*International Journal of Research in Engineering and Science (IJRES)*
*ISSN (Online): 2320-9364, ISSN (Print): 2320-9356*
*www.ijres.org Volume 10 Issue 7 ǁ July 2022 ǁ PP. 80-84*

# SMART DRAGOMAN: Handwritten Digit Recognition and Text to Speech Translation

Aby Motty[1] , Gopika KV[2], Jewel Jacob[3] , Manu Martin[4] , Dr Rajesh K S[5]
[1] *PG Student, Department of Computer Applications –Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala, India*
[2] *PG Student, Department of Computer Applications –Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala, India*
[3] *PG Student, Department of Computer Applications –Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala, India*
[4] *PG Student, Department of Computer Applications –Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala, India*
[5] *Professor and Head of the Department, Department of Computer Applications –Saintgits College of Engineering (Autonomous) Pathamuttom Kottayam Kerala, India*

***ABSTRACT***
*Hand written text recognition is one of the significant areas of research and development with a streaming number of possibilities that could be attained. The handwritten digit recognition is the ability of computers to recognize human handwritten digits. It is a hard task for the machine because handwritten digits are not perfect and can be made with many different flavours.*
*Text to speech translation is one of the fastest-growing engineering technologies. Text to speech is a process to convert any text into voice. It has several applications in different areas, and provides many potential benefits. We design a system that can recognize the human voice, translate from English to other languages. The output is in voice form. Proper implementation of this technique can bridge the gap between countries by allowing to communicate without language barriers.The aim of this work is to review existing methods for the handwritten digit recognition and text to speech translation problem using machine learning algorithm. Going forward, we expect to add more functionality.*
***Keywords: Machine learning, Text to speech translation, Google Translate, Ajax API,CNN***

---

Date of Submission: 24-06-2022 Date of acceptance: 05-07-2022

---

## I. INTRODUCTION

The rapid growth of new documents and multimedia news has created new challenges in pattern recognition andmachine learning. Handwriting character recognition has become a standard research area due to advances in technologies such as the handwriting capture devices and powerful mobile computers. However, since handwriting very much depends on thewriter, building a high-reliability recognition system that recognizes any handwritten character input to an application, is challenging. This work considers the problem of recognizing handwritten digits, i.e., numbers from 0 to 9. Typically, handwritten digit recognition is an essential function in a variety of practical applications.

In text to speech translation, we convert the text into voice with the click of a button. And we can translate it to any language we want. A text to speech converter is a device which takes the text as input and gives output in form of voice. The world is connected community and the only real divide between countries and communities is the language barrier. Machine learning enabled speech translators like this allows enables us to overcome this divide and communicate in ways otherwise thought not possible.

## II. LITERATURE SURVEY

In [1] various pre-processing techniques involved in the character recognition with different kind of imageranges from a simple handwritten form based documents and documents containingcolored and complex background are discussed. In this, different preprocessing techniques like skew detection and correction, image enhancement techniques of contrast stretching, binarization, noise removal techniques, andmorphological processing techniques are discussed. It was concluded that using a single technique for preprocessing, we can't completely process the image.

In [2] a method in which first multi-scale neural training with modifications in the input training vectors is adopted A simulator program (a GUI) is designed in such a way that the characters can be located on

any spot on the blank paper in which the characters are written. The results show that such methods with moderate level of training epochs can produce accuracies of at least 85% and more for handwritten upper case English characters.

In our system we have used CNN algorithm for recognizing the handwritten digits. Instead of uploading scanned image of a digit, we draw the digit on the user interface. The system will recognize the digit and display it on the screen along with its accuracy. In this work we are performing handwritten digit recognition using MNIST dataset.

Text to speech is a process to convert any text into voice with specified Language. Text to speech project takes words on digital devices and converts them into audio with a button click . Text to speech python project is constructive for people struggling with reading.

## III. METHODOLOGY

### 3.1 GOOGLE TRANSLATE

Google translate is a multilingual neural machine translation service developed by Google to translate text, documents and websites from one language into another. It offers a website interface, a mobile app for android and iOS and an API that helps developers build browser extensions and software applications. As of April 2022, Google Translate supports 109 languages at various levels, and as of April 2016, claimed over 500 million total users, with more than 100 billion words translated daily, after the company stated in May 2013 that it served over 200 million people daily.

### 3.2 AJAX API

Ajax is a set of web development techniques that uses various web technologies on the client-side to create asynchronous web applications. With Ajax, web applications can send and retrieve data from a server asynchronously (in the background) without interfering with the display and behaviour of the existing page. By decoupling the data interchange layer from the presentation layer, Ajax allows web pages and, by extension, web applications, to change content dynamically without the need to reload the entire page.
In practice, modern implementations commonly utilize JSON instead of XML.

### 3.3 CNN

Convolutional neural network (CNN, or ConvNet) is an algorithm that belongs to the class of artificial neural network (ANN).This algorithm is most commonly used in analyzing visual imagery. CNN algorithm assign importance to various aspects in the input image and differentiates one from the other. They are fully connected feed forward neural network and are very effective.The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. A collection of such fields overlap to cover the entire visual area.CNN finds its applications in image and video recognition, recommendersystems,image classification, image segmentation, medical image analysis, natural language processing, brain–computer interfaces, and financial time series.

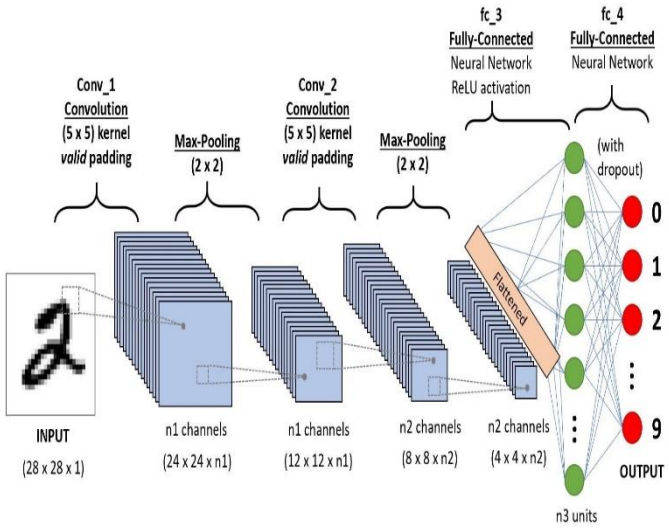


*Figure 1: Architecture of CNN*

**3.3 MNIST DATABASE**

The MNIST database (Modified National Institute of Standards and Technology database) is a large database of handwritten digits that is commonly used for training various image processing systems.The database is also widely used for training and testing in the field of machine learning. It was created by "re-mixing" the samples from NIST's original datasets.The creators felt that since NIST's training dataset was taken from American Census Bureau employees, while the testing dataset was taken from American high school students, it was not well-suited for machine learning experiments. Furthermore, the black and white images from NIST were normalized to fit into a 28x28 pixel bounding box and anti-aliased, which introduced grayscale levels. The MNIST database contains 60,000 training images and 10,000 testing images. Half of the training set and half of the test set were taken from NIST's training dataset, while the other half of the training set and the other half of the test set were taken from NIST's testing dataset. The collection of data and creation of dataset is the most time consuming process in a database and the MNIST database has made it easier by allowing access to a preprocessed dataset can be easily fed into the algorithm enabling the program to learn better, faster and provide faster and error free results.

*Figure 2:Sample images from MNIST test dataset*

**3.4 WORKING**

In Handwritten Digit Recognition, first, we import all the modules that we need for training our model. The Keras library already contains some datasets and MNIST is one of them. So we can easily import the dataset and start working with it. The mnist.load_data() method returns us the training data, its labels and also the testing data and its labels.

The main python library used for implementation of text to speech translation was google trans. The google translate Ajax API was used to make calls such as detect and translate from the application to the library. The system is compatible with Python 3.6 and aboveto meet the organization requirements.

## IV. EXPERIMENTS AND RESULTS

Implementation is the stage of the work where the theoretical design is turned into a working system. The work was carefully planned to ensure maximum efficiency to solve the present problem to improve translation. After development of each module, the system was tested to ensure the functionality of the respective components

In Smart Dragoman, using MNIST dataset for the purpose of digit recognition we haveachieved better performance for the CNN. In ourexperiment, we have found the maximum training accuracy100% and maximum validation accuracy 99.92% both atepoch 15. The overall performance of the network is found99.21%. Moreover, the overall loss ranged from 0.026303 to0.049449. Hence, this proposed method of CNN is moreefficient than the other existing method for digit recognition.In Text to Speech Translator,Text entered is converted to audio.It recognizes both capital as well as small letters. It recognizes numbers as well.


*Figure 3:App Homepage*

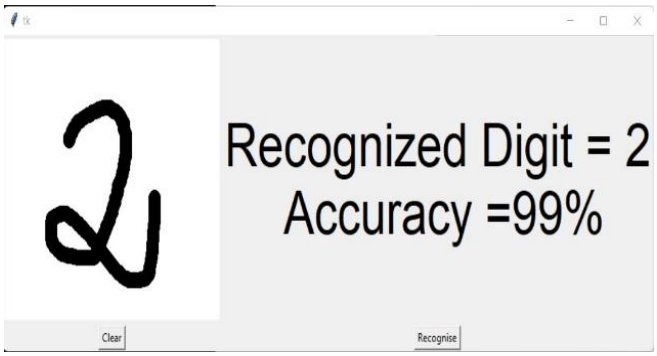


*Figure 4:Handwriting recognition output*


*Figure 5:Speech translation result*

## V. FUTURESCOPE AND CONCLUSION

Recently handwritten digit recognition becomes vital scope and it is appealing many researchers because of its use in variety of machine learning and computer vision applications. Currently, the scope of our engine extends to recognizing one character at a time. We propose to extend this functionality to enable the accurate prediction of multiple characters simultaneously-thereby enabling truly real time character recognition.

In Text-To-Speech, we extend this functionality to add speech to speech recognition.

Smart dragoman is concerned with recognizing handwritten digits. And also include the text translation technique which first recognize and translate the text into preferred language. The challenge of this work was to come up with some basic image correlation and speaker-dependent variation/pronunciation technique, instead of some sophisticated algorithms, and see to what extent we can make this mechanism accurate. The future steps that to go for would be having a closer look at the results of all the versions in order to find new rules. By extracting and implementing them, we will be able to enhance the performance of these versions.

## REFERENCES

[1]. U. Pal, T. Wakabayashi and F. Kimura, "Handwritten numeral recognition of six popular scripts," Ninth International conference on Document Analysis and Recognition ICDAR 07, Vol.2, pp.749-753, 2010.

[2]. Velappa Ganapathy, and Kok Leong Liew, "Handwritten Character RecognitionUsing Multi scale Neural Network Training Technique", Proceedings of World academy of Science, Engineering and Technology, vol. 29, ISSN 1307-6884, May 2007.

[3]. A Dutta and A. Dutta, "Handwritten digit recognition using deep learning", International Journal of Advanced Research in Computer Engineering & Technology, vol.6, no.7,July 2017.

[4]. D. Sasirekha and E. Chandra, "Text To Speech: A Simple Tutorial", *International Journal of Soft Computing and Engineering*, vol. 2, March 2012.

**[5].** Niu n Xiao-Xiao and Y. Suen Ching, "A novel hybrid CNN-SVM classifier for recognizing handwritten digits", *Elsevier*, vol. 45, no. 4, pp. 1318-1325, April 2012.