Sketch Prediction using Neural Networks

Vignesh Iyer
Information Technology (Mumbai University)
Vidyavardhini's College of Engineering and Technology
(Mumbai University)
Nallasopara, India

Dishant Patil
Information Technology (Mumbai University)
Vidyavardhini's College of Engineering and Technology
(Mumbai University)
Nallasopara, India

Milind Purohit
Information Technology (Mumbai University)
Vidyavardhini's College of Engineering and Technology
(Mumbai University)
Virar, India

Dr. Madhavi Waghmare
Information Technology (Mumbai University)
Vidyavardhini's College of Engineering and Technology
(Mumbai University)
Vasai, India

***Abstract*—Recurrent Neural Networks are the state-of-the-art algorithm for sequential data and among others used by Apples Siri and Googles Voice Search. This is because it is the first algorithm that remembers its input, due to an internal memory, which makes it perfectly suited for Machine Learning problems that involve sequential data. It is one of the algorithms behind the scenes of the amazing achievements of Deep Learning in the past few years. We present Sketch Prediction Using Neural Network, a recurrent neural network (RNN) able to construct stroke-based drawings of common objects. The model is trained on a dataset of human-drawn images representing many different classes. We outline a framework for conditional and unconditional sketch generation, and describe new robust training methods for generating coherent sketch drawings in a vector format. Our goal is to train machines to draw and generalize abstract concepts in a manner similar to humans. In this work, as a first step towards this goal, we train our model on a dataset of hand-drawn sketches, each represented as a sequence of motor actions controlling a pen: which direction to move, when to lift the pen up, and when to stop drawing. In doing so, we created a model that potentially has many applications, from assisting the creative process of an artist, to helping teach students how to draw.**

Keywords*—sequential data, stroke-based drawings, Recurrent Neural Network (RNN), Conditional and unconditional drawings.

I. INTRODUCTION

There have been major advancements in generative modelling of images using neural networks as a generative tool. Generative Adversarial Networks (GANs), Variational Inference (VI), and Autoregressive (AR) models have become popular tools in this fast-growing area. Most of work thus far has been targeted towards modelling low resolution, pixel images. Humans, however, do not understand the world as a grid of pixels, but rather develop abstract concepts to represent what we see. From a young age, we develop the ability to communicate what we see by drawing on paper with a pencil or crayon. In this way we learn to express a sequential, vector representation of an image as a short sequence of strokes thus we investigate an alternative to traditional pixel image modelling approaches, and propose a generative model for vector image. We develop a training procedure unique to vector images to make training more robust. Now the research work in the domain can be mainly classified as followings: one such work is portrait drawing by Paul the Robot [6], where an underlying algorithm controlling a mechanical robot arm sketches lines on a canvas with a programmable artistic style to mimic a given digitized portrait of a person. Neural Network-based approaches have been developed for generative models of images, although the majority of neural network-related research on image generation deal with pixel images [5]. There has been relatively little work done on vector image generation using neural networks. In addition to unconditionally generating sketches, we also explore encoding existing sketches in to latent space of embedding vectors. There have been many normalization techniques used in prediction or generating sketch based on pixel images, and we perform layer normalization techniques on raw data to generate modelling of vector images rather than pixel-based images.

II. PROBLEM STATEMENT

Recently, there have been many advancements in Generative Modelling of Images using Neural Networks as generative tool. Most of the work thus far has been targeted towards modelling low resolution, pixel images. Humans however do not understand the world as a grid of pixels, they rather develop abstract concept to represent what we see. We currently do not have any generative model for vector images; hence we seek to create a generative model for vector images than traditional pixel image modelling approaches. The sketch prediction system is used to predict the sketch drawing or picture even completing the system. But the use of pixel images in this method is sub-optimal hence we are vectorize the pixel images, since the vector images remains the same the resultant will either be enlarged or shrunk down. We are using recurrent neural networks since it provides fast processing and support layer normalization.

III. MOTIVATION OF PROJECT

The motivation of this project is regarding the sketch prediction using neural networks which is basically used to predict the images or pictures that you are going to sketch or draw actually without drawing the full picture or image that you are going to draw using the neural networks and the need is to simplify the required image and make a perfect image required by the user in order to obtain the users expectations according to drawings which are usually predicted using the neural networks. Thus, the main motivation for the sketch prediction using neural networks is to identify the best image which is required by the user input and obtaining the best output image back to the user as a main result using neural networks. The main purpose of sketch prediction is to make the accurate images using neural networks and by using the RDP algorithm i.e. (Ramer-Douglas-Peucker) algorithm. And this particular algorithm is an end point fit algorithm i.e. iterative algorithm which makes the pictures or images more perfect and accurate according to the user's assumptions or thinking which has been done to draw the particular images that are going to be drawn using the neural networks that are been identified by the required users which meet their own expectations accordingly. The RNN (Recurrent Neural Network) is required to make the pixel images of a particular picture or image into the vector based images that are been used to get the more accurate and perfect image which is required by the user which has been provided by the RNN to which the user can obtain the pictures images for which the user has given the input to the system. And also, the RNN sometimes also encodes the required images and makes it bidirectional RNN and can obtain the required images from the user input.

Unlikely the feedforward neural networks are a network that allows for a bi-directional flow of data, and the networks between the connected units forms a directed cycle. Allows for dynamic temporal behaviour to be exhibited. And the RNN is capable of using its internal memory to process arbitrary sequences of I/p. Thus, RNN is designed for capturing information from sequences/timeseries data. Therefore, the Recursive formula is as follows: St =Fn (St-1, Xt), where St=time step, Fn = function, Xt=input, St-1=previous state.

A. Maintaining the Motivation of Project

Using the Neural Networks, the whole image or picture will be drawn according to the user requirement by recognizing the required image or picture visible to the user. Therefore, maintaining the required need by the neural networks is to obtain the exact image for the required user to be obtaining the expected results by the user requirements by getting the predicted drawings by using neural networks. The need for the neural networks is increasing day by day in today's world so using the sketch prediction using neural networks the drawings have been made easy for the user to identify the required image within a period of time. The LSTM is the long short-term memory which is used as a recurrent neural network architecture for deep learning. And the RNN cell which is an abstract component of sketch RNN model which will be used as LSTM as an encoder RNN, and the HyperLSTM is used for the RNN decoder. Thus, the LSTM and HyperLSTM is required to encode and decode the required content for which the particular images can be obtained according to the user input data. And for the sequence tasks generations the HyperLSTM is used as a type for the RNN cell. Thus, the RNN is one of the major components in using the LSTM and HyperLSTM to obtain the images or pictures from the neural sense of a particular by using the neural networks for the required user.

B. Objectives

- To get the raw data of sketches from the user.
- To perform normalization on raw data using RDP(Ramer-Douglas-Peucker) algorithm.
- To perform machine learning on normalized data.
- To implement the trained models into our system.
- To give user input of strokes to the trained models.
- To show the final output of sketch drawn by RNN.

C. Unconditional Generation

Figure 1: Unconditional generation of firetrucks, gardens and owls with varying t.

As a special case, we can also train our model to generate sketches unconditionally, where we only train the decoder RNN module, without any input or latent vectors. By removing the encoder, the decoder RNN as a standalone model is an autoregressive model without latent variables. In the use case, the initial hidden states and cell states of the decoder RNN are initialized to zero. The input x_i of the decoder RNN at each step is only S_{i-1} or S'_{i-1}, as we do not need to concatenate a latent vector z.

D. Conditional Generation

Figure 2: Conditional generation of cats.

We qualitatively assess the reconstructed sketch S' given an input sketch S. We sample reconstructions at various levels of temperature *t* using a model trained on the single cat class, starting at 0.01 on the left and linearly increasing to 1.0 on the right. The reconstructed cat sketches have similar properties as the input image, and occasionally add or remove details such as a whisker, a mouth, a nose or the orientation of the tail.

IV. DIAGRAMS

A. *Architecture Diagram*

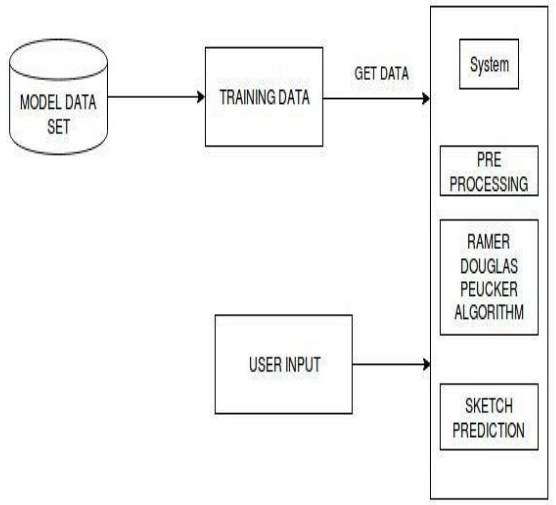

Figure 3: Architecture Diagram

According to Figure 3 the architecture diagram shows that the given data model set provides the obtained data to the system and through the user input the system processes the pre-processing events and performs the RDP (Ramer-Douglas-Peucker) algorithm then further the required sketch prediction is done and then the required results are obtained accordingly.

B. *Sequence Diagram*

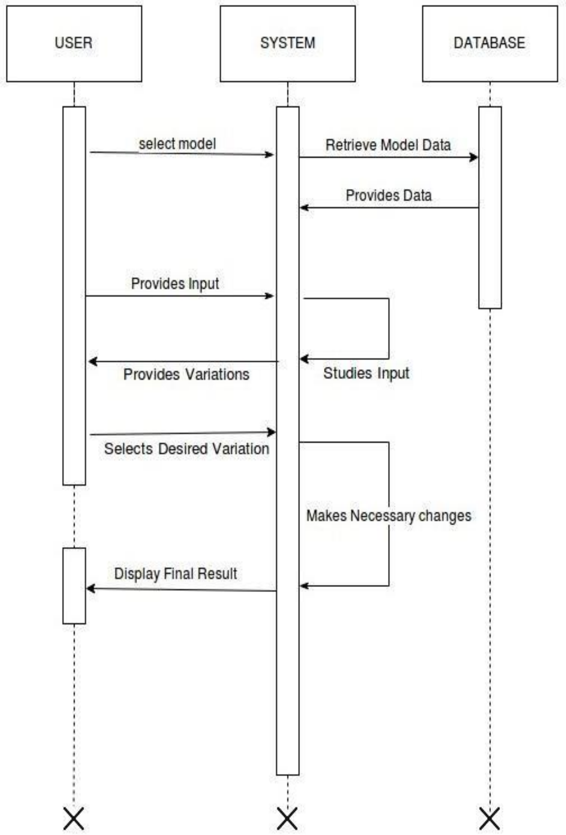

Figure 4: Sequence Diagram

According to Figure 4 the sequence diagram shows a given dataset of user can easily obtain the sequence of required data from the system when a particular model is selected and from the system the data model is retrieved through database and then further the database provides the data to the system .The user provides the input to the system and the system studies the input and provides variations to the user input. And then user selects the desired variation to the system and the system makes necessary changes and then the system displays the final result to the user.

C. *Conclusion*

In this paper we proposed the use of the vector images instead of pixel images instead of pixel images and we introduced a new algorithm named as RDP algorithm. Since this algorithm reduced the number of points in a curve that is approximated by a series of points.

Hence the method proposed in this paper using neural networks predicts the images or pictures without drawing the full picture or image that you are going to draw using bi- directional.

ACKNOWLEDGMENT

The author would like to thank Dr.Madhavi Waghmare for her continued support and guidance she would also like to thank all his friends and family; without their enthusiasm this project would not have been possible.

REFERENCES

[1] X. Zhang, F. Yin, Y. Zhang, C. Liu, and Y. Bengio. Drawing and Recognizing Chinese Characters with Recurrent Neural Network. CoRR, abs/1606.06539, 2016.
[2] R. Plamondon and S. Srihari, "Online and offline handwriting recognition: a comprehensive survey," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63–84, 2000.
[3] C.-L. Liu, S. Jaeger, and M. Nakagawa, "Online recognition of Chinese characters: The state-of-the-art," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 2, pp. 198–213, 2004.
[4] A. Graves, "Generating sequences with recurrent neural networks," arXiv:1308.0850, 2013.
[5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE, vol. 86, no. 11, pp. 2278– 2324, 1998.
[6] P. Tresset and F. Fol Leymarie. Portrait drawing by paul the robot. Compute Graph, 37(5):348-363, Aug,2013.