Handwritten Character Recognition in Handwritten Character Recognition

Subjects: Computer Science, Artificial Intelligence Contributor: Nazmus Sagib, Khandaker Foysal Haque, Ahmed Abdelgawad

Neural networks have made big strides in image classification. Convolutional neural networks (CNN) work successfully to run neural networks on direct images. Handwritten character recognition (HCR) is now a very powerful tool to detect traffic signals, translate language, and extract information from documents, etc. Although handwritten character recognition technology is in use in the industry, present accuracy is not outstanding, which compromises both performance and usability.

Keywords: handwritten character recognition ; English character recognition ; convolutional neural networks (CNNs)

1. Introduction

Handwriting is the most typical and systematic way of recording facts and information. The handwriting of an individual is idiosyncratic and unique to individual people. The capability of software or a device to recognize and analyze human handwriting in any language is called a handwritten character recognition (HCR) system. Recognition can be performed from both online and offline handwriting. In recent years, applications of handwriting recognition are thriving, widely used in reading postal addresses, language translation, bank forms and check amounts, digital libraries, keyword spotting, and traffic sign detection.

Image acquisition, preprocessing, segmentation, feature extraction, and classification are the typical processes of an HCR system, as shown in **Figure 1**. The initial step is to receive an image form of handwritten characters, which is recognized as image acquisition that will proceed as an input to preprocessing. In preprocessing, distortions of the scanned images are removed and converted into binary images. Afterward, in the segmentation step, each character is divided into sub images. Then, it will extract every characteristic of the features from each image of the character. This stage is especially important for the last step of the HCR system, which is called classification ^[1]. Based on classification accuracy and different approaches to recognize the images, there are many classification methods, i.e., convolutional neural networks (CNNs), support vector machines (SVMs), recurrent neural networks (RNNs), deep belief networks, deep Boltzmann machines, and K-nearest neighbor (KNN)^[2].



Figure 1. Representation of a common handwritten character recognition (HCR) system.

A subclass of machine learning comprises neural networks (NNs), which are information-processing methods inspired by the biological process of the human brain. **Figure 2** represents the basic neural network. The number of layers is indicated by deep learning in a neural network. Neurons, being the information-processing element, build the foundation of neural networks that draws parallels from the biological neural network. Weights associated with the connection links, bias, inputs, and outputs are the primary components of an NN. Every node is called a perceptron in a neural network (NN) ^[3]. Research is being conducted to obtain the best accuracy, but the accuracy using a CNN is not outstanding, which compromises the performance and usability for handwritten character recognition. Hence, the aim is to obtain the highest accuracy by introducing a handwritten character recognition (HCR) system using a CNN, which can automatically extract the important features from the images better than multilayer perceptron (MLP) ^{[4][S][6][Z][8][9]}.



Figure 2. Representation of a basic neural network (NN).

CNNs were first employed in 1980 ^[10]. The conception of convolutional neural networks (CNNs) was motivated by the human brain. People can identify objects from their childhood because they have seen hundreds of pictures of those objects, which is why a child can guess an object that they have never seen before. CNNs work in a similar way. CNNs used for analyzing visual images are a variation of an MLP deep neural network that is fully connected. Fully connected means that each neuron in the layer is fully connected to all the neurons in the subsequent layer. Some of the renowned CNN architectures are AlexNet (8 layers), VGG (16, 19 layers), GoogLeNet (22 layers), and ResNet (152 layers) ^[11]. CNN models can provide an excellent recognition result because they do not need to collect prior knowledge of designer features. As for CNNs, they do not depend on the rotation of input images.

A CNN model has been broadly set for the HCR system, using the MNIST dataset. Such research has been carried out for several years. A few researchers have found the accuracy to be up to 99% for the recognition of handwritten digits ^[12]. An experiment was carried out using a combination of multiple CNN models for MNIST digits and had 99.73% accuracy $^{[13]}$. Afterward, for the same MNIST dataset, the recognition accuracy was improved to 99.77%, when this experiment of the 7-net committee was extended to a 35-net committee ^[14]. Niu and Suen minimized the structural risk by integrating the SVM for the MNIST digit recognition and obtain the astonishing accuracy of 99.81% ^[15]. Chinese handwritten character recognition was investigated using a CNN ^[16]. Recently, Alvear-Sandoval et al. worked on deep neural networks (DNN) for MNIST and obtained a 0.19% error rate ^[17]. Nevertheless, after a vigilant investigation, it has been observed that the maximal recognition accuracy of the MNIST dataset can be attained by using only ensemble methods, as these aid in improving the classification accuracy. However, there are tradeoffs, i.e., high computational cost and increased testing complexity ^[18]. A tailored CNN model is proposed which attains higher accuracy with light computational complexity.

Research on HCR technology has been going on for long time now and it is in use by the industry, but the accuracy is low, which compromises the usability and overall performance of the technology. Until now, the character recognition technologies in use are still not very dependable and need more development to be deployed broadly for unfailing applications. On this account, characters of the English alphabet and digit recognition are performed by proposing a custom-tailored CNN model with two different datasets of handwritten images, i.e., Kaggle and MNIST, respectively, which achieve higher accuracies. The important features of these proposed projects are as follows:

- In the proposed CNN model, four 2D convolutional layers are kept the same and unchanged to obtain the maximum comparable recognition accuracy into two different datasets, Kaggle and MNIST, for handwritten letters and digits, respectively. This proves the versatility of the proposed model.
- A custom-tailored, lightweight, high-accuracy CNN model (with four convolutional layers, three max-pooling layers, and two dense layers) is proposed by keeping in mind that it should not overfit. Thus, the computational complexity of the model is reduced.
- Two different optimizers are used for each of the datasets, and three different learning rates (LRs) are used for each of the optimizers to evaluate the best models of the twelve models designed. This suitable selection will assist the research community in obtaining a deeper understanding of HCR.
- To the best of the authors' knowledge, the novelty of this is that no researchers to date have worked with the classification report in such detail with a tailored CNN model generalized for both handwritten English alphabet and

digit recognition. Moreover, the proposed CNN model gives above 99% recognition accuracy both in compact MNIST digit datasets and in extensive Kaggle datasets for alphabets.

• The distribution of the dataset is imbalanced. Hence, only the accuracy would be ineffectual in evaluating model performance, so advanced performances are analyzed to a great extent with a classification report for the best two proposed models for the Kaggle and MNIST datasets, respectively. Classification reports indicate the *F*1 score for each of the 10 classes for digits (0–9) and each of the 26 classes for alphabet (A–Z). In the case of multiclass classification, researchers examined averaging methods for the *F*1 score, resulting in different average scores, i.e., micro, macro, and weighted average, which is another novelty of this proposed project.

2. Development of Techniques for Classifying Handwritten Characters and Numerals or Digits

Many new techniques have been introduced in research papers to classify handwritten characters and numerals or digits. Shallow networks have already shown promising results for handwriting recognition ^{[19][20][21][22][23][24][25][26]}. Hinton et al. investigated deep belief networks (DBN), which have three layers along with a grasping algorithm, and recorded an accuracy of 98.75% for the MNIST dataset ^[27]. Pham et al. improved the performance of recurrent neural networks (RNNs), reducing the word error rate (WER) and character error rate (CER) by employing a regularization method of dropout to recognize unconstrained handwriting ^[28].

The convolutional neural network (CNN) delivered a vast change as it delivers a state-of-the-art performance in HCR accuracy ^{[29][30][31][32][33]}. In 2003, for visual document analysis, a common CNN architecture was introduced by Simard et al., which loosened the training of complex methods of neural networks ^[34]. Wang et al. used multilayer CNNs for end-to-end text recognition on benchmark datasets, e.g., street view text and ICDAR 2003, and accomplished brilliant results ^[35].

For scene text recognition, Shi et al. introduced a new approach, the conventional recurrent neural network (CRNN), integrating both the deep CNN (DCNN) and recurrent neural network (RNN), and announced its superiority to traditional methods of character recognition ^[36]. For semantic segmentation, Badrinarayanan et al. proposed a deep convolutional network architecture where the max-pooling layer was used to obtain good performance; the authors also compared their model with current techniques. The segmentation architecture known as SegNet consists of a pixel-wise classification layer, an encoder network, and a decoder network ^{[37][38]}. In offline handwritten character recognition, CNN has shown outstanding performance for different regional and international languages. Researchers have conducted studies on Chinese handwritten text recognition ^{[39][40][41]}; Arabic language ^[42]; handwritten Urdu text recognition ^{[43][44]}; handwritten Tamil character recognition nuclei scripts ^[47].

Gupta et al. used features extracted from a CNN in their model and recognized the informative local regions in ^[48] from recent character images, accomplishing a recognition accuracy of 95.96% by applying a novel multi-objective optimization framework for HCR which comprises handwritten Bangla numerals, handwritten Devanagari characters, and handwritten English numerals. High performance of the CROHME dataset was observed in the work of Nguyen et al. ^[49]. The author employed a multiscale CNN for clustering handwritten mathematical expression (HME) and concluded by identifying that their model can be improved by training the CNN with a combination of global, attentive, and max-pooling layers.

Recognition of word location in historical books, for example on Gutenberg's Bible pages, is wisely addressed in the work of Ziran et al. ^[50] by developing an R-CNN-based deep learning framework. Ptucha et al. introduced an intelligent character recognition (ICR) system, logically using a conventional neural network ^[51]. IAM datasets and French-language-based RIMES lexicon datasets were used to evaluate the model, which reported a commendable result. The variance between model parameters and hyper-parameters was highlighted in ^[52]. The hyper-parameters include the number of epochs, hidden units, hidden layers, learning rate (LR), kernel size, activation function, etc., which must be determined before the training begins to determine the performance of the CNN ^[53]. It is mentioned that, if the hyper-parameters are chosen poorly, it can lead to a bad CNN performance. The total number of hyper-parameters of some CNN models are 27, 57, 78, and 150, respectively, for AlexNet ^[54], VGG-16 ^[55], GoogleNet ^[56], and ResNet-52 ^[57]. To improve the recognition performance, practicing researchers play an important role in the handwriting recognition field for designing CNN parameters effectively. Tapotosh Ghosh et al. converted the images into black-and-white 28 × 28 forms with white as the foreground color in ^[58] by approaching InceptionResNetV2, DenseNet121, and InceptionNetV3 using the CMATERdb dataset.

References

- Priya, A.; Mishra, S.; Raj, S.; Mandal, S.; Datta, S. Online and offline character recognition: A survey. In Proceedings of the International Conference on Communication and Signal Processing, (ICCSP), Melmaruvathur, Tamilnadu, India, 6– 8 April 2016; pp. 967–970.
- Gunawan, T.S.; Noor, A.F.R.M.; Kartiwi, M. Development of english handwritten recognition using deep neural network. Indones. J. Electr. Eng. Comput. Sci. 2018, 10, 562–568.
- Vinh, T.Q.; Duy, L.H.; Nhan, N.T. Vietnamese handwritten character recognition using convolutional neural network. IAE S Int. J. Artif. Intell. 2020, 9, 276–283.
- Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–25 June 2005; Vol ume 1, pp. 886–893.
- 5. Lowe, D.G. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 2004, 60, 91–110.
- Xiao, J.; Zhu, X.; Huang, C.; Yang, X.; Wen, F.; Zhong, M. A New Approach for Stock Price Analysis and Prediction Bas ed on SSA and SVM. Int. J. Inf. Technol. Decis. Mak. 2019, 18, 35–63.
- Wang, D.; Huang, L.; Tang, L. Dissipativity and synchronization of generalized BAM neural networks with multivariate di scontinuous activations. IEEE Trans. Neural Netw. Learn. Syst. 2018, 29, 3815–3827.
- 8. Kuang, F.; Zhang, S.; Jin, Z.; Xu, W. A novel SVM by combining kernel principal component analysis and improved cha otic particle swarm optimization for intrusion detection. Soft Comput. 2015, 19, 1187–1199.
- Choudhary, A.; Ahlawat, S.; Rishi, R. A binarization feature extraction approach to OCR: MLP vs. RBF. In Proceedings of the International Conference on Distributed Computing and Technology (ICDCIT), Bhubaneswar, India, 6–9 February 2014; Springer: Cham, Switzerland, 2014; pp. 341–346.
- 10. Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol. Cybern. 1980, 36, 193–202.
- 11. Ahlawat, S.; Choudhary, A.; Nayyar, A.; Singh, S.; Yoon, B. Improved handwritten digit recognition using convolutional n eural networks (Cnn). Sensors 2020, 20, 3344.
- Jarrett, K.; Kavukcuoglu, K.; Ranzato, M.; LeCun, Y. What is the best multi-stage architecture for object recognition? In Proceedings of the IEEE 12th International Conference on Computer Vision (ICCV), Kyoto, Japan, 29 September–2 Oc tober 2009; pp. 2146–2153.
- 13. Cireşan, D.C.; Meier, U.; Masci, J.; Gambardella, L.M.; Schmidhuber, J. High-Performance Neural Networks for Visual Object Classification. arXiv 2011, arXiv:1102.0183v1.
- Ciresan, D.; Meier, U.; Schmidhuber, J. Multi-column deep neural networks for image classification. In Proceedings of t he IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; IEEE: Pisc ataway, NJ, USA, 2012; pp. 3642–3649.
- 15. Niu, X.X.; Suen, C.Y. A novel hybrid CNN-SVM classifier for recognizing handwritten digits. Pattern Recognit. 2012, 45, 1318–1325.
- 16. Qu, X.; Wang, W.; Lu, K.; Zhou, J. Data augmentation and directional feature maps extraction for in-air handwritten Chi nese character recognition based on convolutional neural network. Pattern Recognit. Lett. 2018, 111, 9–15.
- 17. Alvear-Sandoval, R.F.; Figueiras-Vidal, A.R. On building ensembles of stacked denoising auto-encoding classifiers and their further improvement. Inf. Fusion 2018, 39, 41–52.
- 18. Demir, C.; Alpaydin, E. Cost-conscious classifier ensembles. Pattern Recognit. Lett. 2005, 26, 2206–2214.
- 19. Choudhary, A.; Ahlawat, S.; Rishi, R. A Neural Approach to Cursive Handwritten Character Recognition Using Features Extracted from Binarization Technique. Stud. Fuzziness Soft Comput. 2015, 319, 745–771.
- Cai, Z.W.; Huang, L.H. Finite-time synchronization by switching state-feedback control for discontinuous Cohen–Gross berg neural networks with mixed delays. Int. J. Mach. Learn. Cybern. 2018, 9, 1683–1695.
- 21. Zeng, D.; Dai, Y.; Li, F.; Sherratt, R.S.; Wang, J. Adversarial learning for distant supervised relation extraction. Comput. Mater. Contin. 2018, 55, 121–136.
- 22. Long, M.; Zeng, Y. Detecting iris liveness with batch normalized convolutional neural network. Comput. Mater. Contin. 2 019, 58, 493–504.
- 23. Huang, C.; Liu, B. New studies on dynamic analysis of inertial neural networks involving non-reduced order method. Ne urocomputing 2019, 325, 283–287.

- 24. Xiang, L.; Li, Y.; Hao, W.; Yang, P.; Shen, X. Reversible natural language watermarking using synonym substitution and arithmetic coding. Comput. Mater. Contin. 2018, 55, 541–559.
- 25. Huang, Y.S.; Wang, Z.Y. Decentralized adaptive fuzzy control for a class of large-scale MIMO nonlinear systems with st rong interconnection and its application to automated highway systems. Inf. Sci. (Ny). 2014, 274, 210–224.
- 26. Ahlawat, S.; Rishi, R. A Genetic Algorithm Based Feature Selection for Handwritten Digit Recognition. Recent Pat. Com put. Sci. 2018, 12, 304–316.
- 27. Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. Neural Comput. 2006, 18, 1527–155
 4.
- Pham, V.; Bluche, T.; Kermorvant, C.; Louradour, J. Dropout Improves Recurrent Neural Networks for Handwriting Rec ognition. In Proceedings of the 14th International Conference on Frontiers in Handwriting Recognition (ICFHR), Herakli on, Greece, 1–4 September 2014; pp. 285–290.
- 29. Lang, G.; Li, Q.; Cai, M.; Yang, T.; Xiao, Q. Incremental approaches to knowledge reduction based on characteristic ma trices. Int. J. Mach. Learn. Cybern. 2017, 8, 203–222.
- Tabik, S.; Alvear-Sandoval, R.F.; Ruiz, M.M.; Sancho-Gómez, J.L.; Figueiras-Vidal, A.R.; Herrera, F. MNIST-NET10: A h eterogeneous deep networks fusion based on the degree of certainty to reach 0.1% error rate. ensembles overview an d proposal. Inf. Fusion 2020, 62, 73–80.
- Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Se gmentation. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 2481–2495.
- Liang, T.; Xu, X.; Xiao, P. A new image classification method based on modified condensed nearest neighbor and conv olutional neural networks. Pattern Recognit. Lett. 2017, 94, 105–111.
- 33. Sueiras, J.; Ruiz, V.; Sanchez, A.; Velez, J.F. Offline continuous handwriting recognition using sequence to sequence n eural networks. Neurocomputing 2018, 289, 119–128.
- Simard, P.Y.; Steinkraus, D.; Platt, J.C. Best practices for convolutional neural networks applied to visual document ana lysis. In Proceedings of the International Conference on Document Analysis and Recognition(ICDAR), Edinburgh, UK, 3–6 August 2003; Volume 3, pp. 958–963.
- Wang, T.; Wu, D.J.; Coates, A.; Ng, A.Y. End-to-end text recognition with convolutional neural networks. In Proceedings of the 21st-International Conference on Pattern Recognition, Tsukuba, Japan, 11–15 November 2012; pp. 3304–3308.
- 36. Shi, B.; Bai, X.; Yao, C. An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Appli cation to Scene Text Recognition. IEEE Trans. Pattern Anal. Mach. Intell. 2017, 39, 2298–2304.
- Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 2018, 40, 8 34–848.
- Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
- Wu, Y.C.; Yin, F.; Liu, C.L. Improving handwritten Chinese text recognition using neural network language models and c onvolutional neural network shape models. Pattern Recognit. 2017, 65, 251–264.
- Xie, Z.; Sun, Z.; Jin, L.; Feng, Z.; Zhang, S. Fully convolutional recurrent network for handwritten Chinese text recogniti on. In Proceedings of the International Conference on Pattern Recognition, Cancun, Mexico, 4–8 December 2016; pp. 4011–4016.
- 41. Liu, C.L.; Yin, F.; Wang, D.H.; Wang, Q.F. Online and offline handwritten Chinese character recognition: Benchmarking on new datasets. Pattern Recognit. 2013, 46, 155–162.
- 42. Boufenar, C.; Kerboua, A.; Batouche, M. Investigation on deep learning for off-line handwritten Arabic character recogni tion. Cogn. Syst. Res. 2018, 50, 180–195.
- 43. Husnain, M.; Missen, M.M.S.; Mumtaz, S.; Jhanidr, M.Z.; Coustaty, M.; Luqman, M.M.; Ogier, J.M.; Choi, G.S. Recognit ion of urdu handwritten characters using convolutional neural network. Appl. Sci. 2019, 9, 2758.
- 44. Ahmed, S.B.; Naz, S.; Swati, S.; Razzak, M.I. Handwritten Urdu character recognition using one-dimensional BLSTM cl assifier. Neural Comput. Appl. 2019, 31, 1143–1151.
- 45. Kavitha, B.R.; Srimathi, C. Benchmarking on offline Handwritten Tamil Character Recognition using convolutional neura I networks. J. King Saud Univ.-Comput. Inf. Sci. 2019, 34, 1183–1190.
- 46. Dewan, S.; Chakravarthy, S. A system for offline character recognition using auto-encoder networks. In Proceedings of the the International Conference on Neural Information Processing, Doha, Qatar, 12–15 November 2012.

- 47. Sarkhel, R.; Das, N.; Das, A.; Kundu, M.; Nasipuri, M. A multi-scale deep quad tree based feature extraction method for the recognition of isolated handwritten characters of popular indic scripts. Pattern Recognit. 2017, 71, 78–93.
- Gupta, A.; Sarkhel, R.; Das, N.; Kundu, M. Multiobjective optimization for recognition of isolated handwritten Indic script s. Pattern Recognit. Lett. 2019, 128, 318–325.
- 49. Nguyen, C.T.; Khuong, V.T.M.; Nguyen, H.T.; Nakagawa, M. CNN based spatial classification features for clustering offli ne handwritten mathematical expressions. Pattern Recognit. Lett. 2020, 131, 113–120.
- 50. Ziran, Z.; Pic, X.; Undri Innocenti, S.; Mugnai, D.; Marinai, S. Text alignment in early printed books combining deep lear ning and dynamic programming. Pattern Recognit. Lett. 2020, 133, 109–115.
- Ptucha, R.; Petroski Such, F.; Pillai, S.; Brockler, F.; Singh, V.; Hutkowski, P. Intelligent character recognition using fully convolutional neural networks. Pattern Recognit. 2019, 88, 604–613.
- 52. Tso, W.W.; Burnak, B.; Pistikopoulos, E.N. HY-POP: Hyperparameter optimization of machine learning models through parametric programming. Comput. Chem. Eng. 2020, 139, 106902.
- 53. Cui, H.; Bai, J. A new hyperparameters optimization method for convolutional neural networks. Pattern Recognit. Lett. 2 019, 125, 828–834.
- 54. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. Commun. A CM 2017, 60, 84–90.
- 55. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the 3rd International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
- 56. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going dee per with convolutions. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Reco gnition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770– 778.
- Ghosh, T.; Abedin, M.H.Z.; Al Banna, H.; Mumenin, N.; Abu Yousuf, M. Performance Analysis of State of the Art Convol utional Neural Network Architectures in Bangla Handwritten Character Recognition. Pattern Recognit. Image Anal. 202 1, 31, 60–71.

Retrieved from https://encyclopedia.pub/entry/history/show/55214