2020 16th IEEE International Colloquium on Signal Processing & its Applications (CSPA 2020), 28-29 Feb. 2020, Langkawi, Malaysia

Handwriting Detection and Recognition Improvements Based on Hidden Markov Model and Deep Learning

Mohammed Hazim Alkawaz
Faculty of Information Sciences & Engineering
Management and Science University
Shah Alam, Selangor, Malaysia
mohammed_hazim@msu.edu.my

Cheng Chun Seong
School of Graduates Studies
Management and Science University
Shah Alam, Selangor, Malaysia
ndrewccs@gmail.com

Husniza Razalli
Faculty of Information Sciences & Engineering
Management and Science University
Shah Alam, Selangor, Malaysia
husniza_razalli@msu.edu.my

Abstract— **The online handwriting detection and recognition has become an important research in … area. An individual's writing can be easily forged and disguised in various ways including freehand simulation, tracing and image transfer, making genuine handwriting recognition a challenging task. With the advent of various online handwriting recognition systems developed, but for English characters recognition these still lack the simplicity and accuracy. While identification approaches were successfully reported, good forgeries are able to outsmart the existing tools. Existing flaws in recognition systems led to more research works in automatic detection and recognition works via computer techniques, feature extraction, classification accuracy comparison, performance evaluation and pattern recognition. To realize simpler and efficient English character recognition, we develop a handwriting detection and recognition system based on the Kohonen Network and deep learning. The system consists of interfaces for the online handwritten character was featured in matrix form of sizes 5x7 pixel and 35x33 pixels represented with binary values. Identifying all occupied character strokes in the series of binary string recognizes the full character. The recognition performance was compared between 35 pixels and 1155 pixels environment, evaluated in terms of accuracy, and consistency. An experiment was conducted with 25 online handwritten input data of straight stroke ('V', 'X', 'Y') and curve stroke ('C', 'O', 'S') characters collected from 25 participants. Findings show an overall improvement of 31% recognition accuracy of using 35x33 pixels against the 5x7 pixels. Handwriting characters featured in 35x33 pixels outperformed the 5x7 pixels accuracy by 37.49% on straight stroke characters and 24.52% on curve stroke.**

Keywords: **Online Handwriting, Detection, Deep Learning, Recognition Accuracy, Pixels, Hidden Markov Model, Kohonen Network**

I. INTRODUCTION

Every individual has different handwriting as unique as the personality traits; even when a similar sentence is written twice by the same person the handwriting may not appear exactly the same [1]. Handwritten characters differ by 12 considerable characteristics: line quality, spacing (line or spaces between character and word), height, width and size of letters, pen lifts and separations, connection strokes, beginning and ending strokes, unusual letter formation, shading (pen pressure), slant, baseline habits, flourishment and embellishments and diacritic placement. External conditions also play a role in affecting the style of handwriting such as the types and colours of ink, pen tip type, smoothness of paper, table surface quality and material, personal emotions, age, gender and speed of the writing process.

II. LITERATURE REVIEW

An online handwriting uses a unique electronic pen as an interface input device on electronic surface writing. The high sensitivity of pen device is important to produce online handwritings of high detection accuracy. In the online handwriting, features can be extracted from either the pen trajectory or the resulting images.

The input wordings extracted directly from online electronic pen device input are merely raw and need to be filtered to retrieve qualitative sample data [2]. In handwriting recognition works performed by Jaeger et al, a word was divided by lines: upper, lower, base, and corpus. Based on those lines, the height of a word was determined. However, the weakness was that the sizes of wordings could not reflect the stroke patterns for reliable recognition analysis.

Nonetheless, according to [3] the standard recognition process is similar in most studies, whereby a sequence of features extracted from the data. The features were then matched to a sequence of labels (usually characters or sub-character strokes) using Hidden

Markov Model (HMM) or HMM-neural network hybrid.

Graves et al, whereas, used three components of recognition system: multidimensional recurrent neural networks, and multidimensional LSTM in particular; the connectionist temporal classification output layer; and the hierarchical structure. The advantage of the approach was that the system could be used in generic and has proven successful for both English and Arabic characters.

Optical Character Recognition (OCR) is another popular technique used in recognizing either scanned or written text characters online [5] work was capable to detect bidirectional wordings especially words read from right to left, the coverage was only limited to standard English alphabets but no other languages especially those with strokes. Apparently, a pre-training process might be required to learn different languages' characters before being able to recognize them. The program must be familiar with the particular language patterns before being able to properly detect them.

Apart from the OCR, another recent study applied the Intelligent Character Recognition (ICR) as reported in [6].

ICR is the advanced form of the OCR with the ability to learn characters as part of its training process [6]. The approach used in showed high performance and highly optimized algorithm to detect characters rather than the OCR. The advantage was that the ICR can extract and recognize texts even from the poor-quality image while the disadvantage was that it has low-resolution image and high complexity background.

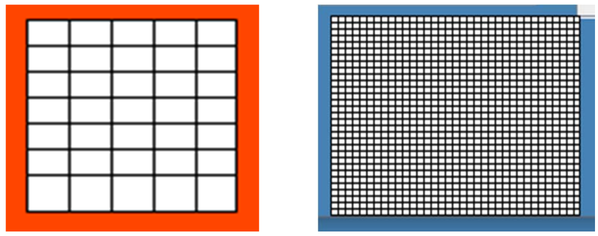
Fuzzy Inference System (FIS), which was originally designed for online characters' recognition, was used in [7]. FIS has generative and discriminative capacities to evaluate handwritten symbols, with respect to the study handwriting feature set. Various other feature sets can be designed to analyze cursive writing with other criteria. Such analysis is performed by initial learning from few data and incremental real-time learning from the run-time data flow; to adapt its model and support class adding during its usage. The evolving nature of the FIS allows incremental learning of a specific model of a child's handwriting as it improves [7]. However, the confidence level computed in the study was merely below 80%.

In handwriting recognition area of research, character binarization has always been the key for common character extraction before recognition takes place [8]. Various methods of extraction can be found. Feature extraction of a handwritten character was done by [10] using the Hidden Markov Model in order to identify the segmentations of data. A majority of researches have been focusing on methods to enhance the data accuracy alone. However, [9] has focused on a unique approach whereby handwritten data is left untouched while works towards improving deep learning algorithm to achieve new handwriting recognition benchmark on ICDAR-2013 database.

III. Methodology

A. Data Collection

In this work, the environment is equipped with deep-learning ready and algorithms that allows every handwritten data to be down-sampled, stored, and learned. However, the stored data in the environment were kept constant to a fixed number of one-record in order to test on the different pixel environments. The beginning phase would be raw handwritten data input. High sensitivity touch pad and pen (or direct touchscreen input will be supported in the test environment. For the first environment, the coordinated pixel size of 5x7 pixel (35 pixels) was used based on the existing experiment (Fig. 1(a)). The second environment consists of 35x33 pixel (1155 pixels). The increase in number of pixels was aimed to expand the area for the raw data to be down sampled. With significant larger number of pixels in the testing environment, the hypothesis aimed with the significantly increased in number of pixels, the pixel-coordination sensitivity will also increase for raw data to cover a much higher span of pixels and thus produced more significant binary pattern from the collected data. The 35x33 pixel (1155 pixels) input environment is the maximum supported size by the existing test environment and is fairly acceptable in terms of data accuracy (Fig. 1(b)).



(a) (b)

Fig. 1. Matrix form of sizes (a) 5x7 pixel (b) 35x33 pixel environment

The environment constant has been fixed with the character database to minimal (single set data of complete English alphanumeric characters). The reason is to use an environment with minimal "knowledge" towards known alphanumeric characters to avoid the influence of deep learning in this experiment. Straight stroke characters (V, X, Y) and curve stroke characters (C, O, S) were selected for this experiment and data collection because of the characters consist of very close writing pattern and creates the potential to "confuse" the character detection among the selected characters. These characters

were chosen in the experiment to increase the recognition detection difficulty levels. The detection outcome would be detected either correctly as the input characters or incorrectly as any other character among the selected character group or even to another character other than the selected characters in this experiment.

A total of 25 people consisting 15 male and 10 females were involved on a voluntary basis in the data collection. In the first test, each person was required to input handwritten data 3 times for each character VXY for non-curve characters and COS for curve characters. The targeted characters were chosen due to the potentially closed similarity to each other and the input environment alone has the capability to differentiate the input character difference.
For the second test, respondents were required to input a handwritten character for five times on both 5x7 and 35x33 pixel environment. These data will be stored into the environment's database. The stored data will be tabulated in binary string sorted in list view as shown in Fig. 2. Based on the hypothesis, any repetitive handwritten character by a same person should show a consistent pattern.

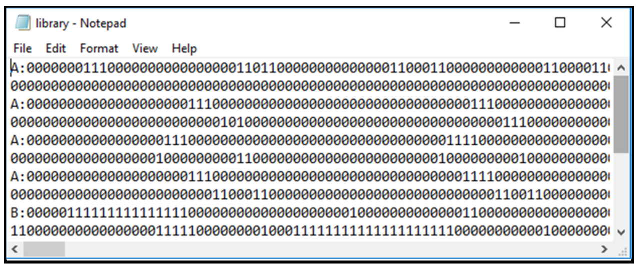


Fig. 2. Sample of stored binary value of a handwritten data

B. Preprocessing

The raw data undergoes data filtration before to be finalised as an acceptable handwritten data. Every handwritten raw input data will be extracted from the input space and down sampled onto the coordinated pixels. For this experiment, inputs were placed onto 5x7 and 35x33 pixels. Then, all the pixels are converted into a binary string, which indicates the value of 1 for occupied coordinated pixel overlapped by part of the raw data and 0 for coordinated pixel that is unoccupied by the raw data. For the 5x7 pixel environment will consists of 35-length binary string while the other environment will contain 1155-length binary string. Theoretically, if a set of data is written by the same person, the binary string values will produce a constant pattern. This similar step is to be repeated several times in order to obtain multiple sets of data from respondents. Character recognition process does not take place in the raw handwriting data input space but solely from the data that is later being down-sampled onto the pixels. The size of the raw handwritten input does not influence the character down sampling process. Apart from down sampling, raw handwritten data were also preprocessed to fit entirely onto the specific pixel space which makes sizes a non-influential factor towards this experiment (Fig. 4(a), (b), (c), (d)).

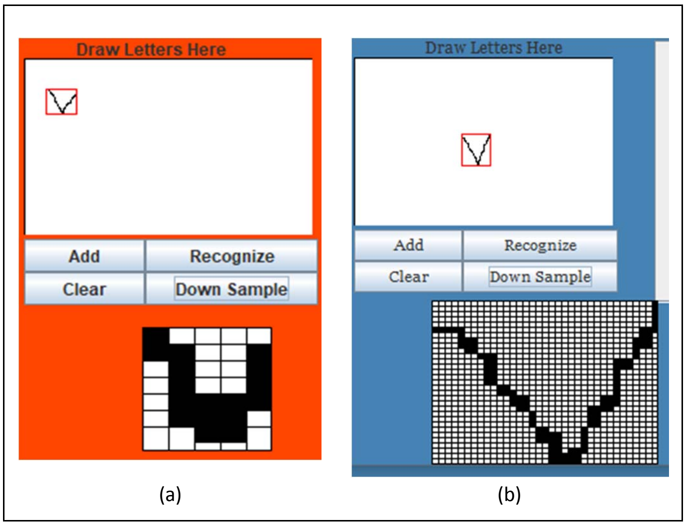


Fig. 3. Small sized input character 'V' on (a) 5x7 pixel (b) 35x33 pixel environment

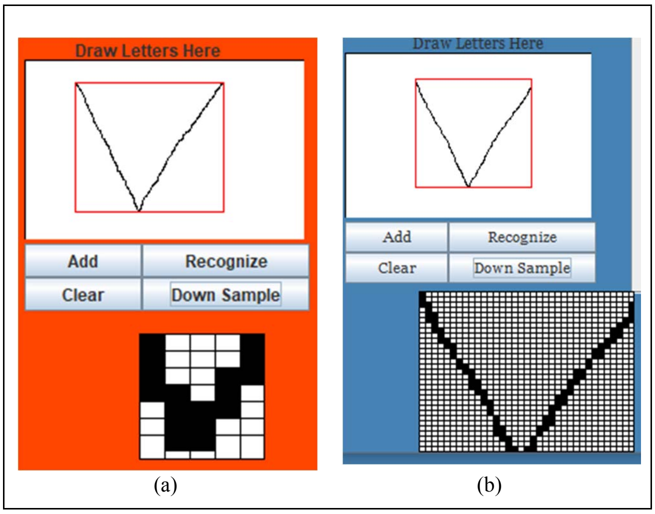


Fig. 4. Large sized input character 'V' on (a) 5x7 pixel (b) 35x33 pixel environment

IV. Results and Discussion

From the experiment conducted on 25 respondents, the hypothesis was accepted that the number of pixels does influence the character detection. The selected stroke characters (V, X, Y) and curved characters (C, O, S) were meant to increase the detection difficulty level in order to proof the possible influence of pixel environment. The stroke and curved characters consist of very close writing pattern and creates the potential to "confuse" the character detection among the selected characters.

In straight stroke character ('V', 'X', 'Y') recognition, the successful detection rate tested in the 5x7 pixel environment was 35.92% to 44.00% (Table 1(a)). While the same experiment conducted on the 35x33 pixel environment, the outcomes were 81.48% to 85.40%. The increase from 35 pixels to 1155 pixels improved the handwritten character detection rate by 49.48% for 'V', 25.52% for 'X' and 37.48% for 'Y'.

With regards to the curve stroke characters ('C', 'O', 'S') detection tested in the 5x7 pixel environment, the accuracies obtained were 64.00%, 60.04% and 61.44% respectively. On the 35x33 pixel environment, the detection results for 'C', 'O' and 'S' show a higher accuracy, with 92.08%, 84.12% and 82.84% respectively. The same effect observed confirms that 1155-length pixel environment increases character recognitions; 28.08% enhancement for character 'C', 24.08% on 'O' and 21.40% on 'S' (Table 1(b)).

Table 1(a): Result of Stroke Character Detection

5x7 Pixel Environment		**35x33 Pixel Environment**
Character	**Success Recognition Rate (%)**	**Success Recognition Rate (%)**
V	35.92	85.40
X	57.32	82.84
Y	44.00	81.48

Table 1(b): Result of Curve Character Detection

5x7 Pixel Environment		**35x33 Pixel Environment**
Character	**Success Recognition Rate (%)**	**Success Recognition Rate (%)**
C	64.00	92.08
O	60.04	84.12
S	61.44	82.84

Throughout the experiment, the selection of specific characters in both straight stroke and curved stroke groups was intentional in order to establish events of repetitive mis-detection of characters within character groups, especially on the 5x7 pixel environment. The highest misrecognition was on character Y while the lowest on character C. misdetection took place for a total of 20 times and the most occurrence characters was Y and W.

Higher accuracy in handwritten down sampled data recognition was observed and achieved in 1155 pixels compared to 35 pixels. The limited number of pixels had potentially created a “confusion” in identifying the down sampled data in the test environment. In 5x7 pixel environment, each individual pixel size was too large that the raw input data formed within the limited 35-pixels environment has caused the difficulty in performing differentiation more precisely. While on the other hand, greater number of pixel environment (33x35 pixel), has significantly improved the down sampled data into a much precise and observable character formed in the entire pixels environment.

Curve stroke characters recognition accuracies were higher than the straight stroke characters. The smaller number of similarity patterns (coverage input area and stroke style) of the curved characters reduces the chances of misrecognition by an average of 20%. The straight stroke characters have higher chances of similarities especially between letters …

Based on the results shown in Fig. 5 compared to Fig. 7 and Fig. 6 compared to Fig. 8, the hypothesis was accepted that the number of pixels does influence the character detection accuracy.

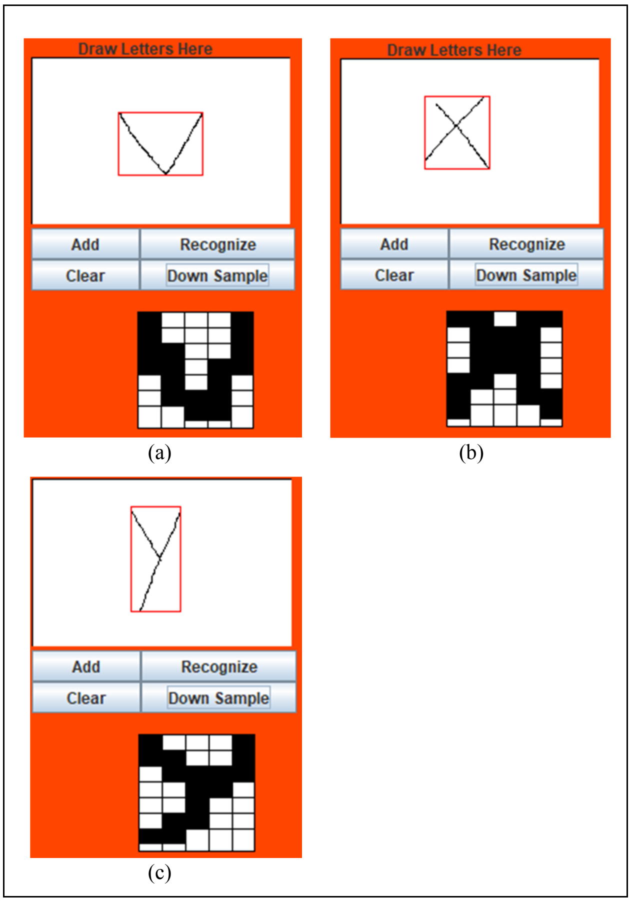


Fig. 5(a), (b), (c) sample of down sampled in 5x7 stroke

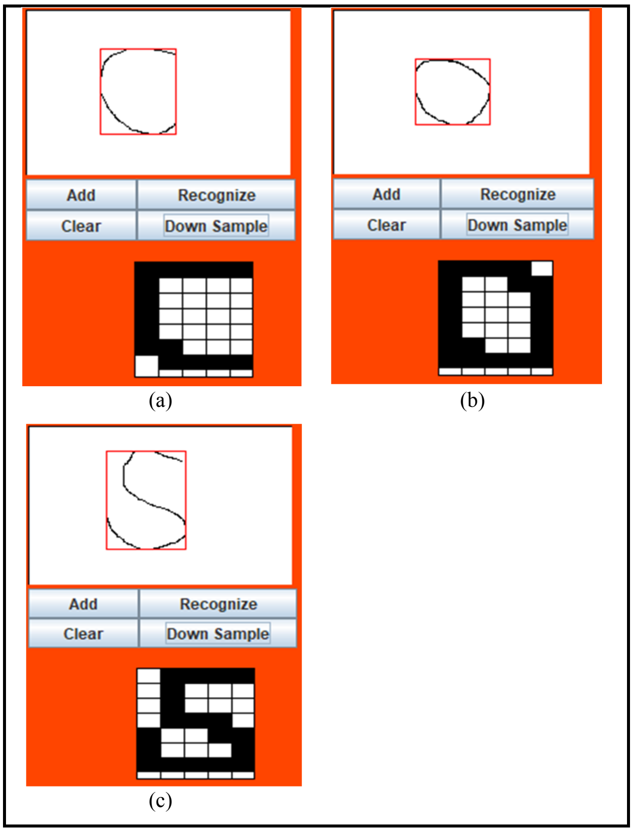


Fig. 6(a), (b), (c). sample of down sampled in 5x7 curved

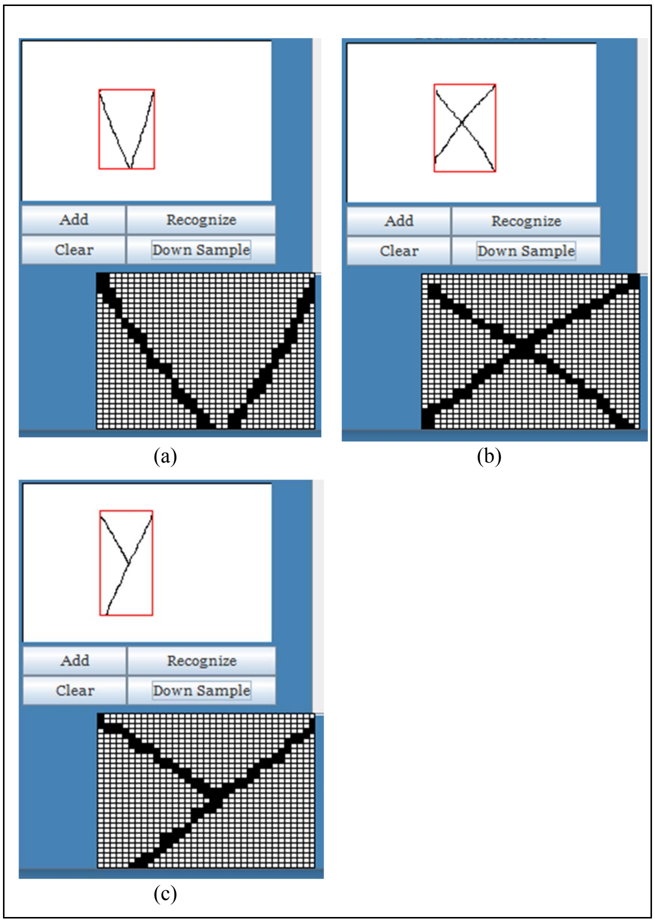
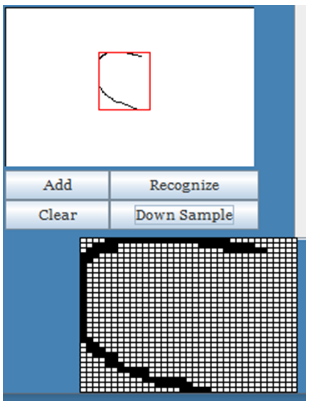
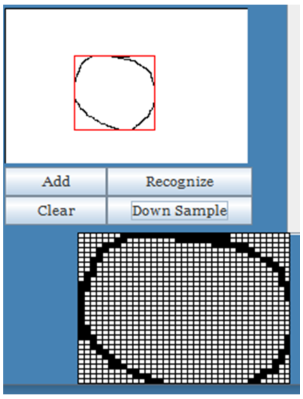


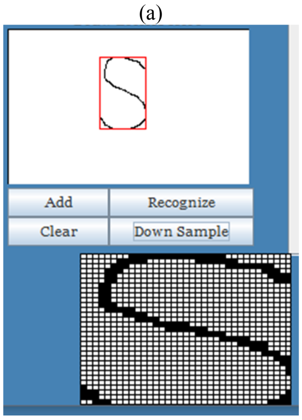
Fig. 7(a), (b), (c) sample of down sampled in 33x55 stroke



(a)



(b)



(c)

Fig. 8(a), (b), (c). sample of down sampled in 33x55 curved

V. Conclusion

In conclusion, the outcome of this experiment has achieved the objective and the hypothesis was proven to be accurate. The significant increase in number of pixels used in data input had further increased the accuracy of down sampled handwritten data. The advantage of relying on huge number of pixels was to produce sharper and clearer handwritten data, which aids the recognition process and outcome. As a future work, the number of pixels can be further increased to a significant amount in order to increase character recognition accuracy. With a very large number of pixels, the recognition capability may be improved to the extent whereby the identification of handwritten data could be distinguished between the respondents.

Acknowledgment

Authors are grateful to School of Graduate Studies (SGS), Research Management Centre (RMC), and Faculty of Information Sciences and Engineering (FISE) of Management and Science University Malaysia for research support.

References

[1] Harralson et. al. Handwriting Examination: Theory, Proficiency, and Methods. A Survey of Forensic Handwriting Examination Research in Response to the Nas Report. Retrieved 2013.

[2] Jaeger, S., Manke, S., Reichert, J., & Waibel, A. (2001). Online Handwriting Recognition: The Npen++ Recognizer. International Journal on Document Analysis and Recognition, 3(3), 169-180.

[3] Graves & Schmidhuber (2009). Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks. Neural Information Processing Systems Conference.

[4] Verma B, Lu J, Ghosh M and Ghosh R 2004 A Feature Extraction Technique for Online Handwriting Recognition Neural Networks 2004 Proceedings 2004 IEEE Int. Jt. Conf. 2 1337-1341.

[5] Jayakumar, A., Babu, G. S., Raman, R., & Nedungadi, P. (2015). Integrating Writing Direction and Handwriting Letter Recognition in Touch-Enabled Devices (Vol. 2, Ser. 380). Springer India.

[6] Ahmed et al. (2015). OCR System for Poor Quality Images Using Chain-Code Representation. The 1st International Conference on Advanced Intelligent System and Informatics (Aisi2015), 407, 151-161.

[7] Bouillon, M., Anquetil, E., & Almaksour, A. (2013). Decremental Learning of Evolving Fuzzy Inference Systems: Application to Handwritten Gesture Recognition.

[8] Sarkar, M., Roy, S., Das, A., & Nath, A. (2017). International Journal of Advance Research in Computer Science and Management Studies. Recognition of Handwritten Characters Using Binarization, Image Segmentation and Correlation Matching Technique, 5(6).

[9] Zhang X Y, Bengio Y and Liu C L 2017 Online and Offline Handwritten Chinese Character Recognition: A Comprehensive Study and New Benchmark Pattern Recognition. 61 348-360

[10] O. Samanta et al (2018). An HMM framework based on spherical-linear features for online cursive handwriting recognition, Information Sciences, Volume 441, Pages 133-151, ISSN 0020-0255