CHARACTER RECOGNITION USING DEEP LEARNING ALGORITHM

Sridhar.S[1,1], Guttula Bhargav Mani Deep[1], Asish Bhoi[1], Danthuluri Swathi[1] , Deepilli Leelarani[1]

1.1Professor,department of Electronics and Communication,Lendi Institute of Engineering and Technology,Vizianagaram.

1 Student,department of Electronics and Communication,Lendi Institute of Engineering and Technology,Vizianagaram.

Abstract-This paper describes deep learning algorithms for character recognition. The huge amount of data and algorithms now makes it easier to train detailed neural networks. Character recognition is another application running using otherapproaches. This paper provides a CNN (Convolutional Neural Network) -based character recognition method. Trainyour dataset and classify it to see the results.A character recognition system was designed using image segmentation. Character recognition is done by the help of the python programminglanguage**.**

Keywords-Python,Deep Learning,Neural Networks, Datasets, Optical CharacterRecognition (OCR)

1 Introduction

At present highly technology depends on the huge data volume and also the requirements, for all purposes demands a large input of data storages into the computers. AI (Artificial Intelligence) is widely used across all over the world with the help of deep learning networks different applications are used. As the DNNS comes with the cost and high complexity. The Deep learning techniques are widely used as to improve their efficiency without percentage loss of accuracy or increasing in hardware cost in AI systems. The performance of DNNs comes as from raw data after learning over large and process it for extracting features. However, the DNNs comes with high complexity to obtain greater accuracy.While for DNNs Computation as at basic need form the general-purpose computers as for DNN processing usage of graphical processing units required.

OCR technology generally tends to recognition of characters, words and to compute them such that determine the information and to change into computerized character.As OCR combines with most of the research topics such signal processing, Computer graphics etc. Optical Character recognition is mechanism of conversion of images of different characters into the machine encoded text or text-imposed image. As OCR is widely used for different applications such as automatic data entry for different uses as for example passports, computerized receipts for different purposes, business cards etc. Asearly versions are needed to trained on different images worked on the single font later different fonts of images are used. The variety of image file format inputs are capable of high-level fonts classification is common nowdays.AS there are some systems that are approximates the original images capable of reproducing formatted output. Early technology involves telegraphy used for

optical character recognition. Emmanuel Goldberg developed a system forreading characters in 1914.Afterthat, those characters are converted into telegraph codes.In this project we take an image which consist of character such as words,digit and further processed such thatdigital character is formed. As the project further processed as it combined of the training the neural network with algorithm that segment the characters images in the given image which processes with neural network. Then add layers to take the form available to the end user, and the full-featured model helps the end user turn various characters into digitized output. There willbe need of the add in the layers as such that the segmentation of words takes place. We approach this issue. CNNs tend to work better with raw input pixels than image features or parts thathave a complete word image. Usethe proposed deep learning techniques to classify and identify different images.

2 About Artificial Intelligence and Machine learning

2.1 Artificial Intelligence

The term artificial intelligence was defined by John McCarthy in1956 as"the science and engineering of creating intelligent machines". AI is a simulation of human intelligence
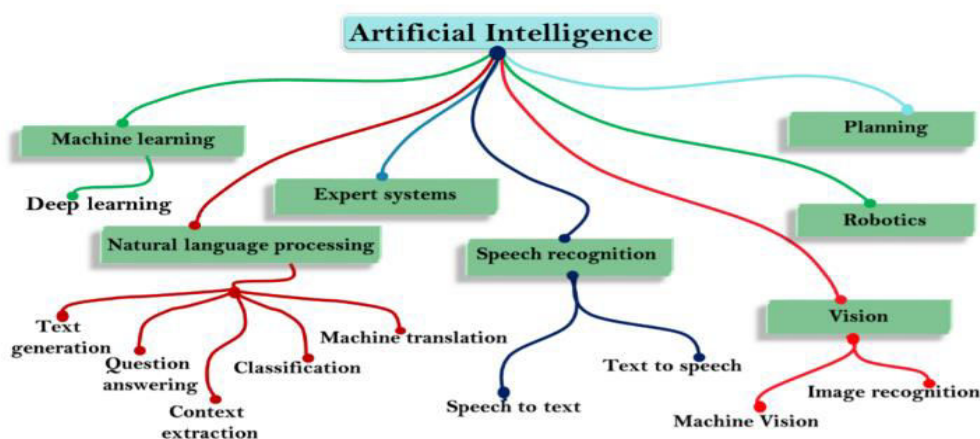

Figure 1: Subfields of Artificial intelligence

processed by computers. In particular,AI's unique applicationsof computer systems include expert systems, natural language processing, and speech.Cognition and machinevision.AI refers to the simulation of various resources to mimic human behaviour. The term may also relateto the human mind, such as learning orproblem-solving that machinesdenote.AI includes a variety oftechnologies and algorithms,including machine learning,neural network expert systems, and deep learning,including robotics, as shown.Currently, AI technology is having a significant impact worldwidein variousfields such as medical, space, robotics, and military. To overcome some of the limitations of symbolic AI, they are a varietyof methodologiescalledcomputationalintelligence,including neural networks, fuzzy systems, evolutionary computations, and other computational models. Although the methodsusedinAI aredifferent, the two main methodologies are top-down and bottom-up. While thetop-down approach believes in creating computer programs accordingto the operation of the human brain, and outlines of the system are established andspecified, the bottom-up approach is realized by providingsimilar electronicreplication to the neural networks of the human brain.

2.2 Machine Learning

Machine learning (ML) is the branch of artificial intelligence (AI), which can learn and improve in the explicit programming of an application by running it automatically or regularly (or) improving over time through machine learning (ML) experiences. The study of computer algorithms. data. It is being considered a component of artificial intelligence. Machine learning algorithms build models based on training data and make predictions and judgments without explicit programming. Machine learning algorithms, medicine, e-mail filtering, computer vision, etc. are being used in a variety of applications where it is difficult or impossible to develop conventional algorithms to perform the necessary tasks. A part of machine learning closely related to computer statistics that focuses on the generation of predictions by computers. However, not all machine learning is statistical learning. Mathematical optimization studies provide methods, theory, and applications to the field of machine learning. Data mining is a related field of research that focuses on exploratory data analysis when applied to business problems, and machine learning is also used.

Machine learning involves discovering ways in which computers can perform tasks without explicit programming. Here you can learn your computer from the data provided and perform specific tasks. For simple operations assigned to a computer, you can program an algorithm that tells the machine how to take all necessary steps to solve the problem. You don't have to learn from the computer part. For advanced tasks, one can manually create the required algorithms. It is more efficient to help develop the algorithm of the machine itself than to actually let a human programmer specify every step needed. Available. If there are many potential answers, one way is to display some of the correct answers as valid. This data can be used as training data to improve the algorithms that computers use to determine the correct answer. For example, the MNIST handwritten digit dataset is often used to train one task digital character recognition systems.

2.3 Deep Learning

Deep Learning is an artificial intelligence (AI) function that mimics the function of the human brain in processing data and generating patterns used in decision making. Also known as deep chi learning or deep neural networks. Deep learning is a class of machine learning algorithms that uses multiple layers to slowly extract high-level features from a primary input. For example, image processing can identify lower layer edgesand upper layers can identify human concepts such as numbers, letters, and faces. The ability to handle many features makes deep learning very powerful when processing unstructured data. However, since deep learning algorithms can access vast amounts of data to be effective, they are likely to be overkill for less complex problems.Deep learning algorithms try to learn high-level features from data. This is a very characteristic part of deep learning, and an important step in conventional machine learning. Thus, deep learning can reduce the work of developing new feature extractors for every problem. The increase in high-performance computing facilities is driving deep learning techniques to implement deep neural networks. Deep learning offers greater power and flexibility when dealing with unstructured data because it can handle multiple functions. Deep learning algorithms pass data to multiple layers. Each layer can extract features step by step and pass them on to the next layer. The first layer extracts low-

level features and combines the features of subsequent layers to form a complete expression. Deep learning continues to evolve rapidly. Still, the problem to be solved can be solved using some deep learning. Fully understanding the structure of deep learning is still a mystery, but deep learning could be used to make computers smarter. You might even be smarter than humans. The current task is to develop deep learning models in conjunction with mobile to make applications smarter and more intelligent. Here are some OCR solutions in the field of machine learning and pattern recognition. OCR technology can be classified into conventional methods, namely neural network models, based on homemade features and deep learning methods. Existing methods involve homemade feature extraction and classify them accordingly. Since these homemade features are not robust, we cannot guarantee a lot of recognition rate and accuracy. It is also computationally intensive due to the higher dimensions of these features. Therefore, discrimination is reduced. Some of the powerful techniques that can be applied to deep learning algorithms to reduce training time and optimize models are CNN (Convolutional Neural Networks), RNN (Recurrent Neural Networks), RCNN (Recurrent Convolutional Neural Network).

3 Different Character Recognition Techniques

Character recognition isprimarily accomplished using a variety of methods that include differentsteps, such as checking images and identifyingcharacters oftext area text characters in each text area. It uses a standarddeep learning model to identify characters and words in images.As a deep learning model, thereare excellent learning models for identifying various characters. Therefore, special deep learning models has taken to supportlocalization and detectionof images. Someof the most commonly used methods are shown.

Different Models are used for character recognition as first model as the RAM model.The RAM model (Recurrent Attention Model) is given such that when a new scene is presented to the human eye, a specific part of the image attracts the line of sight, so as the first model, other models recognize characters used for the eye first gets the information by focusing on the "slip" of the information. Modelimagesare provided in other sizes around a common center, as filtering and glyph vectors are generated for the salient features of each cut version. These glyph vectors are flattened and pass through a "glyph network" based on visual attention. The Glimpse vector is then sent to the location network and uses the RNN to predict the next part of the image to note. This is the next input for a glimpse of the network. Gradually, the model moves additional parts of the image toensurethat thepreviousglimpseof
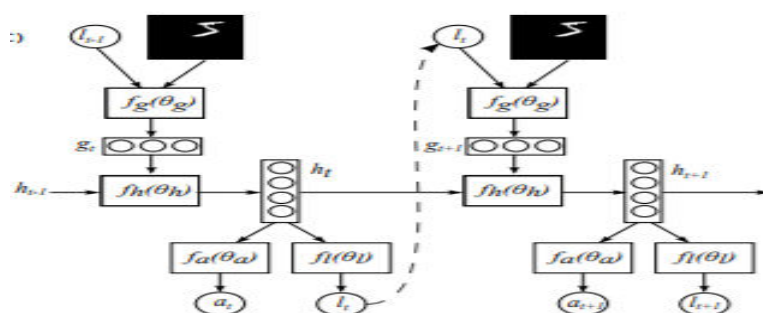
information is sufficient to achieve a high level of accuracy each time the error station radio wave method is performed.

The second method viaAttention is an OCR project which can be used with TensorFlow and is designed for the problem of the original image captions. Caution Include a CRNN followed by a decoder. First, the model uses a complex network layer to extract features from the image. It encodes these functions as a string, passes it to the RNN, and implements a week mechanism borrowed from the Seq2Seq machine translation model. Note The base decoder is used to predict the text of the input image. The above two processes are not only more efficient, but also less accurate.

Last and Most Important Approach The Convolutional Recurrent Neural Network (CRNN). The CRNN approach uses three basic steps to identify a word. The first step is a convolutional neural network (CNN) by processing an image.The first layer divides the image into features and dividesit into"featurecolumns".Later, these columns are given todeep bidirectional LSTM (long-term and short-term memory) cells that provide a sequence for identifyingrelationships between characters. Finally, the LSTM cell output is fed to thetransfer layer, which proceedsa string containing duplicate characters and uses a stochastic approach to organize the output.

4 Proposed CNN Architecture

CNN architecturehas two main parts.First, a fullyconnected layer-previousstepthatuses the convolution tool and the output of the convolution process to separate and identify the features that are present inan image and analysis in a process called featureextraction and predict the class of the image accordingto the features extracted.
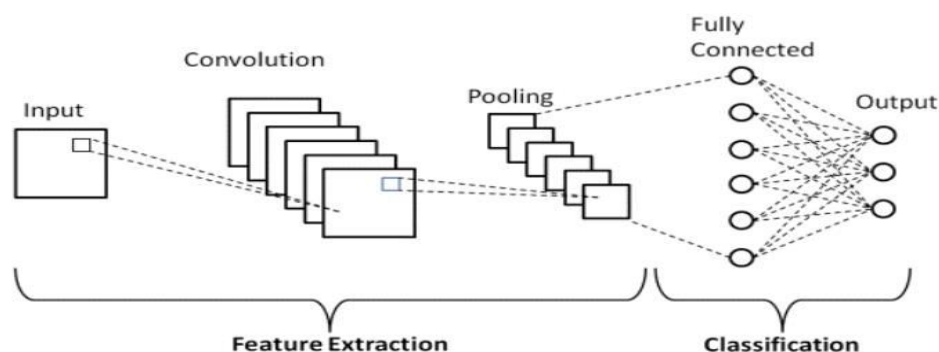
Figure(a). Basic Architecture

There are three types of layers that connects to form a CNN. Convolutional Layer, Pooling Layer, Full Connection (FC) Layer. A CNN structure is formed in which these layers accumulate. In addition to these three layers, there are two important parameters.The dropout layer and activation function are defined below.

1. Convolutional Layer

This layer is the first layer used for extraction of features from input image. This layer is where the convolution operation between the input image and a particular size MxM filter is performed. Slide the filter onto the input image to get the dot product between some of the filter input images for the filter size (MxM). The output is called a functional map and provides information about edges and edge-like images.Later,the feature map is transferred to other layersto learn other features of the input image.

2. Pooling Layer

In most cases, a full layer after the convolution layer continues. The main purpose of this layer is to reduce the computational cost by reducing the size of the convolution feature map. This is achieved by reducing the connections between layers,working independently on each functional map. There are several types of pooling operations, depending on which method you use. The largest element ofmaxpooling is retrievedfrom the function map. The average pool calculates the average of the elements of animage section of a predefined size. Totalpooling calculates the sum of the elements ofa predefined section. The pool layer usually acts as a bridge between the convolution layer and the FClayer.

3. Fully Connected Layer

A fully connected (FC) layer is used to connect neurons between two layers consisting of weight and flexion with neurons. This layer is usually placed before the output layer and forms the last few layers of the CNN architecture. Here, the input image from the previous layer is merged and sent to the FC layer. Next, the flattened vector ispassed through some FC layers where math function operations typically occur. The classification process starts at this stage.

4. Activation Functions

Finally, the most important parameters inthe CNN model is the activation function. Itis used to learn and estimate all kinds of complex relationships between network variables. Simply put, the model determines what istransferred and what istransferred at the edge of the network. Addsnon-linearity to your network. There are several commonly used activation functions, such as the ReLU, SoftMax, tanH and Sigmoid functions. Each of these features has a specific purpose. For binary classification CNN models,we recommend the sigmoid and SoftMax functions for multipleclassifications. SoftMax is commonly used.

Proposed Work

In the proposed system, the EMNIST dataset is extended by adding characters in otherlanguages.First, the input image is provided, converted to a grayscale image, and normalized toshow the same resolution (28 x 28) as the EMNIST dataset. CNNs are trained using the EMNIST dataset and are usedfor categorizers that can give better results when compared to other machine learning algorithms. The featuresare extracted from the input image and provided ina trained Convolution Neural Network model thatconfirms and provides the desired output. Figure showsan architecturaldiagram of character recognition. Various modules of the proposed system include pre-processing,featureextraction,normalization and classification of minimalMaxScaler fitting images of data.
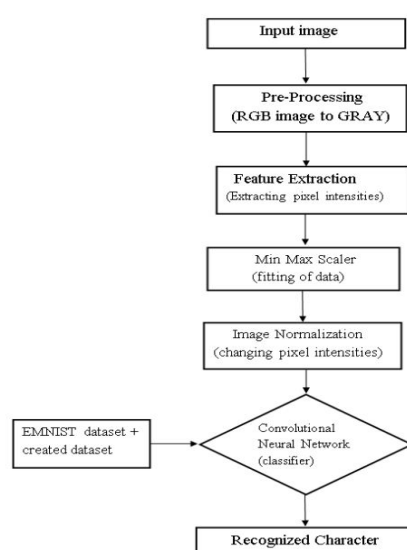
Figure1. Handwritten character Recognition Architecture

PRE-PROCESSING:Pre-processing of the input image is performed by converting given image into a grayscale image. A normal colour image usually consists of a three-channel red channel, a green channel, and a bluechannel, commonly known as RGB. Next, the colour image is converted to a grayscale image consisting of oneblack and white channel to avoid unnecessary noise in the image. Given input imagesare of varyingsizes, accurate predictionsmay be lost when compared to imagesfrom image-trained convolutional neural networks.Therefore, the image is resized so that the resolution of the image matches the resolution of the EMNIST data set and placed as a blank 28 x 28-pixel blank image.

FFEATURE EXTRACTION: Feature extraction is the process of transforming input data into a set of features that can represent the input data well. Feature extraction is related to dimensionality reduction. If the input data is too large to handle, it can be converted to a reduced set of attributes (also known as attribute vectors). It isassumedthat feature selection determines the initial feature subset. The selected function is expected to contain information about the input data, so you can use this reduced representation instead of the full initial data to do what you want. After adjustingthe size of the image, the pixel values are taken in the form of a 1D array representing a value between 255and0, depending on the intensity of the pixels.

IMAGE NORMALIZATION: Normalization is the process ofchangingin the pixel intensity values and it is also called contrast stretching or histogram stretching. By removing the background pixels from this input image, the normalization is left with only the characters in the running image. This can be done using anarbitrary value so that the background pixel isdefinitely less than the color pixel value of the character. In this way, the images are normalizedtobeequal to the values in the EMNIST data set.This image hasa pixel value greater than 0 in the area where the letter'A' is written, and all other areas have a pixel value of 0 after normalization of the image.

CLASSIFICATION:A CNN is used as a classification device to classify handwriting from input images. A CNN consists of an input layer, an output layer, and several hidden layers. The CNN consists of hidden layers as convolutional layer, a pooling layer, a fully connected layer, and a regularization layer. CNN consists of three main components: the convolutional layer, the pooling layer, and the output layer. A commonly used activation function in CNN is ReLU, which isgiven Rectified Linear Unit. The convolutional layer computes the output of the neuron that is connected to the local region of its input and computes the dot product between the small regions connected in the input volume, each weighted. The pooling layer is a nonlinear down sampling form. Maximum pooling is most common one that splits the input image and outputs the maximum for each such subregion. ReLU applies a non-saturating activation function. It improves the determinant function and the nonlinear properties of the network as a whole without affecting the convolutional layer receptors. A rectified linear device has an output of 0 if the input is less than 0, and a raw output otherwise. Its value is
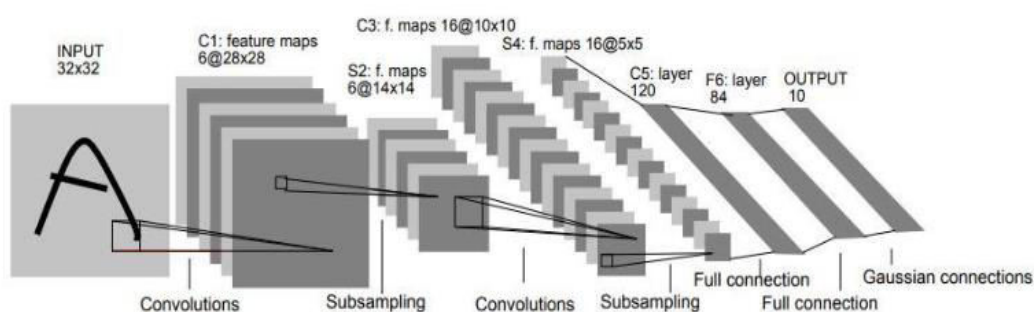

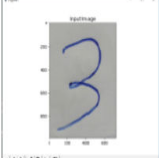
Figure1. CNN input and output layers

calculated based on the following formula. $f(x) = \max(x, 0)$ The SoftMax function is most commonly used in the last layer of neural network-network classifiers. The SoftMax function, like the sigmoid function, reduces the output of each device between 0 and 1. However, split each output so that the sum of the outputs is 1. The output of the SoftMax function is the same as the categorical probability distribution. Therefore, the SoftMax function calculates the event probability distribution for "n" different events.

5 EXPERIMENTAL RESULTS

The input is pre-processed, scaled normalized and fed to the CNN classifier. Because the CNN classifier is modelled on the EMNIST dataset,itcan predict input characters. The database used in our

experimentsconsisted of various characters.

Language	Input image	Output
English Alphabet		The predicted character is A
Number		The predicted character is 3
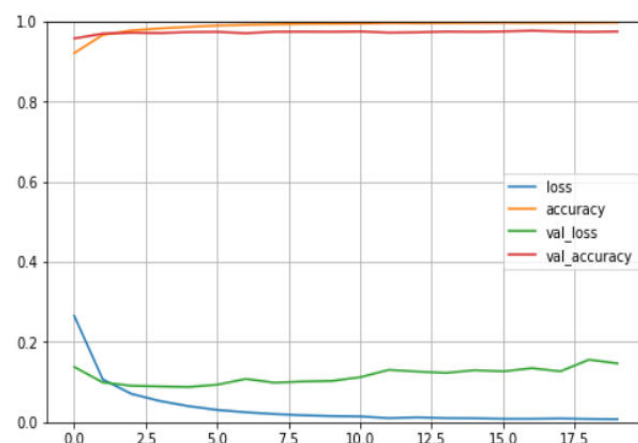
It canbe evenly divided into training , testing and validation sets. Local geometry is defined in the system, such as endpoint cross-points. The membership function extracts the characteristics of the training set and retrieves them using statistics. The project results are listed in Table below andagraphof accuracy is drawn.

6 CONCLUSION

Convolution neural networks consist of multiple blocks such as convolution layers, pool layers, and fully connected layers, which are automatically designed and adaptively learn the spatiallayers of features using backpropagation algorithms. I did.Handwriting recognition has been a daunting task in the last few years. However, with the recent development of the machine learning domain and the huge amount of data generatedindaily life, the image recognition of computer vision has improved significantly.The EMNIST dataset provides approximately 132,000 images of 47 people thatcan be recognized as learning.Convolution neural networkswere used to train EMNIST datasetsfor high accuracy. The input image is pre-processed, standardized,standardized, and providedwith acategorizer to predict characters.

7 REFERENCES

[1] Rohan Vaidya , Darshan Trivedi , Sagar Satra."Handwritten Character Recognition Using Deep-Learning" in Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018) IEEE Xplore Compliant

[2] Wei Lu, ZhijianLi,Bingxue Shi . "Handwritten Digits Recognition with Neural Networks and Fuzzy Logic" in IEEE International Conference on Neural Networks, 1995. Proceedings

[3] B. V. S. Murthy. "Handwriting Recognition Using Supervised Neural Networks" in International Joint Conference on Neural Networks, 1999. IJCNN '99.

[4] J.Pradeep, E.Srinivasan, S.Himavathi. "Neural Network based Handwritten Character Recognition system without feature extraction" in International Conference on Computer, Communication and Electrical Technology ICCCET 2011

[5] Chunpeng Wu, Wei Fan, Yuan He, Jun Sun, Satoshi Naoi." Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network" in 2014 14th International Conference on Frontiers in Handwriting Recognition

[6] Yu Weng1 &Chunlei Xia." A New Deep Learning-Based Handwritten Character Recognition System on Mobile Computing Devices" at https://doi.org/10.1007/s11036-019-01243-5 M

[7] BatuhanBalci, Dan Saadati, Dan Shiferaw." Handwritten Text Recognition using Deep Learning" Through research paper

[8] Mori S (1992) Historical Review of OCR Research and Development. Proc IEEE 80(7):1029–1058

[9] J.Pradeep, E.Srinivasan, S.Himavathi. "Neural Network based Handwritten Character Recognition system without feature extraction" in International Conference on Computer, Communication and Electrical Technology ICCCET 2011.

[10] Rahul R. Palekar , Sushant U. Parab , Dhrumil P. Parikh , Prof. Vijaya N. Kamble. "Real Time License Plate Detection Using OpenCV and Tesseract" in International Conference on Communication and Signal Processing.

[11] "OpenCV" https://en.wikipedia.org/wiki/OpenCV through google

[12] "An open-source machine learning framework for everyone" https://www.tensorflow.org/,[Online] Available: https://www.tensorflow.org/.

[13] "NIST Special Database 19" https://www.nist.gov, [Online]. Available: https://www.nist.gov/srd/nist-special-database-19.

[14]K.Fukushima, "Handwritten Alphanumeric Character Recognition by Neocognitron", IEEE Trans. Neural Networks, V01.2, No.5, May, 1991.

[15]"A-Z Handwritten Alphabets in .csv format" https://www.kaggle.com/sachinpatel21/az-handwritten-alphabets-in-csv-format

[16]Md. Musfiqur Rahman Sazal; SujankumarBiswas;Md.Faijul Amin; Kazuyuki Murase ." Bangla handwritten character recognition using deep belief network" published in 2013 International Conference on Electrical Information and Communication Technology (EICT)

[17] "Artificial Intelligence" Available:https://searchenterpriseai.techtarget.com/definition/AIArtificial-Intelligence.

[18] G. Y. Chen, T. D. Bui and A. Krzyzak, "Contour-Based Handwritten Numeral Recognition Using Multiwavelets and Neural Networks," Pattern Recognition, Vol. 36, No. 7, 2003, pp. 1597-1604. doi:10.1016/S0031-3203(02)00252-2