RESEARCHERID
THOMSON REUTERS

[ICRTCET-2018]

ISSN 2348 – 8034
Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

COMPARATIVE ANALYSIS OF TEXT EXTRACTION USING CLASSICAL AND DEEP LEARNING APPROACH

Geo Johns Antony*1, Karun Varkey[2], Mausumi Goswami[3] & B.S Purkayastha[4]
CHRIST deemed to be University, Assam Central University

ABSTRACT

Digital Image Processing is the only technology which could be practically applied for Classification, Extraction of Features, Projecting Data, Recognition of Patterns and Signal Analysis in a Multi-Scale approach. Automated number plate recognition system is the application of Optical Character Recognition (OCR) by using machine learning algorithms. Machine learning algorithms that have been used for this research are Naive Bayes, K nearest neighbour (KNN), Convolutional Neural Networks (CNN), Support Vector Machine (SVM), Sequential Minimum Optimization (SMO) and Deep Learning. There are other techniques or algorithms which are used by the above mentioned techniques which include: Hidden Markov Models, Pixilation, Neural Networks, Partial Differential Equations, Linear Filtering, Independent Analysis of Components, Anisotropic Diffusion etc. These algorithms and methods will be used to extract the required component (text in this case) from an image.

The various approaches like segmentation of a page, location of the address block, license plate location and indexing of video/audio based on their content are being used. in spite of all these methods, various anomalies like extraction of text from a texture rich or a shaded background, complex or images with low contrast to even slighter variations in simple factors such as orientation, size, font and colour of the image making the process of Text Extraction difficult. The number plate text from the car scene image using the Deep Learning approach (Convolutional Neural Network algorithm) is used. Here, the neural network is trained using sample images from a dataset. The Neural Network is trained to identify number plate images from the image set and then the number plate text is extracted from it.

In this paper comparative analysis of OCR for Automatic Number Plate Recognition and Convolutional Neural Networks approach of Deep Learning is implemented.

Keywords: *Automatic Number Plate Recognition, Optical Character Recognition, Convolutional Neural Networks, deep learning ,image classification.*

I. INTRODUCTION

In the current world information is everywhere, ranging from newspaper, periodicals, pamphlets and pictures to even its digital counterpart in the form of images or videos. The major problem being how to extract the necessary components from these objects with lesser human work. Instead of spending intense man hours into digitizing the entire data, we form an algorithm to develop an application which could digitize this data.

In this project, we will be implementing the classical methodology and the Machine Learning approach. A comparative study between the two methods will also be learnt at the end of the project.

A. Problem Formulation:

Optical Character Recognition (OCR) is the mechanism by which the algorithm converts a handwritten or typewritten text into corresponding machine encoded text even if it's a scanned text, an image or even signs or text from an image. It is having major applications in data collection and analysis (Passport, Aadhar, bank statements etc.). It is the process which aids to digitize paper text so they could be used electronically (helps in digitizing old newspapers) and in turn making it editable.

These applications include digital libraries which is far more convenient to access information pertaining to certain matters rather than searching for data physically in libraries or digging through the old newspaper or periodical collection. The next one will be multimedia systems, which includes storing the vast amount of sounds and videos into digital format and storing them in online databases. Earlier, they were in the form of cassettes, old compact disks used in gramophones or cd players. This could be eased or made easily available by storing inside the computer in digital formats. Information Retrieval Systems is the next sector which could benefit with the introduction of Text Extraction. If the public information pertaining to various issues is made available in the internet rather than spending time at various offices, it will benefit people. GIS (Geographic Information Systems) relates seemingly unrelated data, which helps individuals or an organization as whole to understand the various relationships and the spacial patterns.
So the application made us to develop an algorithm which tries to improve the efficiency of the algorithm by learning the process as it proceeds. Automatic Number Plate Recognition (ANPR) is a technique used to detect the number plate and read the corresponding text. ANPR is implemented using OCR technique. It is used to enhance the characters present in the number plate using a series of image manipulating techniques.

B. Problem Identification:

It is a common occurrence in today's world that vehicles are entering a gated community or a compound without any prior permission or authorization. The gain entry through means like tailgating (unauthorized vehicles follow an authorized user in entering a compound, often they've to open the door for common courtesy, unwittingly putting the tenants, staffs and the building to risk). The other means includes door propping, levering the doors, keys or access card to the community. The biometrics and other type of access machines could be bypassed using this method.

Problems can arise ranging from minor theft to huge crimes. In a community, the major asset is its tenants. Anything that happens to its people by means of an unauthorized entry will be questioned. The security of the compound will be questioned, but it could also happen by means of the above mentioned ways (tailgating, propping, access cards and levering).

In a community, people will be allotted exclusive parking spaces. Here, people will park vehicles in spaces which is not allotted to them. Each tenant is allotted as per flat parking area, he is authorized to park in that particular space. But if a person comes and parks his/her vehicle in that space, in turn causing inconvenience to the tenant.

C. Problem Statement & Objectives
An application will be developed which will help in extracting the number plate text using machine learning methods like CNN and also using the classical methods by using steps like conversion to grey scale, dilation, edge detection and processing.

The problem mentioned for a gated community faces problems like unauthorized entries. Such a problem could be solved by placing a camera at the entrance. The image will be taken and the number plate text will be extracted in text format, the corresponding text will be cross-checked in the database. If the vehicle is unauthorized or not registered, the person won't be granted entry into the compound. This application aims to improve the security in a gated community in turn reducing crimes and unauthorized entries happening there.

The parking lot problem was explained before. This could be prevented by placing a camera in the lot and capturing the number plate, the extracted text would be cross checked in the database. Appropriate messages could be sent to the security officials or the owner if there is a case of unauthorized parking.

D. Limitations
Due to the high end graphical requirements, the Convolutional Neural Networks model was time consuming which in turn required CUDA enabled GPU for parallel processing, was found to be time consuming. There was an increase in time for the extraction of number plate when there was noise in the image (Convolutional Neural

Networks). It was found that accuracy was low for CNN under low light conditions. While classifying the image set, it was found that errors were produced.

There was an ambiguous text during structural component extraction which made it inefficient. Based on this survey, it is addressed to mention the gaps, we would be trying to implement a method which would produce results addressing the gaps mentioned in this work.

The Machine Learning algorithm requires a system which is of higher specifications. The difference between the above mentioned and a laptop with a basic specification is that it will take time to train. The Neural Network requires time to train and this depends on the specification (the higher the better). Since the Neural Network comprises of multiple matrix multiplication, we use GPU instead of a CPU because it has higher number of simpler cores compared to the CPU which has only a few number of simpler cores. The GPU which we were using only had a CUDA rating of 2.1 which is in contrast to the required 3.0 rating GPU, which being the standard for the learning process in Neural Network. There was a time constraint in executing the learning part for larger data since our system was not capable of staying for a stretch of more than 1 week.

The developed program sometimes yielded inaccurate results, it is unable to detect between certain letters and numbers like D and 0.

The primary application can only be applied to a gated community where the number plates can be extracted and to monitor them. Future application is to implement the concept of number plate extraction in surveillance systems which could aid in reducing theft.

II. RELATED WORKS

Optical Character Recognition (OCR) uses various methodologies namely [1] Morphological image processing, feature extraction using Convolution Neural networks and classifiers like Support Vector Machine (SVM). Car plates can be analyzed from different distances, angles using different sample input images. The car plate fed as the input is analyzed and the corresponding location of number plate is obtained and the images are classified into two regions namely number plates and outlier images. Using dilation morphology, the extraction of images by using the height width ratio of the image and the number of connected components in the region is taken as the input. The image feature extractor and the classifiers surveyed were pre trained CNN and SVM. Using the connected components and edge detection, we are able to divide the regions using deep learning methods, in turn obtaining the desired results. The success rate surveyed is 89.7% for 126 given sample images, we will try to extend the sample image set and also improving the success rate and also we will be trying to improve the algorithm in rainy and dark conditions.

Morphological image processing [1] for edge detection using connected component analysis (CCA) technique which does not use machine learning. The edge detection process detects the number plates by its rectangular shape using Hough Transform for the detection of straight lines. The Global Image Features and Rectangular Shapes scans the image which is in binary format and transforms the pixels into various components based on the pixel connectivity (Area and aspect ratio is also included). Texture also matters since the grey scale level of the numbers and the background plate also differs due to the characters present. Also, in some cases the number plates will be coloured which also benefits in easy identification of the plates.

Neural networks is used as a filter for analyzing colour and texture properties of the number plates. The sub components in this part will be Character Recognition and Plate Image Detection. Separated Character Patterns are recognized using a trainable recognition engine. The Plate Image Detection is a 7 layered process involving CNN technique for direct recognition of pixel image which is manually clipped by training it.

ISSN 2348 – 8034
Impact Factor- 5.070

III. NUMBER PLATE EXTRACTION METHODS

A. Classical Approach

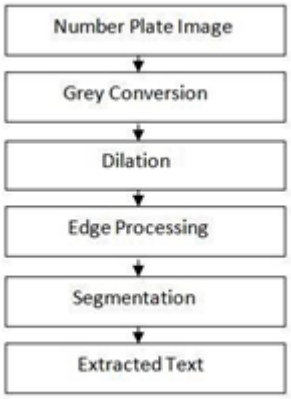

Fig1. A flowchart to represent the image extraction technique using Morphological Image Processing method.

The process classical approach gives a methodology in which the input image is fed into the system and the system follows the morphological image processing method and the result obtained will be a segmented image wherein only the number plate will be segmented out and shown. After obtaining the segmented image, the method applies a simple OCR detection technique to extract the number plate content

The morphological image process comprises of:

1.) *Gray scale conversion*
2.) *Dilating*:
3.) *Edge Detection*
4.) *Segmentation:*

The approach initially takes the input as a normal RGB (Red Green Blue) image which is in turn converted into gray scale image. The gray scale conversion reduces the complexity for further image processing, since the image dimension is changed from RGB to BW (Black White).

The next process is to dilate the image. The image is dilated which adds pixels to the boundaries of an object in an image. It is helpful for Edge detection which is happening in the next step. It darkens the Edges.

The following step is Edge Detection. In this step, the Edges in the figure are detected. The edges are detected using a Sobel filter. The filter takes the difference in intensity between two neighboring pixels and detects the edge if there is a high difference in the intensity level of the two. The neighboring pixels density will be high if an Edge is present. So this method is applied onto the image and the initial segmentation for the image is done. Values are plotted onto a Histogram for further process. Histogram for Intensity level is plotted for both vertical and horizontal values separately

The next step is the final segmentation of the image. In this process, the image after edge detection is taken and based on the Histogram plotted for Intensity level the final segmentation for the number plate is done. A specific threshold is applied for the histogram Intensity plotted and the region obtained is likely to be the Number Plate region

The Image obtained from the previous step is the image of the number plate alone. This image is applied through a character recognition algorithm and the final end result of the number plate text is obtained as the output

B. Deep Learing Approach

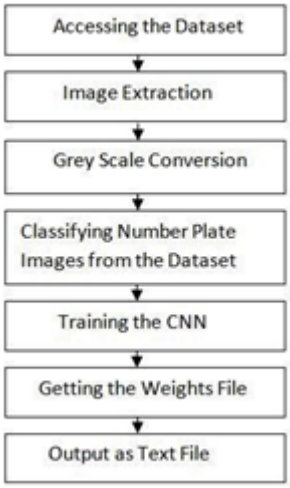

Fig2. A flowchart to represent the image extraction technique using Deep Learning method.

The flowchart drawn above shows how the text is extracted using the Machine Learning approach. Here we have completed the process of extraction in a series of 7 steps:
1.) Accessing the Dataset
2.) Image Extraction from the Dataset
3.) Grey Scale Conversion
4.) Classifying the Number Plate Images from the Dataset
5.) Training the CNN
6.) Getting the weights file
7.) Output as Text File

In implementing step 1 and 2, we will be need to access a database. For our execution, we have used Sun Database. It is a database comprising of 131067 image scenes spanned across 908 categories of scenes, 313884 segmented objects and 4479 categories of objects. It is a database which uses WordNet English dictionary to build the core portion of the dataset, they have counted the entries that corresponds to various scene of images, places and different terrain or environments. We have used Scene Recognition Benchmark in Sun Database as our Dataset (approximately 37 GB). The image with format JPEG/JPG is extracted. Here, the images are extracted based on their size and the image is cropped into 256 X 256.

The next step focuses on converting the image from RGB into two-dimensional Grey Scale format. It is usually done by a simple algorithm where you have to add the r, g and b components of the image and then dividing it by three to get our desired grey scale image. But this method yields black coloured images instead of grey. So, they have used weighted method (luminosity method) to accomplish this the advantage of using this method is that it reduces the processing time and complexity of the program.

The next step, we use a Sobel Filter. This filter classifies images with number plate and the text is taken and given to the CNN as input. This is done using Open CV and Numpi in Python.

The next part is defining the model, there are five layers for CNN. It includes convolutions, pooling and hiding after which we will get the trained program as output. Here, the input image is going through a conv1 process, there

images with similar features are stacked together. The next part is pool1, in this part we will be seeing a reduced stack of images by removing images without a common similarity. This followed by conv2 and pool2, where more images are added to the stack to train the CNN and a resulting pool is formed. Pooling is done using the Max Pooling algorithm. The final pool is then hidden and their properties are determined based on which the program will be executing or how the program is trained to detect a number plate image from a set of images and to extract the text from the number plate.

In the training part, CNN uses Tensor Flow, It connects each node and detects the system based on the model designed. Here, we get a weight file as the output. In the detecting part, it converts an input image to 128X64 bit grey scale images. It detects the number plate of any Input car image and it also uses the weight file obtained from the training to get the text.

IV. GAP ANALYSIS

Literature collection and segregation were taken for Literature Survey and the gap analysis were found for the following papers.

In the paper "Extraction of Number Plate Images Based on Image Category Classification Using Deep Learning", the author used convolution neural network to extract the features. However, the gap analysis found in this paper was that the image should be bright and distinct. The algorithm performed poorly under dimly lit and rainy conditions. The algorithm gave an accuracy of 89.7 percent despite character segmentation being difficult.

K- Nearest Neighbor pattern recognition was used in the paper "Optical Character

Recognition using KNN on Custom Image Dataset" by Tapan Kumar Hazra, Dhirendra Pratap Singh and Nikunj Daga. The gap analyzed in the paper was the high computational cost associated with it. The algorithm was deemed inefficient as the distance increases and there were many classification errors.

In the paper titled "Segmentation of highly unstructured handwritten documents using a neural network technique" by Rathin Radhakrishnan Nair, Bharagava Urala Kota, Ifeoma Nwogu and Venu Govindaraju. The algorithm implemented is Convolutional Neural Network. However, it was found that the algorithm is time consuming and errors were present during some test cases.

In the paper titled, "Entry and exit monitoring using license plate recognition" by A. Yovan Felix, A. Jesudoss and J. Albert Mayan. They have adopted the methodology of adaptive equalization histogram method for several histograms, each corresponding to each individual section of the image for improving the local contrast and to enhance the definitions of edges in each region of the image. The next one is contour method which is a fast pixel following method for image sensing, distinguishing objects from the background. The contour pixels are widely used for smart devices due to its simplicity and its ability in detecting objects. The concept of Deep Neural Networks is also implemented due to its high end performance on problems that outsmarts other solutions in multiple domains. However, the algorithm is limited to gated communities due to its high processing time and parallel processing could not be implemented, which resulted in increase in processing time, the risks associated in running the processor for longer periods.

"Optical character recognition using artificial neural network" by T.K. Das, Asis Kumar Tripathy, Alekha Kumar Mishra, it is found that artificial neural network and edge detection is used in obtaining the results. The neural networks offers less formal statistical training, it has this unique ability to find the non-linear relationship between two types of variables (dependent and independent variables). It has the ability to generalize after obtaining data and the relationships which binds these data together. The edge detection process involves using the Sobel Filter which finds the discontinuities in 1-d signals helps us to obtain a clearer data. However, the following algorithm was inefficient in detecting hand written document.

ISSN 2348 – 8034
Impact Factor- 5.070

In the research paper, "Improving optical character recognition performance for low quality images" by Matteo Brisinello, Ratko Grbic, Matija Pul and Tihomir Andelic. They've used the method Tesseract, the advantage of using tesseract is that by implementing proper pre-processing methods like removing the unwanted lines, applying filters based on the image noises and proper Gaussian filters to smoothen the characters, it will yield high results. However, it cannot recognize the text from the image if the background is colorful.

The following paper uses Structural Components Lee Algorithm. The paper titled, "Algorithm for optical handwritten characters recognition based on structural components extraction" by P. A. Khaustov, V. G. Spitsyn and E. I. Maksimova. The algorithm is used in character processing applications. The algorithm was found inefficient if there is an ambiguity in the text.

In the paper titled, "Designing mobile application for retrieving book information using optical character recognition" by Nana Ramadijanti, Achmad Basuki and G. J. W. Agrippina. They are used the technology of android OCR. Optical Character Recognition (OCR) uses various methodologies namely [1] Morphological image processing, feature extraction using Convolution Neural networks and classifiers like Support Vector Machine (SVM). Car plates can be analyzed from different distances, angles using different sample input images. The gap found in this method is that if there is a noise in the image or if the lighting conditions is bad, then we will obtain undesired or inefficient results.

V. RESULTS & OBSERVATION

We have obtained results for the proposed method.

1) Car Scene Image

2) Conversion to Gray Scale:

ISSN 2348 – 8034
Impact Factor- 5.070

3) The Image  *after dilation process*

4) Edge Processing

5) Segmentation

1)Image input and Conversion to Gray Scale

2)The input Image and later converted to Gray scale

3)The output Image and the number plate output in text format

Here, we have done a comparative analysis of the two different methods which we have adopted. The Classical and the Machine Learning approach.

The Classical approach has smaller advantages over Learning method in some aspects. This method could be implemented in any system irrespective of its specifications, the program is a compact one with few lines of codes. The Classical Method was implemented in Mat lab R2016a, this software can be run in a system with Pentium processor. There is no need for the program to learn or like Neural Networks, there is no learning time for Morphological Image Processing method. This could be helpful for a user who is in need of immediate results.

The Machine Learning approach owing to its time taken for the learning process and the higher CPU requirements produce better and more accurate results among the two methods. This could be due to the training part, the program developed can easily extract the number plate from the input image. In some cases, the algorithm developed using Classical Approach is inefficient. It(Classical Approach) is unable to separate the number plate from the input image, but in Machine Learning such cases are rare or absent to be precise. We have analysed some results for both methods and plotted the same as:

A. Classical Approach

No.of Samples	Accuracy	Correct Classificaion	Incorrect Classification
10	100%	10	0
20	95%	19	1
30	93%	28	2
40	95%	38	2
50	94%	47	3
60	93%	56	4
70	94%	66	4
80	94%	75	5
90	93%	84	6
100	95%	95	5

FIGURE 3.1: Table Classical Method

ISSN 2348 – 8034
Impact Factor- 5.070

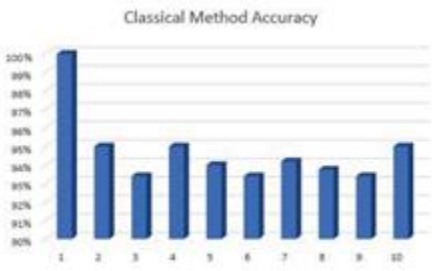

FIGURE 3.2: Bar Chart Classical Method

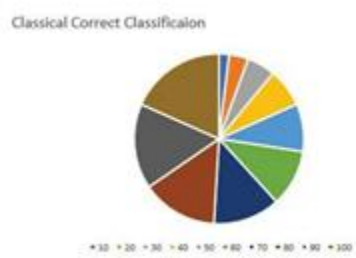

FIGURE 3.3: Pie Chart Classical Method the Machine

B. Deep Learning Approach

No.of Samples	Accuracy	Correct Classificaion	Incorrect Classification
10	100%	10	0
20	95%	19	1
30	97%	29	1
40	100%	40	0
50	96%	48	2
60	98%	59	1
70	99%	69	1
80	96%	77	3
90	99%	89	1
100	97%	97	3

FIGURE 4.1: Table Machine Learning Method

FIGURE 4.2: Bar Chart Machine Learning Method

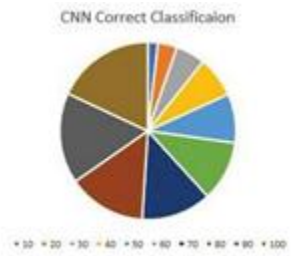

FIGURE 4.3: Pie Chart Machine Learning Method

VI. CONCLUSION

In this paper, we have analysed different techniques of OCR (Optical Character Recognition) and described the process of Convolutional Neural Networks and Classical approach in the extraction of number plates. Techniques like deep learning, Convolutional Neural Networks, Support Vector Machine, Machine Learning, Sequential Minimum Optimization, K Nearest Neighbour and Naive Bayes were surveyed and the gaps were analysed. Apart from these techniques, algorithms like Hidden Markov Model, Viterbi Algorithm, Edge Detection, Contour Methods, Adaptive Equalization, and Structural Components Lee Algorithm were studied and how they contributed to each technique. Due to the high end graphical requirements, the Convolutional Neural Networks model was time consuming which in turn required CUDA enabled GPU for parallel processing were found to be time consuming. There was an increase in time for the extraction of number plate when there was noise in the image (Convolutional Neural Networks). It was found that accuracy was low for CNN under low light conditions. While classifying the image set, it was found that errors were produced. There was an ambiguous text during structural component extraction which made it inefficient [10]. Based on this survey, it is addressed to mention the gaps, we would be trying to implement a method which would produce results addressing the gaps mentioned in this work. The current project is restricted to gated communities, we will be trying to extend our work into real-time surveillance systems like toll booths, traffic signals or interceptors. There will be focus on trying to implement this feature in live stream, currently we are able to process only snapshots or images of a car.

The current algorithm does not work under or does not meet the required accuracy in rainy or dark conditions. This is a limitation for our project, since a gated community must have a working program irrespective of the conditions

REFERENCES

1. *Yoshihiro Shima, "Extraction of Number Plate Images Based on Image Category Classification Using Deep Learning", 2016 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS2016) 17-20 December 2016, Tokyo Japan, 2016.*
2. *Tapan Kumar Hazra, Dhirendra Pratap Singh, Nikunj Daga, "Optical Character Recognition using KNN on Custom Image Dataset", Industrial Automation and Electromechanical Engineering Conference (IEMECON), 2017 8th Annual*
3. *Vanita Jain, Arun Dubey, Amit Gupta, Sanchit Sharma, "Comparative Analysis of Machine learning algorithms in OCR", 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom),2016.*
4. *Pranav P. Nair, Ajay James, C. Saravanan,"Malayalam handwritten charact er recognition using convolutional neur alnetwork", 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), 2017.*
5. *Rathin Radhakrishnan Nair, Bharagava Urala Kota, Ifeoma Nwogu, VenuGovindaraju,"Segmentation of highly unstructured handwritten documents using a neural network technique", 2016 23rd International Conference on Pattern Recognition (ICPR), 2016*
6. *A. Yovan Felix, A. Jesudoss, J. Albert Mayan, "Entry and exit monitoring using license plate recognition", 2017 IEEE International Conference on Smart Technologies and Management fo Computing, Communication, Controls, Energy and Materials (ICSTM), 2017.*
7. *Orhan Bulan, Vladimir Kozitsky, Palghat Ramesh, Matthew Shreve, "Segmentation- and AnnotationFree License Plate Recognition With DeepLocalization and Failure Identification", IEEE Transactions on Intelligent Transportation Systems, 2017.*
8. *T. . K. Das, Asis Kumar Tripathy, Alekha Kumar Mishra, "Optical character recognition using artificial neural network", 2017 International Conference on Computer Communication and Informatics (ICCCI), 2017.*
9. *Matteo Brisinello; Ratko Grbiü; Matija Pul; Tihomir Anÿeliü,"Improving optical character recognition performance for low qualityimages", 2017 International Symposium ELMAR, 2017.*
10. *P. A. Khaustov, V. G. Spitsyn, E. I. Maksimova,"Algorithm for optical handwritten characters recognition based on structural components extraction", 2016 11th International Forum on Strategic*

ISSN 2348 – 8034
Impact Factor- 5.070

11. *Nana Ramadijanti, Achmad Basuki, G. J. W. Agrippina, "Designing mobile application for retrieving book information using opticalcharacter recognition", 2016 International Conference on Knowledge Creation and Intelligent Computing (KCIC), 2016.*
12. *Anisha Priya, Surbhi Mishra, Saloni Raj, Sudarshan Mandal and Sujoy Datta, "Online and Offline Character Recognition: A Survey", International Conference on Communication and Signal Processing, April 6-8, 2016, India,2016.*