NUS CS4243: Computer Vision and Pattern Recognition November 6, 2019

Where's Waldo Detector using Computer Vision

Lecturer: Angela Yao Student(s): Maximilian Fruehauf, David Drews, Choo Wen Xin

Abstract

This report describes our group's implementation of a computer vision algorithm to detect Waldo, Wenda, and the Wizard from a series of "Where's Waldo" books. The goal of this project is to detect the three characters from the provided high-resolution images, which can be very complex with a lot of detail and many other characters. The three characters also may or may not appear in any given image.

Due to the complex nature of the given images, and variation of the characters' appearances, detecting the characters accurately proved to be a challenge. In some cases, we could not identify where the characters are, and a lot of false positives were present as well.

Our proposed solution is to use a histogram over gradients (HoG) feature descriptor, then training a linear support vector machine (SVM) to create our classifier. We were able to detect some instances of Waldo, espcially in the postcard in the top left hand corner of the page.

1 Introduction

"Where's Waldo" is a series of books containing detailed, high-resolution illustrations. For this project, we are given a set of scanned images from the book, and are tasked to design a computer vision algorithm that can detect the three characters from the book: Waldo, Wenda, and the Wizard. Each image may contain one or more of these characters. We are provided with a set of training images with annotations of the bounding box locations of Waldo, Wenda, or the Wizard in each image.

Throughout the course of the project, we have attempted several different methods with varying degrees of success. While there are several existing computer vision algorithms for face and object detection, such as the "You Only Look Once" object detection [1], these do not seem to work too well with illustrated characters like Waldo, in a complex illustrated iamge. The characters' appearance and size also seems to vary across images, sometimes only a smaller part of the character (i.e a part of the face) is visible. This makes identifying the characters more challenging as well.

We proposed using a few methods:

- histogram over gradients (HoG) feature descriptor
- training a linear support vector machine (SVM) to create our multiclass-classifier

We will discuss our proposed solution in the next section.

2 Proposed Solution

Give an overview of your solution, put it in a framework. Then, detail each part in the framework.

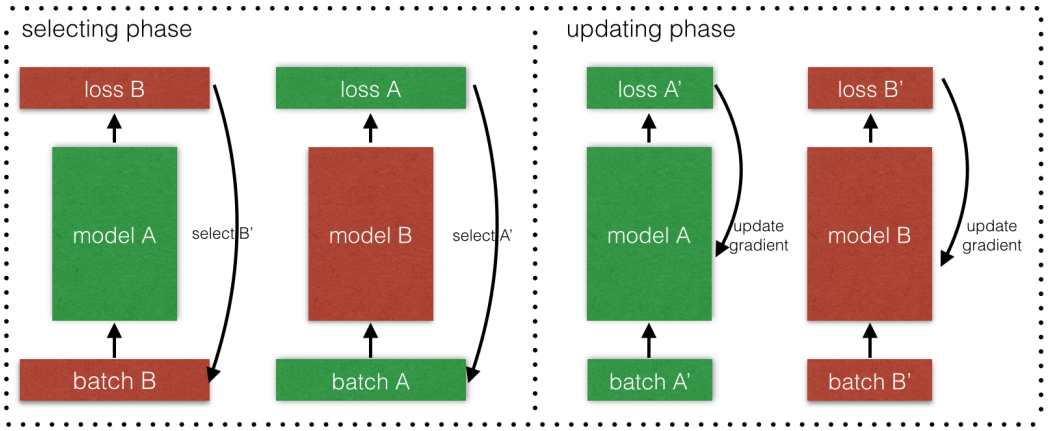

Figure 1: Our proposed solution.

3 Experiments

3.1 Data Preparation and Configuration

Specify how to process the data, how to evaluate the performance (e.g., mAP).

Dataset	#train	#test	#Category
MNIST	60,000	10,000	10
CIFAR-10	50,000	10,000	10

Table 1: Summary of datasets.

3.2 Implementation

Give a figure to illustrate your implementation, then detail each parts.

3.3 Results

Present the results, both qualitatively (visualize) and quantitatively (specific numbers).. Analyze the results

3.4 Discussion

Strengths and weakness in your method.

4 Conclusion

In this project, we ...

5 Group Information

Member	Student ID	Email	Contribution
Maximilian Fruehauf	Axx	e0445541@u.nus.edu	xxx
David Drews	Axx	e0454245@u.nus.edu	xxx
Choo Wen Xin	A0160465H	e0053347@u.nus.edu	xxx

Table 2: Group member information.

References

[1] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. University of Washington, Allen Institute for AI, Facebook AI Research, 2016.