International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 08, 2020
ISSN: 1475-7192

Metallic Monuments Image Character Recognition System Using Deep Learning Algorithm

[1]R.Indra Gandhi , [2] Dr.M.Ponnavaikko

ABSTRACT*---India is known for its wealthy cultural heritage even before prehistoric amount. Recognizing ancient Tamil* ***characters*** *adapt archaeologists to reveal historical events in Cholas amount in twelfth century with vast efforts of the anthropology specialists. The literature were best-known from numerous sources, like inscriptions on temple walls, rocks, pillars, engravings in caves, paintings, copper plates, literature in palm leaves and reports of foreign travelers etc., to write in code their writings. The longer term analysis within the field of archeology can have negative impact thanks to unskillfulness of the manual procedure. Optical Character Recognition practicality is employed to acknowledge ancient Tamil Inscription. OCR module of application i.e. advanced Extreme Deep Learning Machine is especially targeted during this paper. during this paper, we tend to propose a posh extreme Deep learning machine algorithmic rule that this adds some hidden layers to the first ELM network structure, at random initializes the weights between the input layer and therefore the 1st hidden layer also because the bias of the primary hidden layer, utilizes methodology to calculate the parameters of the hidden layers and eventually uses the smallest amount square method to calculate the output weights of the network. The following calculation may be a feature extraction Scale Invariant Feature remodel algorithmic rule to observe and describe native options in input image.*

Keywords*--Copper Plate character recognition (CPCR), Region Of Interest (ROI), Complex Extreme Deep Learning* ***Machine*** *(CEDLM), Extreme Learning Machine (ELM), American Standard Code for Information Interchange (ASCII)*

I. INTRODUCTION

Epigraphically in-scripted monuments plays a mile stone in knowing the civilized past. Many a civilization was recognized solely by the record of data that they left behind with the assistance of their best lingual potential. Copper Plate Character acknowledgment upgrades the handling of copper plate footage by allowing you to naturally understand and disencumber content substance from varied info fields. Tamil, the linguistic communication of a southern state in India has many million speakers across the globe and is an officer language in countries like Srilanka, Malaysia & Singapore. Tamil has twelve vowels and eighteen consonants. These area unit combined with one another to yield 216 composite characters and one special character (aayatha ezhuthu) counting to a complete of (12+18+ 216+1) 247 characters. Vowels in Tamil area unit of 2 varieties like short (kuril)

[1] *Research Scholar , Bharath Institute of Higher Education and Research, Chennai , Tamil Nadu, shambhavi.rajesh@gmail.com*

[2] *Pro-Vice Chancellor, Bharath Institute of Higher Education and Research, Chennai , Tamil Nadu*

Received: 21 Dec 2019 | Revised: 18 Jan 2020 | Accepted: 05 Feb 2020

and long (Nedil) and it's conjointly referred to as as Uyir Ezhuthu. Consonants area unit of 3 varieties like Vallinam, Idaiyinam, and Mellinam.

II. RELATED WORKS

Bhattacharya et al. [1] proposed a two phase approach. In the primary stage an unaided bunching technique was applied to create a base measure of gatherings of written by hand Tamil character classes. In the second stage a regulated characterization procedure was thought of in all of those very little gatherings for definite acknowledgment. The amount of transistion and therefore the chain code histogram are the highlights utilized in the first and second stages one by one. Indra Gandhi et al proposed another methodology of utilizing Kohonen SOM (Self Organizing Map) for perceiving the online Tamil character [2]. The vectors of the twofold image area unit created. At the purpose once the division of the character is finished, at that time the images area unit scaled to exceptional tallness and weight. Some undesirable segments area unit incorporated, nevertheless it tends to be exhausted by sobel edge identification. The center channel is employed to expand the productivity. The SOM isn't material to the cursive characters which are utilized right now. Jagadeesh Kannan et al [4] utilised Octal Graph technique for the acknowledgment of the Tamil hand written characters. Here, the character come back on the positional notation diagram's pel is modified over into the hub of the chart. Each hub has eight fields, that's the rationale referred to as positional notation chart. Each hub is related to the opposite hub keen about the sting esteem. The image is modified over to the positional notation chart by the means that, as an example, standardization, transformation, Identification of gauging elements and highlight extraction. On the off likelihood that the character is repetitive and on the off likelihood that it contains various bends, at that time positional notation diagram technique is not cheap. All of those works largely focus on perceiving the Tamil characters by utilizing distinctive classification. However the planned work understand the old Tamil characters and anticipate the amount by utilizing TSVM to urge the next exactitude

III. IMPLEMENTATION

11th century written by hand Tamil contents are measure often assembled into four categories to be specific Vowels, Consonants, Composite characters and Aydham. These four categories are measure taken for characterization reason. Customary calculations are way more slow than needed in light-weight of the actual fact that the slope based mostly learning calculation and also the parameters should be tuned iteratively. This paper propose associate algorithmic program named complex hidden layers extreme learning machine (CEDLM). The structure of the CEDLM (select the three-hidden-layer ELM for example) is illustrated in Figure 3 . The work flow of the three-hidden-layer ELM is illustrated in Figure.

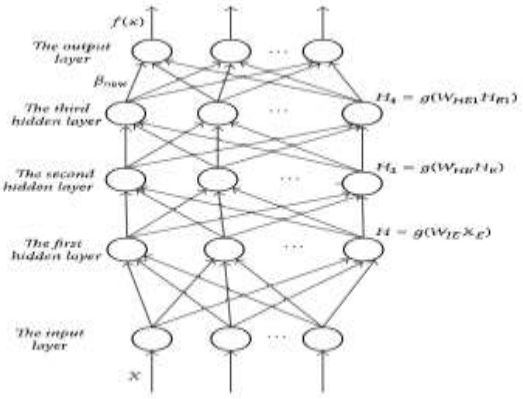

Figure 3: Structure of the three-hidden-layer ELM

Algorithm:.

Step 1: Assume the training sample dataset is $\{XT\}=(x_i,t_i)$ $(i=1,2,3,\ldots..,Q)$, where the matrix X is the input samples and the matrix T is the labeled samples. Each hidden layer has l hidden neurons with the activation function g(x).

Step 2: Randomly initialize the weights between W the input layer and the first hidden layer as well as the bias B of the first hidden neurons

$$W_{IE} = [B \;\; W], X_E = [1 \;\; X]^T.$$

Step 3: Calculate the equation $H = g(W_{IE}X_E).$.

Step 4: Calculate the weights between the hidden layers and the output layer β= $(I/\lambda + H^T H)^{-1}H^T T$ or $\beta = H^T(I/\lambda + HH^T)^{-1}T$.

Step 5: Calculate the expected output of the second hidden layer $H_1 = T\beta^+$.

Step 6: According to formulae (12)–(14) and the algorithm steps (4, 5), calculate the weights between the first hidden layer and the second hidden layer and the bias B_1 of the second hidden neurons $W_{HE} = g^{-1}(H_1)H_E^+.$.

Step 7: Obtain and update the actual output of the second hidden layer $H_2 = g(W_{HE}H_E).$.

Step 8: Update the weights matrix β between the hidden layer and the output layer $\beta_{new} = (I/\lambda + H_2^T H_2)^{-1}H_2^T T$ or $\beta_{new} = H_2^T(I/\lambda + H_2 H_2^T)^{-1}T$.

Step 9: If the number of the hidden layer is three, we can calculate the parameters by recycle executing the above operation from step to step . Now β_{new} is expressed as follows $\beta_{new} = \beta, \; H_E = [1 \;\; H_2]^T$.

Step 10: Calculate the output $f(x) = H_2\beta_{new}.$.

If the number K of the hidden layer is more than three, recycle is executing step 5 to step9 for (K-1) times. All the H matrix (H_1,H_2) must be normalized between the range of −0.9 and 0.9, when the max of the matrix is more than 1 and the min of the matrix is less than −1.

Received: 21 Dec 2019 | Revised: 18 Jan 2020 | Accepted: 05 Feb 2020

IV. EXPERIMENTAL RESULTS

To check the productivity and legitimacy of the proposed framework, the framework was tried for precision and the outcomes were contrasted and the aftereffects of past frameworks on similar databases. So on make sure the nature of images, all photos were gathered from continuous copper plate photos taken by us. and what is more from completely different sources. Getting ready sets contained in more than seventy getting ready photos and in more than fifty testing tests from CPI-A01 [5].

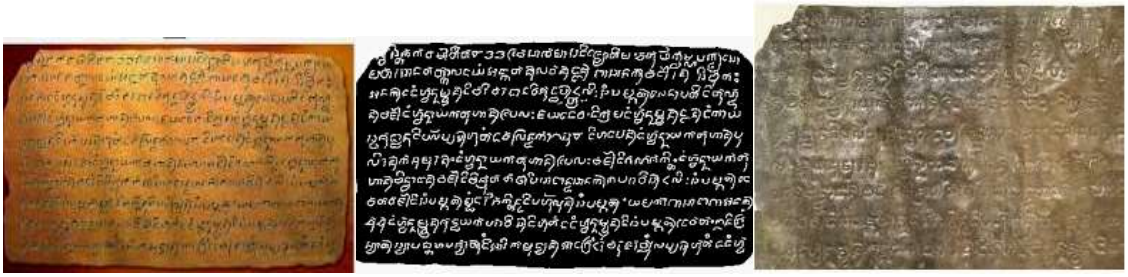

Fig.4: Sample Copper Plate images

The individual contents CPI-A01 information were portioned physically to isolate them into words. Tutorial courses were twenty five distinctive Tamil words in varied sizes, directions, clamour degrees and matter designs. The examinations were led on associate AMD Quad-center, a pair of GHz processor, four GB DDR3 Ram system and Windows8.1 operating framework. The code was written in MATLAB language utilizing MATLAB 2011Rb programming. The projected Copper plate character recognizing technique is tried on the images of varied copper plate engravings gathered from completely different items of Madras, India. Copper plate engravings created by completely different lines that managed over Madras show extraordinary highlights in their vogue regarding the type of stone, cleaning, piece of content, recording on copper plate with shading, etching the content on copper plate, and what is more passionate about matters of raising the copper plate in an exceedingly appropriate spot. a substantial ton of those engravings ar disintegrated therefore gravely that it's exhausting to tell apart the vital data, particularly once the surface is consumption or inscribed. Attributable too many years of decay, dominant a part of these archaic writings are in poor condition, and various content bits are as of currently absent. The hurt went on to such a degree, that either the items do not exist, or segments are nevermore conspicuous and past healing. The presentation aftereffect of the character recognizing method on such photos is likewise declared straight away. At first, the part dataset contains full fourteen highlights, the preparation dataset is of measurements 102x14 and therefore the testing dataset is of measurements 22x14. the 2 sets are applied to EDLM and CEDLM. The classification performance criteria of the issues ar the typical classification accuracy of the testing information. Within the Figure5, the average testing classification correct share for the ELM, EDLM and CEDLM formula is shown clearly

Table 1: RMSE values CPI-A01 dataset

S.NO	Algorithm	Training RMSE	Testing RMSE
1.	Extreme Learning Machine (ELM)	$8.6233E{-}9$	$1.0797E{-}8$
2.	Extreme Deep Learning Machine (EDLM)	$1.8428E{-}6$	$1.5204E{-}5$

3.	Complex Extreme Deep Learning Machine (CEDLM)-Proposed System		
	$f_1(x)$	$8.722E{-}15$	$1.3055E{-}14$
	$f_2(x)$	0.0011	0.0019
	$f_3(x)$	0.2110	0.4177

V. CONCLUSION

Copper Plate Optical Character Recognition (CPOCR) for composed content is an extremely testing and open territory of research. This paper built up a copper plate Tamil OCR for manually written words dependent on a blend of the Complex Extreme Deep Learning Machine (CEDLM) classifier with Multi hidden Layer Feed Forward Network and factual based component determination. Toward the start, the framework utilized a 14 highlights dataset. At that point information was sustained into EDLM organize, which is a quick and basic multi concealed layer feed forward system (MLFN). The framework accomplished high acknowledgment precision of 92.05% for various examples in an extremely brief timeframe. The CEDLM calculation acquires the qualities of customary ELM that haphazardly introduces the loads and predisposition (between the information layer and the main covered up layer), also embraces a piece of the ELM calculation, and utilizations the backwards initiation capacity to compute the loads and inclination of concealed layers (with the exception of the primary shrouded layer). At that point we make the real concealed layer yield inexact to the normal shrouded layer yield and utilize the parameters got above to figure the real yield. In the capacity relapse issues, this calculation decreases the least mean square blunder. In the datasets order issues, the normal precision of the various arrangements is fundamentally higher than that of the ELM and EDLM organize structure. In such cases, the CEDLM can improve the presentation of the system structure. Resultant data gives scope for value addition to hhistorical *copper monuments vital data* to revive and preserve for our future followers.

REFERENCE

1. Vikas J Dongre, Vijay H Mankar(2011),"Devanagri Document Segmentation using Histogram Approach", "International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.1, No.3, pp.68-73.
2. Rafael C. Gonzalez, Richard E. Woods (2007), "Digital Image Processing", 2nd ed., Beijing: Publishing House of Electronics Industry.
3. W. Zhu, J. Miao, L. Qing, and G.-B. Huang, "Hierarchical Extreme Learning Machine for unsupervised representation learning," in Proceedings of the International Joint Conference on Neural Networks,IJCNN2015, Ireland, July2015.
4. Jagadeesh Kannan R and Prabhakar R, "An improved Handwritten Tamil Character Recognition System using Octal Graph", Int. J. of Computer Science, ISSN 1549-3636, Vol 4 (7): 509-516, 2008
5. R. Moreno, D. Puig, C. Julia, and M. Garcia, ―A new methodology for evaluation of edge detectors,‖ in Proceedings of the 16th IEEE International Conference on Image Processing (ICIP), 2009, pp. 2157–2160.
6. H. Voorhees and T. Poggio,‖ Detecting textons and texture boundries in natural images‖ ICCV 87:250-25,198. S.E.Umbaugh, ―Computer Imaging : Digital Image Analysis and Processing‖, CRC Press, 2005

7. A Gayathri, A Srinivasan , “An efficient algorithm for image denoising using NLM and DBUTM estimation”, TENCON 2014-2014 IEEE Region 10 Conference, Page No.1-6, 2014.
8. A Gayathri, A Srinivasan ,” Moving object detection by fuzzy aggregation using low rank weightage representation”, Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014 pp 335-342.
9. A.Gayathri, S.Christy, “Image de-noising using optimized self similar patch based filter”, International Journal of Innovative Technology and Exploring Engineering, 8(12), pp. 1570-157, 2019