

Comprehensive Survey on Machine Learning Application for Handwriting Recognition

Yogesh Kumar Das^{1,*}, Pulkit Jain¹, Sree Kumar K¹

¹*Department of Computer Science Engineering, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur, Kancheepuram District, Tamil Nadu- 603203, India.*

**Corresponding Author*

Abstract

Handwriting is an integral means of information propagation and dissemination in everyday life even with the advent of better methodologies. Due to the multitude in the variation of human transactions, handwriting being recognized by machines has impact in everyday life, as in reading text that has been handwritten in any personal digital assisting devices, in doctor's prescription, in money transactions done through cheques, in signing text fields in web-based applications, etc. This paper basically describes language which is written down and its nature, and the manner it is transcribed as digital information, and the rudimentary theories supporting algorithms which recognize written language. It aims to elucidate various techniques inclusive of the incident which concerns the convenience of direction of information placement throughout the process of writing, cases which operates on scanned pictures (Optical Character recognition) are considered. Algorithms for analyzing, handwritten text recognition, and output dependability with systems with real-world solicitations are specified. Alternative avenues of application, viz. verifying signs, authenticating the author of the particular text, and tools which facilitate text written by hand's learning are also taken into consideration. This technique employs the usage of select classifiers to train the data in order to obtain sustainable and accurate output.

Keywords: on-line, handwriting recognition, off-line; pre-processing, post-processing, segmentation, classification, feature extraction.

INTRODUCTION

Handwriting recognition involves a procedure in which written text is transformed into computer understandable and programmable format. The aim behind the applications which recognize text messages which are written, system is to contrive a convenient application for symbol exemplification that allows positive mining of text from written papers and to digitize and decode the written text message into computer understandable text. Recognition of written documents corresponding to additional smart applications delivers some services in cellular phones, robotics applications and other devices for documentation and corroboration purposes. Active domain research in written documents comprises recognition of both types: online as well as offline, corroboration of signs, postal services related interpretation and cheques at bank processing. Perplexing difficulties in online written corroboration are diverse symbol magnitudes, altered writing

processes and replacement pixels fashioned by hesitation in writing or interposed non-adjacent pixels caused by quick writing.

Material and Methods: The research is based on extensive studies done on pre-existing research papers.

RUDIMENTARY EXPLANATION OF RECOGNITION, INTERPRETATION AND IDENTIFICATION: Processes and methodologies of various varieties can be concomitant with handwriting. Basically, handwriting comprises of annotations made on any empty space medium and through handwriting recognition, it is converted to its corresponding graphical representation. For English vocabulary, the representational depiction is characteristically the 8-bit American Standard Code for Information Interchange sign of characters 16-bit Unicode is now essentially used for the depiction for the representation of the majority of the written language. In the process of documentation of handwritten characters, it is possible to pin-point the author of a certain combination of text on comparison with a pre-existing database which is obtainable from data collection of various individuals with the assumption that each handwriting is distinct. Another technique associated with this method is verification of signature and confirm whether it belongs to the particular individual or not. Credentials and certification are methods that define the distinctive nature of the writing of a unambiguous author, which have solicitations in forensic scrutiny while comprehension of handwriting and analysis are progressions whose aims are to screen out the disparities so as to decide on the appropriateness of the written text.

The ability to comprehend written messages is a unique skill which is integrated into human understanding. The domain expertise is obtained by the accumulation of knowledge. The desired amount of data required for the on-line and the offline cases varies vastly. For online cases, a couple of hundred bytes, classically tested at 100 trials per second which is relatively equivalent to a couple of hundred bytes is required in the case of online recognition of written text and in the off-line case nearly close to a megabyte of data. Considering the larger picture, text messages in general are stored in paper format, this message is essentially analogous in nature, and through the method of digitization and scanning can be transformed into digital form. The resultant output's accuracy hinges on the smallest independent character which is distinguishable in a reliable manner. The range of possible graphical representation and the medium of transmission and

the mannerism of storage also plays a major role in the efficiency of the output. On the comparison of the offline an online handwriting recognition methodology it is concluded that the online method is much more reliable and efficient. However, any consequential acknowledgement of the written text hinges directly on the choice of the matching dataset.

DIVERSIFICATION IN HANDWRITING RECOGNITION:

The digitized image thus obtained is converted to be represented in binary form, in off-line case, the remaining methodologies for classification can be indistinguishable but for two major variances. When the author completes writing the text message the process of offline recognition is initiated and the digitized picture is pre-processed. In the next step, there is no spatial data related with the image which prevents the classifier with any premise in order to find the nature of handwriting. Thus, the domain expertise of the application is limited. So, it can be implied that the algorithm for offline recognition of handwriting only comes into play, once the entire process writing down the message is completed. On the flipside when it comes recognition using on-line recognition the spatial data is also stored hence dynamic information is transmitted which improves the accuracy. The major advantage of online recognition over off-line recognition of handwriting is that dynamic transmission of data takes place and error corrections take place simultaneously. In this process many adaptations are possible while a particular character is being written. In case the handwritten text is not properly recognized by the algorithm then the author has the facility to rewrite the message till it's correctly and accurately recognized. Thus, it is possible to acclimatise the machine. To facilitate the previously mentioned process, recognizers are used which store samples from a large collection of users and help with accurate recognition. This causes the simultaneous adaptation to take place in between the user and the application and vice-versa. For the comparison in-between offline and keyboard-oriented input to the application taking the domain acumen of handwriting recognition, using an electronic pen. To boot, online handwriting corroboration develops the work-flow, the datum is instantaneously obtainable. On the flipside, the convenient method in writing any message successfully decreases effort at the inception point of using such applications for everyday use.

1. **Off-line:** In this method of text recognition, the picture is taken from a scanning device (scanned images of the text written entities), photographic diaphragm or additional ordinal feedback peripherals. The digital pictures are converted into its binary equivalent through inception point method founded on colour arrangement ranging from colour to gray scale, so that the binary pixels store the value 1 or 0.
2. **On-line:** For this method of text recognition, the present information is extracted forth to the application and the process of acknowledgement of the text message is carried out simultaneously. Basically, the input is in the form of a concatenation

of characters having (x, y) coordinate points to an electronic feedback device moving on a receptive device.

3. **Writer Independent and Writer Dependent:** This system identifies with a wide bandwidth of conceivable writing methodologies, whereas an author dependent recognition application is competent enough to recognize solitarily from a particular set of authors. Therefore, this application system executes on information with a reduced inconsistency and thus there exists a fortuitous option of having advanced consistency to be attained in contrast to author sovereign recognition application.

ONLINE BASED HANDWRITING RECOGNITION:

In the process of message recognition using on-line algorithm, involves altering symbolic representation in its temporal form to the writing process. Any system which is utilized for recognition involves the training stage, the fitting stage, the testing stage and finally the prediction stage. In the initial phase, information is developed through pre-processing, mined their significant structures through feature extraction, conjoining comparable symbols through the process of clustering and these characteristics are stored as training samples in process referred to as template management. In the following phase, prediction function is employed to facilitate the identification of comparable structures with test features in the process of classification. All the processes involved with recognition of handwriting using online system are described below:

1. **Pre-processing:** Information obtained from input receptive sources are often not accurate and it is difficult to carry out classification due to presence of distortions and white noise. For the process of analysing to initiate refining of information takes places under the technique of pre-processing. The process of pre-processing involves a varied combination of the following techniques: Sampling, noise elimination, discretization, integration, and transformation. Goal oriented approach and the variety in the nature of scripting helps in the evolution of different system which employs different techniques.
2. **Sampling:** Using the electronic pen as input device, each stroke up and down trajectories is stored as information procurement. These pen-up and pen-down strokes which are used to move the electronic pen over the receptive surface and this data is stored as a concatenation of characters storing the coordinates as a function of time. This process is employed to keep a track on the pen strokes and their depth characteristics. The span of the trajectories and the number of coordinates in the concatenation of characters varies even when authors transcribe with the equivalent rapidity. Re-sampling is conducted in

order to remove any discrepancies associated with white noise.

3. **Noise Elimination:** There are many sources of noise which cause distortion in the strokes of the transcribed character. Finger movement leading to instability during transcribing and electronically converting fault of the feedback systems contribute as the core causes. Such noise discrepancy which seeps into process of information acquisition does not invariably affect the text message in the off-line recognition process, but it has dire consequences when it comes to on-line recognition technique. Filters of different kinds are used to limit the impact of such types of noises.
4. **Normalization:** For the process of recognition to take place it is essential to have a small training set which can be easily tested, and also the characteristics of the text be legible. But this idealistic writing style is not possible in case of every sample set. Thus, normalization becomes a quintessential process which has to be concluded before the process of feature extraction takes place.

Translation, rotation and scaling, etc. are some of the basic techniques employed for the process of normalization. Various altering techniques for middle line, top line or base line are employed to adjust any slant in the written text which might lead to distortion. It is more problematic to alter the alignment of the written character in comparison to a word processing.

5. **Repetition Removal:** Since the impact of the digitizer is quite sensitive that even when the electronic pen is not touching the receptive input screen, it can leave a mark. This causes the subsequent manifestation of points through coordinates. Another issue which occurs is that while writing characters there are corners and curves and it is possible to be re-written over the same set of coordinates referred to as dominant. These redundant coordinates are to be detached by appropriate techniques for improved performance of the application such as, recovering superior pictures of the symbol and augmenting the rapidity by decreasing the span of the character.

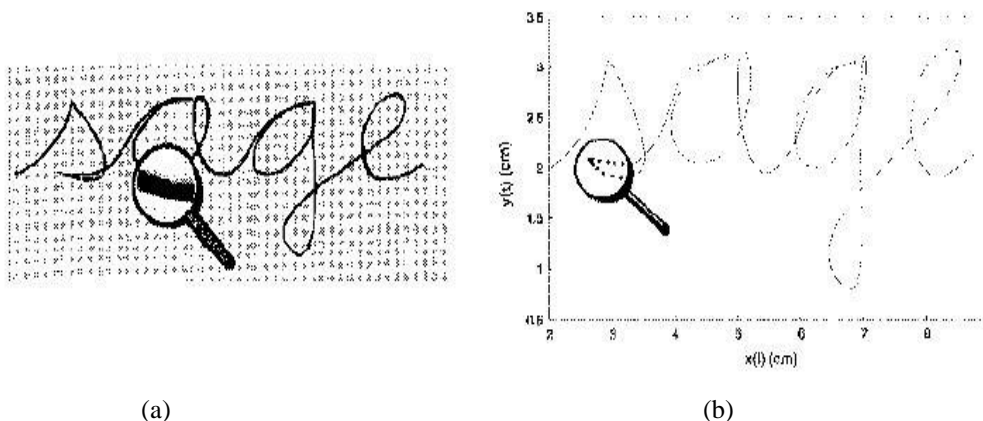


Figure 1. (a) Diagram for offline method. The appearance of the expression is transformed into grayscale pixels using a device for scan ning
 (b)Diagram for online method. The Cartesian coordinates of the electronic pen tip are logged as a function of time with a digitizing device

Applications: Handwriting character recognition has been developed and evolved through extensive exercises in academic domain utilizing practical applications. Majority of the focus is laid on three of the academic territory in the subsequent segment: pen-based computers, signature verifiers, and development tools. Pen-based applications employs recognition using the pen strokes as a central characteristics and commands for gesture controls across multiple platforms. The next application contracts with signatures and works on the principle that every author has a distinct signature and cannot be easily duplicated. The final application of this type encapsulates the course integrates numerous structures that utilize the neural based device's features of writing to enterprise applications for edification and recuperation commitments.

1. **Pen-Based Computers:** In the 1960s, this method was initially researched by Alan Kay. From that point onwards, numerous study groups have been occupied on the execution of this notion, trying to assimilate into something simplistic and practical applications a crystal clear spatial coordinate-sensing contraption with a tangible depiction, under the mechanism of a dominant micro-computer. The definitive objective in this case is to imitate and cover pen and paper representation by the reflexive dispensation of digital ink.
2. **Signature Verifiers:** This is a precise course of reflexive handwriting dispensation, the evaluation of a sign used for testing against a limited orientation cases which are composed as an author registers in the application. Based on this system application, all

the attributes associated with the author's data and other relevant information from the signature thus obtained, without taking into consideration other components. This data has to be virtually time independent and effectually discriminant. Signature verifying application indulges mainly to utilize the particular, private and individualistic symbol of the writing. There is a twin difficulty when it comes to employing signature verifiers. The primary issue is to validate the text that has been signed resembles to the exclusive individualities of the author, without caring for the contents. A false positive in this circumstance is the denunciation of a dependable sign, is what leads to an error referred to as I. Next task is additionally challenging in comparison to the previously mentioned one and comprises circumventing the recognition of counterfeits is trustworthy. The forbearance stages involving error of type II, techniques in which corroboration of signature is essential is lesser than the tolerable recognition of handwriting, for both types of inaccuracies. The bulk of the signature verification applications graft with a miscalculation boundary of less than 5 percent common amongst the two types of faults. Any declining levels of a particular category of error inexorably intensifies the next one. The appraisal using algorithms of corroboration of signature, as for numerous configuration recognition difficulties, increases numerous complications, making any impartial assessment between dissimilar approaches rather subtle. On the flipside, it also differs from system application to the next hinging on the individualistic properties of each person.

3. **Developmental Tools:** In conjugation with the development in the vast domain of signature verifiers as well as recognizers, progress is also being made to develop direct or indirect methodologies for processing in the automatic recognition of handwriting. The awareness of expending a system application to educate handwriting has caused numerous studies about the ergonomic facet of the application and how it is made simplistic for children, but also provides a pleasurable atmosphere for handwriting. The domain for academic apparatuses founded on handwriting is anticipated to occur in the approaching time period of research. Such learning tools to help children exercise writing related activities and it is supposed to not only locate their place in an academic setting, but also in the turf of geriatrics for the reintegration of the disabled and the aged for their improved mechanism for locomotive exercises.

OFFLINE BASED HANDWRITING RECOGNITION:

When the method of recognition of handwriting using offline technique is used, it takes into account the electronic picture or handwritten message. This system application plays a quintessential role in bridging the gap in-between man to machine communication and this process supports automatic

analysis of documents that have been handwritten. The previously mentioned technique is essentially a subcategory of the optical character recognition, whose dominion varies from computer print to handwriting but it is extensively used for machine-print. The acknowledgement of the fact that the English handwriting presents unique challenges and benefits, it is essential that the approach should also be used in the recognition of texts from varied scripts. The paper elucidates the nature of development work in this field. This domain is generally classified into two broad categories namely online and offline.

In case of offline method, the recognition process usually takes into account tangible sources of input like paper documents. Recognition of handwritten text has any hurdles and the most significant of them being unreliability due to human interaction which affects the accuracy levels in output and also dependent on the type of script being used. It is considered offline method if it applies to formerly written message, such as any pictures perused by a suitable peripheral. The issue of online process is typically more convenient than the offline technique as supplementary data is accessible. Regular database is vital for the research of recognition text messages which is written. Digital catalogue enables with the progress, appraisal, and assessment of myriad written acknowledgement algorithms. Numerous databases have been established in the symbol corroboration domain fluctuating from published, handwritten, remote or other types, and the different categories of scripts.

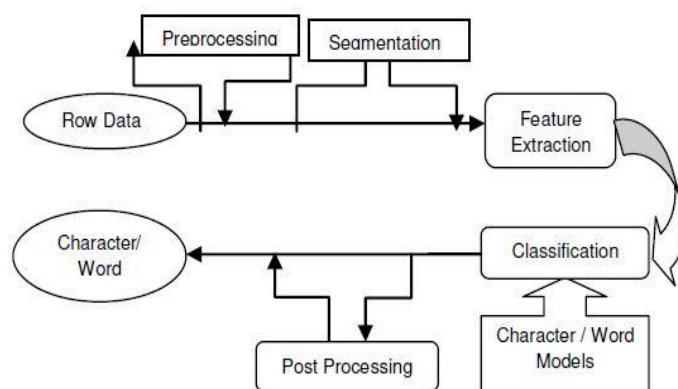


Figure 2. Optical Character Recognition working process

The various techniques required for recognition of written messages can be broken down into the following steps:

1. **Pre-Processing:** This initial phase is an assemblage of techniques that is applied to the concurrent alterations made to the obtained image. The process utilizes an unfiltered picture and augments it by limiting discrepancy and noise related sources, and henceforth streamlines classification, segmentation, feature extraction, prediction and subsequently recognition. But to nature of the text input is also dependent on various causes.
 - a. *Document history:* If any text file gets to be copied multiple times and various imitations

are available then the originality of the document is lost. The width of the strokes of the written character vary, different sources of noise is created, and divergence reduces.

- b. **Printing process:** An electronically typed file lesser distortion compared to a file that is typewritten, but even the typewritten document is more distinct than that printer which uses the dot-matrix mechanism's output. Additional distortions which link to the method for printing are listed below.
 - c. **Font clarity:** Different styles, different sizes, italic or bold characters, subscripts or superscripts, and using collective font sizes and styles which are complicated for recognition.
 - d. **Paper quality:** *Flat*, uniform grain and heavyweight paper is clear to read than lighter weight, low-quality paper.
 - e. **Document condition:** It becomes hard to read if there is existence of unnecessary marking and discolouring.
 - f. **Image acquisition:** The digital imaging by low quality and sampling rate, and often make distortions like zigzags. By positioning difference, pale, blur document, and the low resolution the description of scanned text can be adjusted.
2. **Segmentation:** It occurs in a page image and forms various logical parts, like lines of a paragraph, text from graphics, and characters. After the pre-processing stage, most of the OCR systems confine to every single character before recognizing them. Segmentation of page can be done in different ways: word decomposition and page segmentation, when functioning with pages that have different object types like mathematical formulas, graphics, text blocks and labels, page decomposition divide the various page elements, generating text blocks, sub-words and lines. While page division might recognise group of the logical structure of a page, word separation differentiates the characters of a substring.
 3. **Feature Extraction:** This stage identifies the segments of text and picks the group of features which can help to analyse text segmentation and they are proceed in a form suitable for recognition phase. After getting the isolated pattern by OCR system then it extracts the important features of the pattern and sends to the classifier. It is the important and challenging problem of pattern recognition. The preferred features should be a limited set whose values able to differentiate between arrangements of various classes but are similar for patterns within the same class. The feature extraction step matches with classification due to the type of features extracted

here must match with the expected classifier. The interleaved control versus one-step control is the main control access for feature extraction and classification.

4. **Classification:** This is the decision-making stage for an OCR system. To analyse the text segment the classification makes use of feature extraction. This stage classifies the test data by use of feature models obtain in training modelling stage. Classification extract features from a pattern which are compared to those of the model set. The classification among the features can identify the pattern of a certain class. The classification gives a set of possible solutions instead of generating a unique solution when pattern is classified. The post-processing stage select the correct solution with use high-level information. In classical way classification followed two different paradigms: syntactic and statistical classification.
5. **Post Processing:** This stage is the last stage, in which improvement is done in recognition by filtering the choice taken by the classification stage and using context the word get recognizes. It is mainly gives the better result and is used as a group of approach that depends on context information and character frequencies. To increase word recognition rate is the main function of post-processing. The post-processing uses that higher-level information to analyse the precision of the solutions which is not available to the classifier as it sometimes gives set of feasible solutions instead of a unique solution. Spell checking, and correction are most common post-processing operations.

RECOGNIZING OFFLINE HANDWRITTEN ENGLISH WORDS BY DIFFERENT METHODS: In the research, *Horst Bunke, Jean-Ce'dric Chappelier and Matthias Zimmermann* proposed a consecutive coupling of a Markov Model which identifies for offline handwritten which is to be in English with an uncertain bottom up chart parser using grammars which is stochastic context-free obtain from a text corpus. After doing many experimentations, the result shows that recognition rates can significantly improve by syntax analysis. There were two different recognition tasks which were defined in experimental Setup and System Optimization one is the Multiwriter Task and the other was Writer Independent Task. In the Multiwriter Task by a large set of known writers the recognizer is trained on handwritten texts. While in the Writer Independent Task the way of writing is hidden which means the writers defined in the training set are not characterized in either the validation or the test set of this task. This research was backed by the Swiss National Science Foundation NCCR program "Interactive Multimodal Information Management" in the individual Project "Scene Analysis." the footnotes effortlessly and place them at the lower end of the column on the page on which they are mentioned. Use single-spaced, Times 8-point type. Readers could get benefit by avoiding the use of footnotes taken

together and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence).

A system was developed by *Samy Bengio, Alessandro Vinciarelli, and Horst Bunke* for the offline recognition for large vocabularies with no limits in handwritten texts. In this one assumption was taken that the data is written in English. This results in improvement in performance of our system by the application of Statistical Language Models. Various experiments were carried out using single and multiple writer data. Dictionary of variable size have been used. Which gives the better result of the system using language models (when the Dictionary contains 50,000 words, the error rate is decreased by some 50 percent for single writer data and by some 25 percent for multiple writer data). To properly deal with unconstrained text recognition an experimental setup was proposed.

In the research which described a method that can recognise a handwritten literal amount on different bank checks was given by *Courtellemont P., Paquet T. Lecourtier Y, Oliever C. and Avila M.,*. In this the pre-processing of the original 256 gray-levels image was presented, containing in identical background, and the locating of the handwritten information. The Markovian approach was chosen, which is first applied to the sequences of words, for recognition of the amount. The first results bring an improvement in the recognition rate which was allowed by considering an extension of the method to the sequences of graphemes' in words.

In the research proposed by *Rami Al-Hajj Mohamad, Anne-Laure Bianne-Bernard, Chafic Mokbel, Laurence Likforman-Sulem, Christopher Kermorvant, and Fare's Menasri* [42] which was developed to make a system to efficiently work in recognizing word resulting from the combination of three handwriting recognizers. The HMM based recognizer is important component of this combined system which considers changing nature and contextual information for a better modelling of writing Special Conference Issue: National Conference on Cloud Computing & Big Data 121 units.

In the research by *M.J Bleda, Jorge, Salvador E.B, and Z.-M. Francisco* shows the use of hybrid Markov Model and Artificial Neural Network the limitless offline handwritten texts can be recognise. To figure out the emission probabilities the basic portion of the optical models can be modelled with a Multilayer Perceptron and with Markov chains. A novel segmentation-based and lexicon driven handwritten English recognition systems was presented by *Po yang, Aiquan Yuan, Gang Bai, yanni Guo, Xinting Zgao.* A modified online segmentation method-based rules can be applied for the segmentation. For offline character recognition the convolutional neural networks are introduced. Experiments are examined on UNIPEN lowercase data sets.

Table 1: Word Recognition Techniques developed by noted researchers

Sr.	Authors	Method	Classifier		Dataset (Word)	Accuracy (%)
1.	<i>OLIVER C., AVILA M., COURTELLEMONT P., PAQUET T., LECOURTIER Y.</i> [40]	Image Segmentation	HMM		3000	52.25
2.	<i>ALESSANDRO VINCIARELLI, SAMY BENGIO, HORST BUNKE</i> [39]	Language Resources	HMM and Statistical Model	Multi Writers Task	50000	75.00
				Single Writer Task	50000	50.00
3.	<i>MATTHIAS Z., JEAN-C. CHAPPELIER, HORST BUNKE</i> [41]	Stochastic Context-Free Grammars (SCFG)	HMM	Multi Writers Task	2000	75.60
				Single Writer Task	2000	54.40
4.	<i>Anne-Laure Bianne-Bernard, Fare's Menasri, Rami Al-Hajj Mohamad, Chafic Mokbel, Christopher Kermorvant, and Laurence Likforman-Sulem</i> [42]	Rimes feature extraction	HMM		2130	74.1-78.0
		Sliding Window System	HMM	Context Independent	5334	68.57
				Context Independent	5334	75.53
Rimes feature extraction and Sliding Window	NN and HMM		7742	89.10		
5.	<i>Salvador E.B., M. J. Castro-Bleda, Jorge Gorbe-Moya, and Francisco Z.-M.</i> [43]	MLP Based	Hybrid HMM/ANN Model		30000	84.39
6.	<i>Aiquan Yuan, Gang Bai, Po Yang, Yanni Guo, Xinting Zhao</i> [44]	Lexicon-driven and Segmentation Based	CNN		1791	92.20

Figure 3: Word Recognition Techniques developed by noted researchers

LANGUAGE ANALYSIS AND PROCESSING: From on-line or off-line approach for recognition the language models are essential as after passing through a noisy channel they can recover strings of words, such as handwriting or print degeneration. In written language recognition the lexicon of words segment is the most important model. To rate the lexicon, the algorithm for sentence matching for a lexicon and candidate words are used, frequently using a Levenshtein distance metric variant that include different edition costs into the ranking process. In the training data methods for string matching are generally enhanced by adding lexicon statistics. By linguistic constraints Lexical subsets can be determined, in continuous recognizing text, for each word the lexicon is constrained by the semantics, pragmatics and syntax of the sentence. The language model efficiency is estimated in text perplexity, the average number of following words that can be estimated for each word in a text is measured. To improve the performance of recognition system the statistical information at the word-sequence level must be include. The performance can be increase by selecting lower level words from recognition of word the output makes the entire sentence more probable when the surrounding context signifies such selection. To change nearest word formed by a word recognizer the Lexical techniques can be used. By including re-ranking, deleting or suggesting new word candidates the modifications can be done. If word A is present, there is a high probability that word B also is present this is Collocations which are word patterns that appears regularly in language.

CONCLUSION

In the handwritten character recognition many work, and research have been done. But still, full accuracy has not been attained which can give an extent of further work in this direction. Recognition of separate characters can give better accuracy, but due to different writing style word recognition can be affected. The integrated method removes the complex segmentation, but they use a limited vocabulary. Because of its complexity segmentation acquires less accuracy. The classifier can give good accuracy where extent of words is bound to limited numbers as it has to deal with fix number of variation.

The survey shows that because of wrong character portions of touching or damaged characters there is an error in recognizing handwritten English characters. Many segments of various following lines may also overlap because of upper case or lower-case modifiers of the English text, and to get better result proper portion of such overlay segments are needed. The post-processing of classifier gives results by making the entire language with the handwriting recognition system can somewhat increase the classification in printed as well as word recognition as suggested by many authors.

REFERENCES

- [1] Hamid Mirvaziri, Mohammad Masood Javidi, Najme Mansouri, "Handwriting Recognition Algorithm in Different Languages: Survey", Visual Informatics: Bridging Research and Practice, Part of the Lecture Notes in Computer Science book series (LNCS, volume 5857), 2009, pp 487-497.
- [2] Rejean plamondon, sargur srihari, "On-line and off - line handwriting recognition: a comprehensive survey", IEEE Transactions on Pattern Analysis and Machine Intelligence · January 2000, pp 63-84.
- [3] Aiquan Yuan, Gang Bai, Lijing Jiao, Yajie Liu, "Offline Handwritten English Character Recognition based on Convolutional Neural Network", 10th IAPR International Workshop on Document Analysis Systems 2012.
- [4] Alessandro Vinciarelli, Samy Bengio, Horst Bunke, "Offline Recognition of Unconstrained Handwritten Texts Using HMMs and Statistical Language Models", IEEE Transactions on pattern Analysis and Machine Intelligence, Volume 26 Issue 6, June 2004 Page 709-720.
- [5] Oliver C., Avila M., Courtellemontp., Paquet T., Lecourtier Y. "Handwritten Word Recognition by Image Segmentation and Hidden Markov Models" IEEE 0-7803-0891-3/93.
- [6] Matthias Zimmermann, Jean-Ce'dric Chappelier, and Horst Bunke. "Offline Grammar-Based Recognition of Handwritten Sentences" IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 5, May 2006.
- [7] Anne-Laure Bianne-Bernard, Fare`s Menasri, Rami Al-Hajj Mohamad, Chafic Mokbel, Christopher Kermorvant, and Laurence Likforman-Sulem. "Dynamic and Contextual Information in HMM Modeling for Handwritten Word Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 10, October 2011.
- [8] Salvador Espan˜a-Boquera, Maria Jose Castro-Bleda, Jorge Gorbe-Moya, and Francisco Zamora-Martinez. "Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 33, No. 4, April 2011.
- [9] V. S. Dhaka, Manoj Kumar, Prashant Chaudhary, "Offline Handwritten English Script Recognition: A Survey", Special Conference Issue: National Conference on Cloud Computing & Big Data, pp 114-124.
- [10] Santosh K.C., Cholwich Nattee. A comprehensive survey on on-line handwriting recognition technology and its real application to the Nepalese natural handwriting. Kathmandu University Journal of Science, Engineering, and Technology, Kathmandu University, 2009, 5 (I), pp.31-55.

- [11] Anoop M. Namboodiri, Student Member, IEEE, Anil K. Jain, 2004. Fellow, IEEE, "On-line Handwritten Script Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 1 pp. 124–130.
- [12] Homayoon S. M. Beigi, Krishna Nathan, Gregory J. Clary and Jayashree Subrahmonia, 1994. "Size Normalization in On-line Unconstrained Handwriting Recognition", IEEE, ICIP, pp.163-173.
- [13] Jiang Gao, Xiaoqing Ding, Jing Zheng, 2000. "Image Pattern Recognition Based on Examples– A combined Statistical and Structural-Syntactic Approach", In Proceedings of Joint IAPR International Workshops, SSPR 2000 and SPR 2000., Vol. 1876, pp. 57–66.
- [14] Krupa dholakia, "A survey on handwriting character recognition techniques for various indian languages", international journal of computer applications (0975-8887) volume 115-no.1, April 2015.
- [15] Monica Patel, Shital P. Thakkar, "HandwritTEN character Recognition in English: a survey", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015.