A novel approach to solve the Waldo puzzle

Aswanth P P, Mohammed Ameen, Joe Antony and Dr.Shashidhar G. Koolagudi
Department of Computer Science and Engineering
National Institute of Technology Karnataka, Surathkal

***Abstract*—This paper explains an approach to solve the Waldo puzzle by using machine learning concepts. The Waldo Puzzle consists of a picture or an image with a lot of different characters and objects. The objective is to locate the position of a certain character named Waldo in the image. The difficulty of the puzzle is usually high,and it generally takes a considerable amount of time to find him using the naked eye. Through machine learning, and using object detection strategies, we locate Waldo in the order of a few seconds, making it a much more effortless approach.**

***Keywords*—Machine Learning, Template Matching, Object Detection, Tensorflow**

I. Introduction

Waldo's puzzle can be solved using template matching concepts in open CV or with the machine learning concepts.In template matching user needs to provide the parent image as well as the small template image consists of Waldo character.The problem with this approach is if we are giving the different waldo image as template wont show the accurate answer.Hence in template matching we actually need to identify Waldo puzzle in parent image in order to feed it's cropped image as the template image.

In our approach,We are using RCNN with Inception v2 model which is already trained to detect pets from a image(trained on COCO data sets).While the model could be trained from scratch starting with randomly initialized network weights, this process would probably take weeks. Instead, we used a method called transfer learning.

II. Implementation

A. Preparing the Datasets

While dealing with neural networks is the most notable process in deep learning, it sadly turns out that the step data scientists spend the most time on is preparing and formatting training data.

Target values for the most simple machine learning problems are usually scalars (like for a digit detector) or a categorical string. Tensorflow Object Detection API training data uses a combination of both. It consists of a set of images accompanied with labels of desired objects and locations of where they appear in an image. Locations are defined with two points since (in 2d space) two points are enough to draw a bounding box around an object.

So in order to create the training set, we need to come up with a set of Wheres Wally puzzle images with locations of where Wally appears.

B. Preparing the Model

Tensorflow Object Detection API provides a set of pre trained models with different performances (usually a speed-accuracy tradeoff) trained on several public datasets.

While the model could be trained from scratch starting with randomly initialised network weights, this process would probably take weeks. Instead, we used a method called transfer learning.

It involves taking a model usually trained to solve some general problem and retraining it to solve ours. The idea behind transfer learning is that instead of reinventing the wheel by training our model from scratch we can use the knowledge obtained in the pre trained model and transfer it to out new one. This saves us a lot of time so that the time spend for training can be invested into obtaining only the knowledge specific to our problem.

III. Execution and Result

We created a python script to detect waldo from the source image .Python script "detectWaldo.py" will take trained model,source image and label(here it is Waldo),returns whether Waldo is there or not if there it will generate a png file with square frame around the Waldo character

IV. Analysis

References

[1] Xianghua Fan, Fuyou Zhang, Haixia Wang and Xiao Lu , *The System of Face Detection Based on OpenCV*

[2] Xinyi Zhou, Wei Gong, WenLong Fu and Fengtong Du , *Application of Deep Learning in Object Detection*

[3] K. He, X. Zhang, S. Ren, and J. Sun , *Deep residual learning for image recognition*

[4] Shaoqing Ren, Kaiming He, Ross Girshick and Jian Sun , *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*

[5] Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton , *ImageNet Classification with Deep Convolutional Neural Networks*