IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 7, Issue 4, Aug - Sept, 2019
ISSN: 2320 – 8791 (Impact Factor: 2.317)
www.ijreat.org

Ancient Kawi Character Recognition Using Deep Learning

Abdulrazak Yahya Saleh[1], Sadia Afrin[2]

[1] **Faculty of Cognitive Science & Human Development, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia**

[2] **Faculty of Cognitive Science & Human Development, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia**

Abstract

In document analysis and recognition, recognizing characters is still an important problem. Document processing is gaining popularity in day by day in the field of pattern recognition. This study has been conducted to recognize Kawi character which is an ancient script of Malaysia by deep learning. As recently, deep learning has achieved more attention because of the ability to learn from raw data in character recognition. The Kawi character database has been developed for this research. Convolutional neural network models have recently proved an impressive classification performance towards recognizing characters. So, the Kawi character recognition system has been developed using CNN. The system has the capability to extract features and accuracy improvement. This research achieved 89% accuracy using CNN model.

Keywords: *Artificial Neural Network, Character Recognition, Convolutional Neural Network, Deep Learning, Kawi Script.*

1. Introduction

One of the most research applications in pattern recognition is character recognition [1]. Machine simulation of human reading considered as character recognition [2]. To identify characters by computer programs without the input given by human is the main aim of character recognition [3].

As mentioned earlier, character recognition is a popular and wide field of research and many ancient scripts have been recognized such as English script [4], Arabic script [5], Farsi script [6], Brahmi script [7], Bangla script [8], Malayalam script [9], Telegu script [10] etc. As individual has different handwriting styles, after sufficient research still open research problems remain [11]. However, also many old scripts exist which are not yet recognized. Kawi script is one of the scripts among them which had been used in Malaysia for long period. Due to this, many important documents were written in this script by Malaysians. Recognizing Kawi script will help to discover and understand many unrevealed facts about Malaysian history.

Brahmi script was introduced and spread into northern and southern during first millennium. After Brahmi, Grantha script started to spread and became popular. Malayalam script, Tamil script and Pallava script derived from Grantha script. However, the parent of Kawi script was Pallava [12]. Fig. 1 shows the origin of the Kawi script whereas Fig. 2 shows the Vowels and consonants of Kawi script.

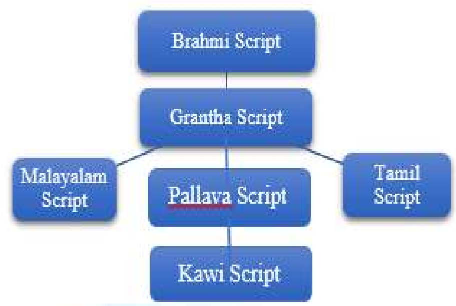

Fig. 1: Origin of Kawi script

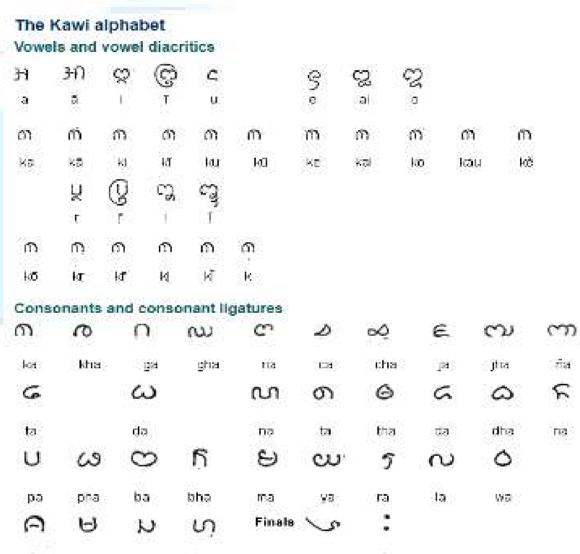

Fig. 2: The Vowels and consonants of Kawi script

This is the only Austronesian language whose earlier forms are attested in written record in the shape of an extensive corpus of texts. This is considered as the great importance of Kawi. Kawi is the ancestor language of modern Javanese. The script is also called 'Old Javanese'; however, this term is rarely used. Kawi was used for writing Sanskrit and Old Javanese. Kawi is descended from southern Brahmi through Pallava. The development of Kawi may be classified into two broad phases: 'Early Kawi' (c.750–925 CE) and 'Later Kawi' (c.925–1250 CE). Kawi is the ancestor of modern Indonesian writing systems, such as Javanese and Balinese, as well as Batak, Buginese, Rejang, Sundanese, and related scripts. Kawi script made its first appearance in Laguna Copperplate Inscription which is a legal document dated AD 804 and with that it was able to reach a rich output of creative and re-creation of Sanskrit originals [13].

2. Related Work

Many research has been done such as Brahmi character recognition [7], Arabic character recognition [5], Chinese character recognition [14], Bengali character recognition [8], Devanagari character recognition [15], Tamil character recognition [16] to make computer recognized the ancient scripts. As mentioned earlier, many researches have been conducted for ancient languages. However, few researches have been conducted for Kawi script when the origin of Kawi script which is Brahmi script and related script such as Malayalam, Tamil script are already recognized. So, this research is recognized the characters of Kawi script.

Deep learning has been applied because of its outstanding performance towards character recognition [8,15,17-18] to recognize Kawi script. The aim of Deep learning is to predict from many types of data such as images, sounds, and biological data [19]. Deep learning maybe loosely defined as an attempt to train a hierarchy of feature detectors – with each layer learning a higher representation of the preceding layer [18]. Computer vision has received a tremendous impact in the performance by deep learning on many fields such as image recognition [20-22], object localization, pose estimation, object tracking, objects detection, image segmentation or image captioning. Many deep learning methods such as Convolutional Neural Network, Recurrent Neural Network (RNN), and Deep Belief Network have received great attention on the classification tasks [23]. Among many networks of deep learning, CNN is the most used network for handwritten character recognition [24]. Many systems and algorithms have been proposed for better accuracy and overcome the limitations of recognizing the character. Fig. 3 presents shows the effectiveness of other character recognition system and Deep leaning.

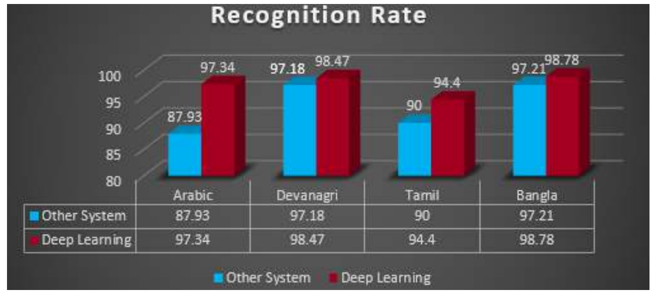

Fig. 3: Accuracy of other character recognition system [25,26] and Deep leaning [8].

Based on that, Convolutional Neural Network has been chosen to conduct this research due to high usages and accuracy. Convolutional neural network has leading accuracy in terms of character recognition [27]. CNN can understand image characteristics comparing to other models which is the main reason of the success of CNN [24]. In other words, to create connections of the image data structure, CNN is appropriate network [24].

3. Methodology

3.1 Architecture of Convolutional Neural Network

A typical CNN consists of convolutional layer, spatial pooling layer and full-connected layer. Extracting features from feature maps at lower layer is the responsibility of convolutional layer [28]. Fig. 4 explains the architecture of CNN which contains convolutional layer, subsampling layer and classification layer.

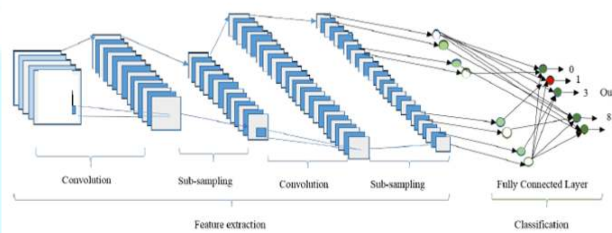

Fig. 4: The architecture of CNN [8]

3.2 Convolution Layer

Convolutional layer is the main part of a CNN. Application of mathematical computations of discrete convolution for the input or feature maps, convolutional layer is highly responsible. Many facts influence the parametrization of convolutional layer such as number of the maps, kernel sizes and connection table etc. [29]. The outcomes from the kernel go through linear or non-linear activation functions such as sigmoid, hyperbolic tangent, rectified linear, and identity functions for output feature maps. In general, it can be mathematically modeled as below.

$$x_j^l = f\left(\textstyle\sum_{i \in M_j} x_i^{l-1} k_{ij}^l + b_j^l\right) \qquad (1)$$

In this equation, x_j^l is the output of the current layer. Then x_i^{l-1} is a previous layer output, k_{ij}^l is kernel for present layer, and b_j^l is the bias for current layer. m_j represents the selection of input maps. For every output map is given an additive bias b.

3.3 Subsampling Layer

After completing the process of convolution, the linear activations process for nonlinear activation function. The further step is subsampling which can be able to exchange the input at certain place along with a summary statistic of the neighboring input values. In this layer, there is no change for the input or output maps. For example, if N is the input maps then exactly N will be the output maps. The mathematical equation of subsampling layer is

$$x_j^l = f\left(\beta_j^l \text{down}\left(x_j^{l-1}\right) + b_j^l\right) \qquad (2)$$

It shows up subsampling function. According to [8], "this function usually sums up over n × n block of the maps from the previous layers and identifies the average value or the highest values among the n × n block maps".

3.4 Classification Layer

Finally, after several convolutional and pooling layers, the high-level reasoning in the neural network is done via fully connected layers. It considered as fully connected layer which can computes the score of the class of objects using the extracted features from convolutional layer. Typically, the most expensive part of convolutional network training is learning the features. The output layer is usually relatively inexpensive due to the small number of features provided as input to this layer after passing through several layers of pooling. When performing supervised training with gradient descent, every gradient step requires a complete run of forward propagation and backward propagation through the entire network [20].

4. Results and Discussion

4.1 Dataset

Standard databases play an important role in handwritten recognition evaluation. In addition, it can provide many training and testing data. Nowadays, more researchers are paying much attention to the database standardization and verify their work by the standard database [30]. Fig. 5 illustrates the flow to process the data. Moreover, Table 1 shows the Number of Training and Testing Data and Table 2 shows the Number of Training and Testing for each Character.

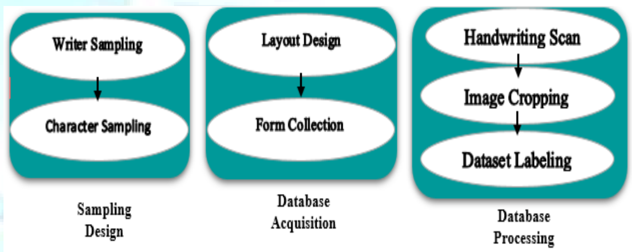

Fig.5: Flowchart of Kawi Database

Writing samples have been collected from 8 writers from four different countries which are Germany, Malaysia, India, and Bangladesh as shown in Fig. 6. According to the theory of handwriting, the handwriting of individuals becomes a consistent state at twenty-five years old, after that few change occurs [30]. So, the age range of the writers is different to explore a dissimilar way of writing. Fig. 7 shows the samples of collected data from writers.

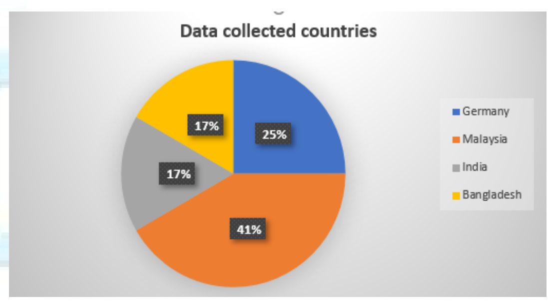

Fig. 6: Number of the writer from different countries

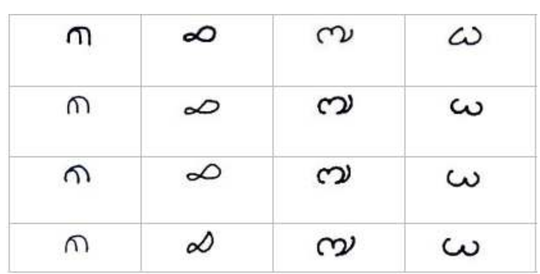

Fig. 7: Samples of collected data from writers

Table 1: Number of Training and Testing Data

Data	Size
Training data	3315
Testing data	390

Table 2: Number of Training and Testing for each Character

Data	Size
Training data	85
Testing data	10

Training data selection has been carried out randomly. Fig. 8 shows the sample of training data. The training has been done on incremental basis. These are the sample of the randomly selected characters for train the model.

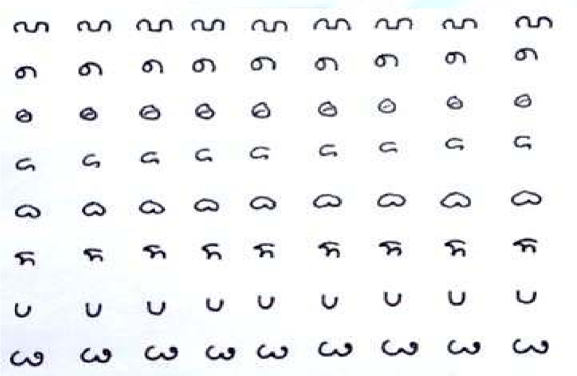

Fig. 8: Randomly selected characters for training the CNN

4.2 CNN Structure

ConvNet is not considered a deep neural network which contains many hidden layers. This deep network imitates the way visual cortex of the brain processes and recognizes images [31]. It is sometimes hard for the experts to understand the concept of their first encounter.

4.2.1 Feature Map

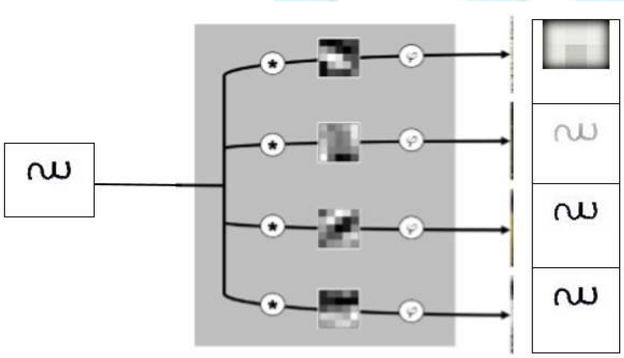

Fig. 9: Feature mapping of characters in Convolutional layer

The convolution layer generates new images called feature maps. Special features of the original images have been accentuated by feature map. Convolutional layer performs differently than other neural network because it does not employ any connection weight and a weighted sum. This layer of the architecture includes filters to convert images which are called convolution filters. In the training process, values are continuously trained. Input is 28×28 pixel black and white images which are (28×28 = 784) input nodes. A feature extraction network includes convolution layer with 20 9×9 convolution filters [17]. Convolution layer generates output which goes through ReLU function and followed by the pooling layer.an example shown in Fig.9 above.

4.2.2 Pooling Layer

The next layer is pooling layer. The pooling layer reduced the size of the image, as it combines neighboring pixels of a certain area of the image into a single representative value [17].

$$p_R = P_{i \in R}(z_i) \tag{3}$$

P is a pooling function over the region of pixel R. In fact, pooling layers made convolution invariant to rotation and shifting [32].

The neighboring pixels are usually selected from the square matrix. The representative value is usually set as the mean or maximum of the selected pixels. As it is a two-dimensional operation, consider the 2×2-pixel input image, which is expressed by the matrix shown in Fig.10

4	6
30	0

Fig. 10: The two-by-two-pixel input image

This research combined the pixels of the input image into a 2×2 matrix without overlapping the elements. Once the input image passes through the pooling layer, it shrinks into a 2×2-pixel input image [17]. The pooling reduced the size of the image which helped to relieve the computational load and prevented overfitting.

4.2.3 Fully-connected Layer

This network is trained by selected weights of the neurons so that the network learns to target output from familiar inputs [33]. To solve the weight iteratively, backpropagation algorithm has been used. Gradient descent which is an optimization method minimize minimum of error. Using loss function, the output of the network has been compared with the desired output and the calculation of the gradient of the loss function has occurred here. The neuron weights are updated by calculating the gradient of the weights and subtracting a proportion of the gradient from the weights which is the learning rate [33].

Classification network consists of a single hidden layer and an output layer. The final output of each neuron has been determined by the activation function. The hidden layer has 100 nodes that use the ReLU activation function. Rectified liner function generates output using this function:

$$\varphi(s) = \max(0, s). \tag{4}$$

Since we have 39 classes to classify, the output layer is constructed with 39 nodes. We use the softmax activation function for the output nodes. Softmax activation function which is applied in the output layer:

$$\varphi(s) = \frac{\exp s_k}{\sum_{k=1}^{K} \exp s_k}. \tag{5}$$

Table 3 summarizes the neural network parameters.

Table 3: the neural network parameters

Layer	Remark	Activation Function
Input	28X28 nodes	-
Convolution	20 convolution filters (9X9)	ReLU
Pooling	1 mean pooling (2X2)	-
Hidden	100 nodes	ReLU
Output	39 nodes	Softmax

4.2.4 Performance Evaluation

This project was done using Keras with Tensorflow environment. The CNN model has been evaluated by Kawi dataset for special characters recognition. In Fig. 12, firstly, grayscale Kawi character image result is obtained.

Fig. 11: Grayscale Kawi character image

Then, a convolutional neural network created with 16 filters with width and height as 2. The convolution set the filter to jump 2 pixels together. Therefore, a convolutional, layer with a filter has built that doesn't pad the images with zeroes. 3315 random selected images have been used to train the CNN. The training accuracy achieved 1. Fig.12 visualizes training accuracy. After that 390 character's images used to test the created CNN. The testing accuracy of recognizing the character is 89%. Fig. 13 proves the testing accuracy using Confusion Matrix.

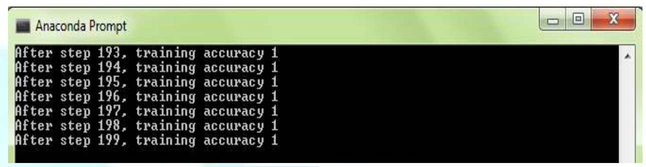

Fig. 12: Training accuracy

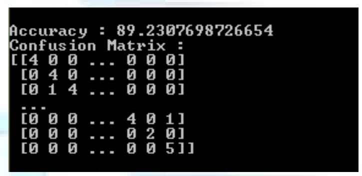

Fig. 13: Testing Accuracy for Kawi character recognition

The concept of confusion matrix is coming from machine learning. It involves data about the predicted classifications conducted by classification system [34]. A confusion matrix contains two-dimensions where one dimension is indexed using actual class of an object and another dimension is indexed by the class which the classifier predicts. In this research, the performance of the classification measured by the confusion matrix.

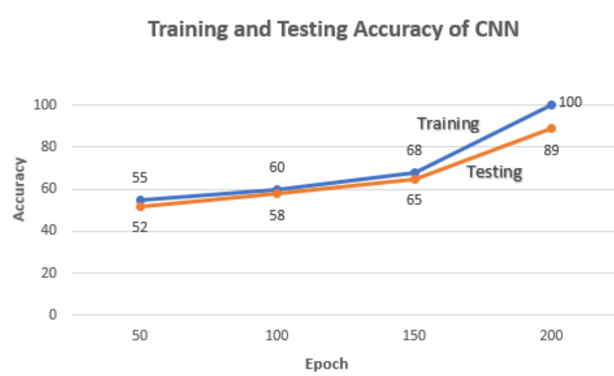

Fig. 14: Graphical representation of training and testing accuracy for CNN Kawi character recognition.

4.2.5 Comparison and Discussion

The researchers from all over the world are implementing new systems for character recognition and analysis [35]. However, the application of Artificial Neural Network played a significant role in this field of research to provide low error system. In ANN, Backpropagation network is the most popular and widely used. So, the created dataset has been applied to Backpropagation network for comparing the accuracy of CNN and Backpropagation network. In Fig. 15, steps of character recognition process using Backpropagation has been presented.

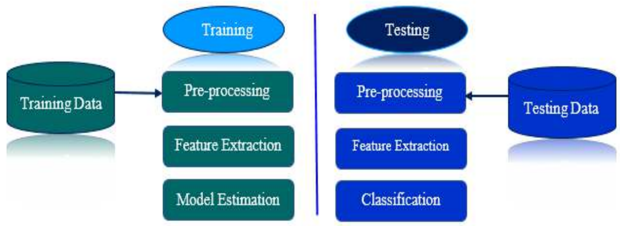

Fig. 15: Flow diagram of Character Recognition Process for ANN

Training sets of such images, each one containing 39 characters is used for training of neural networks. Each set of character images varies for handwriting styles and fonts. Each set of character images contain the last line of 10 characters for testing purpose. Beside these character images, other character images can be used for testing. Training accuracy is 70% for ANN. However, Fig. 16 shows both the training and testing accuracy in ANN.

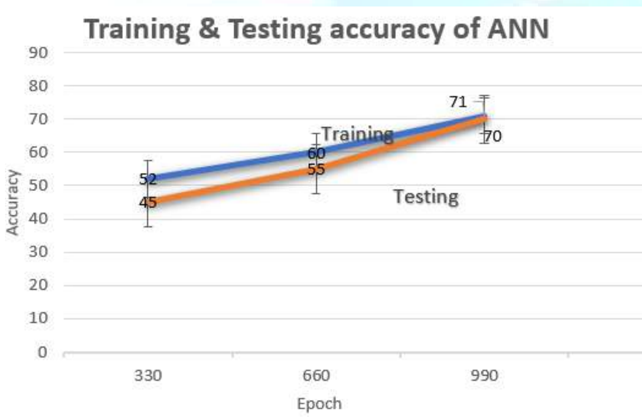

Fig. 16: Training & Testing accuracy of ANN for Kawi character recognition

5. Conclusion

Despite the advances in character recognition technology, Handwritten Kawi Characters Recognition has remained largely unsolved due to the presence of many confusing characters and excessive cursive that lead to low recognition accuracy. On the other hand, the deep learning has provided outstanding performance in many recognition tasks of natural language processing. In this research, the investigation of handwritten Kawi characters (including alphabets, and special characters) recognition approaches used deep learning model Convolutional Neural Network. It is observed that the CNN provides the highest recognition accuracy. This research has achieved recognition rate of 89% for Kawi handwritten characters using CNN. During comparing of the accuracy of CNN, 70% accuracy has been achieved for the Artificial neural network for Kawi character recognition. Between ANN and CNN, this research has found out that CNN has better recognition accuracy. More experiments will be carried to do more comparisons with other standard algorithms to evaluate the performance of CNN algorithm.

Acknowledgment

This research is supported and funded by RIMC of University Malaysia Sarawak (UNIMAS), under the special Grant Scheme (F04/SpGS/1547/2017).

References

[1] Gupta T, Ahuja C & Aich S (2014), Optical Character Recognition. International Journal of Emerging Technology and Advanced Engineering Website: www.Ijetae.Com ISO Certified Journal, 9001(9), 2–5.

[2] Rachana R. Herekar (2014), Handwritten Character Recognition Based on Zoning Using Euler Number for English Alphabets and Numerals\n. IOSR Journal of Computer Engineering (IOSR-JCE) 16(4): 75–88.

[3] Usman Akram M, Bashir Z, Tariq A & Khan Sa (2013), Geometric feature points based optical character recognition. 2013 IEEE Symposium on Industrial Electronics & Applications, pp. 86–89. https://doi.org/10.1109/ISIEA.2013.6738973

[4] Gupta S & Khosla N (2016), Neural network-based rotation, scale and font invariant english character recognition system. International Journal of Imaging and Robotics 16(1): 91–100.

[5] Boufenar C (2017), Investigation on Deep Learning for Off-line Handwritten Arabic Character Recognition Using Theano Research Platform.

[6] Ahranjany SS, Razzazi F & Ghassemian MH (2010), A very high accuracy handwritten character recognition system for Farsi/Arabic digits using convolutional neural networks. Proceedings 2010 IEEE 5th International Conference on Bio-Inspired Computing: Theories and Applications, BIC-TA 2010, pp. 1585–1592. https://doi.org/10.1109/BICTA.2010.5645265

[7] Gautam N, Sharma RS & Hazrati G (2016), Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems: Volume 2, 51, pp. 519–527. https://doi.org/10.1007/978-3-319-30927-9

[8] Alom MZ, Sidike P, Taha TM & Asari VK (2017), Handwritten Bangla Digit Recognition Using Deep Learning. ArXiv Preprint ArXiv:1705.02680.

[9] Saidas SR, Rohithram T, Sanoj KP & Manuel M (2016), Malayalam Character Recognition using Discrete Cosine

Transform, 05(15741), 15741–15743. https://doi.org/10.18535/ijecs/v5i2.14

[10] Prasad SD & Kanduri Y (2017), Telugu handwritten character recognition using adaptive and static zoning methods. 2016 IEEE Students' Technology Symposium, TechSym 2016, (i), pp. 299–304. https://doi.org/10.1109/TechSym.2016.7872700

[11] Sharma A, Khare S & Chavan S (2017), A Review on Handwritten Character Recognition 1 1,2,3, 8491, 71–75.

[12] Kaminsky AP & Long RD (Eds.). (2011), India today: an encyclopedia of life in the Republic (Vol. 2). ABC-CLIO.

[13] Ervasti JM & Campbell KP (1991), Membrane organization of the dystrophin-glycoprotein complex. Cell 66(6): 1121-1131.

[14] Zhang XY, Yin F, Zhang YM, Liu CL & Bengio Y (2016), Drawing and Recognizing Chinese Characters with Recurrent Neural Network, 8828(c), 1–14. https://doi.org/10.1109/TPAMI.2017.2695539

[15] Acharya S, Pant AK & Gyawali PK (2015), Deep learning based large-scale handwritten Devanagari character recognition. SKIMA 2015 - 9th International Conference on Software, Knowledge, Information Management and Applications. https://doi.org/10.1109/SKIMA.2015.7400041

[16] Ashlin Deepa RN & Rajeswara Rao R (2016), An efficient offline Tamil handwritten character recognition system using zernike moments and diagonal-based features. International Journal of Applied Engineering Research 11(4): 2607–2610.

[17] Kim P (2017), MATLAB Deep Learning. https://doi.org/10.1007/978-1-4842-2845-6

[18] Roy S, Das N, Kundu M & Nasipuri M (2017), Handwritten isolated Bangla compound character recognition: A new benchmark using a novel deep learning approach. Pattern Recognition Letters, 90: 15–21. https://doi.org/10.1016/j.patrec.2017.03.004

[19] Bach F & Poggio T (2016), IntroductionSpecial issue: Deep learning. Information and Inference 5(2): 103–104. https://doi.org/10.1093/imaiai/iaw010

[20] Goodfellow IJ, Bulatov Y, Ibarz J, Arnoud S & Shet V (2013), Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks, 1–13. Retrieved from http://arxiv.org/abs/1312.6082

[21] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, … Arbor A (2014), Going Deeper with Convolutions, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

[22] Chen L, Wang S, Fan W, Sun J & Naoi S (2016), Beyond human recognition: A CNN-based framework for handwritten character recognition. Proceedings - 3rd IAPR Asian Conference on Pattern Recognition, ACPR 2015, pp. 695–699. https://doi.org/10.1109/ACPR.2015.7486592

[23] Chen G, Li Y & Srihari SN (2016), Word Recognition with Deep Conditional Random Fields. Retrieved from http://arxiv.org/abs/1612.01072

[24] Chen L, Wang S, Fan W, Sun J & Naoi S (2017), Cascading training for relaxation CNN on handwritten character recognition. Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR, pp. 162–167. https://doi.org/10.1109/ICFHR.2016.0041

[25] Rajashekararadhya SV, Vanaja Ranjan P & Manjunath Aradhya VN (2008), Isolated handwritten Kannada and Tamil numeral recognition: A novel approach. Proceedings - 1st International Conference on Emerging Trends in Engineering and Technology, ICETET 2008, pp. 1192–1195. https://doi.org/10.1109/ICETET.2008.37

[26] Rabi M, Amrouch M, Mahani Z & Mammass D (2016), Recognition of cursive Arabic handwritten text using embedded training based on HMMs. Proceedings - 2016 International Conference on Engineering and MIS, ICEMIS 2016. https://doi.org/10.1109/ICEMIS.2016.7745330

[27] Uchida S, Ide S, Iwana BK & Zhu A. (2017), A further step to perfect accuracy by training CNN with larger data. Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR, pp. 405–410. https://doi.org/10.1109/ICFHR.2016.0082

[28] Wu C, Fan W, He Y, Sun J & Naoi S (2014), Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network. Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR, pp. 291–296. https://doi.org/10.1109/ICFHR.2014.56

[29] Cireșan DC, Meier U, Masci J, Gambardella LM & Schmidhuber J (2011), Flexible, high performance convolutional neural networks for image classification. IJCAI International Joint Conference on Artificial Intelligence, pp. 1237–1242. https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-210

[30] Su T, Zhang T & Guan D (2007), Corpus-based HIT-MW database for offline recognition of general-purpose Chinese handwritten text. International Journal on Document Analysis and Recognition 10(1): 27–38. https://doi.org/10.1007/s10032-006-0037-6

[31] Cao C, Liu X, Yang Y, Yu Y, Wang J, Wang Z, … Huang TS (n.d.). Look and Think Twice: Capturing Top-Down Visual Attention with Feedback Convolutional Neural Networks *. Retrieved from https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Cao_Look_and_Think_ICCV_2015_paper.pdf

[32] Cadène R, Thome N & Cord M (2016), Master's Thesis: Deep Learning for Visual Recognition. Retrieved from http://arxiv.org/abs/1610.05567

[33] Stenroos O (2017), Object detection from images using convolutional neural networks Title: Object detection from images using convolutional neural networks. Retrieved from https://aaltodoc.aalto.fi/bitstream/handle/123456789/27960/master_Stenroos_Olavi_2017.pdf?sequence=1

[34] Deng X, Liu Q, Deng Y & Mahadevan S (2016), An improved method to construct basic probability assignment based on the confusion matrix for classification problem. Information Sciences 340–341: 250–261. https://doi.org/10.1016/j.ins.2016.01.033

[35] Tiwari RR, Aparnavishwanath & Wadhone D (2013), Handwritten Digit Recognition Using Back Propagation Neural Network & K-Nearest Neighbour. International Journal of Electrical, Electronics and Data Communication 1(50): 20-25.