Quick Draw Doodle Recognition

Anubhav Shrimal (MT18033), Vrutti Patel (MT18020)
Advisor: Dr. Mayank Vatsa

Computer Vision (CSE 344)

ABSTRACT

In Quick Draw the AI system tries to classify the hand-drawn doodle into a predetermined category which is quite similar to pictionary. By this project we are trying to achieve the same using various methodologies such as K-Nearest Neighbor, CNN, CNN+LSTM, etc. and compare their performance on different evaluation metric.

USE CASE OF THE PROBLEM

- It is a challenge in Computer Vision & Machine Learning to handle noisy data and dataset with many different representations of the same class. The Quick Draw Doodle Recognition challenge is a good example of these issues because different users may draw the same object differently or the doodles could be incomplete which is similar to noisy data (Example Fig. 2).
- This application can be used as a fast prototyping tool for designers or artists by suggesting them the accurate templates on the basis of the rough doodles made by them.
- It can be extended by replacing the doodles with doodles of alphabets and then convert the hand-written text into digital text format.

DATASET

- The Quick Draw dataset is a collection of millions of drawings of 300+ categories which is created by using doodles of different players of the game. The drawings were captured as timestamped vectors, tagged with metadata of the label which the player was asked to draw.
- The complete dataset is huge (~73GB) and so we'll be using only a subset of the complete data.
- We'll split the dataset in training and validation set with 80-20 ratio.
- Link: https://www.kaggle.com/c/quickdraw-doodle-recognition/data

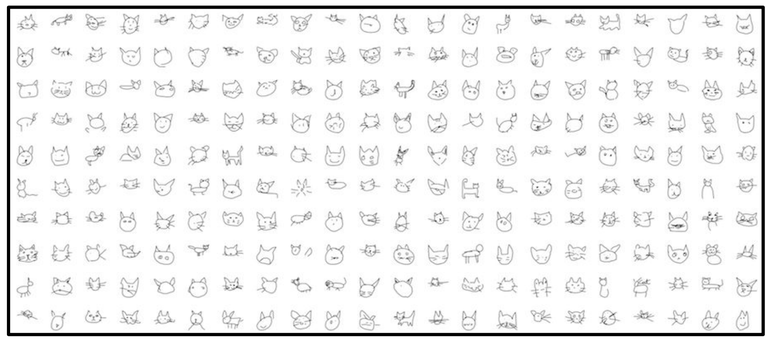

Figure 1. Quick Draw Dataset. (Source: Google Images)

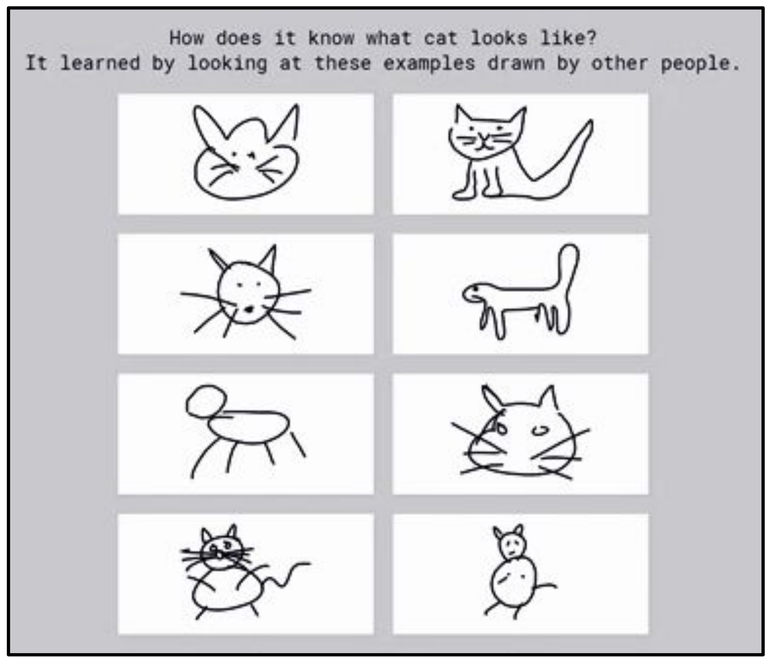

Figure 2. Different cat doodles created by users. (Source: Google Images)

LITERATURE REVIEW

Past work has been done on free hand sketch recognition [1],[5] where the authors have used KNN and SVM based approaches. In the recent times, Deep Neural Network based approaches have been applied using CNN and ResNet architectures [3] and also incorporating LSTM for pen strokes.

EVALUATION METRIC

- Mean Average Precision (MAP@3)
- Receiver Operating Characteristic (ROC)
- Cumulative Matching Characteristic (CMC)

PROJECT TASK & MILESTONES

Mid-term Evaluation:

- Visualizing & understanding the dataset.
- Implementing K-Nearest Neighbor (KNN).
- Self implemented Convolutional Neural Network (CNN).

End-Sem Evaluation:

- Implementing deeper CNN architectures such as ResNet34 and ResNet50.
- Extending CNN with Long Short Term Memory network (LSTM) for pen stroke based recognition.
- Comparing performance of models by evaluating it across different measures such as MAP@3, ROC and CMC.
- Creating OpenCV based doodle recognition application for user interaction.

REFERENCES

1. Lu, W., & Tran, E. (2017). Free-hand Sketch Recognition Classification.
2. M. Eitz, J. Hays, and M. Alexa. How do humans sketch objects? ACM Trans. Graph. (Proc. SIGGRAPH), 31(4):44:1– 44:10, 2012.
3. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition, 2016.
4. Kim, J., Kim, B. S., & Savarese, S. (2012). Comparing image classification methods: K-nearest-neighbor and support-vector machines. Ann Arbor, 1001, 48109-2122.
5. Ha, D., & Eck, D. (2017). A neural representation of sketch drawings. arXiv preprint arXiv:1704.03477.