International Journal of Informatics and Computation (IJICOM)
Vol. 4, No.2, December 2022
ISSN: 2685-8711, E-ISSN: 2714-5263 DOI: 10.35842/ijicom

How to Stepping up Characters Recognition using CNN Algorithm?

Al-Sadi Khaled[1], Putra Wanda[2].

Abstract
Character recognition is essential to the ancient understanding culture of human sociologies. It consists of differences in textures, backdrops, font size, and coloring that causes challenges in distinguishing the characters in the current images. Various papers proposed numerous to deal with handwritten character recognition. However, several traditional remain drawbacks because methods still rely on operations based on visual capabilities. Therefore, to deal with the issue, we propose a novel recognition model using a Convolutional Neural Network to produce an effective result. To build the model, we collect datasets, do pre-processing, training with several different parameters to get the highest accuracy results. Based on experiments, our proposed model can produce a higher accuracy with a slight loss. Therefore, it can be a promising approach to addressing traditional character recognition.

Keywords
Handwriting, Recognition, Deep Learning, CNN.

1. Introduction

Differences in geography and culture bring various language and writing styles to the traditional character. Thus, it is challenging to recognize the conventional characters, including in textures, backdrops, font size, and manual coloring. Characters recognition has been a field of research interest for more than five decades [8]. Character recognition using image analysis techniques is one of the most important concepts related to natural language processing [9]. Because handwriting has been around for an extended period, it has a lot of history. Previously, humans could only interact through vocal or nonverbal means. Script has evolved and varies based on the region [1].

Understanding handwritten characters or typed papers is the basic knowledge of human beings and is a critical part of cultural development. A conventional technique to deal with character recognition is using OCR. This technique recognizes patterns and views images by converting electronic text and pictures into digital characters to enable machines to read them. Optical Character Recognition (OCR) conducts character recognition of printed text images or handwritten picture documents [15]. Segmentation, feature extraction, and feature classification are the three essential phases of character recognition. Various approaches explored automatic police number identification, automatic postal address verification from envelopes, and bank check processing. A modern methodology for recognizing handwritten characters is needed to address the issues. However, using previous works with manually designed or handcrafted features, it is difficult to deal with character recognition based on their structural and morphological appearances [2].

Current methods to deal with character recognition are using learning representation by obtaining overall features of character images and building a classification model. For instance, a paper proposed K-Nearest Neighbors to establish character recognition in many sample images [13].

Corresponding Author: Al-Sadi Khaled, Harbin University of Science & Technology (jti_22@yahoo.com)
1. Al-Sadi Khaled, Harbin University of Science & Technology (jti_22@yahoo.com)
2. Putra Wanda, Universitas Respati Yogyakarta (wpwawan@gmail.com)

Deep learning has recently shown promise in various pattern recognition fields, including handwriting identification, speech and face recognition, and natural language processing. Most advanced studies adopt CNN to create a breakthrough in feature extraction techniques. Because this model has fewer layers than others, it takes significantly less time to complete the task. Therefore, several papers discussed the deep learning approach to character recognition. For example, a study explored the recognition issue using the Devanagari character dataset [8].

A paper proposed a framework using the RNN as a discriminative model for recognizing Chinese characters and a generative model for drawing (generating) Chinese characters. The generated characters (in vector format) are human-readable. The discriminative RNN model can recognize them with high accuracy. The result can produce a practical result of RNN as both generative and discriminative models for drawing and identifying Chinese characters [23]. Another study analyzed character recognition using part of the RNN algorithm called LSTM to obtain the optimal score in sequence character data. LSTM network captures long-term dependency in sequence data, storing information from trained data and using it to classify test data. It can also be combined with CNN to create a Convolutional LSTM neural network model, which can capture spatial and temporal dependency to organize data more accurately [22].

Therefore, this paper proposes a technique to build a handwriting pattern recognition model using CNN to perform character recognition. In handwriting character recognition, we present several significant contributions to this research, particularly in the categorization of handwritten recognition using learning methods as follows:

1. We introduce a new handwritten character recognition technique involving utilizing the CNN algorithm to train dataset features and then exploiting them to develop viable models. We took datasets from respondents and then grouped them according to their respective classes to produce a dataset that could be used in this study.
2. We build a character recognition model to recognize the traditional script pattern. Instead of using conventional techniques in handwriting character recognition, we construct a novel method to deal with character recognition issues using a learning approach. To construct our recognition model, we set several hyper-parameters to produce the optimal model.
3. To measure the model performance, we test the proposed model in the unseen samples to obtain higher accuracy in handwritten character recognition. Afterward, we conduct the evaluation metrics to validate the trained model performance in character classification.

Organization: The following is a breakdown of the journal’s structure: Part II delves further into past findings. Part III discusses the study’s issue description. Section IV explains the experimental design, including a feature learning algorithm, a dataset, and pre-processing, while Section V gives the study’s findings and extensive analysis. Finally, section VI summarizes the research achievement.

2. Literature Review

Deep learning is a growing research area in the current year [26][27][28][29][30]. Several researchers constructed various methods to deal with recognition issues. A study proposed a system to recognize a person's face. Using a novel GOL texture characteristics technology, a face identification system was built for the ORL and GT facial picture databases. To evaluate the proposed system separate investigation was performed using GLCM and LBP for similar databases. Each face recognition system's average specificity, sensitivity, and retrieval time were used to evaluate its performance. After a comparison, the suggested GOL approach had much greater accuracy than the independent GLCM and LBP methods. The GLCM and LBP results were compared to the proposed method's results, and the findings show that the suggested technique is a better way for a face recognition system [3].

A paper compared the standard LSTM model with other deep models on the MNIST dataset. LSTM has been shown to perform well on the MNIST dataset because of its ability to capture long-term dependency in sequence data, store information from trained data, and use it to classify test data. It can be combined with CNN to create a convolutional LSTM to optimize capturing both spatial and temporal dependency to produce more accurate results. Based on the experimental result, the methods can harvest a classification accuracy of 98.46% [22].

A paper presented a CNN method to address Amharic character image recognition. The experiments trained 80,000 Amharic character images with different degradation levels. This study evaluated the performance of the recognition model and achieved state-of-art performance with an average recognition accuracy of 92.71% [4]. Another study presented a character recognition system using CNN for documents written in the Kannada language. They trained 74K to build to achieve an accuracy of 98% for the document containing non-overlapping lines of characters [10].

Another study proposed handwritten Chinese character recognition. Based on the experimental results, the method can produce an accuracy rate of 90.91% with 5000 times training and reduce the MSE score to 0.0079. Moreover, the proposed model can perform well in real-time tests [5]. Another paper proposed the recognition of handwritten Latin characters with diacritics using CNN. The study assesses the effectiveness of CNN-based architectures where a network is trained in recognizing handwritten characters based on Latin script. The proposed architecture produced an accuracy of 96% for the extended character set [9].

Many papers have researched character-level recognition but less on word-level recognition. To recognize the character image, a paper constructed a distinctive learning model for Gujarati Handwriting Character Recognition. The study focuses on developing an artificial intelligence-based offline HCR system for the Gujarati language. The study gathered extensive data collection of up to 10,000 images from 250 people. It constructed a supervised classifier approach based on CNN and MLP to recognize handwritten Gujarati characters. Based on the experiment result, the accuracy generated using CNN is 97.21%, and the accuracy rate using MLP is 64.48% [7].

To build a character recognition system, one of the most critical steps is a feature extraction and a classification algorithm for character recognition. Before introducing deep learning, many papers explore various feature extraction and classification methods to deal with the issues. However, manual feature extraction is costly and time-consumed. Thus, it requires an automatic feature extraction to deal with the problem. Deep learning can conduct feature extraction and classification automatically without requiring different algorithms. The DNN architecture comprises a large number of non-linear hidden layers with a large number of connections and parameters to carry out the training model [2].

Therefore, we propose a model for character recognition using a deep learning architecture with a character image dataset. To conduct our study, we gather various image characters as our dataset. We construct the learning model to recognize Lampung language characters.

3. Proposed Method

This study uses CNN to construct a character recognition model using image features, including diameter, size, and weight. We propose a model for training feature vector x in the equation, and there is a bias b. Data is passed to functions with parameters to complete the recognition process. This function will calculate the weight of each feature in the vector by multiplying it by the parameter. Equation 1 can be rewritten as equation 2, where xi is the vector x. This function has a range [-∞, ∞].

$$(\mathrm{x}) = \mathrm{x} \cdot \mathrm{w} + b \tag{1}$$

$$f(\mathrm{x}) = x_1 w_1 + x_1 w_2 + \cdots + x_N w_N + b \tag{2}$$

This regression function will produce a constant value used for class recognition.

Various research communities have proposed CNN architectures to address computer vision issues [21][24]. In this experiment, we studied character recognition of the Lampung script using CNN. We take a dataset from 100 respondents, and then the dataset will be trained and tested to get the best accuracy. The dataset is in 2D images taken using a smartphone camera.

Convolution and subsampling techniques in CNN are primarily used to extract features from raw input data. Convolution operations are multiplications of tiny kernel matrices, and defined portions of a two-dimensional input matrix are used to achieve this goal. The kernel will be shifted, and numerous multiplications will be performed from left to right and from top to bottom across defined portions of the input matrix to generate a single dimension more minor than the feature map of the input matrix. The equation for a convolution operation is described in Formula 3 as stated below as follows:

$$Q_j = f\left(\sum_{i=1}^{N} I_{i,i} * K_{i,j} + B_j\right) \tag{3}$$

Q_j is an element of a single output matrix from a convolution operation. The output matrix is produced from an activation function f. After computing the total of all multiplications of the kernel matrix$K_{i,j}$ and the input matrix $I_{i,i}$, the bias value B_j is added to the resultant matrix elements. Finally, it becomes the function's input. The activation function f utilized in this investigation is a rectified linear unit (ReLU), which is defined in the Formula below:

$$f(x) = \begin{cases} x \ (x \geq 0) \\ 0 \ (x < 0) \end{cases}$$

Each feature map will be subjected to a subsampling or pooling process after convolution procedures for dimension reduction. This study employed the max-pooling function for subsampling to retrieve notable features. A two-dimensional m x n kernel will choose the most significant value of $(m\ x\ n)$ nearby components and construct a single element for a new feature map matrix to minimize the dimension of a single feature map. The kernel will be moved from left to right and top to bottom to create a new feature map, similar to the convolution process. A dropout regularization procedure is also used in the training phase of the CNN model to reduce overfitting and enhance performance. Some neurons in CNN layers will be deactivated randomly with Bernoulli distribution using the dropout technique.

Following then, all of the neurons in all layers of the CNN model will be stimulated again during the testing phase.

In the CNN process, feature maps will be flattened after many convolution and subsampling processes to be categorized using MLP or a fully-connected neural network. A fully-connected neural network, often known as an MLP, comprises numerous layers. Each layer consists of several neurons that will execute a matrix multiplication between an input matrix x_i and internal weights $w_{j,i}$ as defined in Formula 4 as follows:

$$u_j = f\left(\sum_{i=1}^{n} w_{j,i}\, x_i + b_j\right) \tag{4}$$

Where b_j represents the bias value, n represents the number of neurons in a single layer, and f represents an activation function.

Feature maps will be handled in the output layer after processing several layers. A SoftMax function gives probabilities of classes $p(x)$ to which the CNN input may belong as the output layer of a fully-connected neural network or an MLP. A SoftMax function is defined in Formula 5.

$$p(x) = \frac{e^{x}}{\sum_{k=1}^{K} e^{x}} \tag{5}$$

One neural network model that employs a supervised learning method is CNN. This implies that in the training phase, the model utilizes a cost function to compute the distance between the model's output, the predicted class, and the actual class, the input, to update the internal weight matrix. As a cost function, the CNN employs the cross-entropy error function, which is defined in Formula 6 as follows:

$$\mathrm{E} = -\sum_{\mathrm{i}=1}^{n}(t_i \log(x_i) + (1 - t_i)\log(1 - x_i)) \tag{6}$$

In formula (6), t_i is the target class and x_i is the output of the CNN model.

The characters recognition dataset was divided into 80% for training and 20% for testing. Xavier weight initialization was utilized in each training cycle to initialize internal weight matrices in the CNN. During the testing phase, the performance of each fold in the CNN will assess using a confusion matrix and the classification accuracy value. CNN can correctly map the input dataset to the output dataset by changing the trainable parameters and the number of hidden layers.

4. Experimental Setup

a. Main Idea

The primary purpose of this paper is to create a character recognition model using CNN. We conduct handwriting recognition experiments using the learning methods to produce higher results in image recognition. Given an input image take (RGB color), we conduct the pre-processing data steps as converting the image to grayscale and binary images, finding the contour of tokens in the image, and finding each character's contour in each token, cutting each character and saving it. Several works demonstrated that the CNN model is helpful in image recognition and performs excellent recognition accuracy on other types of handwriting recognition [13][17].

B. Dataset

In this experiment, we collect a 2D dataset from a sample of 100 respondents' handwriting. Then, we build the proposed learning model using the training and test datasets to evaluate the performance model. First, we collect a sample of 2000 datasets divided into 20 classes. This study divides the experiment dataset into two parts, 1600 for training data and 400 for testing. Fig. 1 shows the primary letter of the character.

Fig 1. Script Characters

Table 1 shows details of the training and testing.

Dataset	**Sample**
Data Training (80%)	1600
Data Testing (20%)	400
Total	2000

C. Data Pre-Processing

This study divides the pre-processing operation into five stages. The first step in the proposed approach following the acquisition of the input document, the suggested technique converts the colored picture to greyscale and conducts denoising processes. After denoising, the contrast in the image will increase due to normalization. After data normalization, the provided handwritten and segmentation input document will show the boundary of each line. After line segmentation, words in each line will be extracted using vertical segmentation. The limits of each character in a word are then marked to extract the individual essence. [10].

D. Recognition Method

We gather the datasets in 2D images with length and width to conduct this study. Then the data is resized and cropped to a size of 28x28 pixels. Then we change the features into an array and store them into a vector to be used as input from fully connected. In data pre-processing, Dense was simplified to 128 nodes. These results are used in the hidden layer that sums the values between the input and output nodes using ReLU activation. Then, following the pre-processing stage, we compute the acquired features to train the prediction model. We divide the dataset into training and testing sections during the feature extraction process. Thus, we use training datasets to develop or train models, whereas performance or accuracy models use test datasets [13].

The next process of CNN is the pooling layer process that utilizes MaxPooling2D to the maximum value in each output feature. In this study, we set the pool size of 2x2 in the hidden layer of the neural network. The layers are randomly arranged to reject 20% of the neurons to avoid overfitting. The fifth layer is the flattened layer, which converts the 2D matrix data into a vector called Flatten. It enables a conventional, fully connected layer to handle the output fully. Finally, the output layer has 20 neurons for each of the 20 classes. We also calculate a SoftMax activation function to generate probability-like predictions for each category.

5. Result & Analysis

A. Recognition Test

This experiment obtains a trade-off between accuracy and performance time by adjusting various hyperparameters to acquire the best network performance. We set epoch = 100 and batch size = 64 with optimizer Adam throughout the training and testing phase. Based on the classification test, our proposed model can produce an accuracy = 97.50%. Table 2 shows training loss and training accuracy with Adam optimizer with hyperparameter setting.

Table 2. Training loss and training accuracy

Hyperparameter	Optimizer	Training Loss	Training Accuracy
Epoch = 100 Batch size = 32	Adam	0.1250	0.9750

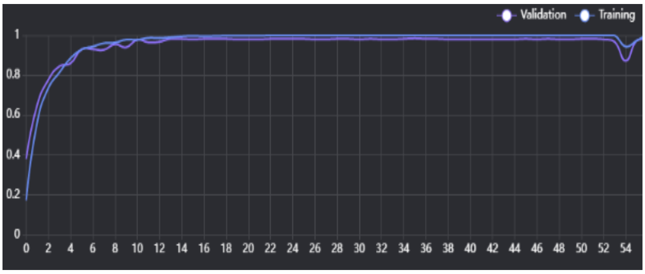

Fig 2. Training and Validation Accuracy

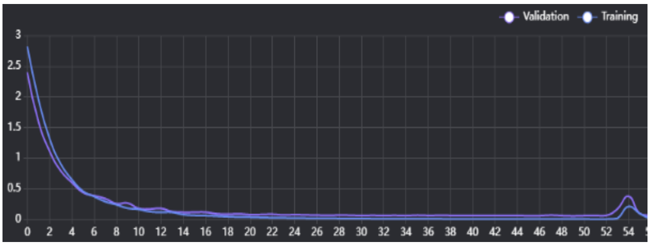

Fig 3. Training and Validation Loss

Fig. 2 shows the training and validation accuracy, with a blue line as training and purple as validation. The results of the training accuracy show a significant increase in the accuracy graph. The training accuracy can produce a higher score by tuning several hyperparameters. Fig. 3 shows the training and validation loss, where the training loss shows a decreased value and makes a decreasing error rate.

B. Evaluation Metric

We calculate the evaluation metrics of a character recognition model to measure the prediction quality of our model. in this phase, we calculate the precision score as the ratio of correctly positive predictions to the overall positive predicted results, and recall is the ratio of correct positive predictions compared to all correctly positive data. Obtaining the F1-score requires calculating the average value of precision and recall that describes a harmonic average of precision and recall. Table 3 describes accuracy, precision, recall, and F1 score values.

Table 3. The result of the classification report

Classification Report	**Precision**	**Recall**	**F1-Score**
0	0.955	0.995	0.975
1	0.995	0.955	0.975
Accuracy	-	-	0.975
macro avg	0.975	0.975	0.975
weighted avg	0.975	0.975	0.975

6. Conclusion

Traditional character recognition techniques rely on visual abilities or traditional image processing methods. However, it remains weak in introducing many characters. Thus, to deal with the issues, we construct a character recognition model using CNN to introduce Lampung characters quickly and efficiently. In this experiment, we collect multiple datasets, perform pre-processing, train our model by setting parameters to get high accuracy values, then test the model using unseen samples.

The proposed model can make a trade-off between accuracy and performance time by adjusting various hyperparameters. We set some hyperparameters to improve neural network performance by setting epoch = 100, batch size = 32, and split validation = 0.02. Based on the classification test, our approach can obtain an accuracy rate of 97.5% with a loss accuracy value of 0.125%. Therefore, the proposed model can be a promising solution for character recognition problems.

As future work, the next research can adopt other algorithms to improve this model, such as using GAN or GCN architecture. Dynamic learning architecture such as GAN can produce higher accuracy values because GAN can generate images from specific image datasets, create high-quality data, and manipulate data properly.

References

[1] Rosalina Rosalina, J.P. Hutagalung, Genta Sahuri. "Hiragana Handwriting Recognition Using Depp Neural Network Search." iJIM. 2020.

[2] R. Jana and S. Bhattacharyya. "Character Recognition from Handwritten Image Using Convolutional Neural Networks." Advances in Intelligent Systems and Computing, vol 922. 2019.

[3] Kumar K, Wagan AA, Khuhro MA, Umrani A, Chhajro A, Hafeez A, Laghari AA. "Texture based FACE recognition using GLCM and LBP schemes." Indian Journal of Science and Technology. 2020.

[4] B.H. Belay and T.A. Habtegebrial. “Amharic Character Image Recognition.” IEEE International Conference on Communication Technology. 2018.

[5] Yuan Zhuang et al. “A Handwritten Chinese Character Recognition based on Convolutional Neural Network and Median Filtering.” Journal of Physics: Conference Series. 2021.

[6] Chandra. K.D, Amanda. L.F, Afiahayati. “Convolutional Neural Network for Handwritten Javanese Character Recognition.” IJCCS. 2018.

[7] J. Pareek, D. Singhania, R.R. Kumari, S. Purohit. “Gujarati Handwritten Character Recognition from Text Images.” Third International Conference on Computing and Network Communications. 2020.

[8] Riya Guha et al. “DevNet: an efficient architecture for handwritten Devanagari character recognition.” International Journal of Pattern Recognition and Artificial Intelligence. 2019.

[9] Edyta LUKASIK et al. “Recognition of Handwritten Latin character with diacritics using CNN.” Bull. Pol. Ac.: Tech. 69(1). 2021.

[10] Asha. K, Krishnappa. H.K. “Kannada Handwritten Document Recognition using Convolutional Neural Network.” IEEE International on Computer Systems and Information Technology for Sustainable Solutions. 2018.

[11] A. Hazra, P. Choudhary, S. Inunganbi & M. Adhikari. “Bangla-Meitei Mayek scripts handwritten character recognition using Convolutional Neural Network.” Appl Intell 51. 2020.

[12] Ahamed. P, Kundu. S, T. et al. “Handwritten Arabic numerals recognition using convolutional neural network.” J Ambient Intell Human Comput 11, 5445–5457. 2020.

[13] A.H. Mawaddah, C.A. Sari, De .R.I. Moses Setiadi. “Handwriting Recognition of Hiragana Characters using Convolutional Neural Network.” iSemantic. 2020.

[14] Neha Nilesh Jadhav Sarnik. “Human Activity Recognition using CNN.” International Journal of Scientific and Research Publication (IJSRP). 2020.

[15] Megha Agarwal, Shalika, Vinam Tomar, Priyanka Gupta. “Handwritten Character Recognition using Neural Network and Tensor Flow.” IJITEE. 2019.

[16] P. Marzuki, A. R. Syafeeza et al. “A design of license plate recognition system using convolutional neural network.” IJECE. 2019.

[17] T.H. Phan, D.C. Tran, M.F. Hassan. “Vietnamese character recognition based on CNN with reduced character classes.” Bulletin of Electrical Engineering and Informatics. 2021.

[18] Rajalakshmi J. Kumar P. “Gesture Recognition using CNN and RNN.” International Journal of Recent Technology and Engineering (IJRTE). 2020.

[19] Ankit S, Vyas. H.B, Prajapati. V.K, Dabhi. “Survey on Face Expression Recognition using CNN.” International Conference on Advanced Computing & Communication Systems (ICACCS). 2019.

[20] Najadat, H.M. Alshboul. A. Alabed, A.f, “Arabic Handwritten Characters Recognition using Convolutional Neural Network.” International Conference on Information and Communication Systems (ICICS). 2019.

[21] Diqi, Mohammad., Muhdalifah, Mujastia. “Design and Building Javanese Script Classification in The State Museum of Sonobudoyo Yogyakarta.” International Journal of Informatics and Computation (IJICOM). 2020.

[22] Sarita Yadav, Ankur Pandey et al. “Handwriting Recognition using LSTM Networks.” International Journal of New Technology and Research (IJNTR). 2018.

[23] Zhang, X.-Y., Yin, F., Zhang, Y.-M., Liu, C.-L., & Bengio, Y. “Drawing and Recognizing Chinese Characters with Recurrent Neural Network.” IEEE Transactions on Pattern Analysis and Machine Intelligence. 2018.

[24] Abidin, Adnan., Hamzah., Endah, Marselina. “Efficient Fruits Classification Using Convolutional Neural Network.” International Journal of Informatics and Computation (IJICOM). 2021.

[25] Ronaldo, Antomy David. “Effective Soil Type Classification Using Convolutional Neural Network.” International Journal of Informatics and Computation (IJICOM). 2021.

[26] Wanda P, Jie HJ, DeepProfile: finding fake profile in online social network using dynamic CNN. J Inf Secur Appl 52: 102465.2020.

[27] Wanda P, Marselina Endah H, Jie HJ, DeepOSN: Bringing deep learning as malicious detection scheme in online social network. IAES Int J Artif Intell (IJ-AI) 9(1):146, 2020.
[28] H. J. Jie and P. Wanda, “RunPool: A Dynamic Pooling Layer for Convolution Neural Network,” vol. 13, no. 1, pp. 66–76, 2020.
[29] Wanda, P., Jie, H.J. DeepFriend: finding abnormal nodes in online social networks using dynamic deep learning. Soc. Netw. Anal. Min. 11, 34 (2021).
[30] Wanda, P. RunMax: fake profile classification using novel nonlinear activation in CNN. Soc. Netw. Anal. Min. 12, 158 (2022).