

Intelligent Feature Extraction for Handwritten Tamil Mantra Recognition in Slogans by using PCA based Deep Convolutional Neural Network



M. Sornam, C. Vishnu Priya, Poornima Devi. M

Abstract: Offline handwritten identification of characters is a core problem in pattern matching. The main challenge for researchers in the identification of handwritten characters is inscribing individual styles. Tamil mantras identification is a challenging job due to many missing features in the mantras ' complex structure. The problem at hand is to break traditional hand-designed features. A new venture has been undertaken to automatically extract the complex features for recognition and classification from the complex structure bypassing the individual Tamil characters into the Convolutional Neural Network, a special type of deep learning network. The best convolutional model was chosen to improve efficiency by comparing different convolutional models that vary in activation functions, classifiers, and pooling functions. Principal Component Analysis (PCA) was used to select the top n eigenvectors from the image for better efficiency. So with the above trained best model with PCA for independent Tamil character images, handwritten Tamil fonts in the slogans (a group of characters) have been well recognized.

Keywords: Convolutional Neural Network, Deep Learning, Mantras, PCA, Pattern Recognition, Slogans, Tamil Characters.

I. INTRODUCTION

Machine learning is a computer science field that focuses on choosing a model of a network that trains a machine in the best way to perform as a human being [1]. Besides selecting a network model, the big challenge facing conventional neural networks is the extraction of features. It gives the developer a big challenge in complex problems including recognition of artefacts and recognition of handwritten characters. Therefore, for the above complex problems, the developer must choose accurate function information such that it helps the system make an effective decision.

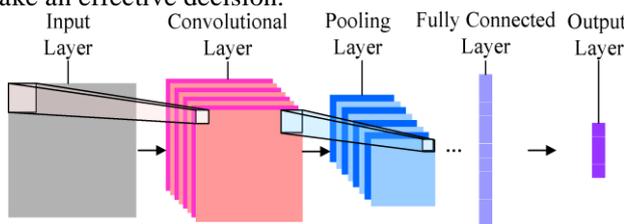


Fig.1. Sample architecture of the Convolutional Neural Network.

A deep learning technique is a state-of-art in machine learning models, which overcomes the problem and reduces a programmer's burden.

Revised Manuscript Received on April 13, 2020.

* Correspondence Author

M. Sornam *, Department of Computer Science, University of Madras, Chennai, India. Email: madasamy.sornam@gmail.com

C. Vishnu Priya, Department of Computer Science, University of Madras, Chennai, India. Email: priyachengal12506@gmail.com

Poornima Devi. M, Department of Computer Science, University of Madras, Chennai, India. Email: poornima160492@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

With little programmer guidance, the model used in deep learning is capable of extracting the correct features by itself. Thus this theory of deep learning brought machine learning into a new era. With this benefit, it has put its footprints in various real-time applications, including Deep face from Facebook, SIRI from Apple, Microsoft Cortana and DeepMind from Google. This has resulted in deep learning as an active research area where researchers focus primarily on making effective use of these models and advancing artificial intelligence to human level. A Deep Learning (DL) algorithm includes various networks, including Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), the Deep Belief Network (DBN), Deep Boltzmann Machine (DBM) and Auto Encoder (AE).

Convolutional Neural Network (CNN) is mostly used for image classification [2] and pattern recognition among those networks. The Convolutional architecture of the neural network consists of a layer of convolution, a pooling layer followed by a fully connected layer.

A sample of architecture of the Convolutional Network appears in Fig. 1. The input layer holds an image's raw pixel and then passes it on to the convolutional layer that applies the precise amount of convolution filters to the image [3, 4]. Then the output is passed to the ReLU activation function. After that, the output from the convolutional layer is downsampled within the pooling layer to reduce the feature map's dimensionality. One of the most important fields of machine learning is handwritten character recognition. A pen and paper are more convenient than a keyboard in case of many situations. Yet identifying this handwritten character is a most challenging task for Pal and Chaudhuri (2004), because it differs from individual persons to persons and also the character varies because of different moods of the same person [5]. The automation of character recognition would be added advantageous for visionless persons to read the letters, to robotically recognize postal documents, cheque, banknotes, license plate identification, etc.[6-8]. Several strategies were developed for identifying the character written in paper and palm leaf [9]. Thus, identification of character provided its specific research study in pattern recognition.

Tamil is one of the world's oldest South Indian languages. Currently, in countries such as India, Sri Lanka and Malaysia, Tamil languages were in use. So the Tamil language is the language that millions of people speak the most. There are 30 basic characters in Tamil alphabets which include 12 vowels and 18 consonants.



Also, besides the Tamil language character contains one ayudhaezhuthu and 6 Grantham characters. The vowel and consonant combinations get 216 alphabets. So in the Tamil language, the total number of characters is 247 except 6 Grantham characters as shown in Fig. 2.

This article is systematized as follows: The existing Tamil character recognition work is discussed in Section 2, Proposed work is elaborated in Section 3, Experimental results and discussion were presented in Section 4, followed by a conclusion and future work in Section 5.

ஃ	அ	ஆ	இ	ஈ	உ	ஊ	எ	ஏ	ஐ	ஓ	ஔ	ஔள
க	கா	கி	கீ	கு	கூ	கே	கை	கொ	கோ	கெள		
ங	ஙா	ஙி	ஙீ	ஙு	ஙூ	ஙே	ஙை	ஙொ	ஙோ	ஙெள		
ச	சா	சி	சீ	சு	சூ	சே	சை	சொ	சோ	செள		
ஞ	ஞா	ஞி	ஞீ	ஞு	ஞூ	ஞே	ஞை	ஞொ	ஞோ	ஞெள		
ட	டா	டி	டீ	டு	டூ	டே	டை	டொ	டோ	டெள		
ண	ணா	ணி	ணீ	ணு	ணூ	ணே	ணை	ணொ	ணோ	ணெள		
த	தா	தி	தீ	து	தூ	தே	தை	தொ	தோ	தெள		
ந	நா	நி	நீ	நு	நூ	நே	நை	நொ	நோ	நெள		
ப	பா	பி	பீ	பு	பூ	பே	பை	பொ	போ	பெள		
ம்	மா	மி	மீ	மு	மூ	மே	மை	மொ	மோ	மெள		
ய	யா	யி	யீ	யு	யூ	யே	யை	யொ	யோ	யெள		
ர்	ரா	ரி	ரீ	ரு	ரூ	ரே	ரை	ரொ	ரோ	ரெள		
ல்	லா	லி	லீ	லு	லூ	லே	லை	லொ	லோ	லெள		
வ	வா	வி	வீ	வு	வூ	வே	வை	வொ	வோ	வேள		
ழ	ழா	ழி	ழீ	ழு	ழூ	ழே	ழை	ழொ	ழோ	ழெள		
ள்	ளா	ளி	ளீ	ளு	ளூ	ளே	ளை	ளொ	ளோ	ளெள		
ற்	றா	றி	றீ	று	றூ	றே	றை	றொ	றோ	றெள		
ன்	னா	னி	னீ	னு	னூ	னே	னை	னொ	னோ	னெள		

Fig.2. Tamil characters without Grantham character.

II. EXISTING METHODS OF TAMIL CHARACTERS

Siromoney et al.[10] described a model for the recognition of Tamil characters through computer printed material. The method involves transforming each character into a binary matrix and extracting the features by scanning row and column-wise. And those characteristics have been encoded to identify the characters.

In a Genetic algorithm, Vellingiriraj et al.[11] proposed to identify the Tamil handwritten characters in palm leaves script. Binarization and thickening preprocess were performed at the beginning of the cropping process in the images followed by the extraction process of features that extracted vertical, horizontal, and zigzag strokes and loops.

Sundar et al.[12] suggested the identification of printed characters in Tamil script using the features Histogram of Oriented Gradients (HOG). Such features were characterized using a backpropagation classifier followed by the Fisher Linear Discriminant Analysis (FLDA) classifier to overcome the backpropagation classifier character misclassification.

Aparna et al. [13] described the identification of Tamil printed document characters. First, the input images were correctly skewed and then the characters were segmented into individual characters using horizontal and vertical projections. Then the character traits were extracted from each character and then the images were divided into three groups. The characters were categorized at the first level based on height and classified at the second level based on extensions, and recognition was done at the final level based on Euclidian distance and neighboring K-Nearest respectively.

In combining CNN and Support Vector Machine (SVM) classifier for the recognition of MNIST dataset digits, Niu et al.[14] had given a good concept. Here the CNN structural design was used to extract the features, and

these features were defined by replacing the last layer in CNN with SVM using a radial basis kernel. The method for recognizing digits and English alphabets was suggested by Subbuthai et al.[15] by mining the features using Principal Component Analysis (PCA) and classifying the digits and alphabets using Euclidean distance.

Using the IWFHR-10 dataset, Vijayaraghavan et al.[16] have proposed CNN architecture for Tamil character identification. With spatial pooling and local contrast normalization, the traditional convolutional neural network design was enhanced here. The testing was carried out using stochastic gradient descent with an adaptive learning rate, and classification was carried out using the Softmax classifier.

Many techniques have been suggested based on the identification of a handwritten character [17-21]. Chinese, Arabic, English, Numerical digits were effectively categorized using convolutional neural networks [22,23]. Yet there is no attempt to identify Tamil characters by integrating PCA and CNN using separate convolutional models on this specific dataset.

Even, based on these qualified models the Tamil mantras in the slogans have been recognized. So the paper focuses on applying the aforementioned concept and effectively obtaining the outcome.

III. PROPOSED WORK

A. Dataset

The HPLabs dataset[24] includes Tamil characters written by various regional writers that include school students, graduates, and adults from India's various cities such as Bangalore, Karnataka, and Tamil Nadu. Those data were collected in standard UNIPEN format using HP TabletPCs. It includes versions offline as well as online. This dataset has been used for competition for the Tamil character identification in IWFHR 2006. Of a large number of samples, only 47,160 HPLabs [24] samples have been used as a network reference. These were brought to a standard size of 100x 100 pixels, because the pictures were in a different size.

B. Principal Component Analysis

Principal Component Analysis is a statistical method by which data is simplified by extracting best features. Such features highlight the input image similarities and differences that are important for classifying the images [15]. It is also used in the reduction of dimensionality by ignoring the image's irrelevant characteristics. Eigenvectors were obtained from the covariance matrix dimensions n. The most important features of PCA are the primary p eigenvector of the n eigenvalues. Since the covariance matrix's p-eigenvector points to the direction of maximum variance, and so on, these p-eigenvector function as a vector of data features.



PCA Algorithm:

Imagine an n-dimensioned grayscale image X.

Step 1: The picture is translated to nx1-size vectors using(1).

$$i, k=1,2,3,\dots n \quad (1)$$

$$X_k = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{bmatrix} \quad \text{Step 2: Calculate the vector image mean by (2).}$$

$$X_{ak} = \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{bmatrix} \quad i, k=1,2,3,\dots n \quad (2)$$

Step 3: Image vector is converted into a column vector and then standardized by using (3) to find the difference between a mean of an image vector and the original image.

$$X = (X_1 - X_{a1} \quad X_2 - X_{a2} \quad \dots X_n - X_{an}) \quad (3)$$

Step 4: Calculation of the covariance matrix Q is performed using (4).

$$Q = XX^T \quad (4)$$

Step 5: Eigenvalues and eigenvectors are calculated using (5) and then the eigenvalues are arranged in increasing order. Now choose top p eigenvalues and corresponding eigenvectors to reconstruct the images.

$$[Q - \lambda I]x = 0 \quad (5)$$

Step 6: Reconstruct the image using top p eigenvectors respectively.

From the above steps, the PCA algorithm is implemented on each Handwritten Tamil character images of standard size 100x100 pixels. The image is reconstructed using eigenvectors, with an interval of 10 at each iteration. As shown in Table 1, after 4 iterations the image formed using the top 40 eigenvectors has less noise. So this image is passed to the network for automatic feature extraction and classification.

Table-I: PCA Image

Images	Eigenvector value
	Original Image
	Eigenvalue=10
	Eigenvalue=20
	Eigenvalue=30
	Eigenvalue=40

C. Feature Extraction and Training

The extraction of features is the biggest problem in handwritten recognition of Tamil characters. The output of the feature quality is directly proportional to machine

decision. The best features will get the machine to take a specific decision. So the process of extraction of the function was done automatically by the convolutional neural networks [25,26,16] to solve the complication.

The proposed convolutional neural network used 37,728 training samples, 9,432 validation samples, and 1000 test samples, respectively. The proposed architecture contains one input layer, one output layer, two convolution layers, two pooling layers and three fully connected layers, totally nine layers.

The Tamil character image dataset's extracted PCA features were resized to 40x 40. The resized images were regularized to [-1, 1] and passed as an input to the stack with various types of layers.

According to the basic model, the initial layer is the convolution layer which contains 20 features detectors from 5x5 size filters and to bring non-linearity among the feature detectors [27], a non-linear activation function called ReLU is exercised to each feature maps. Now max-pooling under 2x2 windows is performed in the next layer in demand to shorten the dimensionality of the feature maps. Convolution (Conv) is again performed to extract 32 feature maps using a 3x3 filter followed by max-pooling, as described in the previous layer of max-pooling (MP). Such image matrices have been flattened and transferred for classification onto the fully connected (FC) layers. Table 2 shows the model's description.

To increase the efficiency different models were chosen and trained for classification. Table 3 details the System Description II, in which LeakyReLU replaced the activation function ReLU in the basic model and trained for 50 epochs.

Similarly, Table 4 shows the Model summary III in which the activation function in basic function is replaced with Leaky ReLU and then it is passed into Softmax classifier for classification among Tamil characters.

At last, Table 5 exhibits Model IV in which the model is trained with Average Pooling (AP), Leaky ReLU activation function and Softmax Classifier and the accuracy results are noted to choose the best model among four models including basic model. The above-said models were trained with the learning rate 0.0001, along with the batch size 5. On the whole, the parameters used in the model were shown in Table 6. So among these models, the best model is chosen based on the efficient recognition rate and the weights of the model were saved. With the trained weights, the characters in the slogans were detected. Here the characters are written in the paint with each character 40x40 in a pixel area. Therefore the slogans with n characters occupy 40x40xn pixels. The slogans were segmented into the individual character of 40x40 pixels. Segmented characters were passed into the network for identification. Slogans from Tirukkural and Naaladiyar were taken for testing as shown in Fig. 3 and Fig. 4. After training the individual characters and testing the slogans the model has obtained an accuracy of 89.85% approximately.

Table-II: Model Summary I

Layer	Filter Size	#Filters	Activation	Output
Conv1	5x5	20	ReLU	36x36x20
MP2	2x2	-	-	18x18x20
Conv3	3x3	32	ReLU	16x16x32
MP4	2x2	-	-	8x8x32
FC5	-	-	Tanh	(None,2048)
FC6	-	-	Tanh	(None,146)
FC7	-	-	Sigmoid	(None,146)

Table-III: Model Summary II

Layer	Filter Size	#Filters	Activation	Output
Conv1	5x5	20	Leaky ReLU	36x36x20
MP2	2x2	-	-	18x18x20
Conv3	3x3	32	Leaky ReLU	16x16x32
MP4	2x2	-	-	8x8x32
FC5	-	-	Tanh	(None,2048)
FC6	-	-	Tanh	(None,146)
FC7	-	-	Sigmoid	(None,146)

Table-IV: Model Summary III

Layer	Filter Size	#Filters	Activation	Output
Conv1	5x5	20	Leaky ReLU	36x36x20
MP2	2x2	-	-	18x18x20
Conv3	3x3	32	Leaky ReLU	16x16x32
MP4	2x2	-	-	8x8x32
FC5	-	-	Tanh	(None,2048)
FC6	-	-	Tanh	(None,146)
FC7	-	-	Softmax	(None,146)

Table-V: Model Summary IV

Layer	Filter Size	#Filters	Activation	Output
Conv1	5x5	20	Leaky ReLU	36x36x20
AP2	2x2	-	-	18x18x20
Conv3	3x3	32	Leaky ReLU	16x16x32
AP4	2x2	-	-	8x8x32
FC5	-	-	Tanh	(None,2048)
FC6	-	-	Tanh	(None,146)
FC7	-	-	Softmax	(None,146)

Table-VI: Parameters

Parameter	Values
Learning Rate	0.0001
Batch Size	5
Optimizer	Adam

அகரகுதவாருக்கெவ்வாம் ஆதி
பகவன்முதற்பே உவகு

Fig.3. A Slogan from Tirukkural.

நிரவரும் என்வண்ணி நவ்வறி வாளர்
குழவி மடந்தே துறந்தார் யரைதீரா
மள்ளா இளமை யிகுந்தாரே கோவ்வன்றி
இன்னாங் கெழுந்தீர்ப் பார்

Fig.4. Slogan from Naaladiyar

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The network extracts the features automatically by going through different intermediate layers and displaying the output in Fig. 5. The accuracy curve for the proposed network was obtained by training various models has been shown in Figs. 6, 7, 8 and 9 and these curves prove that the Leaky ReLU activation function with Softmax classifier had given better results with the small number of epochs.

The characters from the slogans were segmented as shown in Fig. 10 and tested with the models which gave better accuracy. Table 7 shows the results of character detection for the slogan from Tirukkural.

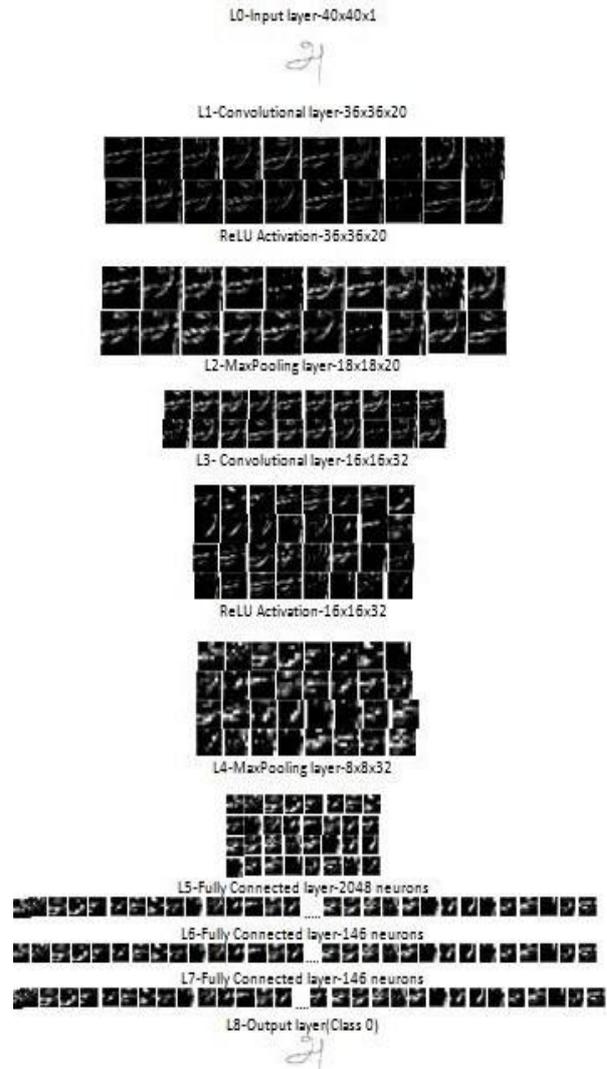


Fig.5. Network architecture with feature maps.

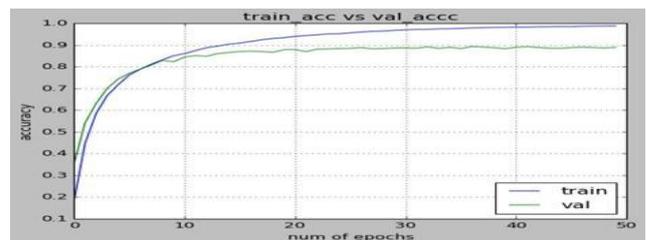


Fig.6. Accuracy curve of network model I with Max Pooling, ReLU activation function and Sigmoid Classifier.

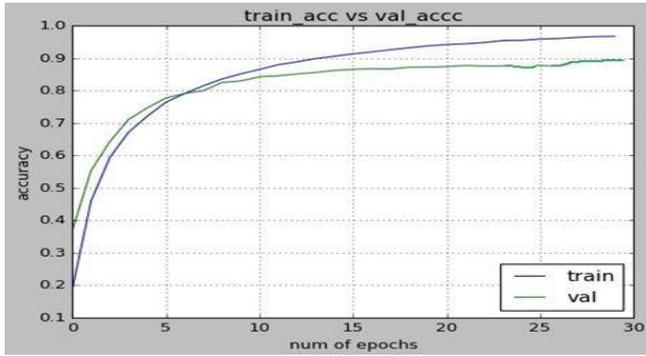


Fig.7. Accuracy curve of network model II with Max Pooling, Leaky ReLU activation function and Sigmoid Classifier.

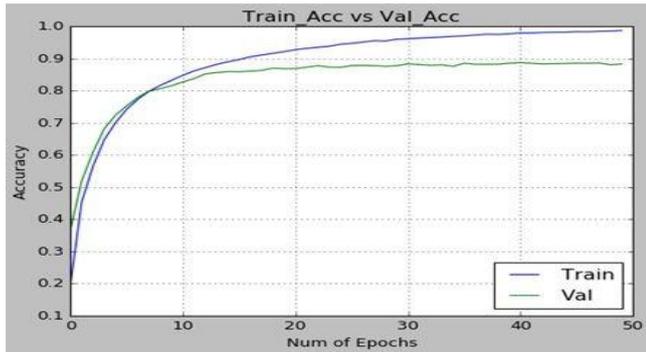


Fig.8. Accuracy curve of network model III with Max Pooling, Leaky ReLU activation function and Softmax Classifier.

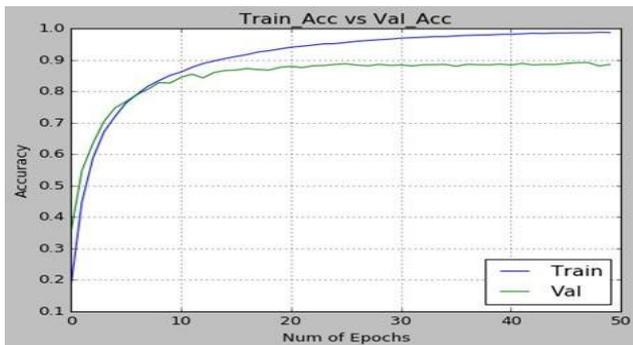


Fig.9. Accuracy curve of network model IV with Average Pooling, Leaky ReLU activation function and Softmax Classifier.



Fig. 10. A sample segmented character from slogan Tirukkural

Table VII- Slogan Results of Tirukkural

Parameter	Values
Total characters	29
Positive detection	26

Apart from the error and accuracy curve, the most commonly used performance metrics to check the model performance is shown in Table 8. To calculate the metrics, the True Positive Rate (TPR), True Negative Rate (TNR),

False Positive Rate (FPR), False Negative Rate (FNR) [28], the confusion matrix plays the basic role.

Table VIII- Performance metrics formulae

Metrics	Formula
Accuracy	$\left(\frac{TPR + TNR}{TPR + FPR + FNR + TNR} \right)$
Precision	$\left(\frac{TPR}{TPR + FPR} \right)$
Recall	$\left(\frac{TPR}{TPR + FNR} \right)$
F1 Score	$\left(\frac{2 * (Recall * Precision)}{Recall + Precision} \right)$

Table IX- Comparison of performance measures with proposed work

Metrics	Accuracy	Precision	Recall	F1 Score
CNN – Chinese characters [29]	90%	-	-	-
CNN-GRU-MNIST [30]	99.21%	-	-	-
CNN-Tamil characters [31]	95.16%	94%	95%	95%
PCA MP ReLU (Proposed)	88.82%	88%	88%	88%
PCA MP LeakyReLU (Proposed)	87.23%	86%	87%	86%
PCA MP LeakyReLU Softmax (Proposed)	89.85%	89%	88%	88%
PCA AP LeakyReLU Softmax (Proposed)	87%	88%	87%	87%

Using the above-mentioned values, model output was estimated and the average result for 146 classes in Table 9 illustrates that model PCA performance with Max pooling, Leaky ReLU, Softmax classifier delivered better results with 89.85% accuracy.

V. CONCLUSION AND FUTURE WORK

The proposed research presented a Tamil character recognition system by hybridizing PCA and convolutional neural networks using various architecture models that detected the slogans with better performance for the offline IWFHR-10 datasets.

This hybrid combination made the classification of the data set fair, and the performance measures were encouraging. The proposed model PCA with Max Pooling, LeakyReLU, Softmax classifier achieved 89.85% accuracy. The work has been implemented in Intel(R) Pentium(R) CPU B950 @ 2.10GHz for the device with 2 GB RAM and processor. So this work can be expanded further by increasing the optimization by combining the evolutionary algorithm with GPU use.

REFERENCES

1. Sornam. M. and Poornima Devi. M, "A survey on backpropagation neural network," *IJCNES* 5, 2016, pp. 70–74.
2. Sornam. M, Kavitha. M. S. and Vanitha. V, "A survey on image classification and activity recognition using deep convolutional neural network architecture," *IEE, ICoAC*, 2017, pp.121–126.
3. Goodfellow. I.J, Bulatov. Y, Ibarz. J, Arnaud. S. and Shet. V, "Multi-digit number recognition from street view imagery using deep convolutional neural networks," *Pattern Recognition*, 2014, pp. 147-159.
4. Sornam. M. and Shalini. R., "A survey on selected algorithm-evolutionary, deep learning, extreme learning machine," *IJCA*, 5, 2016, pp. 50–54.
5. Sundaram. S. and Ramakrishnan. A.G, "An improved online Tamil character recognition engine using post-processing methods," *ICDAR-IEEE*, 2009, pp. 1216–1220.
6. Feng. B, Ren. M, Zhang. X. and Suen. C, "Automatic recognition of serial numbers in bank notes," *Pattern Recognition*, Vol.48, Issue-8, 2014, pp. 2621-2634.
7. Thadchanamoorthy. S, Kodikara. N.D, Premaretne. H.L, Pal. U. and Kimura. F. "Tamil handwritten city name database development and recognition for postal automation," *IEEE-ICDAR*, 2013, pp. 793–797.
8. Sornam. M, Kavitha. M.S. and Poornima Devi. M, "An efficient morlet function based adaptive method for faster backpropagation for handwritten character recognition," *ICACA-IEEE*, 2016, pp.135–139.
9. Sornam. M. and Poornima Devi. M, "Tamil palm leaf manuscript character segmentation using GLCM feature extraction," *IJCSE* 6, 2018, pp. 167–173.
10. Siromoney. G, Chandrasekaran. R. and Chandrasekaran. M, "Computer recognition of printed Tamil characters," *IEEE*, 1978, pp. 243–247.
11. Vellingiriraj. E.K. and Balasubramanie. P, "Recognition of ancient Tamil handwritten characters in palm manuscripts using genetic algorithm," *IJSET*, 2, 2013, pp. 342–346.
12. Sundar. K.A. and John. M, "A high precision printed character recognition method for Tamil script," *IEEE ICCIC*, 2013, pp.41–49.
13. Aparna. K.G. and Ramakrishnan. A.G, "A complete Tamil optical character recognition system," In: Lopresti D, Hu J, Kashi R. (eds) Document Analysis Systems V. DAS 2002, *Lecture Notes in Computer Science*, Springer, 2002.
14. Niu. X. and Suen. C, "A novel hybrid CNN-svm classifier for recognizing handwritten digits," *Science Journal of Pattern Recognition*, 2012, pp. 1318–1325.
15. Subbuthai. P, Periasamy. A. and Muruganand. S, "Identifying the character by applying PCA method," *IJCA* 60, 2012, pp. 8–11.
16. Vijayaraghavan. P. and Sra. U, "Handwritten Tamil recognition using a convolutional neural network," *MIT Media Lab*, 2009, pp. 156-161. doi:web.media.mit.edu/~sra/tamil_cnn.pdf.
17. Kussul. E. and Baidyk. T, "Improved method of handwritten digit recognition tested on MNIST database," *Science Journal of Image and Vision Computing* 22, 2004, pp. 971–981.
18. Sornam. M. and Vishnupriya. C, "A comparison of deep learning neural networks for image processing applications," *IJETA*, 4, 2017, pp.60–64.
19. Sornam. M. and Manimozhi. S, "License plate detection and character recognition system," *ICCCMIT*, 2014, pp. 28–31.
20. Kannan. R.J. and Subramanian. S, "An adaptive approach of Tamil character recognition using deep learning with big data-a survey," In: Satapathy S, Govardhan A, Raju K, Mandal J(eds) Emerging ICT for Bridging the Future - Proceedings of the 49th Annual Convention of the Computer Society of India (CSI), *Advances in Intelligent Systems and Computing*, Springer, 2015.
21. George. C, Podhumani. S, Monish. K.R. and Pavithra. T, "Survey on handwritten character recognition using artificial neural network," *IJETA*, 10, 2016, pp. 287–290.
22. Ciresan. D. and Meier. U, "Multi-column deep neural networks for offline handwritten Chinese character classification," *IEEE*, 2015, pp. 2161–4407.

23. Razzazi. F. and Ghassemian. M, "A very high accuracy handwritten character recognition system for Farsi/Arabic digits using convolutional neural networks," *IEEE*, 2010, pp. 1585–1592.
24. HPLabs, "Hplabisolated handwritten Tamil character dataset", 2010, doi:http://shiftright.com/mirrors/www.hpl.hp.com/india/research/penhw-resources/tamil-iso-char.html.
25. Alwzawy. H.A, Albehadili. H.M, Alwan. Y.S. and Islam. N.E, "Handwritten digit recognition using convolutional neural networks," *IJRCCCE*, 4, 2016, pp. 1101–1106.
26. Sornam. M. and Vishnupriya. C, "Deep convolutional neural network for handwritten Tamil character recognition using principal component analysis," *Smart and Innovative Trends in Next Generation Computing Technologies*, NGCT 2017, *Communications in Computer and Information Science*, Springer 827, 2018, pp. 778–787.
27. Karun Verma and Sharma. R. K, "An efficient writing-zone identification technique for Online Handwritten Gurumukhi character recognition," *Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci.* 88 (2), 2018, pp. 297–307.
28. Poornima Devi. M. and Sornam. M, "Systematic classification of historical handwritten Tamil palm leaf manuscript using CART algorithm and RBF network," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8(11), 2019, pp. 185-194.
29. Qiuyun Hao, Xiaoming. Wu, Sen Zhang, Peng Zhang, Xiaofeng Ma and Jingsai. Jiang, "Research on Offline Handwritten Chinese Character Recognition Based on Deep Learning," *ICIST*, 2019, pp. 470-474.
30. Vantruong Nguyen, Jueping Cai and Jie Chu, "Hybrid CNN-GRU Model for High Efficient and written Digit Recognition," *ACM Proceedings of the Second International Conference on Artificial Intelligence and Pattern Recognition*, 2019, pp.66-71.
31. Kavitha. B.R. and Srimathi. C, "Benchmarking on offline Handwritten Tamil Character Recognition using convolutional neural networks," *Journal of King Saud University – Computer and Information Sciences*, 2019, pp. 1- 8.

AUTHORS PROFILE



Dr. M. Sornam obtained her masters degree in Mathematics from the RIASM, University of Madras in the year 1987, masters in Computer Applications from the University of Madras in the year 1991 and her PhD in the year 2013 from the University of Madras. She started her career as a Lecturer in Computer Science at Anna Adarsh College, Chennai, since 1991. Later, since 1996 to 2000 she worked as a Lecturer in Computer Science at T.S. Narayanasami College of Arts and Science, Chennai. Since 2001, she has been working in the Department of Computer Science, University of Madras. At present, she is working as a Professor in Computer Science at University of Madras. Her area of interest includes Artificial Intelligence and Artificial Neural Network, machine learning. Deep learning, Natural language processing, Image Processing applications, Data Mining, Pattern Recognition and its Applications.



C.Vishnu Priya pursued her Masters in Computer Science with distinction and first rank at University of Madras, India in the year 2016 and Master of Philosophy at University of Madras, India in the year 2017. Her area of research includes Artificial Intelligence, Artificial Neural Network, Image Processing, Deep Learning Neural Network



Poornima Devi. M pursued her Bachelor of Science at Soka Ikeda College for Women affiliated to University of Madras, India in the year 2013, Master of Science at Queen Mary's College for Women affiliated to University of Madras, India in the year 2015 and Master of Philosophy at University of Madras, India in the year 2016. She is currently pursuing research at the Department of Computer Science, University of Madras, India. Her main research focuses on Artificial Intelligence, Artificial Neural Network, Image Processing, Natural language processing and Deep Learning Neural Networks