MSBD5001 Group Project Proposal: Doodle Recognition

CHONG Ka Hung *20543539* | GAO Han *20565848* | XU Yanhui *20549583* | WANG Yutong *20541402* | YANG Yue *20549600*

Abstract

The doodle recognition is a fun experimental game which aims to make a guess to the users' drawing. Drawings are marked as time-stamped vectors, and including other information like the content the plyer draws and the country in where the player is located. A good classifier among the dataset is created to estimate the drawing in a particular category. In this project, the convolutional neural network and the ensemble learning provide ways to get the training data, and the modify the model by cross validation.

1 Introduction

There is a trend that everything in the world tends to be intelligent in recent years, and pattern recognition is popular as well. Doodle recognition is an education experimental game shows the way the artificial intelligence works. When the game starts, the host provides you with object tags one by one, and you have 20 seconds to draw each object tag. If the AI guesses your sketch successfully, you will get a check mark and move on, otherwise failed the game. Method to accomplish the game is to do data analysis using reasonable techniques. In the project, some particular categories are discussed, and the whole dataset should do the data cleaning first, and the training data is available by the convolution neural network and the ensemble learning. Besides, the parameters need to be modified by cross validation. The doodle recognition can be a good solution for the pattern recognition based on the real-world datasets.

2 Dataset

The dataset is available on kaggle[1]and is collected by the player "Quick Draw!",which is a famous app to let users draw within a time limitation to match a description. Here we choose categories that belong to transportation means, such as airplane, bicycle and so on.

Attributes	Data Type	Description
`country code`	Classified(string)	A two-letter country code (ISO 3166-1 alpha-2) of where the player was located.
`drawing`	Vector	A JSON array representing the vector drawing
`key_id`	64-bit unsigned integer	A unique identifier across all drawings.
`recognized`	Classified	Whether the word was recognized by the game.
`timestamp`	datetime	When the drawing was created.
`word`	Classified(string)	Category the player was prompted to draw. Here all of them belong to transportation means.

Table 1: Data Specification

Drawing array:

```
{[x0, x1, x2...],[y0, y1, y2...],[t0, t1, t2...]}
x & y: pixel coordinates(real value)
t: time in ms since the first point(integer)
```

A sample drawing record is below:

```
"key_id":"5530503534870528",
"word":"canoe",
"countrycode":"US",
"timestamp":"2017-01-27 18:51:13.930020",
"recognized":True,
"drawing":[[[24, 30, 36, 46, 94, 107, 119, 1...]]]
```

3 Problem Explored

According to the background and dataset provided, here are several main challenges that we meet. Firstly, different from actual items in real world, manual drawings have strong subjective understanding which is distinctive among individuals. So it is hard to capture some general features of a specific category. One possible way is to provide some alternatives.

Secondly, just because of strong subjectivity and collection from the real application, the dataset has a lot of noisy data. Some unreasonable drawings are placed in a certain category, which impedes correct feature extraction.

Thirdly, even if the classification of some samples seems rational, the main mechanism of training data derives from the app "Quick draw!", which is artificially designed. It may not reflect the true pattern. Besides, the pattern might vary among different areas in the world.

4 Expected Outcome

An outcome is expected that the recognizer can not only effectively learn from manually-labeled human drawings in different distribution but also perform well on noisy data like incomplete sketches and label mismatching works, so that it can successfully predict the label of input drawings with high accuracy, precision and recall rate.

5 Methodology

5.1 Data Preprocessing

Before the begin of training the model, it is necessary to decrease the possibility that the model learn the nonsense noise of the dataset. Generally, preprocessing can do help to remove the noise from the dataset. In this project, for example, some unrecognizable pictures are actually labeled as certain kind of object. If the trained model learn the features of these pictures, it will eventually impact the prediction accuracy of this model.At current stage, two types of unrecognizable pictures are considered and displayed below. Pictures like them would be removed from the dataset.

5.2 Model Choosing

In this project, Convolutional Neural Network (CNN) is chosen to take care of the image classification. In terms of image recognition and classification, CNN is currently widely used and its performance is widely recognized. With CNN, it is possible to make the model training progress more efficient and the eventual model more accurate than using other models

(a) Noisy Data 1 (b) Noisy Data 2

Figure 1: Unrecognizable Pictures

5.3 Cross Validation

The goal of cross-validation is to test the model's ability to predict new data that was not used in estimating it, in order to flag problems like over-fitting. In neural network, it can help with controlling over training weights, and measuring the performance under current parameter settings. In this project, K-fold cross validation is chosen to achieve model selection and parameter optimization.

5.4 Ensemble Generation

Ensemble learning helps improve machine learning results by combining several models. This approach allows the production of better predictive performance compared to a single model. It is effective to construct diverse models by varying architectures, hyper-parameter setting, and train techniques. It is able to choose one of some typical ensemble learning techniques, e.g. Boosting(majority vote) and Stacking, in this project.

6 Conclusion

To concluded, this project aims to build a recognizer that can learn from labeled doodle data set and then predict the label of human drawings in some predefined categories. The techniques involved to complete this project includes data preprocessing, model choosing, cross validation and ensemble generation. An expectation is made that the predicted result of classifier will have high accuracy, precision and recall rate so that the input sketch can be successfully labeled.

References

[1] kaggle. `https://www.kaggle.com/c/quickdraw-doodle-recognition/data`, 2018.