Fig.2.

2.1 KNN

2.2 SVM

2.3 Convolution Neural Network

3 Experiments and Results

3.1 KNN

3.2 SVM

3.3 Convolution Neural Network

4 Conclusion

References

[1] Greg Griffin, Alex Holub and Pietro Perona.
Caltech-256 Object Category Dataset.

[2] Hao Zhang, Alexander C. Berg, Michael Maire and Jitendra Malik
SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition.

[3] Knuth: Computers and Typesetting,
`http://www-cs-faculty.stanford.edu/~uno/abcde.html`