DESIGN AND IMPLEMENTATION OF A HANDWRITING TEXT RECOGNITION SYSTEM

A. Musa[1,2]
[1]*Dept. of Electrical and Computer Engineering, Kwara State University, Malete.*
[2]*Institute for Intelligent Systems, University of Johannesburg, South Africa.*
*twhid2001@yahoo.com

Keywords: –
Handwriting recognition,
Image processing,
Machine learning,
Deep learning,
Neural network.

Article History: –
Received: July, 2021
Reviewed: August, 2021
Accepted: September, 2021
Published: September, 2021

ABSTRACT

Handwritten characters are difficult to recognize due to diversity of human handwriting style, variation in angle, size and shape of letters. In this work, an extensively employed method of handwritten text recognition is used to transform handwritten data into electronic format or digital form, preprocess, extract distinctive features, and classify written characters. The proposed approach is to design a Graphical User Interface (GUI) that corresponds to the ability of human beings to identify and verify handwritten characters by training datasets of 80 labeled digit images with Machine Learning, Deep Learning and Convolution Neural Network (CNN) for character recognition. The proposed recognition system performs excellently for the cursive handwriting with 80-90% accuracy.

1. INTRODUCTION

Handwriting recognition is one of the most interesting and challenging research areas in the field of Image Processing, Artificial Intelligence (AI), computer vision, and pattern recognition. A computer performing handwriting recognition is said to be able to acquire and detect characters in paper documents, pictures, and other sources and convert them into machine-encoded form. Its application is found in optical character recognition and more advanced intelligent character recognition systems.

Most of these systems nowadays implement machine learning mechanisms such as neural networks. Machine learning is a branch of AI inspired by psychology and biology that deals with learning from a set of data and can be applied to solve wide spectrum of problems. A supervised machine learning model is given instances of data specific to a problem domain and an answer that solves the problem for each instance. When learning is complete, the model is able to, not only, provide answers to the data it has learned on, but also to get unseen data with high precision [1].

Related works are presented in Section 2, while the methodology is discussed in Section 3. In Section 4, the results obtained are presented and discussed, while conclusion is drawn in Section 5.

2. RELATED WORKS

Some past related works are contained in [2] to [16]. [2] carried out full survey of various techniques used for character recognition. However, a proficient solution is required to resolve associated difficulty so that overall performance can be amplified. [3] is a review on comparison among different kinds of feature extractor and classifier. Various attitudes of hand written character recognition are discussed with their performance. The work, however, needed an efficient solution to solve the problem of low accuracy in some languages so that overall performance can be increased.

[4] explains the various phases of image processing being used in character recognition such as Image Restoration, Image Enhancement, Segmentation, Feature Extraction and Classification with Recognition. These techniques are used along with Neural Network due to its high tolerance to noise. This

helps in removing all unwanted signals in images that are distorted over years. Successful character recognition becomes possible for such documents as well and the system generate perfect results.

The work in [5] entails acquisition of probability distribution of the features related to the character structure and generating character images of various handwritings using the probability distribution by generating strokes based on the probability distribution of each stroke and assembling them into a character. It becomes possible to generate character images of various handwriting samples not influenced by the original images.

[6] demonstrated the use of Feed Forward Algorithm along with Back Propagation algorithm. The paper successfully recognizes characters using a multilayer perception with a hidden layer. Higher performance can be achieved in Back Propagation once the number of hidden nodes to be used is successfully determined. The recognition of characters is proved to be better and gives an accuracy of 70% and above for English language handwritten characters.

[7] successfully implemented both online and offline handwriting recognition, particularly for isolated hand printed characters and words. Although the paper mostly focused on the processing of English language, there are numerous projects going on in many countries to process and recognize specific languages. It is hoped that many of these projects will succeed to maintain the diversity and richness of the global human experience.

[8] accomplished the procedure for identification of characters utilizing JAVA and assessments were carried out on quite a lot of data units available. No improvement on outcomes to solve problems that were confronted with the training of a bigger set of images.

In [9], [10] and [11], different character recognition systems have been discussed. While [9] used Neural Network, deep learning was used by [11]. Similarly, [12], [13] and [14] have also worked on handwritten character recognition using different methods.

3. METHODOLOGY

In order to achieve the analysis of the learning model, implementation of the algorithms must be developed with consideration to existing problems of character recognition some of which have been reviewed in Section 2. This section explains the research methodology which involves the use of machine learning. Machine learning code is written with the use of python programming language. Large amount of dataset from three individuals with at least 20 sets of writings is then tested for good accuracy. The images are processed, useful features extracted, then classified with the use deep learning algorithm and CNN. The result of the model can be visualized with the use of python library called Tkinter and Tkthemes.

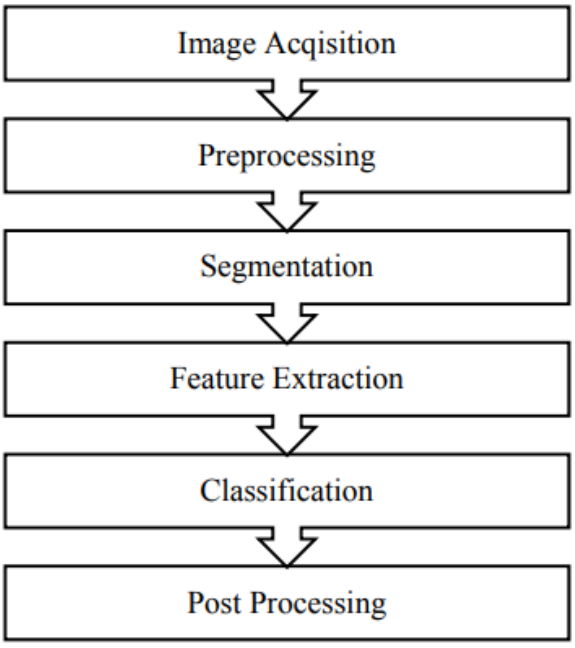


Fig. 1. Block diagram of Handwriting Recognition System

Fig. 1 is the block diagram of the Handwriting Recognition System (HRS). The algorithms used in character recognition may be divided into three categories, namely: image preprocessing, feature extraction, and classification. They are normally used in sequence image. Preprocessing helps make feature extraction a smoother process, while feature extraction is necessary for correct classification.

3.1 Datasets

These are images of 20 sets of three individual writings that the HRS will learn and train with the use of different patterns. The datasets include when each individual is in a Happy State and Tired State likewise in two different inks. In machine learning, the higher the datasets the more accurate is the model. Fig. 2 to Fig 4. shows the data sets of two individuals.

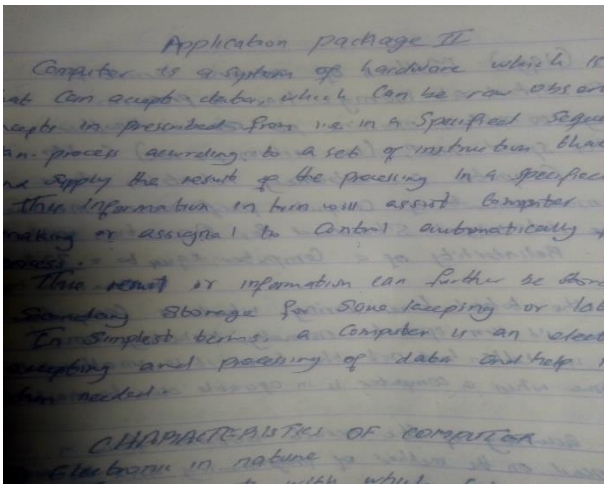


Fig. 2. Student writing in a Happy State

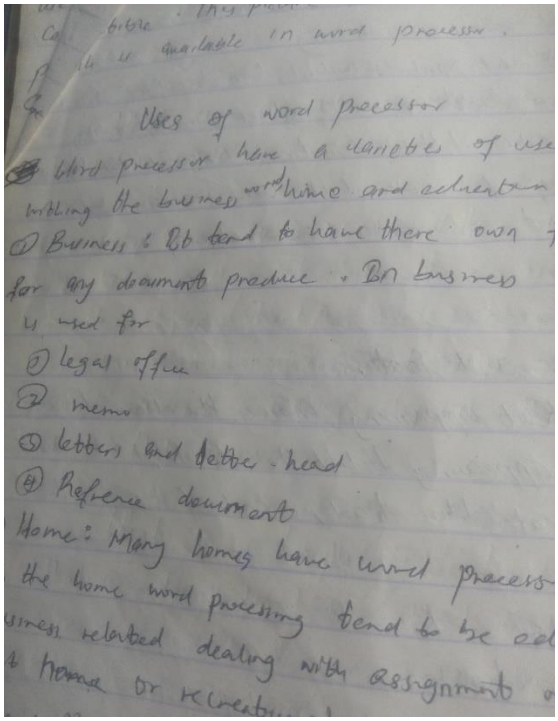


Fig. 3. Student writing in a Tired State

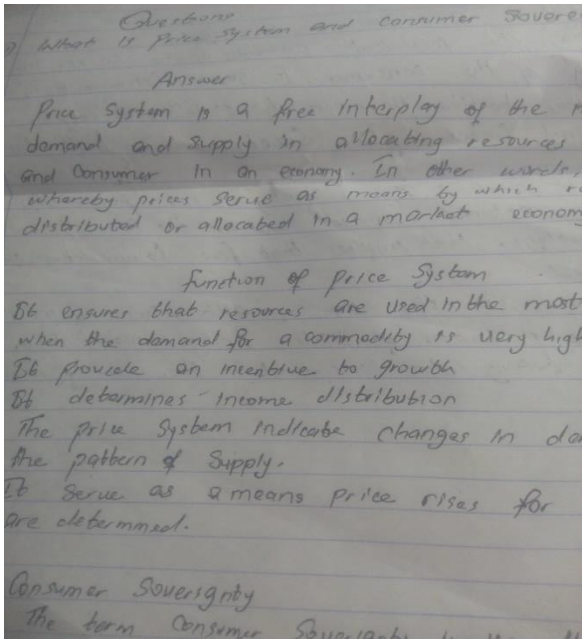


Fig. 4. Student in a happy state with a different ink

3.2 Image Acquisition

Digitized or digital image is initially taken as input. The most common of these devices is the electronic tablet or digitizer. These devices use a digital pen. Input images for handwritten characters can also be taken by using other methods such as scanners, photographs or by directly writing in the computer.

3.3 Image Preprocessing

Pre-processing is the basic phase of character recognition and it is crucial for good recognition rate. The main objective of pre-processing steps is to normalize strokes and remove variations that would otherwise complicate recognition and reduce the recognition rate. These variations or distortions include the irregular size of text, missing points during pen movement collections, jitter present in text, left or right bend in handwriting and uneven distances of points from neighboring position. Image preprocessing is crucial in the recognition pipeline for correct character prediction. These methods typically include noise removal, image segmentation, cropping, scaling, and more.

3.4 Segmentation

Segmentation is done by separation of the individual characters of an image. Generally, document is processed in a hierarchical way. At first, level lines are segmented using row histogram. From each row, words are extracted using column histogram and finally characters are extracted from words. The task of image segmentation is to split an image into parts with strong correlation with objects or the real-world properties the image represents.

3.5 Feature Extraction

Features of a character depict the morphological and spatial characteristics in the image. Feature extraction is a method of extracting of features of characters from the sample image.

The functions that are used in feature extraction are:

- Indexing and Labeling
- Boxing and Cropping
- Reshaping and Resizing

3.6 Classification

Classification is done with Convolution Neural Network (CNN) in order to have learning model analysis of the writings. To create a model of the trained handwritings, each image must go through the different layers of the CNN as shown in Fig. 5 by selecting the pixels three times because the depth is three. There are four types of layers for a CNN, namely:

- The Convolutional layer
- The Pooling layers
- The ReLU correction layer and
- The Fully connected/ Flatten layer

An algorithm of how the model is created is given in Fig. 5.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 200, 200, 32)      896
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 100, 100, 32)      0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 100, 100, 32)      9248
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 50, 50, 32)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 50, 50, 64)        18496
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 25, 25, 64)        0
_________________________________________________________________
flatten (Flatten)            (None, 40000)             0
_________________________________________________________________
dense (Dense)                (None, 128)               5120128
_________________________________________________________________
dropout (Dropout)            (None, 128)               0
_________________________________________________________________
dense_1 (Dense)              (None, 3)                 387
=================================================================
Total params: 5,149,155
Trainable params: 5,149,155
Non-trainable params: 0
_________________________________________________________________
```

Fig. 5. Creation of the model

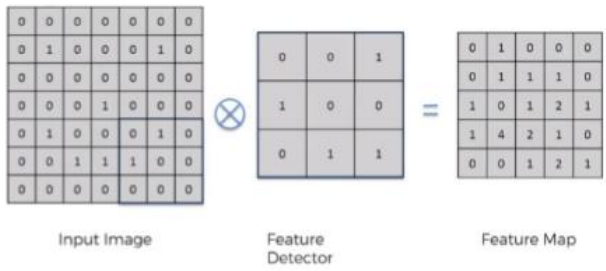


Fig. 6. Convolution operation

There is an input image of 7 x 7 in Fig. 6. The feature detector is of the size 3 x 3. The feature detector is placed over the image at the top left of the input image and then the pixel values in the feature detector are multiplied by the pixel values in the input image. The result is then added. The feature detector then moves to N step towards the right. Here, N refers to stride. A stride is basically the number of steps that a feature detector takes from left to right and then from top to bottom to find a new value for the feature map.

The ReLu operation is performed to introduce nonlinearity in the image. In the ReLu operation, all the negative values in a feature map are replaced by 0. All the positive values are left untouched. Suppose we have the feature map shown in Fig. 7.

-4	2	1	-2
1	-1	8	0
3	-3	1	4
1	0	1	-2

Fig. 1. Before ReLU

When the ReLU function is applied on the feature map in Fig. 7, the resultant feature map will look like what is shown in Fig.8.

0	2	1	0
1	-0	8	0
3	0	1	4
1	0	1	0

Fig. 2. After ReLU operation

Pooling does exactly that. In pooling, we have a feature map and then a pooling filter, which can be of any size. Next, we move the pooling filter over the feature map and apply the pooling operation. There can be many pooling operations such as max pooling, min pooling and average pooling. In max pooling, we choose the maximum value from the pooling filter.

Pooling not only introduces spatial invariance, but also reduces the size of an image. Consider the third and the fourth rows as well as the first and the second columns in Fig. 9, we have four values 1, 0, 1, and 4. When we apply max pooling on these four pixels, the maximum value will be chosen, i.e., 4 can be seen in the pooled feature map.

The pooled feature maps are flattened to form a one-dimensional vector to find more features from an image, as shown in Fig. 10.

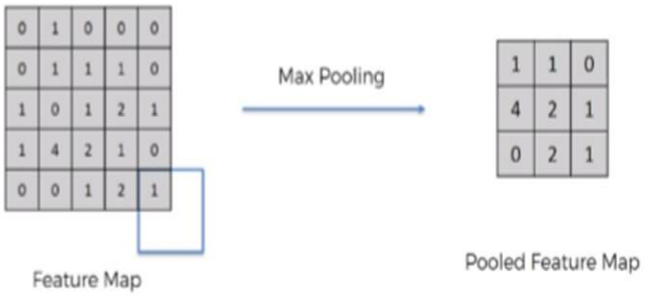


Fig. 9. Pooling operation



Fig. 10. Flatten layer

The one-dimensional vector is then used as input to a densely or fully connected neural network layer which gives out the image in Fig. 11 if dense.

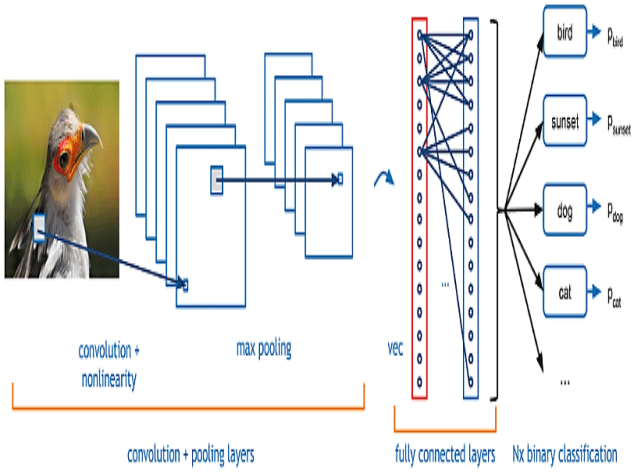


Fig. 11 CNN image after dense

4. RESULTS AND DISCUSSION

Fig. 12 and Fig. 13 show the graphical accuracy of both the accuracy and loss of the train and validation datasets. The accuracy of the train dataset is 92%, while the loss is 0.2%. As far as the results are concerned, the trained dataset has high accuracy which is very advantageous to our prediction rate.

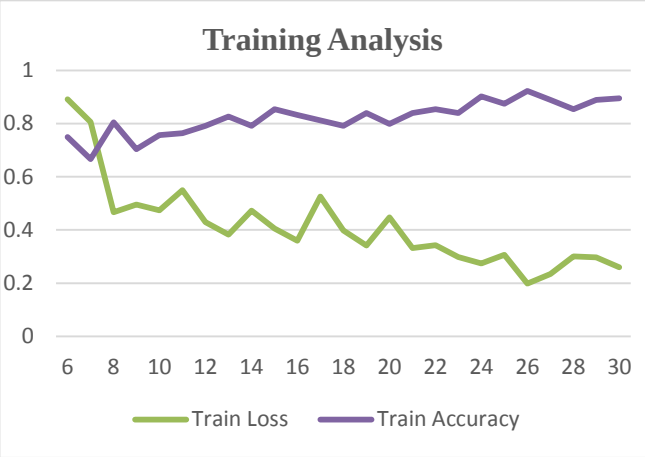


Fig. 12. Handwriting train analysis

The accuracy of the validation dataset is 98.1% while the loss is 0.7%. The results indicate that the validation dataset has high accuracy which is also very advantageous to our recognition rate.

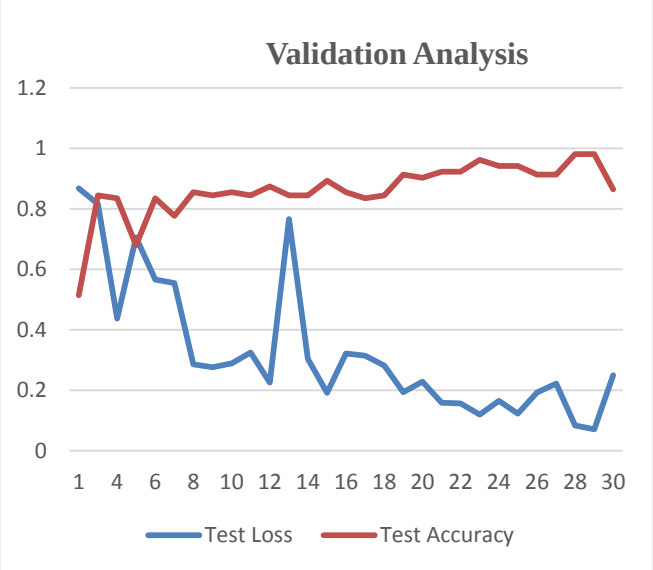


Fig. 13. Handwriting validation analyses

5. CONCLUSIONS

The structure behind identification of physical writing characters has been exhibited through this research. The span of this research involved functions, i.e. acquisition of images, selection of characters, categorization, pre-processing process, and creation of outcomes. A proposed handwriting text recognition system has been successfully designed, trained with 80 datasets, tested thereby achieving about 90% accuracy using CNN. The results obtained show that the trained and validation datasets have high accuracy. The results are good for the recognition rate.

In order to get up to 100% accuracy, more data must be added. A further recommendation is the use of Neural Network for handwriting detection on more challenging cases like prediction of individual's emotional state through handwriting.

REFERENCES

[1] J. Pradeep, S. Himavathi and E. Srinvanasan, "Diagonal Based Feature Extraction System Using Neural Network," 3rd International Conference on Electronics Computer Technology, pp. 364-368, 2012.

[2] M. Sonkusare and N. Sahu, "A Survey on Handwritten Character," Advances on Vision Computing, p. 12, 2017.

[3] P. Karishma and G. Mikita , "Offline Handwritten Character Recognition," A Survey Review, p. 5, 2017.

[4] M. Krishna, G. Satyanarayana and V. Sri, "Digital Image Processing Techniques in Character Recognition - A Survey.," 2018.

[5] H. Taihei, G. Keiji, O. Hidehiro and T. Toshiya, "A Study of Data Augmentation for Handwritten Character Recognition using Deep Learning," 2018 16th International Conference on Frontiers in Handwriting Recognition, p. 557, 2018.

[6] S. K. Girish and M. P. Singh, "Performance analysis of the feed forward, neural networks with gradient descent learning algorithms and structure similarity algorithm for recognition of handwritten "marathi" characters," Journal of Information, Knowledge and Research in Computer Science and Applications, p. 176, 2016.

[7] R. Plamondon and S. N. Sargur, "On-Line and Off-Line Handwriting Recognition: A Competitive Survey," IEEE Transactions On Pattern Analysis and Machine Intelligence, 2000.

[8] A. Mustafaoglu, "Offline handwriting recognition using neural network," Manchester, 2018.

[9] S. Ankit and C. R. Dipti, "Character Recognition Using Neural Network," Character Recognition Using Neural Network, p. 667, 2015.

[10] S. N. Bhavikatti, P. A. Desai and R. Patil, "Design & Simulation of Handwriting Text Recognition System," International Journal of Current Engineering and Technology, p. 259, 2013.

[11] B. Batuhan, D. Saadati and D. Shiferaw, "Handwritten Text Recognition using Deep Learning," 2018.

[12] D. Singh and P. Anita, "Handwritten English Character Recognition Using Neural Network," p. 144, 2010.

[13] M. Abayanarh and L. Zenkouar, "Offline Handwriting Character Using Moment Features and Neural Network," Computer Technology and Application, p. 29, 2017.

[14] I. C. Education, "Deep Learning," in IBM Cloud Learn Hub, 2020.

[15] K. Gurney, An Introduction to Neural Network, London: Berne Convention, 1997.

[16] J. Frankenfield, "Investopedia," 28 August 2020. [Online]. Available: https://www.investopedia.com/terms/a/artificial-neural-networks-ann.asp.