2022 4th International Conference on Artificial Intelligence and Speech Technology (AIST) | 978-1-6654-9902-6/22/$31.00 ©2022 IEEE | DOI: 10.1109/AIST55798.2022.10064860

# Recognition Of Handwritten English Character Using Convolutional Neural Network

Sapna Katoch
*Department of Computer Science*
*Lovely Professional University*
Phagwara,India
sapnaktch@gmail.com

Manik Rakhra
*Assitant Professor, Department of Computer Science*
*Lovely Professional University*
Phagwara,India
rakhramanik786@gmail.com

Dalwinder Singh
*Assitant Professor, Department of Computer Science*
*Lovely Professional University*
Phagwara,India
dalwinder.singh@gmail.com

***Abstract*— In the domain of computer vision and image processing, one of the most active and difficult study fields is handwritten character recognition. It may be used as a reading tool for bank checks, for identifying characters on forms, and for a slew of other purposes. The optical character recognition of the papers is similar to documents produced by hand by a human. This OCR is put to use to improve the simplification of the process of character translation, which may be obtained from a broad range of file types, such as image and word document files. Researchers have made tremendous progress in HCR by making use of vast amounts of raw data and new breakthroughs in Deep Learning and Machine Learning algorithms. The fundamental purpose of this research paper is to give a solution for several techniques of handwriting recognition. These methods include the usage of touch input through a mobile screen as well as the use of an image file. CNN is used to identify characters in a test dataset in this work. Work on CNNs' capacity to detect characters from a picture dataset and their accuracy of recognition will be examined. Characters are recognized by CNN by comparing and contrasting their shapes and distinguishing characteristics. The dataset A_Z Handwritten was used to test our CNN implementation's handwriting accuracy and model gives the 100% result to recognize the character.**

***Keywords*— Handwritten character recognition, A_Z Dataset, Deep Learning, CNN (Convolutional neural network), Feature extraction, Classification***

## I. Introduction

When it comes to pattern recognition, we are aware that deep learning models are highly effective classifiers and that they have been widely utilized for a variety of picture identification tasks, including human face recognition, location estimation, and character recognition. [10][12][13][14] Deep learning models have revolutionized pattern recognition and achieved significant advances in a wide range of applications. Figure 1 shows how the deep learning model can automatically extract features and classify them, as opposed to the previous approaches. It may be seen as a "black box" by employers, who merely need to enter a picture to get a recognition result. Traditional approaches, on the other hand, often need the creation of artificial features and the human adjustment of the classifier. In particular, the empirical elements of the classical technique have a significant impact on its success. Traditional research methodologies have run out of steam after decades of use. Deep learning's development offers a new path to breaking the human performance ceiling and perhaps going beyond it.

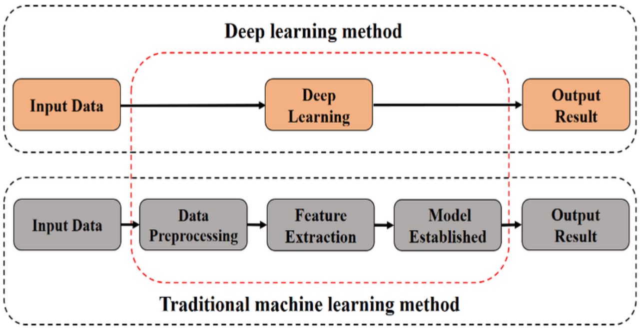


Figure 1. Deep Learning vs. Traditional Learning

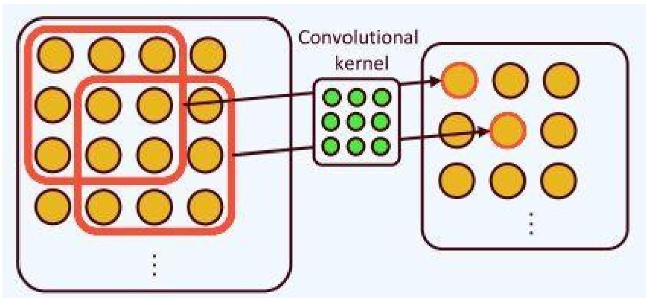


Figure 2. Convolutional kernel process of CNN

The deep learning model that is used most often for the purpose of image recognition is known as a convolutional neural network, or CNN. As a result, CNNs are a great fit for representing picture structures. To begin, the pixels in the picture have a significant link with their immediate neighbours, but only a weak relationship with pixels that are far away. Similar to the CNN's "local connection" technique, in which each neuron is only linked to a limited number of other neurons, this trait is shown in Figure 2. (local receptive Fifield). It also guarantees that comparable features like as texture and brightness may be shared throughout various parts of the picture in CNN. As a result of the aforementioned advantages, CNN is a top option for image identification.

Identification of handwritten characters is a challenging task due to the fact that the appearance of a character may vary significantly based on the writer, the writing style, and the noise. A considerable deal of work has been done to enhance recognition, including the development of improved features and the optimization of classifiers. However, as previously said, conventional techniques' enhancements pale in comparison to CNN's. A promising recognition rate may be achieved simply by using a CNN. However, there are additional difficulties that must be addressed if CNN is to improve its performance. In this research, a CNN-based framework was proposed for handwritten character recognition; in terms of performance, it surpassed other CNN-based algorithms that were previously available. Furthermore, our approach outperformed humans on the A Z Handwritten dataset, a well-known benchmark for handwritten character recognition. There has never been an occasion when a machine has come out on top in this kind of assignment. This ground-breaking study demonstrated that robots can now do certain tasks better than trained humans thanks to advances in deep learning technology.

## II. LITERATURE REVIEW

Professor Sathish and graduate student Yasir Babiker Hamdan [1] have developed a proposal for a research project on the topic of character recognition in handwriting. The benefit of this article is that it provides the user with two options for character identification: first, using an offline character recognition system, and second, using an online character recognition system. The fact that the study demonstrates this is a valuable contribution to the field. Software that is capable of reading and interpreting characters from a broad range of data types, such as picture and text files, was employed. This software is known as OCR (optical character recognition). The primary objective of this article is to demonstrate a solution for several methods of handwriting recognition, including the use of input method through a smartphone and the use of an image file. For nonlinearly divisible problems, we have developed recognition strategies based on a variety of methodologies drawn from artificial neural networks, statistical methods, etc. Among these techniques are those that we have specifically selected to use in our performances. This research collects a number of alternative ways for distinguishing and recognizing handwritten characters that may be viewed in visual documentation. In addition to this, the research compares and contrasts the SVM classifiers network technology with several other methods, including as image compression, structural pattern recognition, and graphical methodologies. In optical character recognition (OCR) systems, the statistical support vector machine, often known as SVM, has been found to offer satisfactory results when optimized using a machine learning technique. The recognition rate of this technique is higher than that of the other methods considered in this research work. In order to improve the proposed model's accuracy, it was put through its tests on a training section consisting of variously styled letters and digits. During our tests, we found that we were 91% accurate in identifying characters on paper.

Mueen Uddin, Jamshed Memon, Rizwan Ahmed Khan and Maira Sami and wrote a review paper to suggest study directions and offer a synopsis of existing research on the topic of character recognition of handwritten documents [2]. The authors conducted a Systematic Literature Review (SLR) by collecting, synthesising, and analysing academic papers produced on handwritten OCR and related topics from 2000 to 2019. In order to conduct the review, they made use of computerized databases that were readily accessible as well as a predetermined process. We found all of the essential articles by a combination of keyword searching, searching forward reference, and searching backward reference. Thorough adherence to the research selection procedure resulted in the acceptance of 176 articles in this SLR. Their review article aims to provide the most up-to-date results and approaches on OCR, as well as to suggest research directions by making reference to the places where further investigation is required.

Recent research suggests that HDR is concentrating its efforts mostly on deep learning (DL) techniques, which has resulted in impressive breakthroughs. However, correct identification and categorization of numbers is still a challenging topic to solve due to the fact that individuals' writing styles vary and the input sample comprises blurring, distortion, brightness, and size variances. S aleh Albahli, Ali Javed, Marriam Nawaz and Aun Irtaza [3] provide an effective and efficient high dynamic range (HDR) system by adding a tailored quicker regional convolutional neural network. This is done in order to compensate for the constraints that have been described (Faster-RCNN). This methodology is broken down into three primary stages. They start by creating handwritten notes in order to obtain the subject matter of interest. After that, an improved version of Faster-RCNN is put to use. Within this version, DenseNet-41 is included in order to carry out the calculation of the deep feature sets. At the very end of the procedure, the regressor and classification layer is used to localize and categorize the digits into a total of ten various groups. The effectiveness of the recommended technique is assessed using the MNIST database, which is the industry standard. This database provides a number of obstacles in terms of the lighting conditions, colours, form and size of the digits, the incidence of blurring and noise effects, and so on. In addition to this, they have shown our method's viability by testing it on many datasets at once. The strategy has been shown to be superior to other contemporary approaches in terms of its ability to competently identify and precisely categorise numbers, as shown by experimental assessments.

According to the research conducted by Professor Vaibhav. V. Mainkar and Mr. Ajinkya B.Upade [4] storing handwritten notes in an image format on a mobile phone makes them much simpler to access in the future. The technique they used to save such information in an electronic format was called "Optical Character Recognition." This procedure includes pre-processing, the extraction of features, and recognitions. There have been quite a few scholars that have made use of OCR in order to recognise the Character. In this study, the authors made use of an Android phone to take the picture, and then they made use of OCR to transform the picture into an electronic format. The model's accuracy for handwritten documents is 90%, making it the simplest method to detect the data and communicate it.

I Khandokar, M. N. Kabir F. Ernawan, Md. M. Hasan and Md. S. Islam [5] identified the handwritten character by using the deep learning technique. After it has been implemented, CNN is used to do analysis on the data characters included inside a test dataset. The primary objective of this research is to determine whether or not CNN is capable of recognizing characters in the picture collection, as well as the degree to which recognition is correct after being trained and tested. CNN is able to recognise the characters by taking into consideration the forms of the characters and analysing, contrasting, and comparing the qualities that differentiate one character from the others. Studies were run using the NIST dataset so that we could establish how well our CNN implementation can recreate handwritten characters. The NIST dataset was used for these experiments. The results of the tests indicate that a level of accuracy of 92.91% may be attained while utilising a training set of 1000 photographs provided by NIST while only using 200 images. This can be done with just 200 photos.

Huaxiang Lu, Zhiyuan Li, Yi Xiao, Qi Wu and Min Jin [6] all made effective use of the CNN (Convolutional neural network) to accurately recognize the Chinese character written by hand. They suggest the creation of a matching network that will build a connection between handwritten characters and template characters. The process through which people learn to write Chinese characters provided the inspiration for this relationship. The parameters that were previously used in the softmax regression layer will be replaced by the matching network with the recovered features from the template character photos. These settings were used before. Once training is complete, the strong individuals in different ways will allow us to generalize the prediction capability to not just fresh data, but also whole new Chinese characters that were never a part of the training set. This will allow us to predict the outcomes of events that have never happened before. This enables us to generalise the capacity of prediction to a considerably larger variety of scenarios than was previously possible. Experiments conducted on the ICDAR-2013 download HCCR datasets showed that the proposed approach achieves performance on scale with that of current CNN-based classifiers. This was discovered after the experiments were carried out on the datasets. In addition, the matching network has a very encouraging potential for generalisation to new Chinese characters that are never seen in the data set that it is currently being trained on. In the training set, you won't find these particular characters.

## III. BACKGROUND STUDY

Because of its wide range of applications, the identification of handwritten characters has garnered a lot of attention in the field of machine learning and pattern recognition. Character recognition has been aided by a variety of handwriting recognition systems. There are several studies and publications out there that explain how to turn paper documents into machine-readable language. In the near future, the digitalization and processing of current paper documents may be a crucial aspect in achieving a paperless world [16]. Because of the unique arrangement of neurons in the visual cortex, CNN architecture closely resembles that seen in the human brain. Nerve cells simply respond to stimuli that are visible in the spectrum. A wide variety of these spectrum fields overlap to occupy the whole display area. Using a CNN, an image's spatial dependencies may be correlated with the image's content in order to recognize it. In order to classify an image, the differentiating characteristics must be found in the input picture, and this is where CNN comes in. To get a better understanding of the image's characteristics, the training process involves determining the image's weights, parameters, and biases [17]. Figure 3 depicts the CNN process, which includes convolution, pooling, and a fully-connected neural network, and is characterized as such in the following manner: Maintain the integrity of your textual and graphical files until after the formatting and styling of the content has been completed. Hard tabs should not be used, and the usage of hard returns should be limited to only one at the conclusion of each paragraph at most. It is essential that the whole paper be free of any and all forms of printing. You don't need to worry about numbering the text heads—the template will take care of it for you.

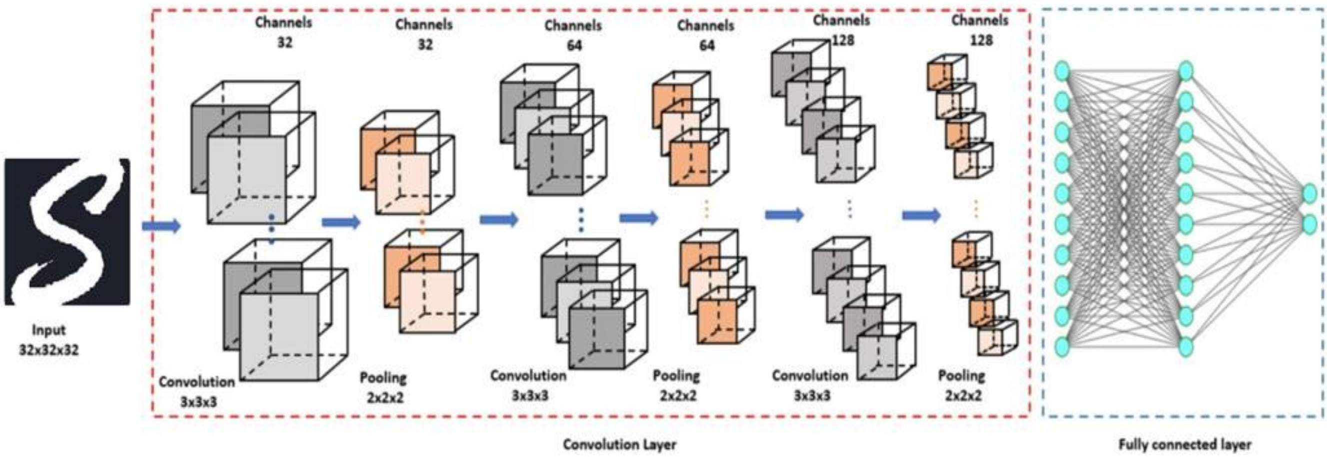


Figure 3. Process of Convolutional Neural Network

### A. Convolution

Image manipulation techniques like sharpening, enhancing, and smoothing the camera's image are all examples of how convolution is often utilised. Convolution is utilised in a variety of machine learning methods, including CNNs. In order to modify the picture, it is first passed through a singular matrix of integers, which may be referred to as a kernel or filter. It is necessary to multiply each picture submatrix pixel value by the matching kernel pixel value, then add the results to get the filtered image's single pixel value. On a micro-scale, the whole picture is processed in this manner[18][20]s. There are a few methods for capturing low-level elements such as borders and colour gradients. Additional layers in the architecture allow for a network with a deep knowledge of a picture, akin to what humans are capable of. Using a convolution algorithm is a way to eliminate the image's high-level information like edges. Reshaped characteristics are less dimensional than they were before the procedure was used [11].

### B. Pooling

Pooling is a second important component of the CNN system. In order to lower the computational complexity inside the network, the spatial size of the picture description must be gradually increased. Pooling may take place in a variety of ways depending on the application, including maximum pooling, minimum pooling, average pooling, and adaptive pooling. In the CNN technique, the maximum pool (also known as Max-Pooling) is a prominent strategy. As a result of this, it reduces the amount of computing power that is necessary by keeping just those significant traits that are translational and rotational versions. This helps to keep the model's appropriate training cycle intact. The image part of the kernel that is encompassed by Max-Pooling is analyzed in order to get the maximum amount of value possible from that section. On the other hand, as compared to single pooling, average pooling provides a greater reduction in noise. In addition to removing the noisy activations, it also performs de-noising while simultaneously reducing the dimensionality [17].

### *C. Fully-connected neural network*

Towards the end of the CNN, a fully linked neural network is implemented. They are nothing more than artificial neural networks that are completely interconnected. While the training session is taking place, the weights that will be used in the network are being calculated. The final output of the convolution/pooling process is sent into a fully connected neural network, which then computes the label that is the most closely matches the picture. Within this part, a connection between the image feature vector and the picture's categorization is developed. In order to get the ultimate outcome, the outputs of the convolution and pooling process need to be multiplied by the weights that are linked with the network connection route. An activation function [18] is applied to the result after it has been computed.

## IV. METHODOLOGY

Handwritten character recognition is one of the topics that has received the greatest attention and study in the subfields of machine learning and computer vision. [16]. We will now go through our methods for recognizing characters from texts in picture format inside a document that we have created. In order to achieve classification and detection, the suggested model, as indicated in Figure 4, comprises four steps, which are as follows: pre-processing; feature extraction; fitting the data; and classification. Authors and Affiliations.

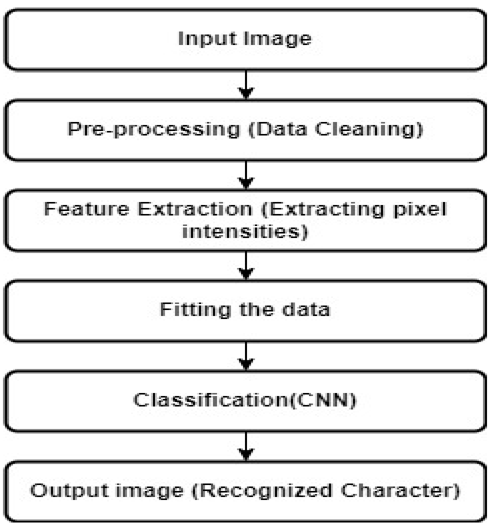


Figure 4. Steps to recognize the handwritten character

### *A. Pre-Processing*

A Pre-processing is the process of cleaning up a picture when it has been received. It substantially improves the picture by removing noise from it. It is also possible that photographs in greyscale or binary formats may be needed, which will be completed at this step [19].

### *B. Feature Extraction*

H It is the process of changing the input data into a collection of features that are highly representative of the input data, which is referred to as feature extraction. The process of reducing the number of dimensions is related to the method of extracting features. The input data may be changed into a more manageable collection of features when it is too huge to be handled in its whole (also named a feature vector). Feature selection [9] is the process of identifying a subset of the original characteristics and determining them. The selected features are supposed to contain all of the key information from the input data, which enables the desired task to be finished by using this extracted features rather than the full starting data, rather of the whole beginning data[26,27].

### *C. Fitting the data*

We are building the model at this point, and we are defining the optimising function and the loss function that will be utilised in the fitting process. Adam is an optimising function that is a mix of the RMSprop and Adagram optimising methods, and it is utilised in this example. Due to the vast size of the dataset, we are just training for a single epoch; however, if necessary, we may train for many epochs (which is recommended for character recognition for better accuracy).

### *D. Classification*

Using a CNN, it is possible to determine if characters in a picture were written by hand. All of a CNN's layers, including the input and output layers as well as any hidden layers, are included here.[21-25] Convolutional, pooling, fully connected, and normalizing layers commonly make up the CNN's hidden layers. A convolutional neural network (CNN) is made up of three primary layers: the output layer, the pooling layer, and the convolutional layer. ReLU, which stands for Rectified Linear Unit, is a typical activation function for CNNs. [7] [8] Neurons that are linked to local areas in the input volume will individually calculate a dot product between their weights and a tiny region that they are related to in the input volume. Non-linear downsampling is used in the pooling layer. With max pooling, the picture is divided into a number of distinct regions with no overlap, and the maximum value is generated for each one. The non-saturation activation function is used by ReLU. Decision function and overall network nonlinearity are improved without harming the convolution layer's ability to process information. Reverse rectification of a linear unit results in an output of 0 when the input is less than zero, and 0 otherwise. The formula for determining its value is as follows:

$$f(x) = \max(x,0) \qquad (1)$$

When building a neural network-based classifier, the softmax function is often used in the final layer. Outputs from each individual unit are compressed by the softmax function, which is similar to the sigmoid function in that it produces values between 0 and 1. However, it splits each output in such a way that the entire sum of the outputs equals one. It is equal to a categorical probability distribution to get the output of the softmax function. [15] As a result, the softmax function computes the probability distribution of one event across a set of n distinct occurrences.

## V. RESULT & DISCUSSION

Python 3 is used to implement the approach that has been presented. The character that was entered in English. The input is pre-processed, scaled, and normalised before being sent on to the CNN classifier for classification. Because the CNN classifier is trained on the A-Z dataset, it can anticipate the character that is being input. We make nine random plots in the form of a third and show the thresholded pictures of nine alphabets on each plot. Figure 5.

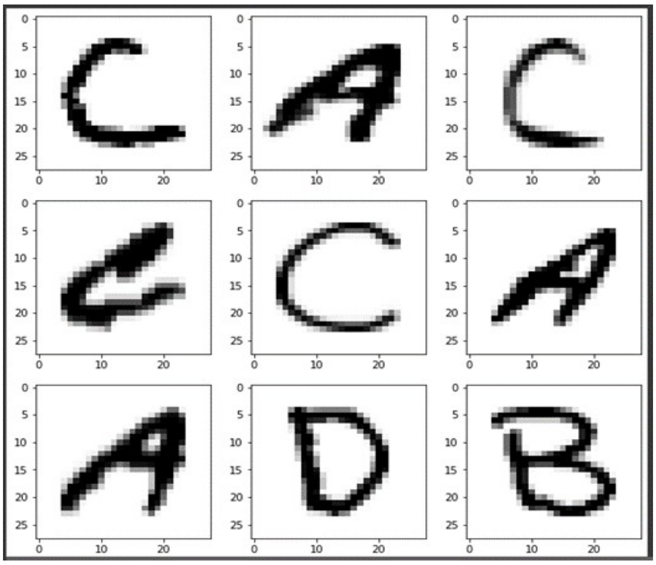

Figure 5. Randomly Shuffle Character

In this section, we'll create nine subplots of (3,3) form and use them to illustrate some of the test dataset alphabets and their predictions. Figure 6.

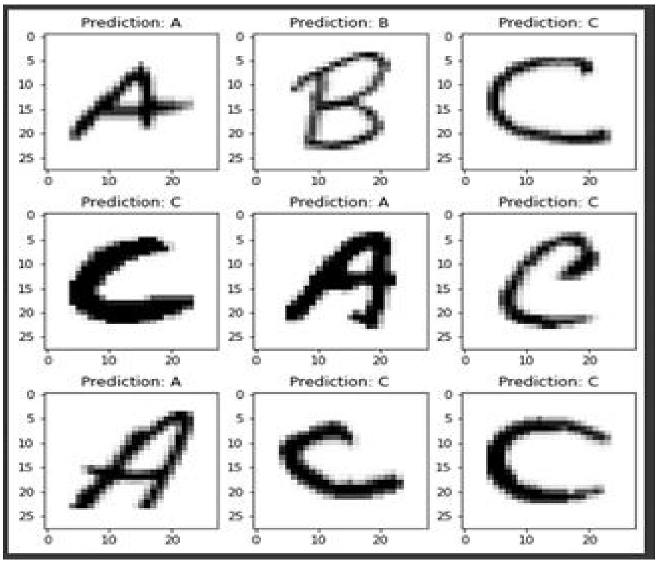


Figure 6. Prediction of character

Following the prediction of the model, we may predict the picture of any alphabet, allowing us to identify the letter in the image predicted. Figure 7.

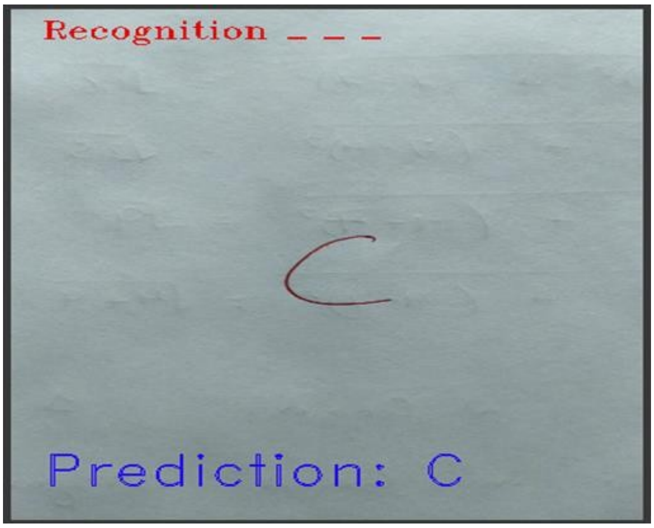


Figure 7. Recognition of Character

## VI. CONCLUSION

In this study, we offer a handwritten character recognition framework that is based on the CNN algorithm. Despite the fact that CNN is already quite strong for classification tasks, it still need an appropriate framework in which to operate in order to attain state-of-the-art performance on recognizing the handwritten english character. According to the handwritten character characteristic, we have established appropriate techniques in the framework, including sample creation, training the sample, testing the sample, and predicting the handwritten character characteristic. As a consequence, the suggested framework reached results that were above and beyond human capability. Because people's expectations of machine learning are almost limitless, the performance of humans is not the end of our study. We will continue to make improvements to the framework that has been presented in the future. The most straightforward method is to increase the CNN scale or the size of the input picture. However, such a straightforward technique necessitates the use of much greater processing resources. Developing improved sample production techniques, training schemes, and network structure for CNN is a more sophisticated approach. In addition, we will look at the feasibility of applying the suggested framework to other pattern recognition tasks, such as picture classification, if this is possible.

## VII. REFERENCES

[1] Y. B. Hamdan, "Construction of Statistical SVM based Recognition Model for Handwritten Character Recognition," Journal of Information Technology and Digital World, vol. 3, no. 2, pp. 92-107, 2021.

[2] M. S. R. K. M. J. Memon, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," IEEE Access, vol. 8, pp. 142642-142668, 2020.

[3] M. N. A. J. A. I. S. Albahli, "An improved faster-RCNN model for handwritten character recognition," Arabian Journal for Science and Engineering, vol. 46, no. 9, pp. 8509-8523, 2021.

[4] J. K. A. U. P. P. V. Mainkar, "Handwritten Character Recognition to Obtain Editable Text," Proceedings of the International Conferenceon

Electronics and Sustainable Communication Systems, ICESC 2020, no. Icesc, pp. 599-602, 2020.

[5] M. H. F. E. S. I. M. K. I. Khandokar, "Handwritten character recognition using convolutional neural network," Journal of Physics: Conference Series, vol. 1918, no. 4, 2021.

[6] Q. W. Y. X. M. J. H. L. Z. Li, "Deep Matching Network for Handwritten Chinese Character Recognition," Pattern Recognition, p. 107, 2020.

[7] D.Ciresan, U. Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification. In 2012 IEEE Conference on Computer Vision and Pattern Recognition, pages 3642-3649. IEEE, 2012.

[8] D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber. Convolutional Neural Network Committees for Handwritten Character Classifification. pages 1135-1139. IEEE, Sept. 2011.

[9] I. J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, and V. Shet. Multi-digit number recognition from street view imagery using deep convolutional neural networks. arXiv preprint arXiv:1312.6082, 2013.

[10] P. Y. Simard, D. Steinkraus, and J. C. Platt. Best practices for convolutional neural networks applied to visual document analysis. In 2013 12th International Conference on Document Analysis and Recognition, volume 2, pages 958-958. IEEE Computer Society, 2003.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classifification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097-1105, 2012.

[12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. arXiv preprint arXiv:1409.4842, 2014.

[13] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf. Deepface: Closing the gap to human-level performance in face verififi-cation. In 2014 IEEE Conference on Computer Vision and Pattern Recognition, pages 1701-1708. IEEE, 2014.

[14] J. J. Tompson, A. Jain, Y. LeCun, and C. Bregler. Joint training of a convolutional network and a graphical model for human pose estimation. In Advances in Neural Information Processing Systems, pages 1799-1807, 2014.

[15] T. Wang, D. J. Wu, A. Coates, and A. Y. Ng. End-to-end text recognition with convolutional neural networks. In International Conference on Pattern Recognition, pages 3304-3308. IEEE, 2012.

[16] Plamondon R and Srihari S N 2000 IEEE Trans. on Patt. Anal. and Mach. Intelligence 22 63

[17] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J and Chen T 2018 Patt. Recognition 77 354

[18] Ptucha R, Such F P, Pillai S, Brockler F, Singh V and Hutkowski P 2019 Patt. Recognition 88 604

[19] Khoon L L and Chuin L S 2016 Int. J. of Soft. Eng. and Comp.Systems 2 10

[20] Sallam A A, Kabir M N, Alginahi Y M, Jamal A and Esmeel T K 2020 Proc. Int. Colloq. on Sign. Proc. & Its Applications 3 225.

[21] T. Fadhaeel, P. C. H, A. Al Ahdal, M. Rakhra and D. Singh, "Design and development an Agriculture robot for Seed sowing, Water spray and Fertigation," 2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), 2022, pp. 148-153, doi: 10.1109/CISES54857.2022.9844341.

[22] Rakhra, Manik, Partho Deb, Omdev Dahiya, Sahil Sonu Chandel, Brinderjit Bhutta, Sumit Badotra, Sunny Kumar, Aman Shaukat, and Dalwinder Singh. "An Analytical Study of the Types of Implements used by Farmers in Mechanized Agriculture." In 2022 International Mobile and Embedded Technology Conference (MECON), pp. 683-687. IEEE, 2022.

[23] Shruti, R. Soumya, M. Rakhra, N. S. Chintagunti, B. Singh and D. Singh, "Modern Data Mining Approach to Handle Multivariate Data and to Implement Best Saving Services for Potential Investor," 2022 International Mobile and Embedded Technology Conference (MECON), 2022, pp. 611-616, doi: 10.1109/MECON53876.2022.9752101.

[24] M. N. Gowda, D. Singh and M. Rakhra, "Machine Learning – Based Diagnosis of Covid-19 using Clinical Data," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), 2022, pp. 910-916, doi: 10.1109/ICIEM54221.2022.9853083.

[25] Chhabra, Abhishek, Tenzin Woeden, Dalwinder Singh, Manik Rakhra, Omdev Dahiya, and Aditya Gupta. "Image Steganalysis with Image decoder using LSB and MSB Technique." In 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), pp. 900-905. IEEE, 2022.

[26] M. Rakhra and R. Singh, "Economic and Social Survey on Renting and Hiring of Agricultural Equipment of Farmers in Punjab," 2021 9th Int. Conf. Reliab. Infocom Technol. Optim. (Trends Futur. Dir. ICRITO 2021, pp. 1–5, 2021, doi: 10.1109/ICRITO51393.2021.9596343.

[27] Takkar, Sakshi, Anuj Kakran, Veerpal Kaur, Manik Rakhra, Manish Sharma, Pargin Bangotra, and Neha Verma. "Recognition of Image-Based Plant Leaf Diseases Using Deep Learning Classification Models." Nature Environment & Pollution Technology 20 (2021).