

# Ancient Text Character Recognition Using Deep Learning

ShikhaChadha<sup>1</sup>, Dr. SonuMittal<sup>2</sup> and Dr. Vivek Singhal<sup>3</sup>

<sup>1</sup>Department of Computer and System Sciences, Jaipur National University, Rajasthan, India.

<sup>2</sup>Department of Computer and System Sciences, Jaipur National University, Rajasthan, India.

<sup>3</sup>Department of Information and Technology, JSS Academy of Technical Education, Noida, Uttar. Pradesh, India.

ORCID: 0000-0002-5664-8522 (Shikha Chadha)

## Abstract

Ancient scripts provide a captivating insight into the knowledge of ancestors which needs to be preserved for future generations. Therefore, there is a need to convert the digital script available in degraded format into textual format. To accomplish this model is being proposed in the paper that comprises of binarization using selection encoder decoder techniques. The results indicate the binarization accuracy as 74.24% approximately and F-measure is 75% (approximately) which comes out to be greater than other previously developed model. The binarized images are being further segmented using Seam Carbel method at character level and are manually compared with the vocabulary, the segmentation accuracy ( $A_s$ ) comes out to be 70% approximately. Further, characters are recognized using a three layer Convolutional Neural Network and the recognition accuracy ( $A_r$ ) is found to be 73% approximately, the recognized images are further converted into text using one to one mapping, to be further used for translation into universally acceptable language like English.

**Keywords:** Ancient Script, Convolutional Neural Network, Deep Learning, Image Segmentation, Machine Recognition System.

## I. INTRODUCTION

Historically the manuscripts were customized in the form of scrolls or books. The scripts could only be studied and understood by archeologists and historians. As the scripts which are available in the form of images, serves as a repository of knowledge that needs to be preserved for further generations. The quality degradation and complexity of ancient scripts due to aging and various climatic conditions proved to be an exhaustive challenge for researchers [1],[2] so there is a need of digitizing and deciphering the literature in order to preserve the knowledge present in script for future centuries.

Text recognition and feature extraction from ancient manuscript has been a great challenge due to various problems like ink bleed, faint ink strokes, background images and unwanted impurities[3]. It includes handwritten characters and extinct language so the motive behind the approach would be

to create a novel system which includes the Machine Recognition System (MRS) for converting the script in a text format which is universally acceptable so as, the information flowing in the form of Vedas and Upanishads being a course of attraction for the archaeologist and foreigners can be used [4].

The paper basically focuses on a model for ancient character text recognition and translating it into text that could be used further for translating into recognizable language[5]. Selection auto encoder decoder technique is used for the purpose of character level recognition [6].

The contribution of the paper is to perform image recognition of the low resource Sundanese language and convert it into text language and further convert it into universally accepted language. To accomplish the following things, need to be stated.

- To collect and analyze the ancient script images available in the palm leaf literatureformat.
- To pre-process i.e. Red Green Blue (RGB) to Grey scale, binarize and segment the ancient script images using adaptive threshold and san corbel method respectively.
- To recognize the segmented ancient script images into text format using Convolutional Neural Network(CNN).
- To Verify and Validate (V&V) the developed integrated ancient script text recognition system by performing comparative analysis with various existing system.

The proposed method for character recognition is compared on the basis of accuracy, F-measure, Peak-Signal to Noise Ratio (PSNR) and binarization time with previously developed methods like Otsu, Bernsen, Niblack, Wolf, Su, T. Singh etc.

The rest of paper is being organized as - In section II, some related research and findings of various character binarization methods are being discussed whereas in section III a short detail of data collection of the Sundanese text has been summarized, Section IV contextualizes this models with respect to experiment done using auto encoder-Decoder. Section 5 explores results and analysis done on the Sundanese dataset, and then enclosed by a conclusion and some futureworks.

## II. RELATEDWORK

The approach in [7] proposes Niblack based approach for implementing the filter for binarization. The efficiency achieved by the proposed technique is 85% approximately. Approach uses word Level and character level annotation is done in approach used in [8], which includes square coordinates that consists of coordinates of column-top-left, row-top-left, column-bottom-right, and row bottom-right. Ground truth binarized image is constructed using Otsu global thresholding method.

The approach used in [9] encompasses of three-layer architecture. First layer comprises of binarization by normalizing Red Green Blue (RGB), second layer consists of calculating Hue-Saturation-Value (HSV), further the background image estimation is done using direct subtraction and retrospective correction method and results are compared with various other methods. The Methodology suggested in [10] works in two phases, in which first phase neural network is used for recognition of isolated characters and second phase consist of recognition of word and character of variable length using Recurrent Neural Network (RNN) and CNN integrated technique.

The study done in [11] proposed a Khmer Character Recognition (KCR) system implemented with MATLAB environment using Self Organization Map (SOM) network. and multilayer feed-forward neural network using back propagation-learning. The recognition accuracy was achieved to 94.13%. The study done in [12] proposed a handwritten text recognition system on Khmer text using model which is divided into two stages firstly glyph class map generator is created in it class map is created using annotated information of glyph components, whereas in second module of the network the output from first module are encoded and transform it into a context vector. It is further decoded to produce the final word transcription.

The research proposed in [13] developed a dataset containing binarized image, ground truth dataset, word level annotated dataset which is publicly available for users. Otsu's global threshold is used for binarization. The technique proposed in [14] uses a hybrid model for the segmentation of line for 44 old Sundanese manuscripts. It uses the binarization free Seam Carving method and is able to separate small text located at bottom and the top of mail character, it is implemented on smallest energy function shown with evaluation matrix using which the accuracy has improved by 50%. The model proposed by [15] presents a scheme for Bi-directional Maximal Matching (BIMM) on Khmer clusters; it also focuses on Khmer word segmentation on both plain text and Microsoft Word document. For Word document, the implementation is done on currently active Word document and also on files one. The scheme compares the implementation of BIMM with Forward Maximal Matching (FMM) and Backward Maximal Matching (BMM) and with various similar algorithms. The result of accuracy found to be 98.13% whereas time spent was 2.581 seconds for Khmer contents.

In [16] proposed technique firstly transforms the width of the stroke for extracting connected components. Therefore, no of medial positions of text line are estimated using modified piece-wise projection profile technique Furthermore positions are modified adaptively according to the curvature of the actual text lines. Finally, a path finding approach is used finally to construct text line boundary for separating touching components and also to mark the boundary of the text lines. F-Measure result was estimated to be 92.92%. In [17] a model proposed uses Constrained Conditional Random Fields Model (CCRFM), with series of segmentation, POS tagging and Name Entity Recognition. In [18] proposed a Balinese character recognition system using RNN. Otsu's thresholding method is adapted to remove noise and convert it into grey scale. RNN faces the capacity to hold the data for short time so Bidirectional Long Short-Term Memory (BLSTM) is used for solving the problem of vanishing gradient problem, accuracy has been found to be 98.75%.

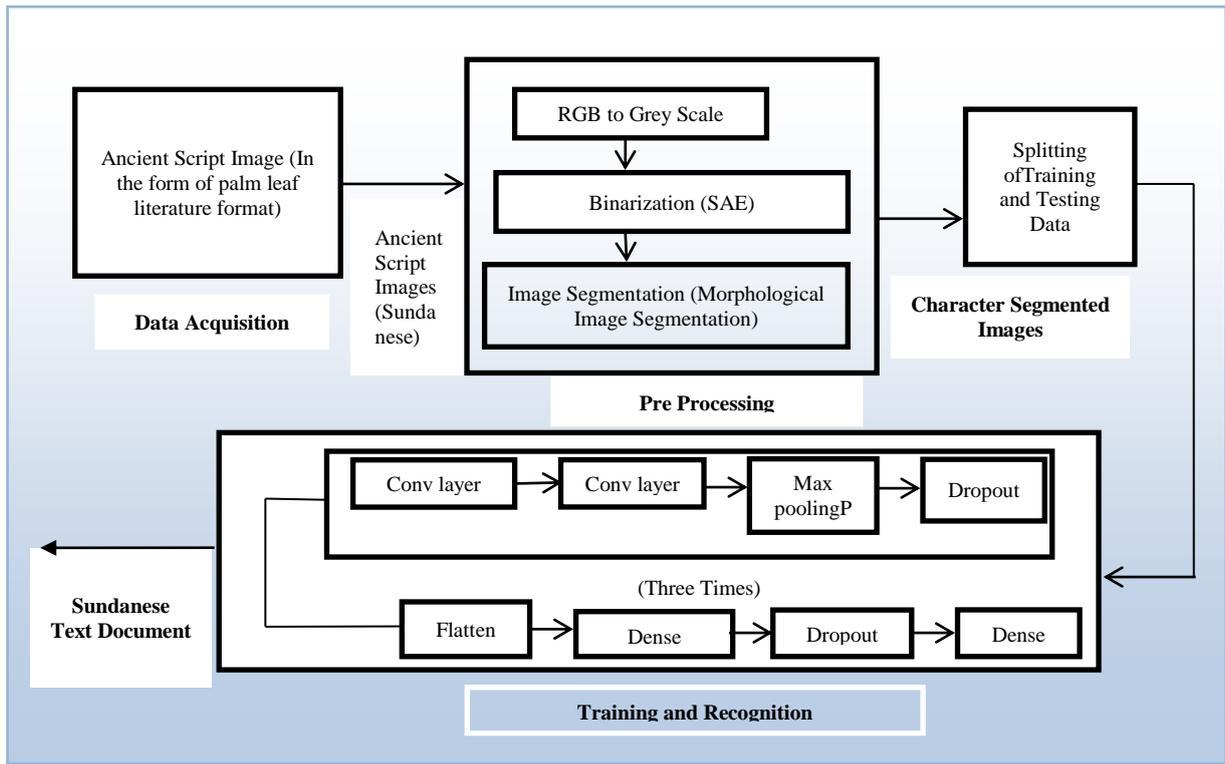
In [18] a technique for Balinese text segmentation is being proposed that includes pre- processing i.e converting it into grey scale after which text segmentation is done using Linear Discriminant Array (LDA) algorithm and the accuracy achieved is increased when compared. In technique [19] a method for historical document analysis has been proposed. Firstly, various neural networks are used like CNN, LSTM and RNN, secondly it focuses on word or text image of different length using both one and two dimensional RNN and error rate drops to 0.42. The method proposed in [20] uses a method by which the available Khmer text is firstly annotated which is further pre- processed using binarization and line, word, character, wise segmented. Three consecutive layer CNN model is used for recognition and the last layer is activated using ReLU activation function. The result into achieve an accuracy of 94.6 %.

## III. MATERIALS AND METHOD

As mentioned in problem statement, it is felt that a model needs to be developed for prevention and storing the culture and information present in the form of ancient palm leaf in degraded format, to be further used by future generations. The model consists of recognition part which comprises of

- Data acquisition that consists of secondary data collection from various sources.
- Pre-processing that needs to remove noise and convert it into grey scale then further segmenting the text in character level is done.
- Recognition module that recognizes the text into image and converts it into text format.

The network then recognizes the language by understanding the vocabulary of the language. The flow of the proposed recognition system is given in Fig 1.



**Fig1.** Architecture For Text Recognition

### III.I DataAcquisition:

In the proposed study, secondary data collection has been done. The Sunda dataset is used, consisting of character level

images of size 6.7 MB, word level annotated dataset of size 231 MB with 66 images of original Sudanese language [21] as shown in Fig 2, it consists of 66 classes with 27 consonants, 7 vowels and 10numerals.



**Fig. 2** Image of Character Level and Line level Dataset

### III.II Methodology

After Data acquisition the Image of ancient manuscripts have been collected, available in distress condition are pre-processed [22]. To preprocess the image it is firstly converted into grey scale so as to make it ready for digitization.

Once the image is converted into grey scale it needs to be binarized. For the proposed method Convolutional Auto Encoder (CAE) is used for binarization that encompasses of an activation functions used for differentiating between foreground and background images the binary value is donated to each pixel in the image. Once model is trained the document is binarized by passing through the model using adaptive threshold.

Adaptive threshold is used for image- to-image processing i.e label for each pixel is not computed independently but a taking an account label assigned to neighboring [23], [24] pixel as shown in algorithm.

1. Local *extrema* of  $S(t)$  are identified.
2. Incorporate the local maxima of the image for calculating the upper-envelope  $emax(t)$  and the local minima of image for getting lower-envelope  $emin(t)$ .

3. local mean (m) is calculated

$$M : \frac{emax(t) + emin(t)}{2} \quad (2)$$

4. Threshold function is calculated as

$$T = M + K.s \quad (3)$$

Where  $s$  is the standard deviation,  $K$  is parameter to tune and  $s$  is standard deviation.

Selection Auto Encoder (SAE) is used [25] as it is able to learn an end-to-end transformation to binarize the image. Image of a fixed size is given as an input and the model outputs a selection value for each pixel of the image depending on the confidence whether the pixel belongs to the foreground or background of the document. These values are eventually threshold to yield a discrete binary result. Auto-encoder consist of feed forward neural network considering that input and output shape is exactly same [26] and is divided in two functions  $f$  and  $g$  where they are encoder and decoder functions. The comparison between original Sundanese Palm leaf image and binarized image is shown in Fig.4

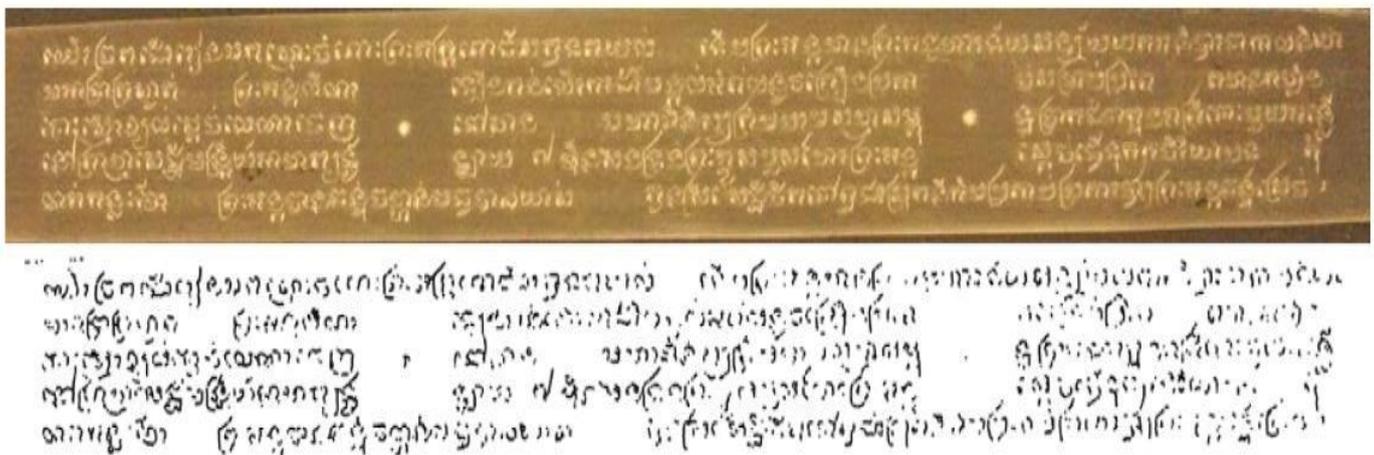


Fig 4 Binarized Image Using Selection Encoder Decoder Technique

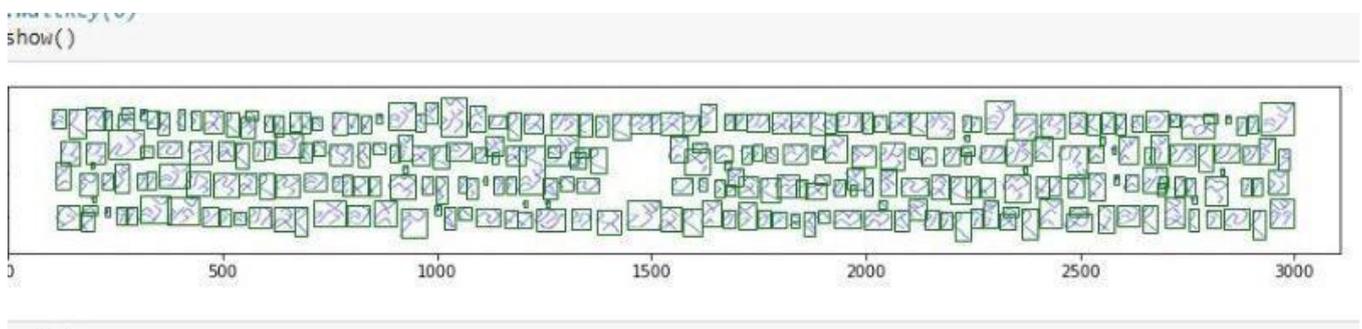


Fig 5 Segmented Image Using Seam Carbel Approach

The binarized images are segmented to change the image representation into the format, so as the image becomes more meaningful and can be easily analyzed. Image segmentation is used to locate objects and boundaries (lines, curves, etc.) in image. Basically, the process of assigning a label to every pixel in an image is the objective behind image segmentation, such that pixels with the same label share certain characteristics.

The proposed study uses morphological segmentation that partitions an image based on the topographic surface of the image. The image is separated into various non-overlapping regions with each region containing a unique particle [27] as shown in Fig. 5. The approach computes medial seams by splitting the input page image into columns whose smoothed projection profiles are then calculated. The positions of the medial seams are obtained based on the local maxima locations of the profiles. The goal of the second stage of the approach is to compute separating seams with the application on the energy map within the area restricted by the medial seams to retarget the size of the image, with preserving the prominent content. The technique carves paths that traverse the image from left to right. The path with the minimum cumulative energy is then chosen.

Thereafter, the segmentation and recognition is done using three layers of CNN. The proposed CNN framework consists of three convolutional layers, maxpooling layer, and dropout layer, flatten layer and dense layer as shown in Table.1.

**Table 1:** Summary of Training Model

Layer (Type)	Output Shape	Parameters
Conv2d_1 (Conv2D)	(None.32,32,32)	895
Conv2d_2 (Conv2D)	(None.30,30,32)	9248
max_pooling2d_1 (MaxPooling2)	(None.15,15,32)	0
dropout_1 (Dropout)	(None.15,15,32)	0
Conv2d_3(Conv2D)	(None.15,15,64)	18496
Conv2d_4(Conv2D)	(None.13,13,64)	36928
max_pooling2d_2 (MaxPooling2)	(None. 6,6,64)	0
dropout_2 (Dropout)	(None. 6, 6, 64)	0
Conv2d_5(Conv2D)	(None. 6, 6, 64)	36928
Conv2d_6(Conv2D)	(None. 4, 4, 64)	36928
max_pooling2d_3 (MaxPooling2)	(None. 2, 2, 64)	0
dropout_3 (Dropout)	(None. 2, 2, 64)	0

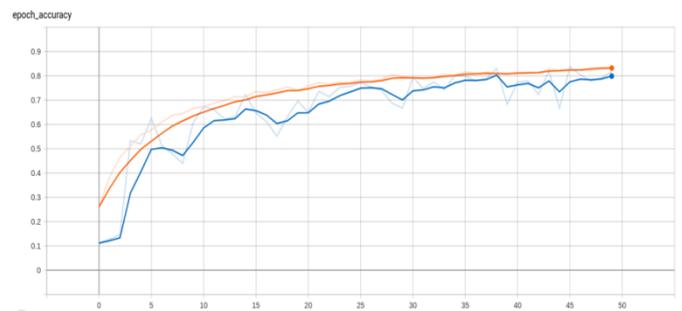
flatten_1(Flatten)	(None. 256)	0
dense_1(Dense)	(None. 512)	131584
dropout_4(Dropout)	(None. 512)	0
dense_2(Dense)	(None. 10)	5130

The recognition model consists of three layers with each layer consisting of two convolutional layers, max pooling layer and one dropout layer. The image is firstly passed in convolutional layer two times. Convolutional layer uses set of detectable features to be applied on the filters after which max pooling layer is used to reduce the spatial size of the image as shown in Table 1 which is further passed to dropout layer for preventing over fitting.

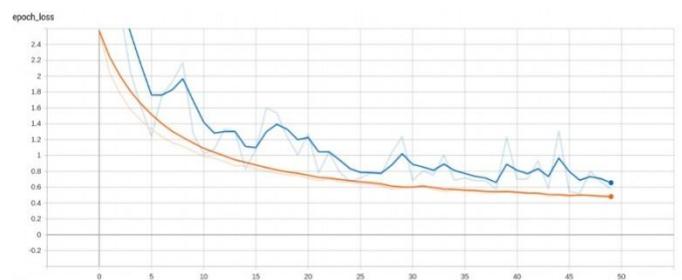
Thereafter, the third layer the output image is passed to the flatten layer. It specifies a function mapping from the given filters to a vector so as the errors can be back propagated through convolutional layers. Thereafter the image is again passed to dropout layer in order to avoid over fitting. Finally, it is again passed to the last layer of model i.e. flatten layer to map the condensed filters to the vectors.

There are total 28 characters used in the dataset with 200 image each for every charter so the total dataset consist of 5600 characters which is further split into training and testing with the 3920 characters are used for training and 1680 for testing. The model has been trained on 50epochs.

The training and validation loss as shown in Fig 6 indicates the rise of accuracy with increase in number of epochs. Whereas the Fig 7 indicates the decrease in training and validation loss as the number of epochs increases.



**Fig. 6** Graph Showing Training and Validation Accuracy with 50 Epochs.



**Fig. 7** Graph Showing Training and Validation Loss with 50 Epochs.

#### IV. RESULT AND CONCLUSION

Once the Images are binarized the precision and recall is being calculated using equation and compared with various methods as shown in Table 2.

$$Precision = \frac{T_p}{T_p + F_n} \quad (4)$$

Where  $T_p$  is the total number of images that are true positive and  $F_n$  is number of True positive and  $F_n$  is number of False negative. Whereas the Recall is calculated using equation.

$$Recall = \frac{T_p}{T_p + F_p} \quad (5)$$

Where the  $T_p$  is no of True positive and  $F_p$  is number of false positive.

**Table 2:** Comparison Table of Precision and Recall

Method	Otsu	Niblack	Sauvola	SAE
<b>Precision</b>	0.57	0.51	0.59	0.71
<b>Recall</b>	0.63	0.58	0.55	0.69

$$Recall = \frac{T_q}{T_d} \quad (6)$$

The binarization accuracy, F-measure, PSNR, Binarization time is being calculated and compared with other previously techniques like Otsu, Niblack, Bernsen etc as shown in Table 2. *F-measure* is shown is calculated using equation 1 given below.

$$F - Measure = \frac{(2 * Precision * Recall)}{(Precision + Recall)} \quad (7)$$

Where  $T_q$  is majority as associated images and  $T_r$  is number of related images retrieved and  $T_d$  is number of total related images.

**Table 3:** Comparison Table of Precision and Recall

Method	Accuracy	F-Measure	PSNR	Binarization time
Otsu	67.4893	62.4618	16.0021	5 sec
Bernsen	65.3366	54.6576	13.3129	15 sec
Niblack	68.4544	64.5834	16.9420	5 sec
Sauvola	67.6766	63.2703	16.3388	6 sec
Wolf	67.6808	67.6808	16.3466	9 sec
N.I.C.K	67.6099	63.0482	16.2159	10 sec
Su	65.4219	54.9573	13.3931	30 sec
T.R Singh	65.9315	56.9001	13.9057	8 sec
Bataineh	67.0797	61.2912	15.3458	9 sec
Sauvola	67.0204	61.6551	15.2584	7 sec
Wan	67.1711	62.0681	15.4839	12 sec
SAE	74.2458	75.57	17.6964	3 min

Since there is no pre-developed software to test the accuracy for the segmentation ( $A_s$ ) of Sudanese script, the accuracy had to be calculated manually by checking each character on the script  $N_t$  and then checking if the character was segmented correctly  $N_s$ . On checking the accuracy, it comes out to be 70%.

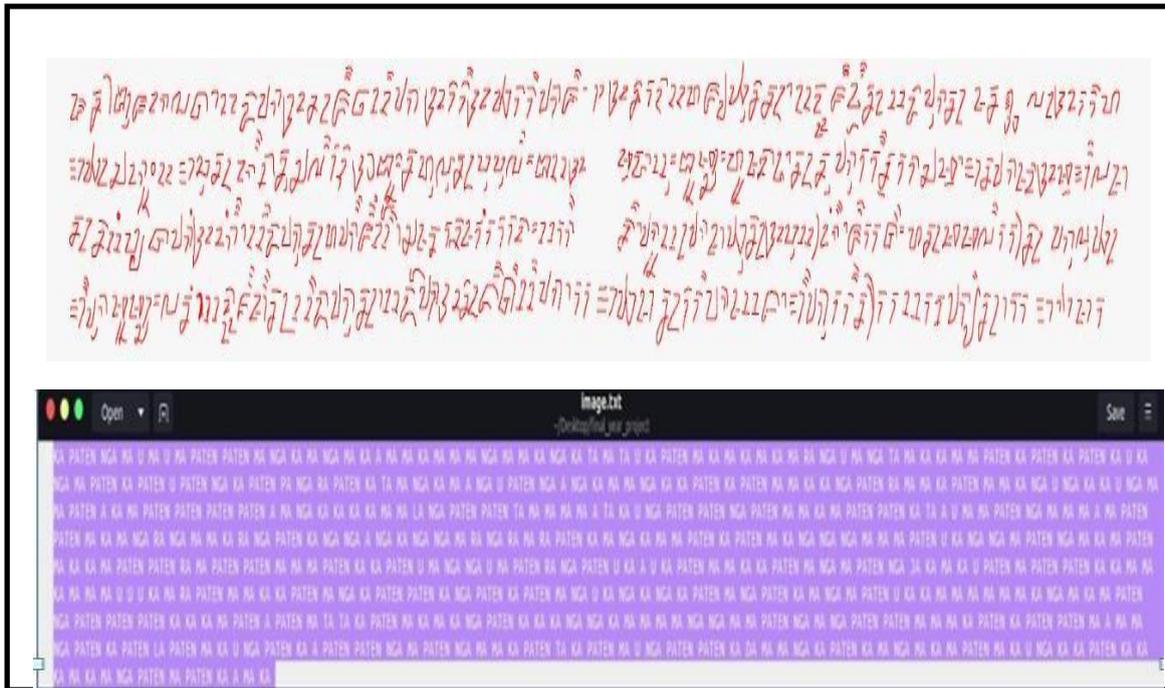
$$A_s = \frac{N_s}{N_t} * 100 \quad (8)$$

The CNN model used for the recognition of the characters gave the accuracy of 73% (approx.). The testing and the

validation accuracy had to be calculated manually as there's no pre-developed software for Sudanese script and the accuracy came out to be 68%.

$$A_r = \frac{N_r}{N_t} * 100 \quad (9)$$

where  $A_r$  is recognition accuracy and  $N_r$  is no of recognized images and  $N_t$  is the total number of images used. After the recognition the image is converted in text by mapping with 27 classes and is converted into text file as shown in Fig 8.



**Fig.8** Recognized Image Converted Into Text File

## V. CONCLUSION

In this paper character level recognition approach is proposed for Sudanese text written on palm leaf. The script available is present in degraded format, needs to be preserved and converted into text format for future generations. The approach used in the paper proposes an encoder decoder technique for binarization and results shows binarization accuracy as 74.24% and F-measure to be 75% that exceeds the other previously developed model. Further character level segmentation of the images are done using Seam Carbel method manually compared with the vocabulary, the segmentation accuracy  $A_s$  comes out to be 70%. Further, characters are recognized using three layers Convolutional Neural Network and the recognition accuracy comes out to be 73%. The recognized images use one to one mapping for further translation into recognizable text i.e. English.

In future the work may be conducted on multilingual ancient scripts recognition system, as the approach proposed is limited to one language i.e. Sundanese.

## REFERENCES

- [1] E. Paulus, M. Suryani, S. Hadi, and I. N. Yulita, "A novel scheme for handwritten binarization method on sundanese palm leaf document images," *Proc. 2nd Int. Conf. Informatics Comput. ICIC 2017*, vol. 2018-January, pp. 1–5, 2018, doi:10.1109/IAC.2017.8280586.
- [2] J. Calvo-Zaragoza and A. J. Gallego, *A selectional auto-encoder approach for document image binarization*, vol. 86. Elsevier Ltd, 2019.
- [3] V. Amrutha Raj, R. L. Jyothi, and A. Anilkumar, "Grantha script recognition from ancient palm leaves using histogram of orientation shape context," *Proc. Int. Conf. Comput. Methodol. Commun. ICCMC 2017*, vol. 2018-January, no. Iccmc, pp. 790–794, 2018, doi:10.1109/ICCMC.2017.8282574.
- [4] E. Paulus, M. Suryani, S. Hadi, R. Sopian, and A. Hidayat, "Binarization and Segmentation Framework for Sundanese Ancient Documents," *J. Sains Dasar*, vol. 6, no. 2, pp. 133–142, 2017, doi:10.21831/j.sainddasar.v6i2.15314.

- [5] Suryani.M ,Paulus.E, Hadi.S, Darsa.U and Burie (2017), "The Handwritten Sundanese Palm Leaf Manuscript Dataset From 15th Century", 14th IAPR International Conference on Document Analysis and Recognition, p.p 796- 800.
- [6] S. Chadha, S. Mittal, and V. Singhal, "An insight of script text extraction performance using machine learning techniques," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 1, pp. 2581–2588, 2019, doi: 10.35940/ijitee.A5224.119119.
- [7] E. Paulus, M. Suryani, and S. Hadi, "Improved Line Segmentation Framework for Sundanese Old Manuscripts," *J. Phys. Conf. Ser.*, vol. 978, no. 1, 2018, doi: 10.1088/1742-6596/978/1/012001.
- [8] E. Paulus, M. Suryani, S. Hadi, and F. Natsir, "An initial study to solve imbalance sundanese handwritten dataset in character recognition," *Proc. 3rd Int. Conf. Informatics Comput. ICIC 2018*, pp. 1–6, 2018, doi: 10.1109/IAC.2018.8780496.
- [9] S. Hadi, U. A. Darsa, E. Paulus, and M. Suryani, "Quality enhancement of degraded sundanese lontar images using direct subtraction and retrospective correction methods in several color space," *Proc. 2nd Int. Conf. Informatics Comput. ICIC 2017*, vol. 2018-January, pp. 1–6, 2018, doi: 10.1109/IAC.2017.8280632.
- [10] H. Meng and D. Morariu, "Khmer character recognition using artificial neural network," *2014 Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA 2014*, 2014, doi:10.1109/APSIPA.2014.7041824.
- [11] D. Valy, M. Verleysen, and S. Chhun, "Text recognition on Khmer historical documents using glyph class map generation with encoder-decoder model," *ICPRAM 2019 - Proc. 8th Int. Conf. Pattern Recognit. Appl. Meth.*
- [12] M. W. A. Kesiman, S. Prum, I. M. G. Sunarya, J. C. Burie, and J. M. Ogier, "An analysis of ground truth binarized image variability of palm leaf manuscripts," *5th Int. Conf. Image Process. Theory, Tools Appl. IPTA 2015*, pp. 229–233, 2015, doi:10.1109/IPTA.2015.7367134.
- [13] Narin Bi and NguonlyTaing2014, "Khmer word segmentation based on Bi- directional Maximal Matching for Plaintext and Microsoft Word document", *Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, p.p1-9.
- [14] Sudarma.M, Ariyani.S and Artana.M (2016), "Balinese Script's Character Reconstruction Using Linear Discriminant Analysis", *Indonesian Journal of Electrical Engineering and Computer Science*, Vol-2,p.p479-485.
- [15] D. Valy, M. Verleysen, S. Chhun, and J. C. Burie, "A New Khmer palm leaf manuscript dataset for document analysis and recognition – SleukRith set," *ACM Int. Conf. Proceeding Ser.*, pp. 1–6, 2017, doi: 10.1145/3151509.3151510.
- [16] J. Bai, Z. Chen, B. Feng, and B. Xu, "Image character recognition using deep convolutional neural network learned from different languages Institute of Automation, Chinese Academy of Sciences , Beijing 100190 , China," pp. 2560–2564, 2014.
- [17] A.A.Suryani,I.Arieshanti,B.W.Yohanes,M.Subair,S.D.B udiwati,andB. S. Rintyarna, "Enriching English into Sundanese and Javanese translation list using pivot language," *Proc. 2016 Int. Conf. Inf. Commun. Technol. Syst. ICTS 2016*, pp. 167–171, 2017, doi: 10.1109/ICTS.2016.7910293.
- [18] K. Cho et al., "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *EMNLP 2014 - 2014 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf.*, pp. 1724–1734, 2014, doi: 10.3115/v1/d14-1179.
- [19] S. Bin Ahmed, S. Naz, M. I. Razzak, R. Yusof, and T. M. Breuel, "Balinese character recognition using bidirectional LSTM classifier," *Lect. Notes Electr. Eng.*, vol. 387, pp. 201–211, 2016, doi:10.1007/978-3-319-32213-1\_18.
- [20] M. H. Mohamed Dyla and F. Morain-Nicolier, "Text line segmentation and binarization of handwritten historical documents using the fast and adaptive bidimensional empirical mode decomposition," *Optik (Stuttg.)*, vol. 188, no. April 2018, pp. 52–63, 2019, doi:10.1016/j.ijleo.2019.04.128.
- [21] H. Fishwick, "Three leaves eaten for the Holy Sacrament," *Notes Queries*, vol. s4-IX, no. 211, p. 39, 1872, doi:10.1093/nq/s4-IX.211.39-a.
- [22] A. P. Vartak, V. Mankar, A. Prof, and H. V. P. M. College, "Morphological Image Segmentation Analysis," *Int. J. Comput. Sci. Appl.*, vol. 6, no. 2, pp. 161–165, 2013.
- [23] M. W. A. Kesiman et al., "Southeast Asian palm leaf manuscript images: a review of handwritten text line segmentation methods and new challenges," *J. Electron. Imaging*, vol. 26, no. 1, p. 011011, 2016, doi: 10.1117/1.jei.26.1.011011.
- [24] A. Issac, M. ParthaSarathi, and M. K. Dutta, "An adaptive threshold based image processing technique for improved glaucoma detection and classification," *Comput. Methods Programs Biomed.*, vol. 122, no. 2, pp. 229– 244, 2015, doi:10.1016/j.cmpb.2015.08.002.
- [25] An improved infrared image processing method based on adaptive threshold denoising," *Eurasip J. Image Video Process.*, vol. 2019, no. 1, 2019, doi: 10.1186/s13640-018-0401-8.
- [26] S. Wang, Z. Ding, and Y. Fu, "Feature selection guided auto-encoder," *31st AAAI Conf. Artif. Intell. AAAI 2017*, pp. 2725–2731, 2017.
- [27] J. Calvo-Zaragoza and A. J. Gallego, "A selectional auto-encoder approach for document image binarization," vol. 86. Elsevier Ltd, 2019.
- [28] M. W. A. Kesiman et al., "Benchmarking of document image analysis tasks for palm leaf manuscripts from southeast Asia," *J. Imaging*, vol. 4, no. 2, 2018, doi:10.3390/jimaging4020043.