Nandhini V, 2021, 9:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

A Literature Survey on Handwriting Recoginition Using Deep Convolutional Neural Networks

Nandhini V, Pandi Geetha, Thammineni Pavithra, Asst. Prof. Dr. L. Malathi
Department of CSE,
Vivekanandha College of Engineering Women,
Namakkal, Tamilnadu.
nandhininandhini0966@gmail.com, pavithrachowdary0511@gmail.com, geetha112698@gmail.com,
malathi@vcew.ac.in.

Abstract- Handwritten character recognition is a translational problem of human writings into machine-editable text format. In this paper, Convolutional Neural Networks (CNN) is presented for handwritten character recognition. Handwritten character was transformed into graphs based on its underlying skeleton structure. Edges of the extracted graph were categorized into shape types and vertices were extracted from each of the edges and their layer wise evaluation using deep learning. Matching procedure of the graph was performed in Convolutional Neural Networks (CNN) approach. Performance evaluation of the proposed method was conducted using validated kaggle dataset which exclude ambiguous and unidentified writing samples. The use of neural network can improve the quality of recognition while achieving good performance and encouraging.

Keywords:- Deep learning, Hand writing detection, Artificial Neural Network, Convolution Neural Network.

I. INTRODUCTION

Handwriting digits and character recognitions have become increasingly important in today's digitized world due to their practical applications in various day to day activities. It can be proven by the fact that in recent years, different recognition systems have been developed or proposed to be used in different fields where high classification efficiency is needed.

Systems that are used to recognize Handwriting letters, characters, and digits help people to solve more complex tasks that otherwise would be time-consuming and costly. A good example is the use of automatic processing systems used in banks to process bank cheques. Without automated bank cheque processing systems, the bank would be required to employ many employees who may not be as efficient as the computerized processing system.

The handwriting recognition systems can be inspired by biological neural networks, which allow humans and animals to learn and model non-linear and complex relationships [1,2]. That means they can be developed from the artificial neural network [4]. The human brain allows individuals to recognize different Handwriting objects such as digits, letters, and characters [5]. However, humans are biased, meaning they can choose to interpret Handwriting letters and digits differently [8]. Computerized systems, on the other hand, are unbiased and can do very challenging tasks that may require humans to use a lot of their energy and time to do similar tasks. There is a need to understand how human-read under writing [10].

The human visual system is primarily involved whenever individuals are reading Handwriting characters, letters, words, or digits. It seems effortless whenever one is reading handwriting, but it is not as easy as people believe. A human can make sense of what they see based on what their brains have been taught, although everything is done unconsciously. A human may not appreciate how difficult it is to solve handwriting.

The challenge of visual pattern recognition is only apparent to develop a computer system to read handwriting [6,17]. The artificial neural networks approach is considered as the best way to develop systems for recognizing handwriting.

Neural networks help to simulate how the human brain works when reading handwriting in a more simplified form. It allows machines to match and even exceed human capabilities at reading handwriting. Humans have different handwriting styles, some of which are difficult to read. Besides, reading handwriting may be time-consuming and tedious, especially when individuals are required to read several Handwriting documents by different individuals [25].

A neural network is the most appropriate for the proposed system due to its ability to derive meaning from complex data and detect trends from data that are not easy to identify by either other human techniques or human [23]. The main aim of this paper is to develop a model that will be used to read Handwriting digits, characters, and words from the image using the concept of Convolution Neural Network. The next sections will provide an overview of the related work, theoretical background, the architecture, methodology, experimental results, and conclusion.

1. Research Objectives:

The main objective of this research is to design an expert system for Handwriting character recognition using neural network approach.

Other objectives include:

- To address the issue of accuracy in Handwriting character recognition systems by developing a system that will use efficient technology for recognizing Handwriting characters and words from image media.
- To investigate and demonstrate the usefulness of neural network technology in development of efficient Handwriting character recognition systems.

2. Research Questions:

This research is aimed to answer the following questions:

- What are the different techniques and methods used in Handwriting character recognition?
- How can the performance of Handwriting recognition systems be improved using artificial neural networks?

3. Target Group:

This paper will be targeting university students and instructors who want to convert their Handwriting notes and papers into electronic format. Despite the increased adoption of digital technology in institutions of higher education, handwriting remains part of students' and instructors' daily lives. Students take Handwriting notes while listening to their lectures and take notes while reading from different sources. Some also note down their thoughts, plans, and ideas on their notes.

Likewise, lecturers have Handwriting notes that they would want to communicate to students. Hence, this paper will be targeting students and lecturers to develop a system that will allow them to convert their Handwriting works into electronic works that can be stored and communicated electronically. The central assumption of this paper is that students and lecturers need to have copies of their works that are stored electronically in their personal computers. Further, handwriting with pen and paper cannot be entirely replaced by digital technology.

II. RELATED WORKS

Handwriting digit recognition has an active community of academics studying it. A lot of important work on convolutional neural networks happened for handwritten digit recognition [1, 6, 8, 10]. There are many active areas of research such as Online Recognition, Offline recognition, Real-Time Handwriting Recognition, Signature Verification, Postal-Address Interpretation, Bank-Check Processing, Writer Recognition.

III. METHODOLOGY

1. Neural Networks:

Neural network is a system inspired by human brain function, consists of neurons connected in parallel with the ability to learn. A basic design of neural network has input layer, hidden layer, and output layer. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract pattern and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been to analyze. Neural networks are particularly useful for solving problems that cannot be expressed as a series of steps, such as recognizing patterns, classifying them into groups, series prediction and data mining.

2. Network Layers:

The commonest type of artificial neural network consists of three groups, or layers, of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. (Figure 1)

- The activity of the input units represents the raw information that is fed into the network.
- The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units.
- The behaviour of the output units depends on the activity of the hidden units and the weights between the hidden and output units. This simple type of network is interesting because the hidden units are free to construct their own representations of the input. The weights between the input and the hidden units determine when each hidden unit is active, and so by modifying these weights, a hidden unit can choose what it represents.

3. Deep Neural Network:

The neural network has layers of units where each layer takes some value from the previous layer. That way, systems that are based on neural networks can compute inputs to get the needed output [29]. The same way neurons pass signals around the brain, and values are passed from one unit in an artificial neural network to another to perform the required computation and get new value as output [17].

The united are layers, forming a system that starts from the layers used for imputing to layer that is used to provide the output. The layers that are found between the input and output layers are called the hidden layer. The hidden layers refer to a deep neural network that is used for computation of the values inputted in the input layer. The term "deep" is used to refer to the hidden layers of the neural network [25] as shown in Fig. 6.

In Handwriting character recognition systems, the deep neural network is involved in learning the characters to be recognized from Handwriting images [33].

With enough training data, the deep neural network can be able to perform any function that a neural network is supposed to do. It is only possible if the neural network has enough hidden layers, although the smaller deep neural network is more computationally efficient than a more extensive deep neural network [19].

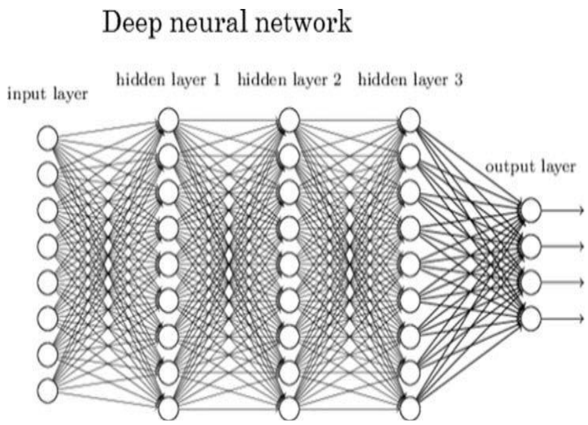


Fig 1. Deep Neural Network.

III. CONVOLUTIONAL NEURAL NETWORKS

In this paper, Convolutional neural network (CNN) framework was used for handwritten character recognition. In that framework, proper sample generation, training scheme and CNN network structure was employed according to the properties of handwritten characters. That CNN-based framework achieved better performance compared with other CNN-based recognition methods.

That CNN-based framework mainly consists of three parts: The Sample generation, CNN models and voting. Sample generation used distortions such as local and global distortion. CNN model was for better training and Voting can significantly improve recognition rate.

The error-rate by this CNN-based framework for character recognition for MNIST data set was 0.18%. So we can still improve this framework by enlarging the CNN scale or input image size. Also we can find better sample generation methods, training scheme and network structure of CNN.

A special type Neural Networks that works in the same way of a regular neural network except that it has a convolution layer at the beginning. A convolutional neural network consists of an input and an output layer, as well as multiple hidden layers.

The hidden layers of a CNN typically consist of convolutional layers, RELU layer i.e. activation function, pooling layers, fully connected layers and softmax layers.

1. Input Layer:
The input layer is used to feed the system with the image with the handwriting. The layer can be colored image (RGB values) or grayscale. It can have dimension W*H*W, depending on the input image. The W*H refers to the width and height of the image, while D refers to the depth of the image.

2. Convolution Layer:
The convolution layer is the building block of the whole network. Most of the computational work that is required to recognize characters from the input is done in this layer (Aggarwal, 2018). The layer consists of a set of learnable filters known as parameters of the convolution layer.

3. Pooling Layer:
The pooling layers are found between the convolutional layers in the CNN architecture. They are responsible for progressively reduce the spatial size of computational work in the network. They help to streamline the underlying computation. They do so by reducing the dimension of the input data by combing the outputs of the neuron clusters. They operate independently. That way, the system can achieve the intended outputs.

4. Fully Connected Layer:
Neurons in a fully connected layer are fully connected to all activations in the prevision layer. Hence, this layer, activations, can be computed with matrix multiplication. Based on the architecture, a system can have multiple fully connected layers. In summary, CNN can be used to achieve a solution to every pattern recognition issue.

The architecture demonstrated above shows how OCR systems using neural networks can read handwriting. The convolutional networks work in the hierarchy and can be used to solve complex structures found in handwriting inputs. Humans inspire the whole idea can recognize writing objects and process what they see in their brains.

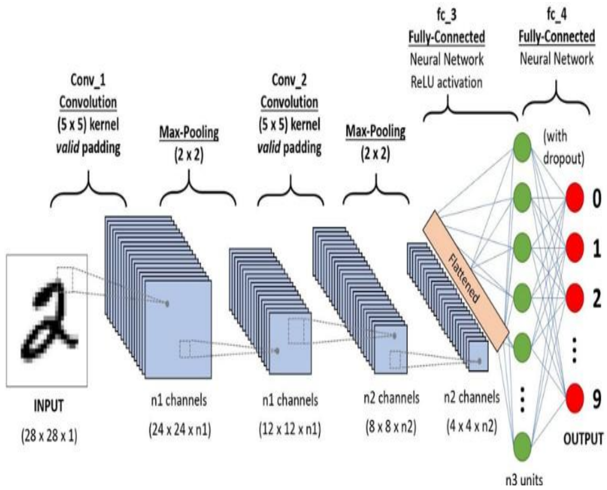


Fig 2. Fully connected layer.

IV. MODULES

The current OCR system will consist of five phases. The phases are image acquisition and digitization, preprocessing, segmentation, feature extraction, and recognition. Fig. 11 shows the methodology that will be used to read handwriting.

1. Image Acquisition and Digitization:
The image acquisition step involves acquiring an input image that contains handwriting. The image, in this case, should be in specific formats such as PNG and JPEG. The image is acquired through a digital camera, scanner, or any other suitable input device. The digitization step, on the other hand, involves converting the input paper into electronic format [20].

The conversion is achieved by first scanning the original document and representing it in the form of an image that can be stored on a computer. The digital image is essential for the pre-processing phase.

2. Preprocessing:
Preprocessing is the second phase of OCR after the digital image has been made as shown in Fig. 12. The digitized image is pre-processed to remove noise, and then it is checked for skewing. Preprocessing is essential for developing data that that are easy for optical character recognition systems. The main objective of pre-processing is to remove the background noise, enhance the region of interest in the image, and make a clear difference between foreground and background.

2.1 Image Enhancement Techniques: To modify attributes of the image to make it more suitable and to improve the quality of the image by reducing noise, increasing contrast, image blurring, and providing more details.

Hence, to process an image so that result is more suitable than the original image and providing better input for automated image processing techniques.

2.1.1 Noise removal: Addictive noises of different types can contaminate images. Hence there is a need to remove noise to improve the quality of the image.

2.1.2 Binarization: This method is used to transform the grayscale image and converting it to black and white, substantially reducing the information contained within the image from different shapes of gray into a binary image.

2.1.3 Normalization: This process in image processing that changes the range of pixel intensity values. Its common purpose of converting an input image into a range of pixel values that are more familiar to the senses. Normalization involves converting images into a standard size.

2.1.4 Skew Correction, Thinning: This is one of the first operations to be applied to scanned documents when converting data to digital format. This process helps to get a single-pixel width to allow easy character recognition. Preprocessing for handwriting characters of current approach is shown in Fig. 13.

3. Segmentation:

Segmentation can be argued to be the most critical process in character recognition techniques. Segmentation of images is done in the testing stage only. It checks for any error point inclusion by checking all points against the average distance between segmentation points incomplete image. The process involves separating individual characters from an image, as shown in Fig. 14.

The process results in multiple segments of the image known as super pixels. The main aim of segmentation is to simplify the representation of an image into something that can be analyzed easily. Hence it has a positive impact on the recognition rate of the script.

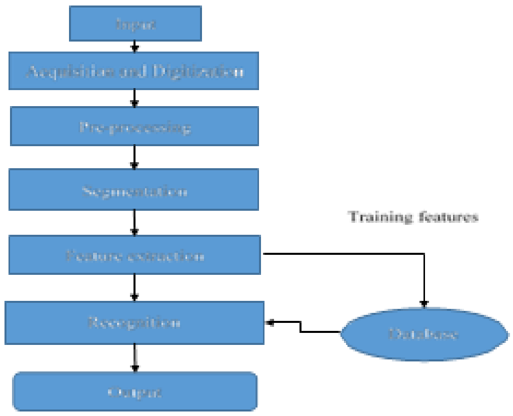


Fig 3. OCR System.

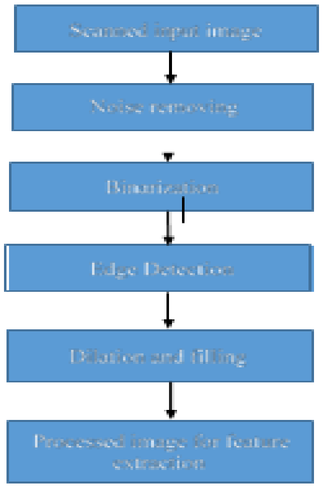


Fig 4. Preprocessing Techniques.

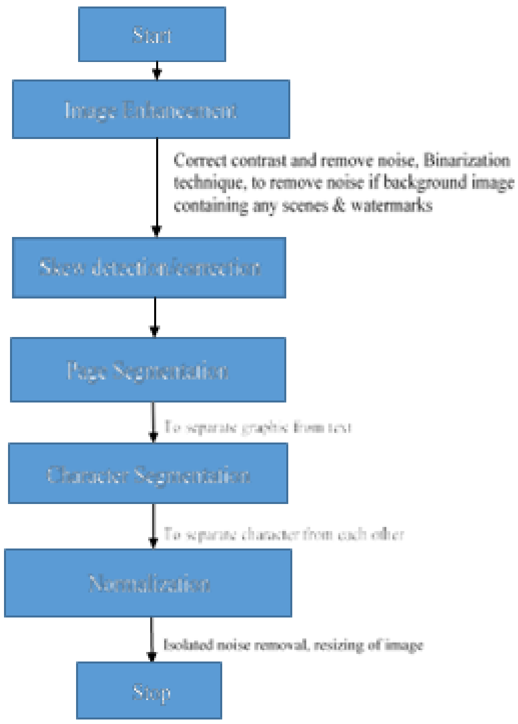


Fig 5. Preprocessing handwriting characters.

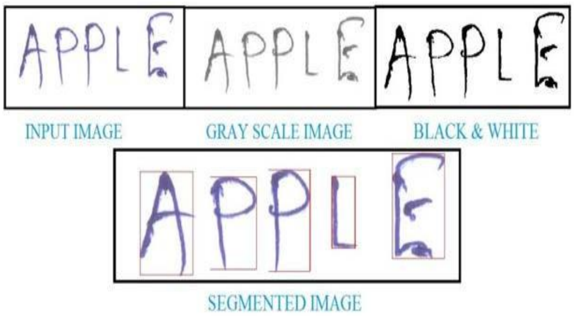


Fig 6. Example of a Segmented Image.

4. Feature Extraction:
In this phase, features of the image are extracted and are defined based on the following attributes: height of the character, numbers of horizontal lines, widths of the character, number of circles, pixels, position of different features and number of vertically oriented arcs, to mention a few.

5. Recognition:
In this phase, the neural network is used for classification and recognition of the characters from the image. The most neural networks that are used by optical character recognition systems are the multiplayer perception (MLP) and Korhonen's Self Organizing Map.

V. DATABASE DETAILS

Here we have used six standard databases available freely for research work. Practices for convolutional neural networks applied to Visual Document Analysis There are mainly two most important things: first one is getting a training set as large as possible; we expand the training set by adding a new form of distorted data. Second is; Convolutional neural networks are better suited for visual document tasks than fully connected networks.

This paper illustrates claims on the MNIST set of English digit images. This paper shows that neural networks achieve the best performance on a handwriting recognition task (MNIST). And the optimal performance on MNIST was achieved using two essential practices. First, Authors created a new general set of elastic distortions and second, they used convolutional neural network.

They also compute a grey level of an object with an example. And the algorithm for evaluating grey level is "bilinear interpolation". Affine distortions greatly improved results on the MNIST database but the best results were obtained when they used elastic deformation. At the end results shows that they achieved the highest performance known to date on the MNIST data set, using elastic distortion and convolutional neural network.

VI. IMPLEMENTATION

To implement our CNN architecture, we will use MatConvNet. MatConvNet is an implementation of Convolutional Neural Networks (CNN) for MATLAB [12]. It exposes the building blocks of CNN as easy-to-use MATLAB functions, providing routines for computing linear convolutions with filter banks, feature pooling and many more. In this manner, MatConvNet allows fast prototyping of new CNN architectures; at the same time, it supports efficient computation on CPU and GPU allowing to train complex models on large datasets such as Image Net ILSVRC.

Convolutional Neural Networks (CNN) are the current state-of-art architecture for the image classification task. Our proposed 2-D Convolutional Neural Network (CNN) model is designed using MatConvNet backened for the well known MNIST digit recognition task. The whole workflow can be to preparing the data, building and compiling of the model, training and evaluating the model and saving the model to disk to reuse.

Preparing the data is the first step of our approach. Before we build the network, we need to set up our training and testing data, combine data, combine labels and reshape into the appropriate size. We save the dataset of normalized data (single precision and zero mean), labels, and miscellaneous (meta) information.

Building and compiling of the model is the second step. To create the CNN, we must initialize MatConvNets DagNN (Directed acyclic graph neural network) network and then define important initialization parameters.

Batch size determines the number of samples for the training phase of the CNN. The CNN will process all the training data, but only in increments of the specified batch size. We use batch size for computational efficiency, and its value will be dependent on the user's available hardware. An epoch is a successful forward pass and a backward pass through the network. It's usually beneficial to set its value high. Then we can reduce the value once if one is satisfied with the convergence at a particular state (chosen epoch) in the network. Learning rate is a very sensitive parameter that pushes the model towards convergence. Finding its best value will be an empirical process unless one invokes more powerful techniques such as batch normalization. In our experiment, we use batch size 40, a number of epochs 8 and learning rate 0.001 for maximum accuracy.

Now we can build our CNN by creating each layer individually. Afterward, we will invoke objective (log loss) and error (classification loss) layers that will provide a graphical visualization of the training and validation convergence after completing each epoch. After building the network layers, we initialize the weights. MatConvNet does this with Gaussian distribution.

The third step is the training and evaluating the model. Training a CNN requires computing the derivatives of the loss concerning the network parameters. We use back propagation algorithm for computing derivatives. It is a memory-efficient implementation of the chain rule for derivatives. We have use Stochastic Gradient Descent (SGD) training algorithm to adjust the weight of the connection between neurons so that the loss reaches a minimum value or stops after several epochs. Only log loss is used to adjust weights. We have used CPU training. It is important to note that GPU training will dramatically help training time for CNN.

Lastly, we can begin the training of CNN by supplying the training data, the constructed model and the current 'batch' of data. When training the CNN, only the data specified for training plays a role in minimizing error in the CNN. The algorithm uses training data for the forward and backward pass. The algorithm uses validation data to see how the CNN responds to new similar data, so it only is fed through the forward pass of the network. Afterward, we save the trained CNN and prepare for the testing phase.

During the training phase of the CNN, each epoch will produce up to two plots (error and objective) shown in Fig.6. For the DagNN network, MatConvNet minimizes the loss, while the error plot allows one to compute more statistical inference. The blue curve represents the training error and the orange represents Global Journal of Computer Science and Technology Volume XIX Issue II Version I the validation error of our CNN model during training.

VII. RESULTS

In this project we have given image as an input then it predicts the output by loading the model which is already previously created and saved.

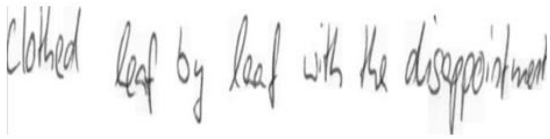
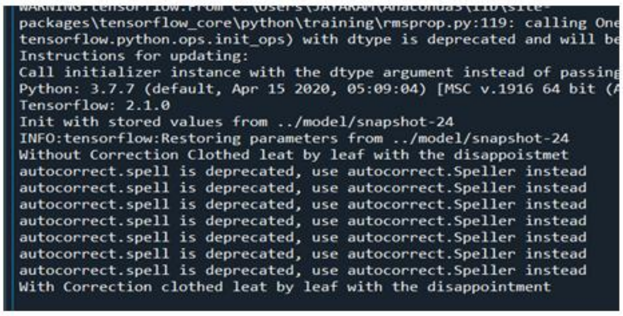


Fig 7. Input Image.

The above image is the input given to the neural network to predict the solution.



```
packages\tensorflow_core\python\training\rmsprop.py:119: calling One
tensorflow.python.ops.init_ops) with dtype is deprecated and will be
Instructions for updating:
Call initializer instance with the dtype argument instead of passing
Python: 3.7.7 (default, Apr 15 2020, 05:09:04) [MSC v.1916 64 bit (A
Tensorflow: 2.1.0
Init with stored values from ../model/snapshot-24
INFO:tensorflow:Restoring parameters from ../model/snapshot-24
Without Correction Clothed leat by leaf with the disappoistmet
autocorrect.spell is deprecated, use autocorrect.Speller instead
autocorrect.spell is deprecated, use autocorrect.Speller instead
autocorrect.spell is deprecated, use autocorrect.Speller instead
autocorrect.spell is deprecated, use autocorrect.Speller instead
autocorrect.spell is deprecated, use autocorrect.Speller instead
autocorrect.spell is deprecated, use autocorrect.Speller instead
autocorrect.spell is deprecated, use autocorrect.Speller instead
With Correction clothed leat by leaf with the disappointment
```

Fig 8. Output Image.

This is image which shows the output to the above input image.

VIII. CONCLUSION AND FUTURE SCOPE

In this project classification of characters takes place. The project is achieved through the conventional neural network. The accuracy we obtained in this is above 90.3%. This algorithm will provide both the efficiency and effective result for the recognition. The project gives best accuracy for the text which has less noise. The accuracy completely depending on the dataset if we increase the data, we can get more accuracy. If we try to avoid cursive writing then also its best results.

Future Work: In future we are planning to extend this study to a larger extent where different embedding models can be considered on large variety of the datasets. The future is completely based on technology no one will use the paper and pen for writing. In that scenario they used write on touch pads so the inbuilt software which can automatically detects text which they writing and convert into digital text so that the searching and understanding very much simplified.

REFEREENCES

[1] Fischer, A., Frinken, V., Bunke, H.: Hidden Markov models for off-line cursive handwriting recognition, in C.R. Rao (ed.): Handbook of Statistics 31, 421 – 442, Elsevier, 2013.

[2] Frinken, V., Bunke, H.: Continuous handwritten script recognition, in Doermann, D., Tombre, K. (eds.): Handbook of Document Image Processing and Recognition, Springer Verlag, 2014.

[3] S. Günter and H. Bunke. A new combination scheme for HMM-based classifiers and its application to handwriting recognition. In Proc. 16th Int. Conf. on Pattern Recognition, volume 2, pages 332–337. IEEE, 2002.

[4] U.-V. Marti and H. Bunke. Text line segmentation and word recognition in a system for general writer independent handwriting recognition. In Proc. 6th Int. Conf. on Document Analysis and Recognition, pages 159– 163.

[5] M. Liwicki and H. Bunke, "Iam-ondb - an on-line English sentence database acquired from the handwritten text on a whiteboard," in ICDAR, 2005.

[6] A. Graves and J. Schmidhuber, "Offline handwriting recognition with multidimensional recurrent neural networks," in Advances in neural information processing systems, 2009, pp. 545–552.

[7] Nafiz Arica, and Fatos T. Yarman-Vural, ―Optical Character Recognition for Cursive Handwriting,‖ IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no.6, pp. 801-113, June 2002.

[8] Anita Pal and Davashankar Singh,"Handwritten English Character Recognition Using Neural Network", International Journal of Computer Science and Communication, pp: 141- 144, 2011.

[9] Sandhya Arora, "Combining Multiple Feature Extraction Techniques for Handwritten Devnagari Character Recognition", IEEE Region 10 Colloquium and the Third ICIIS, Kharagpur, INDIA, December 2008.

[10] Om Prakash Sharma, M. K. Ghose, Krishna Bikram Shah, "An Improved Zone Based Hybrid Feature Extraction Model for Handwritten Alphabets Recognition Using Euler Number", International Journal of Soft Computing and Engineering (ISSN: 2231 - 2307), Vol. 2, Issue 2, pp. 504- 508, May 2012.

[11] N. Venkata Rao and Dr. A.S.C.S.Sastry - Optical Character Recognition Technique Algorithms‖- 2016 Journal of Theoretical and Applied Information Technology.

[12] Ganapathy, V., & Liew, K. L. (2008). Handwritten Character Recognition Using Multiscale Neural Network Training Technique. World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering, 2(3), 638-643.

[13] Grady, J. O. (2010). System Requirements Analysis. Amsterdam, Netherlands: Elsevier.

[14] Gribaudo, M. (2013). Theory and Application of Multi-Formalism Modeling. Hershey, PA: IGI Global.

[15] Gunawan, T. S., Noor, A. F. R. M., & Kartiwi, M. (2018). Development of english handwritten recognition using deep neural network. Indonesian Journal of Electrical Engineering and Computer Science, 10(2), 562-568.

[16] Hertz, J. A. (2018). Introduction to the theory of neural computation. Boca Raton: CRC Press.

[17] Hamid, N. A., & Sjarif, N. N. A. (2017). Handwritten recognition using SVM, KNN and neural network. arXiv preprint arXiv:1702.00723.

[18] Kumar, P., Saini, R., Roy, P. P., & Pal, U. (2018). A lexicon-free approach for 3D handwriting recognition using classifier combination. Pattern Recognition Letters, 103, 1-7.

[19] Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L., & Maglogiannis, I. (2018). Artificial Neural Networks and Machine Learning – ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings. Basingstoke, England: Springer.

[20] Larasati, R., & KeungLam, H. (2017, November). Handwritten digits recognition using ensemble neural networks and ensemble decision tree. In 2017 International Conference on Smart Cities, Automation & Intelligent Computing Systems (ICON-SONICS) (pp. 99-104). IEEE.

[21] Lee, J. H., Shin, J., & Realff, M. J. (2018). Machine learning: Overview of the recent progresses and implications for the process systems engineering field. Computers & Chemical Engineering, 114, 111-121.

[22] Lu, L., Zheng, Y., Carneiro, G., & Yang, L. (2017). Deep Learning and Convolutional Neural Networks for Medical Image Computing: Precision Medicine, High Performance and Large-Scale Datasets. Basingstoke, England: Springer.