Identification of Tamil Ancient Characters and Information Retrieval from Temple Epigraphy Using Image Zoning

Giridharan. R[1]
Student
Department of Computer Technology
Kongu Engineering College
Perundurai – 638 052, Erode, Tamil Nadu, INDIA

Vellingiriraj.E.K[2]
Assistant Professor
Department of Computer Technology
Kongu Engineering College
Perundurai – 638 052, Erode, Tamil Nadu, INDIA
vellingiriraj@kongu.ac.in

Dr. Balasubramanie.P[3]
Professor
Department of Computer Science and Engineering
Kongu Engineering College
Perundurai – 638 052, Erode, Tamil Nadu, INDIA
balu_p@kongu.ac.in

Abstract - **The aim of this paper is to develop a system that involves character recognition and information retrieval of Brahmi, Vattezhuthu and Grantha letters from temple epigraphy and their conversion to the present Tamil digital text format. Though many researchers have implemented various algorithms and techniques for character recognition in different languages, Ancient letter conversion still poses a big challenge. Because Image recognition technology has reached near-perfection when it comes to scanning English and other language text. But optical character recognition (OCR) software capable of digitizing printed Tamil text with high levels of accuracy is still elusive. Only a few people are familiar with the ancient characters and make attempts to convert them into written documents manually. If this continues, all the precious information given by our forefathers will not be known to the future generations. The proposed system overcomes such a situation by converting all the ancient characters from inscriptions and palm manuscripts into Tamil digital text format. After converting into the Tamil digital text, the words will not be correct spelling. Because the ancient words are Etymological word, the meaning is not known. So this modal approach is to solve these types of problem and convert the Tamil digital text with meaning.**

Keywords: Character Recognition, Information Retrieval, Ancient Letters, Segmentation, Image Zoning, Multi Level Perceptron, Neural networks.

I. INTRODUCTION

Tamil is one of the oldest languages in the world with rich literature. In the ancient days, the writers, especially in Tamilnadu, have used palm leaves and inscriptions to encrypt their writings. A very good example of the usage of Palm leaf manuscripts is to store the history of Tamil grammar book named Tholkappiyam which is written during 4th B.C. The ancient literature includes many palm leaf manuscripts that contain rare commentaries on Sangam works, unpublished portions of classics, Saiva, Vaishnava and Jain works, poetry of all descriptions, medical works of exceptional values, food, astronomy & astrology, vaastu & Kaama shastra, jewellery, music, dance & drama, medicine, Siddha and so on. Palm manuscripts are utilized for 3 different categories. Document registration of land and building which are donated by the kings to people are encrypted in palm manuscripts. The literary works, grammar, astrology, science and technology, etc., are encrypted in palm manuscripts. Historical moments of places and dominion are also encrypted. Character recognition is one of the most difficult tasks in the pattern recognition system. There are a lot of difficulties in image processing techniques. To solve these, one should know how to i) separate the characters in the segmentation process, ii) to recognize unlimited character fonts and sculpting styles in noisy image and iii) distinguish characters that have the same shape, but have different pronunciation in characters like (Ng) and (iT). Many researchers try to apply many techniques for breaking through the complex problems of character recognition. The most difficult thing is to recognize the sculpting of Brahmi characters in stone compared to the other character recognition from different sources. The Brahmi characters were sculpted in stones, clay pot, copper plate etc. Even though the researcher converts the ancient characters upto 90% accuracy, but the conversion words cannot be understood with the proper meaning in the sentence. In this paper, a character recognition system is used and it is based on statistical approach and convert etymological word to equivalent Tamil word from temple epigraphy.

II. ANCIENT BRAHMI, GRANTHA & VATTEZHUTTHU SCRIPTS

Brahmi script is one of the most important writing systems in the world by virtue of its time depth and influence. It represents the earliest post-Indus corpus of texts, and some of the earliest historical inscriptions found in India. The best-known Brahmi inscriptions are the rock-cut edicts of Ashoka in North-central India, dated to 250–232 BCE. This elegant script appeared in India most certainly by the 5th century BCE, but the fact is that it had many local variants even in the early texts which suggest that its origin lies further back in time. There are several theories on the origin of the Brahmi script. Brahmic scripts are descended from the Brahmi script. The most reliable of these are short Brahmi inscriptions dated to the 4th century BCE and published by Coningham et al.,[1] but scattered press reports have claimed both dates as early as the 6th century BCE and that the characters are identifiably Tamil Brahmi though these latter claims do not appear to have been published academically.

The earliest epigraphic records are found on rock edicts and hero stones date from around the 5th century BC. The oldest dated Tamil inscription written in the Tamil-Brahmi script has been found in Palani in Southern India, scientifically dated to 540 BCE-the oldest known Brahmi inscriptions on the Indian sub-continent. More than 55% of the epigraphically inscriptions about 55,000 found by the Archaeological Survey of India are in Tamil language. Tamil language inscriptions written in 1st century BCE and 2nd century AD have been discovered in Egypt, Sri Lanka and Thailand. In the Department of Archaeology, a separate epigraphy wing is started, the primary function of which is to copy inscriptions from boulders, stone pillars, stones, temple walls and copper plates. These inscriptions are deciphered, edited and published. So far, about 14,000 inscriptions are copied and their estampages, ink impression are preserved in this wing. Some original copper plates and old palm leaf manuscripts are also under the custody of this department. Only a few of the inscriptions are published while huge stone inscriptions lie published. Recent Survey on Indian Epigraphy places inscriptions of Tamil Nadu at the top of the list. Over the 50,000 palm manuscripts are available in Tamil University, Tanjore. And these palm manuscripts are manually converted by the people from ancient characters to current Tamil text, which takes a long time to convert only a few manuscripts. These manuscripts also contain the information about the land settlement is encrypted by Pallava period during 16th Century. In Government Oriental Manuscripts Library and Research Centre (GOML & RC) is a veritable treasure house of ancient knowledge housing 50,180 palm leaf manuscripts. More than 10,000 Tamil manuscripts are available in Institute of Asian Studies, Chennai, and the 4,730 Tamil manuscripts are available in Saraswathi Mahal Library, Tanjore. There are Millions of Palm Manuscripts available all over the world in renowned library and museums like India, London, Scotland, Japan etc. Only a few palm leaf manuscripts are digitalized manually and many are to be digitalized to enable quick reference in the future. The Brahmi character set, vowels and consonants, is given below.

The Grantha alphabet is a descendent of the Brahmi alphabet and started to emerge during the 5th century AD. Most of the alphabets of southern India evolved from Grantha, and it also influenced the Sinhala and Thai alphabets.

The Grantha alphabet has traditional been used by Tamil speakers to write Sanskrit and is still used in traditional Vedic schools. The script is thought to be based on the stone-age cave inscriptions from the Chera dynasty, and to have gone through three stages of development; Early, Middle and Modern. The only examples of Early or Middle Grantha are from odd words interspersed with Tamil in inscriptions. The earliest inscription written entirely in Grantha has been dated to 1383, and is written in the most modern form of the script.

The Vatteluttu (or Vattezhuttu) script was a syllabic alphabet used in the southern part India, in what is now the states of Kerala and Tamil Nadu. It grew from the southern form of Brahmi script from around the 6th century CE, and was employed to write Tamil and Malayalam languages. One historical development is that the signs inherited from Brahmi unnecessary to represent the Tamil language were eliminated from the Vatteluttu script, a feature also found in the Tamil script and most likely was an influence from Vatteluttu.

The categories of language, alphabet and number of inscriptions on both stone and copper plates discovered Tamil Nadu as the first among Indian States. From this survey, it can easily be understood that Tamil Nadu has the bulk of inscriptions found in India. It has been estimated with a fair degree of accuracy that the inscriptions written in Tamil occupy the first position in volume, amounting nearly to 20,000. Inscriptions in Tamil language are noticed from the third century BCE onwards. Tamil-Brahmi inscriptions are classified as Early Tamil-Brahmi and Late Tamil-Brahmi and dated between 3rd century BCE and 3rd century CE.

III. LITERATURE REVIEW

Historically, handwritten character recognition applications use three major approaches: statistical approach, structural or syntactic approach and neural network-based approach.

A. Statistical Approach

Statistical pattern recognition uses statistical or probability functions for building a recognition algorithm. The input features are extracted from a set of characteristic pattern measurements. A limitation of this approach is the difficulty that exists in expressing pattern classification in terms of structural information [2, 3, 4, 5, 6, 7 and 8].

B. Structural or Syntactic Approach

Syntactic pattern recognition uses syntactic or structural information of patterns to generate knowledge related to patterns. This approach extracts the similarity of patterns and builds pattern syntax or structural rules. The information about pattern syntax rules is used to explain, classify and recognize unknown patterns. This approach is suitable for building a handwritten character recognition system because it uses structural approach to build unlimited handwritten character pattern syntax, but it is not easy to build learning structural rules [9, 10, 11, 12 and 13].

C. Neural Network Based Approach

Neural pattern recognition emulates knowledge of how a biological neural system stores and manipulates information. This artificial neural system is called "neural networks". The notion is that an artificial neural network can solve all problems in an automatic reasoning, including a pattern recognition problem. This approach classifies patterns through predictable properties of neural networks. However, it has only a little amount of semantic information from a network [14, 15 and 16].

IV. METHODOLOGY

All the details of the proposed system design are presented below. First, the overall framework of the ancient Brahmi character recognition system is started. Then the components details are given. The figure 2 shows that framework of character recognition system. This system gets the input image of Brahmi characters and converts into equivalent current Tamil digital format.

Line 1 - Iva kunratuu raiyul patantan eri aritan

Figure 2. Framework of character recognition system

A. System Architecture Overview

The figure 3 shows that the overall architecture of Tamil Brahmi character recognition system for stone inscriptions. The major components of this sytem is image preprocessing, feature extraction and character recognition. The details of architecture diagram is given below.

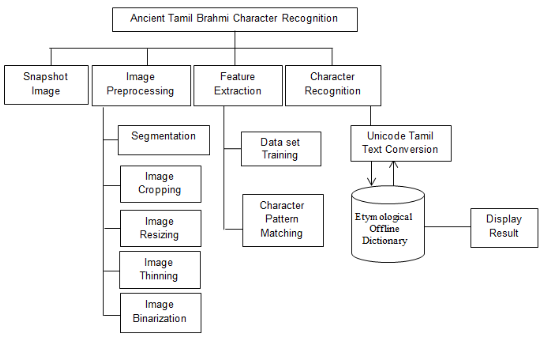


Figure 3. Architecture of Ancient Tamil Brahmi character recognition System for Stone Inscription

B. System Structure

Based on the system framework given above, the Brahmi image is converted into Tamil text format [17]. This framework includes i) image capturing, ii) image preprocessing, iii) feature extraction, iv) character recognition and v) text conversion.

C. Image Capturing

In the first stage, the Brahmi Inscription images are collected from various places. The snapshot images are captured by high quality or high resolution HD / DSLR (High Definition / Digital Single Lens Reflex) camera and stored in Jpeg format.

Figure 4. A sample of Tanjore Epigraphy

D. Image Preprocessing

In the image preprocessing module, the system prepares a Brahmi character image for the feature extraction. This stage consists of four sub-processes: a) image cropping, b) segmentation, c) image resizing, d) image thickening and e) image binarization. Each of these sub-process details is below.

a) Image cropping: The Brahmi Inscription image has the white space from each character, which means that in the inscription image, the Brahmi characters have spaces in between characters. In the stone image, the characters are easily identifiable from each pixel color. The image is changed to gray scale or black and white image and the noisy and unwanted spaces are removed.

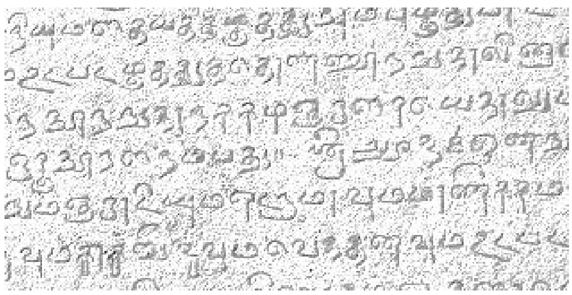


Figure 5. Using Sobel's edge detection method for segmentation

b) Segmentation: Using the grapheme extraction [18] technique, each character can be segmented by the individual characters. Segmentation is divided into three groups: line segmentation, word segmentation and character segmentation. In the present approach, the line (Fig 6) and character segmentation are applied (Fig 7).

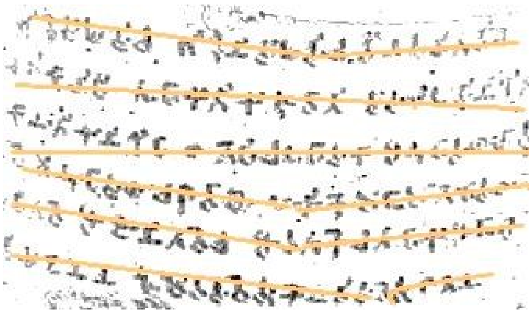


Figure 6. Line Segmentation

Figure 7. Character Segmentation

c) Image resizing: After cropping the particular character, each character may be of different size. So, each character has to be changed to be of equal size. The character image is resized as 100X100 pixels image.

d) Image thinning: The dark pixel (i.e. a character) is converted to be the thinning character. A thickened character extracts easily the thin character by using a nearest pixel darken to lighten color change to particular range.

e) Image binarization: A character stores the Boolean matrix used to store either 0's or 1's. Using the image zoning technique, the dark pixel is stored as 1's and light pixel is stored as 0's.

E. Data set training

Different class of Brahmi characters from various writers of all the 237 Brahmi characters are collected and stored in the training data sets.

F. Character Recognition

A data set of Brahmi characters is matching with the current user data set. The edge detection algorithm is to identify the each character in the training datasets using image zoning.

G. Unicode Text

After matching character, the equivalent current Digital Tamil text is converted using Unicode.

H. Retrieve from Database

After converting the Unicode text, spelling and meaning is checked in the dictionary data base. If the word does not give the meaning in the sentence, nearest meaning is to be searched from the data base.

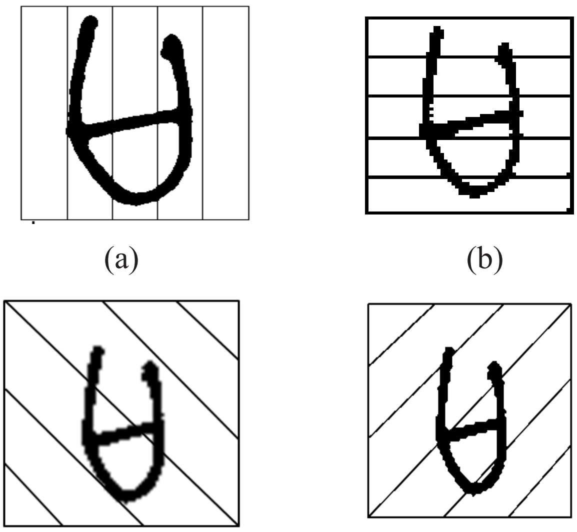
V. PROPOSED ALGORITHM

According to the definition of Zoning, let $Z_M = \{z_1, z_2, z_3 \ldots z_M\}$ be a zoning method. An aspect for zoning based classification concerns the way in which each feature detected in a pattern x which has an influence on each zone of Z_M. Let us consider the classification of a pattern x into one class of $\Omega = \{C_1, C_{,2} \ldots.. C_K\}$ using the feature set $F = \{f_1, f_{,2} \ldots.. f_T\}$. In this case, x can be described by the feature matrix A_x of T rows (features) and M columns (zones).

$$A_x = \begin{bmatrix} A_x(1,1)A_x(1,2) \ldots \ldots. A_x(1,j) \ldots A_x(1,M) \\ A_x(2,1)A_x(2,2) \ldots \ldots. A_x(2,j) \ldots A_x(2,M) \\ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots .. \\ A_x(i,1)A_x(i,2) \ldots \ldots. A_x(i,j) \ldots A_x(i,M) \\ \ldots \quad \ldots.. \quad \ldots.. \quad \ldots.. \quad \ldots. \quad \ldots.. \\ A_x(T,1)A_x(T,2) \ldots \ldots. A_x(T,j) \ldots A_x(T,M) \end{bmatrix}$$

With $A_x\ (i,j) = \sum_{instances\ of\ fi\ in\ x}^{wij}$

where w_{ij} is the weight that defines the degree of influence of an instance of feature f_i on zone z_j. To create the first feature subset, different decompositions of the character image and the density of foreground pixels are calculated in each zone to obtain some features. The density of each zone is calculated by dividing the number of black pixels by the total number of pixels in this zone.



(a) (b)

(c) (d)

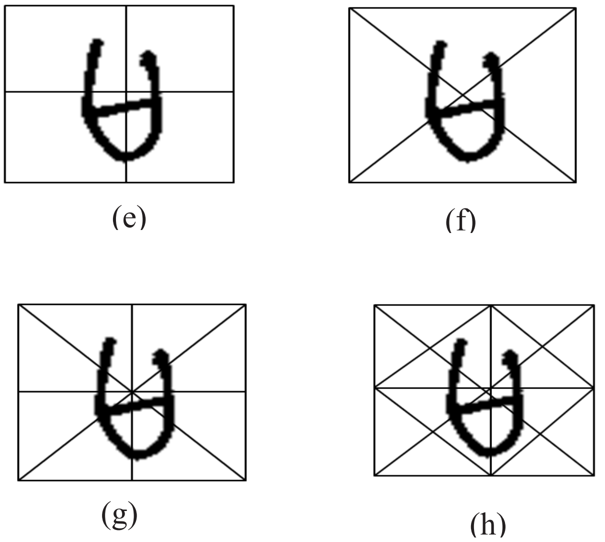


Figure 8. Decomposition of character image
(a) Decomposition to 5 vertical equal zones;
(b) Decomposition to 5 horizontal equal zones;
(c) Decomposition to 6right diagonals;
(d) Decomposition to 6left diagonals;
(e) Decomposition to 8 octants;
(f) Decomposition to 4 diagonal quadrants;
(g) Decomposition to 4 quadrants; and
(h) Decomposition to 4X4 quadrants

The first decomposition, as shown in Figure 8, involves the division of the character image vertically (Fig 8.a) and horizontally (Fig 8.b) to five equal zones, then into 6 right diagonals (Fig 8.(c)) and 6 left diagonals equally (Fig 8.(d)). The second decomposition is obtained by 8 octants (Fig 8.(e)) and 4 diagonal quadrants (Fig 8.(f)), then divided into 4 quadrants (Fig 8.(g)) and the last one into 4X4 quadrants (Fig 8.(h)).

This decomposition of diagonal features increases recognition accuracy and reduces misclassification. Following these decompositions, 54 zones are obtained and the density for each zone is calculated. The perceptron computes a single output from multiple real-valued inputs by forming a linear combination according to its input weights and then possibly putting the output through some non-linear activation function. Mathematically, this can be written as

$$y = \varphi\left(\sum_{i=1}^{n} w_i x_i + b\right) = \varphi(W^T x + b)$$

where w denotes the vector of weights, x is the vector of inputs, b is the bias and φ is the activation function. A signal-flow graph of this operation is shown in Figure 9.

A typical multilayer perceptron (MLP) network consists of a set of source nodes forming the input layer, one or more hidden layers of computation nodes, and an output layer of nodes. The input signal propagates through the network layer-by-layer. The computations performed by such a feedforward network with a single hidden layer

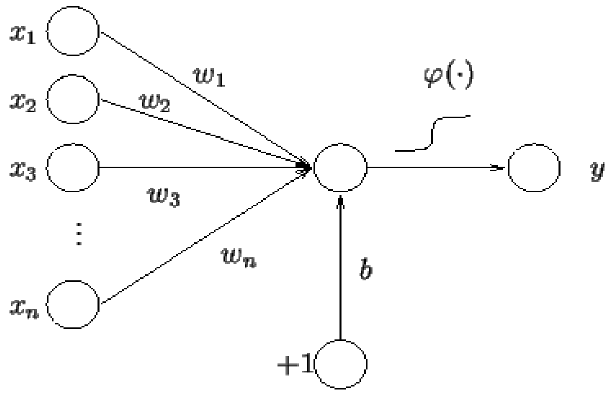


Figure 9. Signal-flow graph of the perceptron

with nonlinear activation functions and a linear output layer can be written mathematically as

$x = f(s) = B\varphi(As + a) + b.$

Where s is a vector of inputs and x a vector of outputs. A is the matrix of weights of the first layer, a is the bias vector of the first layer. B and b are, respectively, the weight matrix and the bias vector of the second layer. The function φ denotes an elementwise nonlinearity.

Neural networks are also an interesting and promising technique, but it seems like it has to mature a bit more before it can be used generally. The current methods are based on back-propagation networks, using one output node for each word in the dictionary and an input node for every possible n-gram in every position of the word, where n usually is one or two. Normally only one of the outputs should be active, indicating which dictionary words the network suggests as a correction. This method works for small (< 1000 words) dictionaries, but it does not scale well. The time requirements are to big on traditional hardware, especially in the learning phase.

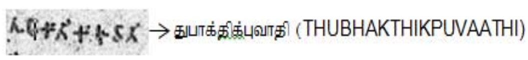


Figure 10. Etymological Spell Checking

This fig.10 Brahmi Etymological spell checker is used as a tool to check the spelling of Tamil Etymological words. It provides possible suggestion for erroneous words. User has the provision to select the suggestion among the list, ignore the suggestion or add the particular word to the dictionary. This module extract the route word from the given word (Noun / Verb) with the help of morphological analyzers and the route word is checked in dictionary and is found, the word is termed as correct word. Otherwise, the correction process is activated. The correction process includes error handling and suggestion generation module. After finding the types of error, the right form of suffix Noun or Verbs are given as input to the suggestion generation module. With the help of morphological generator the correct word is generated. And this module also handles the operation like – Select, change or Ignore the suggested word and adding the word to the dictionary.

VI. RESULTS AND DISCUSSIONS

A. Database

In order to evaluate the performance of the proposed OCR system, the Brahmi database is used as a source of training and testing. The database consists of 5000 characters stored for training and 3000 Brahmi handwritten characters produced by 30 writers who wrote 4 samples of each 25 classes for testing. This Brahmi characters are old one and so the writers have not written the test characters clearly and correctly. After some practice, the writers have written 4 samples of X 25 characters, totally 100 characters from each writer. The training set is composed of 11 vowels, 18 consonants and 198 consonantal vowels, totally 227 characters of all class and the testing set is composed of approximately 25 characters of each class.

The fully connected three-layer perceptron neural network is trained using a sigmoidal activation function, a learning rate of 0.1, a momentum of 0.25 and all weights are randomly initialized in interval [-1,1]. Several runs of back propagation algorithm are performed 1200 times for different architectures by varying the number of nodes in the hidden layer. This is executed in a compatible DELL Laptop, Intel (R) Core (TM) i5-3230M CPU 2.6 GHz and 2 GB of RAM through Java.

B. Discussion

According to experiences, the recognition performance of MLP is obtained when the number of hidden neurons is set to 90. Hundred neurons are not exceeded in the hidden layer to select a number of neurons that provides a compromise between performance and the time taken for the recognition. Thereby, MLP, which allows to achieve a recognition rate of 84.57% and a reasonable time in the recognition phase is 6 milliseconds for each character. Basing on this architecture, the individual accuracy for each class out of the Brahmi characters in the test data are computed

The recognition accuracy of vowel is 90.70% and that of consonant is 90.03%. The remaining 198 consonantal vowel characters are grouped and the character set is matched with the training data set, but the character recognition performance is low compared to that of the vowels and consonants. The recognition accuracy is 83.74% for consonantal vowel characters. Overall character recognition rate is 84.57%. The confusion matrix is shown in Table 3. The results of Amazighs character recognition rate of the Nabil Aharrane et al [19] is 96.47%. The same method is applied in the present research for Brahmi characters with the recognition result of 84.57%. The obtained results show that some characters have a relatively low recognition rate compared to the others.

The lack of character recognition is due to three reasons. One is that the consonantal vowel characters are similar to consonants as shown in Figure 11.

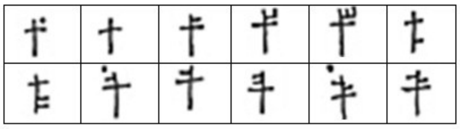


Figure 11. Similarity of Consonantal Vowels

The second one, the writers do not to write the Brahmi characters properly. Moreover, the characters are stored in different strokes and styles with ambiguity as revealed in Figure 12.

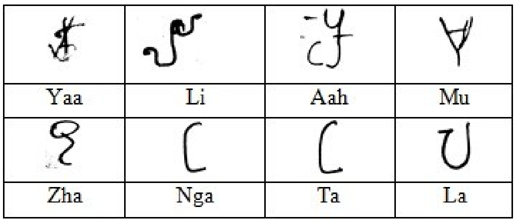


Figure 12. Unknown characters in the data set

The same character with different style and stroke from different writers is given in Figure 13.

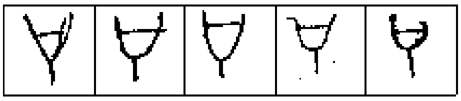


Figure 13. Maa character from different writers

Third one, the data set is stores only a minimum number of data, i.e. only 5000 characters are stored in the database for training set.

TABLE IV. COMPARISON OF OUR ALGORITHM RESULTS WITH OTHER APPROACHES

Approaches	Recognition Rate (in %)	Training Set Size	Testing
Our Approach	84.57	5000	3000
Nabil Aharrane et al [19]	96.47	16926	7254
Giuseppe Pirlo[5]	91.00	29770	7800
Tiji M Jose et al [20]	89.00	7500	2500
Y. Es Saady andal [21]	96.32	18135	2015

In this table 4, our algorithm result is commendable compared with other approaches. And we are comfortable with the result, because existing researchers in this field were not inculcated in Brahmi characters to this extend. Thus, this approach will be helpful for the future researchers to develop the system of recognizing ancient characters.

VII. CONCLUSION AND SCOPE FOR FUTURE WORK

In this paper on Phase I, an approach for character recognition of Brahmi letters using the image zone based on classification and recognition is proposed. The next phase implements this approach to Grantha and Vattezhutthu. This is the new way of approaching Ancient Brahmi character recognition, the result of which is found to be above 90% for consonant and vowel letter recognition, whereas the consonantal vowel recognition is somewhat low due to the lack of similarity of letters. The overall output of our algorithm is 84.57% accuracy which higher than that of the existing systems. Nevertheless, the Brahmi character recognition systems still have many problems that need more complex algorithms to solve the three reasons mentioned above. In future, the system can be modified with the incorporation of better feature extraction methods to improve the accuracy rate. The writers shall get trained for writing the Brahmi letters to create the data set. By that, the error rate of ambiguous letters will be reduced. In addition, the system in future will recognize the Brahmi letters from the stone inscriptions for improving the result in using some hybrid algorithms.

REFERENCES

[1] Coningham, R.A.E.; Allchin, F.R.; Batt, C.M.; Lucy, D. (1996), "Passage to India?Anuradhapura and the Early Use of the Brahmi Script", *Cambridge Archaeological Journal* 6 (1) 73–97,doi:10.1017/S0959774300001608.

[2] G. Nagy, S. Seth, and M. Viswanathan, "A Prototype Document Image Analysis System for Technical Journals," *Computer*, vol. 25, no. 7, Jul1992, pp. 10–22.

[3] F. Shafait, D. Keysers, and T. M. Breuel, "Performance Evaluation and Benchmarking of Six-Page Segmentation Algorithms," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 30, no. 6, Jun 2008, pp. 941–954.

[4] Y.Liang, M.C.Fairhurst and R.M.Guest, "A Synthesisd Word Approach to Word Retrieval in Handwritten Documents", *Elsevier Pattern Recognition*, Vol.45, Jun 2012, PP 4225-4236.

[5] Giuseppe Pirlo, DonatoImpedovo, "Adaptive Membership Functions for Handwritten Character Recognition by Voronoi-Based Image Zoning", *IEEE Trans on Image Processing*, Vol 21, No 9, Sep 2012,pp. 3827-3836.

[6] Chomtip Pornpanomchai,Verachag Wongsawangtham, Satheanpong Jeungudomporn and NannaphatChatsumpun, "Thai Handwritten Character Recognition by Genetic Algorithm (THCRGA)", *IACSIT Journal of Engineering and Technology*, Vol 3, No 2, Apr 2011.

[7] Qiu-Fend Wang, Fei Yin, and Cheng-Lin Liu, "Handwritten Chinese Text Recognition by Integrating Multiple Contexts, *IEEE Trans on Pattern Analysis and Machine Intelligence*, Vol 34, No 8, Aug 2012.

[8] Tiji M Jose and Amitabh Wahi, "Recognition of Tamil Handwritten Characters using Daubechies Wavelet Transforms and Feed-forward Back Propagation Network", *International Journal of Computer Applications*, Vol 64, No 8, Feb 2013, pp 0975-8887.

[9] Jin Chen, Daniel Lopresti, " Model Based Ruling Line Detection in Noisy Handwritten Documents", *Pattern Recognition Letters, Elsevier*, 2012.

[10] A Bharath and Sriganesh Madhvanath, "HMM-Based Lexicon-Driven and Lexicon-Free Word Recognition for Online Handwritten Indic Scripts", *IEEE Trans on Pattern Analysis and Machine Intelligence,*Vol 34, No 4, Apr 2012.

[11] Chomtip Pornpanomchai, Dentcho N. Batanov and Nicholas Dimmitt, "Recognizing Thai handwritten characters and words for humancomputer interaction", *International Journal of Human-Computer Studies*, 2001, pp 259-279.

[12] Chomtip Pornpanomchai, Pattara Panyasrivarom, Nuttakit Pisitviroj and Piyaphume Prutkraiwat, "Thai Handwritten Character Recognition by Euclidean Distance", *The 2nd International Conference on Digital Image Processing (ICDIP 2010)*, 2010, pp 53-58.

[13] Chomtip Pornpanomchai and Montri Daveloh, "Printed Thai Character Recognition by Genetic Algorithm", *The International Conference on Machine Learning and Cybernetics*, 2007, pp 3354-3359.

[14] Boontee Kruatrachue, Nattachat Pantrakarn and Kritawan Siriboon,"State Machine Induction with Positive and Negative for Thai Character Recognition", *The International Conference on Communications, Circuits and Systems*, 2007, pp 971-975.

[15] Parinya Sanguansat, Widhyakorn Asdornwised and Somchai Jitapunkul, "Online Thai Handwritten Character Recognition Using Hidden Markov Models and Support Vector Machines", *The International Symposium on Communications and Information Technologies*, 2004, pp.492-497.

[16] Rud Budsayaplakorn,Widhayakorn Asdornwised and Somchai Jitapunkul, "On-line Thai handwritten character recognition using hidden Markov model and fuzzy logic", *The IEEE 13th Workshop on Neural Networks for Signal Processing,*2003, pp. 537-546.

[17] E.K.Vellingiriraj, P.Balasubramanie, "Recognition of ancient Tamil handwritten characters in palm manuscripts using genetic algorithm", *International Journal of Scientific Engineering and Technology*, Volume 2 Issue 5, 2013, pp. 342-346.

[18] Kritawan Siriboon, Apirak Jirayusakul and Boontee Kruatrachue, "HMM topology selection for on-line Thai handwriting recognition", *The First International Symposium on Cyber Worlds*, 2002, pp. 142-145.

[19] Nabil Aharrane, Karim El Moutaouakil, Khalid Satori, "Recognition of handwritten Amazigh characters based on zoning methods and MLP", *WSEAS transactions on Computers,*Volume 14, 2015, Art. #19, pp. 178-185.

[20] Y. EsSaady, A. Rachidi, M. El Yassa, and D.Mammass: "Amazigh Handwritten CharacterRecognition based on Horizontal and Vertical Centerline of Character", *Inte rnational Journal of Advanced Science and Technology*, Vol. 33,2011, pp. 33-50.