Topic -

An Attempt to Deduce a Convolutional Neural Network Which Can Process Both Real Life Images and Hand Drawn Images.

Project By

Yudhajit Chowdhury
Enrollment Number : 12017002002190
Section - A
Roll - 3

Saheb Ganguly
Enrollment Number : 12017002002206
Section - A
Roll - 67

Under The Supervision Of :

Prof. Sumanta Chakraborty

Department of Computer Science and Engineering
Institute of Engineering and Management West Bengal, India

02.09.20

THIS IS THE FIRST ABSTRACT OF THE PROJECT.

AS THIS PROJECT IS BASED ON DATA SCIENCE AND DUE TO LIMITED DATA, AND TIME FOR MINING AND CLEANING NEW DATA FROM DIFFERENT SOURCES ,SO THIS ABSTRACT IS ONLY A FIRST GLANCE OF FUTURE IDEAS AND IMPLEMENTATIONS. THIS ABSTRACT IS ALSO DESIGNED TO ADAPT THE INITIAL IDEA FOR THIS PROJECT.

INTRODUCTION :

Humans have a lot of senses, and yet our sensory experiences are typically dominated by vision. With that in mind, perhaps it is unsurprising that the vanguard of modern machine learning has been led by computer vision tasks. Likewise, when humans want to communicate or receive information, the most ubiquitous and natural avenue they use is language. Language can be conveyed by spoken and written words, gestures, or some combination of modalities, but for the purposes of this project, we'll focus on the best combination of hyper parameter for a nicely accurate deep learning that can predict a tree by showing a picture of a real tree and implementing that same algorithm on a hand drawn model.

Deep learning has absolutely dominated computer vision over the last few years, achieving top scores on many tasks and their related competitions. The most popular and well known of these computer vision competitions is ImageNet. The ImageNet competition tasks researchers with creating a model that most accurately classifies the given images in the dataset. Nearly every year since 2012 has given us big breakthroughs in developing deep learning models for the task of image classification. Due to its large scale and challenging data, the ImageNet challenge has been the main benchmark for measuring progress.

PROBLEM FORMULATION:

The problem of Image Classification goes like this: Given a set of images of trees that are all labeled with a category, we are asked to predict these categories for a novel set of test images and measure the accuracy of the predictions. There are a variety of challenges associated with this task, including viewpoint variation, scale variation, intra-class variation, image deformation, image occlusion, illumination conditions, background clutter etc.

ALGORITHM(S) :

How might we go about writing an algorithm that can classify images into distinct categories?

Computer Vision researchers have come up with a data-driven approach to solve this. Instead of trying to specify what every one of the image categories of interest look like directly in code, provide the computer with many examples of each image class and then develop learning algorithms that look at these examples and learn about the visual appearance of each class. In other words, they first accumulate a training dataset of labeled images, then feed it to the computer in order for it to get familiar with the data.

With this project we can find the best combination of all the parameters and the hyperparameter to find the best model and accuracy.

THEORETICAL IMPLEMENTATION :

Training CNN Model:

I will create a variety of different CNN-based classification models to evaluate performances on Fashion MNIST. I will be building our model using the Keras framework. For more information on the framework, you can refer to the documentation here. Here are the list of models I will try out and compare their results:

1. CNN with 1 Convolutional Layer

2. CNN with 3 Convolutional Layer

3. CNN with 4 Convolutional Layer

4. VGG-19 Pre-Trained Model

And the steps that we are following are:

·Split the original training data (60,000 images) into 80% training (48,000 images) and 20% validation (12000 images) optimize the classifier, while keeping the test data (10,000 images) to finally evaluate the accuracy of the model on the data it has never seen. This helps to see whether I'm over-fitting on the training data and whether I should lower the learning rate and train for more epochs if validation accuracy is higher than training accuracy or stop over-training if training accuracy shifts higher than the validation. ·

Train the model for 10 epochs with batch size of 256, compiled with categorical_crossentropy loss function and Adam optimizer. ·

Then, add data augmentation, which generates new training samples by rotating, shifting and zooming on the training samples, and train the model on updated data for another 50 epochs. After loading and splitting the data, I preprocess them by reshaping them into the shape the network expects and scaling them so that all values are in the [0, 1] interval. Previously, for instance, the training data were stored in an array of shape (60000, 28, 28) of type uint8 with values in the [0, 255] interval. I transform it into a float32 array of shape (60000, 28 * 28) with values between 0 and 1.

INITIAL REFERENCES :

`images.google.com` - Real Life Image Data Set Will Be Mined From Google Images.

`Kaggle.com` - HandDrawn Image Data Set Will Be Gathered.

`github.com` - Project Archive and Resource Gathering.

METADATA :

Real Life Image :

Provider : Google

License : NA

Hand Drawn Image :

Provider : Google

License : CC BY 4.0 - https://creativecommons.org/licenses/by/4.0/

SIGNATURE OF STUDENTS:

1.

2.

SIGNATURE OF MENTORS:

1.

DATE :