Quick Draw! Predicting the Country

Arthur K. Jakobsson*
Carnegie Mellon University
Pittsburgh, PA 15213
ajakobss@cmu.edu

Prathik Guduri
Carnegie Mellon University
Pittsburgh, PA 15213
pguduri@andrew.cmu.edu

Abstract

Drawing is a skill that humans have used since the beginning of humanity to convey ideas and stories. However, the similarity between a 3rd grader's cartoon of a deer, a caveman's drawing of one and a photo of a deer vary drastically and yet we understand all three to represent the same subject. In this paper we seek to create a network to predict the category of drawings and we propose a network to abstract images into drawings (including the series of strokes required to derive the images).

1 Introduction

We have collected data from a publically available dataset produced by Google's creative lab.

https://github.com/googlecreativelab/quickdraw-dataset

This dataset was created through an online game, "Quick Draw!" in which participants are given a category to draw and 20 seconds to represent the topic. The game then uses a neural network to try and predict the subject of the drawing. Each drawing made, it's associated topic, the drawer's place of origin, the strokes used to create the drawing and some other information are stored in the dataset and are used to improve Google's performance for the game.

We intend to take this dataset preliminarily to recreate and improve upon the predictive model created by Google in order to find the categories. However, our end goal described in the Future Work section describes how we intend to use the data combined with an image dataset of references for the categories to predict how humans would draw an unseen image.

2 Background Literature

There are major variations between categories and other atributes within our dataset. Fernandez-Fernandez et. al. [1] describes a brief statistical analysis of the Quick Draw! dataset. They primarily focus on the level of complexity of the drawings as described by the number of strokes used to represent the drawing. They suggest that as a result of having much fewer strokes, some categories may lack the complexity to be discriminated from others.

The ability for the model to create a drawing has been extensively looking into. Ha et. al. [2] provides an example to create a robust method of sketch generation in a vector format. Similarly Xie et. al. [3] from 2012 provides a baseline method based on painting to try to estimate the brush strokes required to paint an image using policy iteration and RL. Our final product is inspired by these methods to take input images to draw the image however we intend to use a new model architecture to better discriminate the images and to train our generative model. This is futher described in the future works section.

*https://arthurjakobsson.github.io/

10807 Class Project (2016).

Similar models to ours have been created online to guess the identity of a drawing based on the Quick Draw dataset. For example, [4] provides a baseline convolutional framework that demonstrates similar, but slightly lower accuracy (but on a larger dataset) to ours.

3 Methods/Model

4 Preliminary Results

5 Evaluation of Preliminary Work

6 Future Work

Our intended final product flips the current model on its head. We intend to connect the current Quick Draw! dataset to a singular or multiple image datasets that represent the categories drawn. From this we intend to use a Generative Adversarial Network to create drawings from the images provided by category. We hypothesize that our model will be able to learn features such as the keypoints that humans jump to such as whiskers on a cat, eyes on a human and other important features that are put into a recognizable sketch. The GAN will of course have a discriminator which will take in the source image and the drawing produced (or the true drawings from Quick Draw!) and try to detect if it was a drawing we created or a source image. Through this we will train out generative side. We also intend on adding other information that may help the model train such as an edge detection map of the images.

We hope to produce a product that is not just a drawing algorithm but instead learns the features that humans consider recognizable in a drawing. The incentive behind this is that a drawing of a cat is composed of two eyes and 6 whiskers which arguably look nothing like a cat, but is universally recognized as a drawing of one, similarly a pig has two circles as a nose. Currently existing models, try to create accurate representations of drawings from the image but do not try to describe how we as humans "sketch" and break an image into its basic topics.

7 Teammates and Work Division

Past work division:

We worked together to create the network, coding live with each other and therefore put in the same time and work on the project. The writing of this paper is divided equally between us. We both performed background research and brainstorming for this and then we discussed and developed the idea and motivation behind the project together. The background research and project setup was led by Arthur and the development of the training loop and code structure was led by Prathik.

Future work division:

We intend to continue working side-by-side heavily to develop our final model. We will likely divide the work into topics of GAN development, evalution, ablation study and writing. Our current proposal for division is Arthur leads the generative side of the GAn and the ablation while Prathik leads the discriminatory network and the evaluation.

8 Citations

We developed our code using the Quick Draw! datset and the helper code provided within it such as the code used to create the 28x28 images which we downloaded directly.

We took inspiration from Arthur's team project in the robotics institute with Rishi Veerapaneni <rveerapa@andrew.cmu.edu> to help us develop a pytorch framework.

We also used

`https://machinelearningmastery.com/building-a-convolutional-neural-network-in-pytorch/`

as an example of a working CNN. We developed the structure of the code heavily to meet our needs but we took the general torch CNN framework as inspiration from this article.

References

[1] Fernandez-Fernandez, R. & Victores, J.G. & Estevez, D. & Balaguer, C. (2019) Quick, Stat!: A Statistical Analysis of the Quick, Draw! Dataset. *Provides statistical background to Quick Draw! Dataset* Universidad Carlos III de Madrid (UC3M)

[2] Ha D. & Eck, D. (2017) *A Neural Representation of Sketch Drawings.* Google.

[3] Xie, N. & Hachiya, H. & Sugiyama, M (2012) Artist Agent: A Reinforcement Learning Approach to Automatic Stroke Generation in Oriental Ink Painting. *Proceedings of the 29th International Conference on Machine Learning* Department of Computer Science, Tokyo Institute of Technology

[4] kolszewska. Github. https://github.com/kolszewska/QuickDrawClassifier