The Omniglot Challenge

Som Tambe[1], Nikita Chauhan[2], Anmol Pabla[3], Mohit KulKarni[4], and Vaibhav Thakkar[5]

[1]Department of Aerospace Engineering [2]Department of Computer Science and Engineering [3]Department of Mechanical Engineering [4]Department of Mathematics [5]Department of Electrical Engineering

Objective and Overview

Despite remarkable advances in artificial intelligence and machine learning, machine systems have lacked the capability of human conceptual knowledge to learn new concepts with handful examples that too with richer representations than machines do. The Omniglot data set contains 1623 different handwritten characters from 50 different alphabets.
The Omniglot Challenge focuses on developing more human-like algorithms like One-Shot Learning. The Project worked towards developing and comparing BPL and several ML techniques for tasks like generation and classification on the Omniglot Dataset.

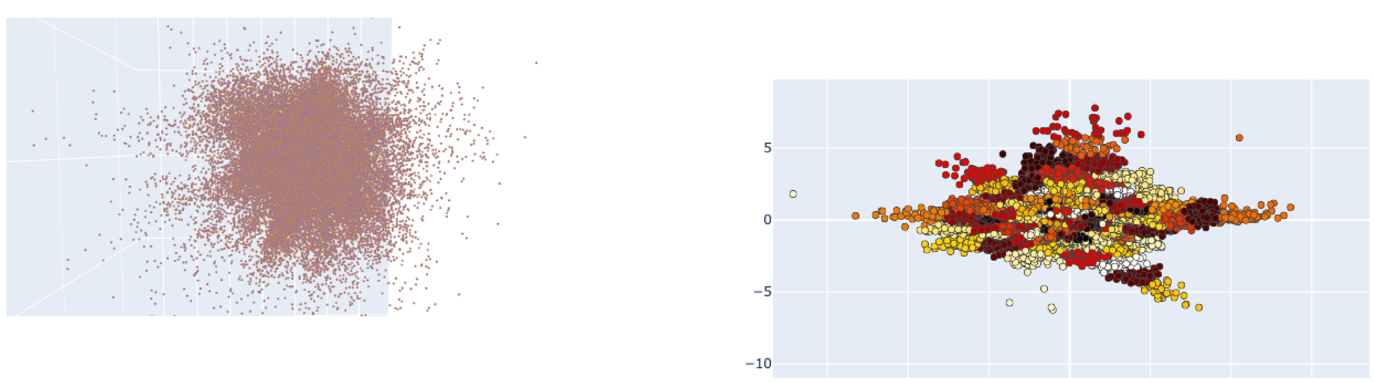
Omniglot: Own Model

This section focuses on developing new custom models for the Omniglot challenge, which can then be scaled to any one-shot learning tasks. Prior models, like Bayesian Program Learning [1] focused much more on hand engineered parts to derive results. We have tried utilizing deep learning techniques to help us universalize the model, so that it can be adapted to a wider range of tasks.

Results

Enhanced Learning based on Stroke Data

The Omniglot dataset supplied us with another set of useful data - the stroke data for each character. This included the x,y,time coordinates for each characters' stroke. This gives us a whole lot of flexibility which we did not have with only the images. We used a Variational Autoencoder to get a lower dimensional representation of the strokes. The idea was to cluster them on the basis of this 2-d or 3-d representation. Here are some images of plots.

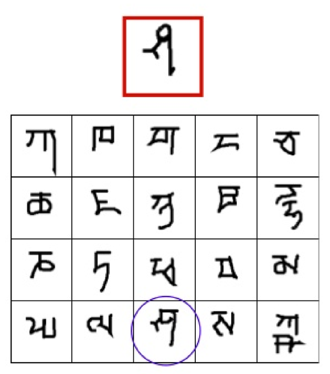


2-d and 3-d plots of stroke data.

This will help us in finding what stroke lies in what character. We can then form a one-hot vector using this stroke data for each character which will denote the strokes present in the character. Perform supervised learning with characters and their respective vectors will give us a network which can derive stroke data given an image. We tried to train a network to do this. It did not train successfully. The reason for this is mentioned in detail in the project report.

One-shot Learning using Siamese Network

We trained a neural network to do same-different pairwise classification on symbols. We show that this approach gets a reasonable accuracy in 20-way one-shot learning on symbols from unseen alphabets. My network achieved 48% accuracy compared to 93% in the original paper. This can be reasoned to techniques like momentum, data-augmentation and hyperparameter optimization and less training than the original paper



20-way One-shot learning

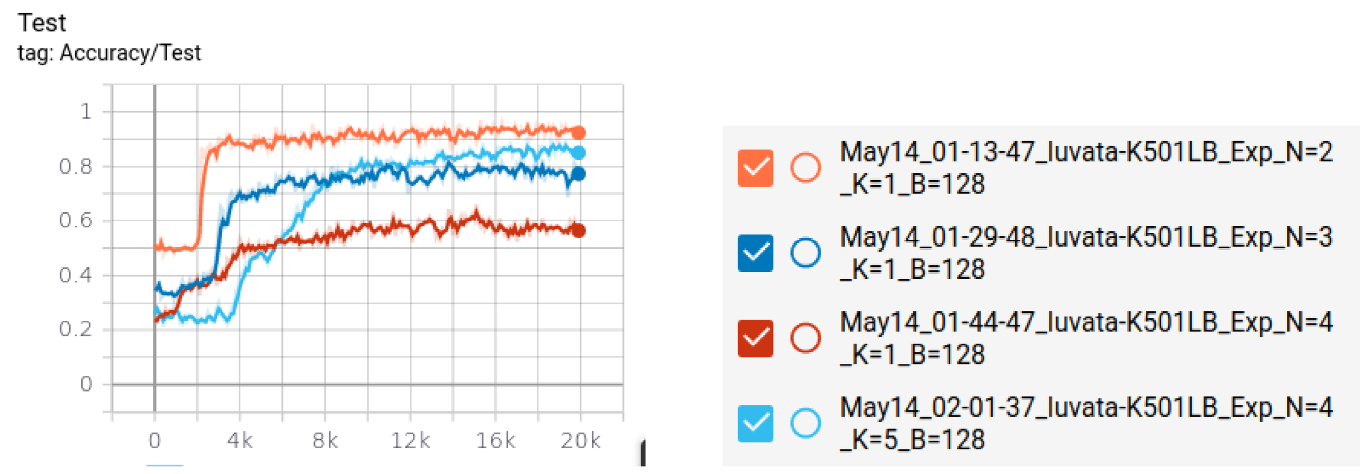
Omniglot: Memory Augmented

This section focuses on few-shot classification using Memory Augmented Neural Networks. Using a external memory module is shown to be significantly better than neural memory modules such as LSTM and GRUs. Here we try to achieve those results in few-shot learning paradigm.messeged

Results

MANN module Few-Shot Classification which was implemented with the help of a main network made up of LSTM network. We tried implementing this on the omniglot dataset and achieved similar reults to the paper(5% error). We trained the network on various N(number of classes per step) and K(number of images per class) per episode.

Tensorboard Visualisation of the results



Omniglot: Text Generation

This section focuses on comparing Bayesian Program Learning Model with SOTA Machine Learning models for tasks other than classification on the 'Omniglot' dataset. The major demerit of the BPL model is that it is very specific to the Omniglot dataset and would not work for other images. The target of the sub-team is to create SOTA generative ML models. that give similar results to the BPL model. In the last few years, deep learning based generative models have gained more and more interest. Among these deep generative models, two major families stand out and deserve a special attention:

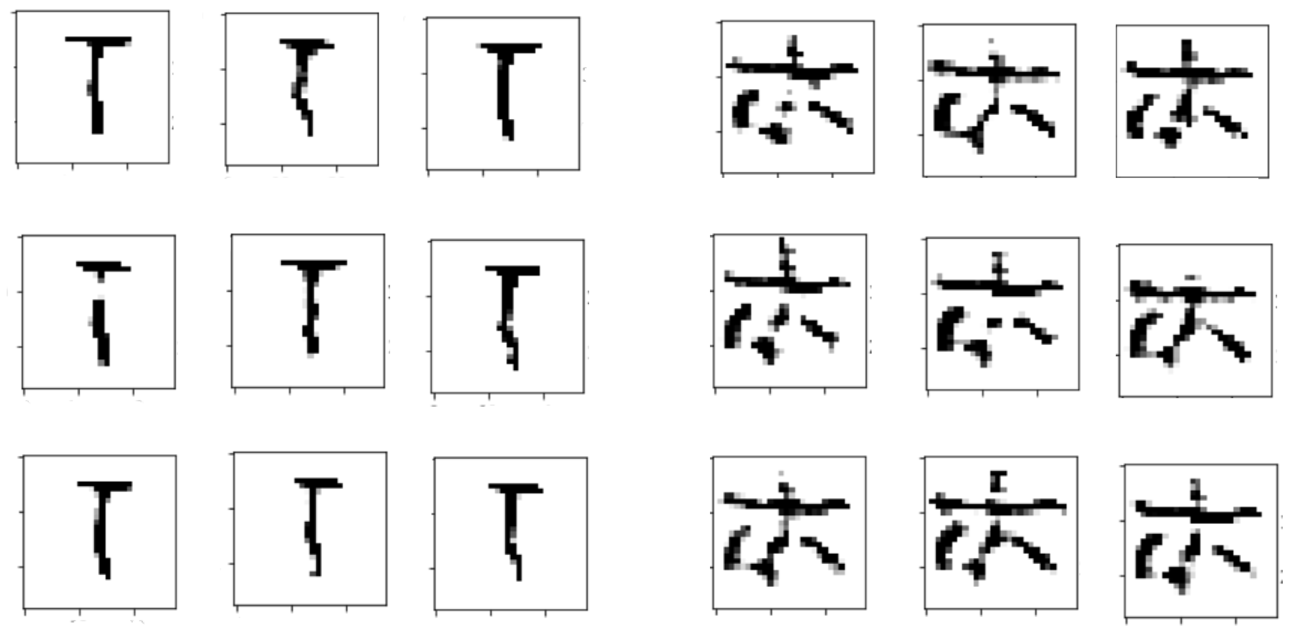
Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs).

The team implemented VAE and GAN models to try and generate images with appreciable accuracy and variability to yield satisfactory results when the task conducted in the original Omniglot Challenge is performed with human judges.

Results

The VAE model was first implemented on the MNIST dataset and then on the Omniglot dataset, the model uses ELBO loss. Several steps were taken to improve the results and also add variability to the output generated. The output images are 28x28px and have decent accuracy and variability. Some of the images generated by the model for three characters are given below:

Images Generated using VAE Model



DCGAN was first implemented to the MNIST dataset and then tried on the Omniglot Dataset. We found that after about 5 epochs, the network starts overfitting rapidly.

Discussion and Future Work

The team would try and improve the models developed and perform experiments with human judges. We would also try and implement other models in the direction of human-like algorithms.
A more detailed version can be found in the report.

Contributions

- **Own Model:** Som Tambe, Nikita Chauhan
- **Text Generation:** Anmol Pabla, Mohit Kulkarni
- **Memory Augmented:** Mohit Kulkarni
- **Mentor:** Shashi Kant