Offline Writer Identification using Convolutional Recurrent Neural Network

Deepa Bendigeri[1]* Gopal A. Bidkar[2] Jagadeesh D.Pujari[3]

[1]Department of Information Science and Engineering, SDM College of Engineering and Technology, Dharwad, India
[2]Department of Electronics and Communication Engineering, SDM College of Engineering and Technology, Dharwad, India
* Corresponding author's Email: deepabendigeri12@gmail.com

Abstract:
Writer identification can be used to find fakes and help with forensic science. Handwritten text is the only way to identify the author offline. Recently, convolutional neural networks (CNNs) have become the best way to classify images on a large scale. The recurrent neural network (RNN) models the spatial relationship between the fragment sequences in order to improve the local fragment feature's capacity for discrimination. To get the best of both models, combine CNN and RNN to create a Convolution Recurrent Neural Network (CRNN). Therefore, the CRNN model for offline writer identification is suggested in this paper. The proposed method achieved 96% accuracy in classifying 690 writers using CRNN. The suggested method can make efficient and strong writer identification based on different sizes, different orientations, or both. The efficiency of the proposed system has been demonstrated. Finally, an accuracy comparison of CRNN with CNN and RNN has been conducted.

Keywords: Writer identification, Offline analysis, Convolution Neural Networks, Deep learning, Recurrent Neural Network, Convolutional Recurrent Neural Network, Long short-term memory.
DOI Number: 10.14704/nq.2022.20.12.NQ77054 **NeuroQuantology 2022; 20(12): 704-711**

1. Introduction

Handwriting is a natural ability that children acquire and which cannot be duplicated. Therefore, no two people's handwriting is identical, and even an individual cannot reproduce their own handwriting perfectly [1]. Last is "variation," which is when a person's writing varies in a natural manner. It is an extremely reliable method for determining a person's identity, and forensic document professionals use it often to verify that someone is who they claim to be. Currently, researchers are focusing heavily on the automated analysis of handwritten papers, particularly in the area of analyzing historical documents. This is because it takes a considerable amount of time for a forensic specialist to manually examine a document and compare it to all the other papers in a database in order to detect a forgery. So, making an automated way to find the author could be very helpful. It would make it easy to find the text written by a suspected author in a large collection of documents, which would make the work of forensic experts easier.

Handwriting is a behavioural biometric identification, unlike fingerprints or iris scans, which are physical. It is altered by factors such as going to school and ageing. Formally, writer identification is the search for a single author within a huge data corpus. It is mostly used in sectors like security and forensics. In recent years, though, it has become more interesting to try to figure out who wrote a piece of paper. [2], [3].

The writer identification is divided into two categories: a) online author identification, which includes information about when the text was written; and b) offline author identification, which uses simply the handwritten text, i.e., a scanned picture of handwritten data is used to identify who authored it [9].

Author identification is used to select the person whose handwriting most closely resembles the genuine author from a pool of registered candidates. Therefore, humanities professionals may determine who authored a certain paper by using techniques for determining who produced handwritten text. Graphology suggests that a person's handwriting is a wonderful way to learn about them since it reveals their personality and feelings. Because the brain teaches the nervous system, hand, arm, and fingers how to move the writing tool, handwriting is also referred to as "brain writing." Therefore, the way your brain functions reveals the kind of person you are [4].

Depending on the topic of the writing, there are two types of offline writer identification: those that do not depend on the writing and those that do. For identifying the author of a text without referring to its content, there are characteristics for work-independent author identification [15]. Contrarily, text-dependent techniques require an input image with fixed text and contrast it with previously saved templates. It also goes by the names "script-based identification" or "content-based identification. Text-dependent strategies operate at the letter or word level, whereas text-independent strategies operate at the line or paragraph level. The first step in the framework for author identification is to decide if the method will be done online, offline, based on the text, or not based on the text. Text-independent offline author identification involves data collection, preprocessing, feature extraction, and classification [10] or identification.

CNN and RNN may be integrated into a single network known as a Convolution Recurrent Neural Network (CRNN) to get the advantages of both. In this network (CRNN), there are several layers of convolutions followed by multiple recurrent layers.

The majority of recent research has centred on creating convolutional neural networks (CNN) and assessing them against a simple classification task for datasets without sequential information, such as the identification of specific handwritten digits or characters [6][16]. This method stands out because it tackles more difficult issues like sequence learning and recognizing handwritten words or sentences. Recurrent neural networks and convolutional neural networks (RNN) must be linked together using flexible and evolving deep learning algorithms in order to overcome this problem (RNN). The sequential text is fed into the language-optimization-based, bespoke deep learning algorithm (CRNN) which has been developed. In the above study, RNN unit evolution was used on its own instead of convolutional neural networks.

2. Literature review

An attention-paying global regular network (GRN) was proposed by Schomaker et al. in 2021. The GRN network has two branches: a branch extracts features from the world from page handwriting, and the second branch extracts local features from word handwriting. To create the overall features of handwriting, global and local aspects combine in a global residual approach. A branch is added to the proposed GRN to extract page-specific features, and a residual attention network is used to extract local features. Experiments show that both tactics are effective. On the CVL dataset, the model outperforms state-of-the-art structures with an impressive top1 accuracy of 99.98% and a top5 accuracy of 100% while requiring less training time and network parameters. The experiment demonstrates the network's strong capabilities in the area of writer identification [14].

He and Schomaker (2020) offered brand-new benchmark research for author recognition using single-word words or text block visuals. Using a deep neural network known as FragNet, the main properties of these word images are recovered. The feature maps on the feature pyramid, which are used to generate feature maps from images, and the fragments recovered from the input image, are used to train the fragment route to predict the identity of the writer. When tested on four benchmark datasets [7], the results show that the authors' proposed technique may be able to create effective and reliable deep representations for author identification based on both word and page imagery.

Rehman et al. (2018) provided a thorough analysis of writer identification techniques and planned to offer dataset taxonomy, feature extraction techniques, and classification for author identification based on conventional and deep learning. The script divided the conversation into English, Arabic, Western, and other languages for the reader's convenience. From the perspective of the algorithm and methods, they divided the discussion according to the order in which the implementation processes should be carried out. In the end, they talked about the problems and unanswered research questions in the field of identifying writers [8].

Hafemann et al. (2017) proposed utilizing convolutional neural networks to discover writer-independent signature representations from image signatures. To capture visual signals that distinguish real signatures from forgeries regardless of user, they propose a novel formulation of the problem that incorporates knowledge about competent forgeries from a subset of users into the process of feature learning. Four datasets, including the GPDS, MCYT, CEDAR, and Brazilian PUC-PR datasets, were thoroughly analyzed. Using cutting-edge approaches, they significantly improved performance on the GPDS-160, achieving an equal error rate of 1.72 percent as opposed to 6.97 percent in the literature. They also showed that the characteristics work outside of the GPDS dataset, with results that are better than the best in the other datasets without having to change the way the characteristics are shown for each dataset [9].

In 2016, Xing and Qiao showed Deep Writer, a deep multi-stream CNN that finds a powerful deep representation for finding writers. Deep Writer, which is trained via soft maximum classification loss and receives local handwritten patches, For the writer identification assignment, they build and optimise a multi-stream structure; for Deep Writer's performance, they introduce data augmentation learning; and for handling text pictures of varying lengths, they introduce patch scanning. They also learn that combining training can enhance performance and that writers in other languages, including English and Chinese, may have some

characteristics. Experiments on the IAM and HWDB datasets show that the models have a high identification accuracy rate of 99.01 percent on 301 authors, 97.03 percent on 657 writers, and 93.85 percent on 300 writers with one Chinese character input. These outcomes perform notably better than those of earlier methods. Furthermore, their systems achieve 98.01 percent accuracy on 301 writers with only four English alphabets as input [5].
This paper presents a deep learning technique for offline writer identification combining CNN and RNN, i.e., CRNN. The paper is organized as follows: Section 2 contains a review of literature, Section 3 discusses about proposed method for writer identification, Section 4 discusses about Results and Discussion. Section 5 discusses concluding remarks.

3. Proposed Method

The proposed research work is given below:

3.1 Convolutional Recurrent Neural Network (CRNN)

The convolutional recurrent neural network model outperforms the traditional CNN for recognizing sense text when end-to-end training employs sequence labelling of picture words as opposed to annotation for each character image [7], [11], [12]. Importantly, the CRNN can automatically extract valuable characteristics without relying on manually generated features.

There is no need for a segmentation step because the recurrent layer of the CRNN model offers significant flexibility for handwritten image sequence data, which is one type of sequence data. CRNN models are based on a convolutional neural network with many layers. Each convolutional layer may contain hyper parameters like the activation function, batch normalisation, pooling operation, number of kernels, kernel size, and skip connections. Convolutional neural networks are widely employed to extract features in CRNNs. The pixel-intensity data is sent into the CNN's first convolutional layer, which then transmits them to the succeeding layers. A CNN sends a collection of feature sequences to a recurrent neural network (RNN). In the final step of the transcription layer, Connection list temporal classification converts the result prediction into a label sequence (CTC). Fig.1 demonstrates the CRNN framework.

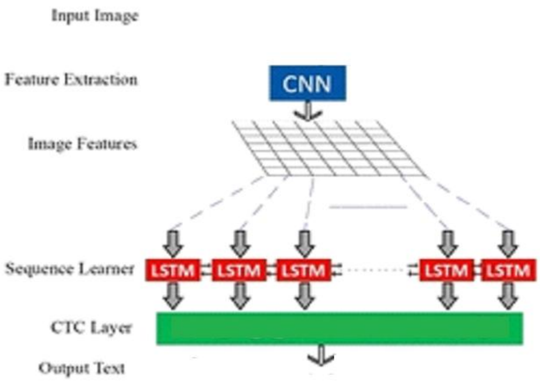

Figure.1 CRNN Frame work

The differences between long short-term memory (LSTM) and bidirectional long short-term memory RNNs (BLSTM) are investigated in this paper. The core of the LSTM network structure is a gate mechanism with long-term dependability for solving problems [19]. There is an input gate, a forget gate, and an output gate in the gate mechanism. Fig.2 depicts the procedure for the LSTM gate mechanism.

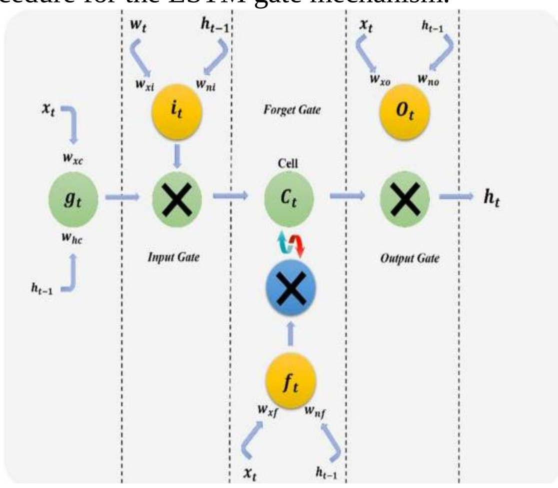

Figure. 2 The standard LSTM cell

3.2. CRNN Configuration

The three primary elements in this study that are suggested to establish the ultimate design of the CRNN are general hyper parameters, convolution hyper parameters, and long short-term memory network hyper parameters [13],[17],[18]. The hyper parameters for the entire CRNN model and configuration network are provided by general hyper parameters. Consideration is given to the batch size, optimizer, learning rate, number of LSTM layers, and number of convolutional layers. The following gives a description of these hyper parameters:

- **Batch size**: There are B training samples collected in a single batch (16, 32, 64, and 128).
- **Optimizer**: Based on the training data, the CRNN network weights are iteratively updated using the adaptive learning rate optimization technique.
- **Learning rate**: The learning rate is the rate at which the weights vary while you train. One decision variable in the solution, LR, represents the learning rate. It can take one of seven values: (1 105, 5 105, 1 104, 5 103, 5 102, 1 102, 5 102).
- **Number of convolution layers**: How many convolutional layers should be added to the CRNN will depend on this decision variable. The first three layers are required in all created people since handwriting recognition is thought to be a complicated classification problem, ensuring the automatic detection of the key traits. The complexity of the model and the number of weights may both rise as the number of convolution layers increases. As a result, the number of convolutional layers is restricted in CRNN to 10, and the DC-search CRNN space allows us to select that number as the maximum.
- **Number of LSTM layers**: How many LSTM layers should be added to the CRNN will depend on this decision variable.

For the rest of the hyper parameters, which can be different for each convolution layer in the network, here are some more factors to consider:

- The number of kernels in each convolution layer is known as the convolution kernels (ck).

(4, 8, 16, 32, 64, 128, 265, 512).

- Convolution kernel size (cs) is the size of the kernel utilised in each convolution layer (2, 3, 4, 5, 6, 7, 8, and 9).
- Convolution batch normalisation (cb) is a method for applying batch normalisation, which is widely used to boost a neural network's efficiency and speed. It is used between nonlinearity layers like max pooling and ReLU and the convolution layer. The decision variable for batch normalisation in this solution has a range of 0 to 1 (0 or 1).
- Convolutional neural networks adopt convolutional activation function (ca), the most common and default activation method in deep learning networks. A more suitable function has to be chosen for this network, such as "relu," "linear," "elu," "selu," or "tanh."
- A pooling layer called convolution pooling size (cp), which requires fewer parameters and less computation from the network, reduces the size of the representation of the input handwriting image. Despite the fact that pooling layers are required to ensure an acceptable computation time during the optimization procedure that determines the optimal network architecture, employing an excessive number of them frequently results in the elimination of vital characteristics or a reduction in the size of the representation (1, 1). The likelihood is set using pooling at the end of each convolution layer to 50%, the pool size to 2, and the stride to 2, 2, and 1 in the decision variables. A skip connection is a connection that is used to increase training convergence and performance (cr).

When setting the hyper parameters for each LSTM layer, the size of the hidden layer (Rh) and the kind of LSTM layer, which could be a bidirectional layer, are also taken into consideration (Rb). The character error rate (CER), which must be minimized, is utilized to rank a section of the dataset's data utilizing all of these hyper parameters in a CRNN structure [20]. To optimize the solution, the settings in Table 1 and the suggested HSSA are used.

Table 1. Solution representation for CRNN configurations

Decision Variable Sectors	Bits Used for Sector-Wide Decision Variables	Hyper parameters	No. of Bits for Each Hyper Parameter
General parameters	10bits	Bs (batch size)	2
		Op (optimizer	1
		Lr (learning rate)	3
		Nc (number of convolution layers)	2
		Nr (Number of LSTM layers)	2
Convolution layer parameter	11bits x7 layers=77 bits	Ck (number of Kernels	3
		Cs(Kernel size)	3
		Cb(batch normalization)	1
		Ca (activation function)	1
		Cp (polling size)	2
		Cr (skip connection or not)	1
Recurrent network parameters x4	3 bits x 4 layers =12 bits	Rh (size of hidden layer)	2
		Rb (bidirectional)	1

4. Results and Discussion

The experimental results of offline writer identification using convolutional recurrent neural network models are shown below. Table 2 shows the different parameters and variables that were used to identify the writer using CRNN for this experiment.

Table 2. Various parameters/variables used for writer identification

Parameters /Variables	Values
Data Set	IAM Data Set
Number of writers/Classes	680
Number of documents per writers	25
Each documents is divided into Number of image documents per writers	8 sub image 25 x8=200 image/writers
Total Data Set	680 x200 =136000 Samples
% of Training and Testing from total Data set	70% -Training ,30%-Testing
Number of Samples in Training Data set	680 x140x5x4 =1904000 Samples
Number of Samples in Testing Data set	680 x60x5x4 =816000 samples
Different Orientation	0° , 10° , 20°, 30° ,40°
Scale	Zoom in=.5,.75 Zoom out=1.5,2
Max No. of epoch	70
Num of layers	16 layers(including 2 LSTM Layers)
Convolution filters size	3x3
Pooling	Max polling
Pooling filters size	2x2
Classifier	Soft max

4.1 Writer identification using CNN:

Table 3 depicts experiments conducted on the IAM dataset with 680 writers, where each writer writes 25 documents, and each document is divided into 8 sub-images. The maximum number of image documents per writer is 200 image documents per writer. Each document is rotated in five different orientations and four scale factors. 70% of the 200 image documents per writer are used for training and 30% of the 200 image documents per writer are used for testing.

The average accuracy and standard deviation are computed scale-wise, orientation-wise, and overall combination-wise. From this, the result is analyzed for three inferences, i.e., the system is sensitive to scale, sensitive to orientation, and sensitive to both scale and orientation. According to Table 3, the proposed method is robust and invariant to scale, orientation, and both because the standard deviation is small. For every training session, as epoch increases, accuracy will go on increasing. Fig. 3 depicts the accuracy graph of CRNN-based writer identification. At a value of 70 in Epoch, CRNN is 96% accurate at figuring out who the writer is. A comparison of CNN, RNN, and CRNN for offline writer identification is shown below in Table 4 and Fig. 4.

Table 3. The average accuracy of all writers with different orientations and scales

Orientation → Scale ↓	**0°**	**10°**	**20°**	**30°**	**40°**	**%Avg. Acc**	**Std Dev**
Zoom in 0.5	96	95.8	95.9	96	95.9	95.92	0.04
Zoom in .75	95.9	95.6	95.9	95.9	95.7	95.8	0.1
Zoom out 1.5	95.7	95.7	95.6	95.8	95.8	95.72	0.08
Zoom out 2	95.8	95.6	95.8	95.7	95.9	95.76	0.11
Average	95.8	95.7	95.8	95.9	95.8	-	-
Std Dev	0.12	0.01	0.12	0.13	95.8	-	-

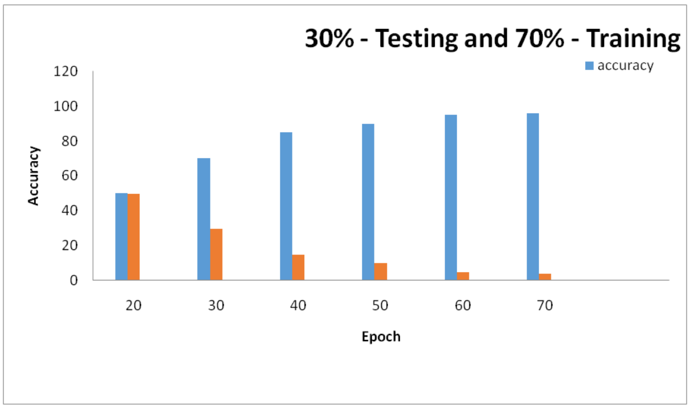

Figure. 3Accuracy Graph of CNN based Writer Identification

Table 4. Accuracy comparison of CNN, RNN and CRNN models for offline writer identification

Sl. No	**Methods**	**Accuracy in percentage**
1	Convolution Neural Networks	90%
2	Recurrent Neural Network	93%
3	Convolutional Recurrent Neural Network	96%

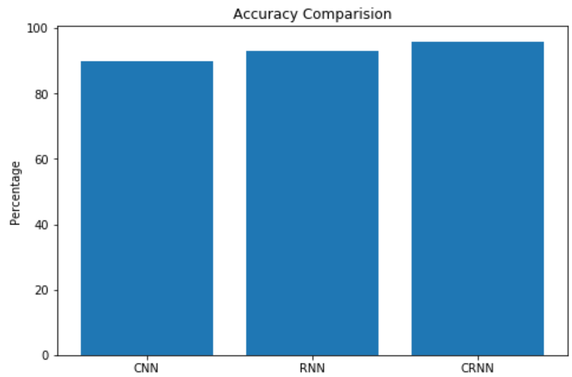

Figure. 4 Accuracy comparisons of CNN, RNN and CRNN models for Offline writer identification

5. Conclusion

This work presents a CRNN-based technique for offline author identification. The experiments were conducted on the IAM data set with 680 writers. The evaluation result demonstrates that CRNN achieves excellent accuracy and efficacy, which may be attributed to CRNN's capabilities. Consequently, this neural network is effective. Based on the size, orientation, and both handwritten documents, the suggested method is a quick and reliable way to figure out who wrote something. The CRNN presents a new classification model for offline writer identification that is efficient and of good quality. From the CRNN, 96% of accuracy is achieved for writer identification. In the future, this experiment could be conducted using different languages and multiple scripts. Finally, the comparison of CRNN with CNN and RNN is given, with the conclusion that CRNN performs better and achieves good accuracy. The future work for the present research work can be carried out for different languages with multiple scripts.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing-original draft preparation, writing-review editing and visualization, have been done by 1st author. The supervision and project administration have been done by 2nd and 3rd authors.

Acknowledgments

I would like to express my deep gratitude to Dr. Gopal A Bidkar and Dr. Jagadeesh D Pujari, my research supervisors, for their patient guidance, enthusiastic encouragement and useful critiques of this research work.

I would also like to thank Dr. P.S Hiremath sir, for his advice and direction which helped me to keep my research progress on schedule. Finally, I wish to thank my family and my colleagues for their support and encouragement throughout my study.

References

[1] Y.Zhu, T.Tan and Wang Y.W.Y, "Biometric Personal Identification Based on Handwriting", In: *Proc.of 15th International Conf. on Pattern Recognition (ICPR*), Barcelona, Vol. 2, pp. 2–5, Sep 2000.

[2] Brink, J.Smit, M. Bulacu and L. Schomaker, "Writer Identification Using Directional Ink-Trace Width Measurements", *Pattern Recognition*, Vol.45, No.1, pp.162–171, Jan 2012.

[3] T. Gilliam, R. Wilson and J.Clark, "Scribe Identification in Medieval English Manuscript", In: *Proc. of 20th International Conf. on Pattern Recognition (ICPR*), Istanbul, pp. 1880–1883, Aug 2010.

[4] G.X.Tan, C.Viard-Gaudin and A.C. Kot, "Automatic writer identification framework for online handwritten documents using character prototypes", *Pattern Recognition*, Vol.42, No.12,pp.3313-3323,2009

[5] L.Xing and Y.Qiao, "Deep Writer: A Multi-Stream Deep CNN for Text independent Writer Identification", In: Proc.of *15th International Conf. on Frontiers in Handwriting Recognition (ICFHR*), pp. 584-589, June 2016.

[6] R. Anil, K. Manjusha, S. Sachin Kumar, and K. P. Soman. "Convolutional Neural Networks for the Recognition of Malayalam Characters", In: *Proc. of of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA)*, pp.493-500, 2014.

[7] S.He and L.Schomaker, "Writer identification using deep fragment networks",*IEEE Transactions on Information Forensics and Security*, vol.15,pp.3013–3022, March 2020.

[8] A. Rehman, S.Naz and M.I. Razzak, "Writer identification using machine learning approaches: a comprehensive review", *Springer, Science+ Business Media, LLC, part of Springer Nature*, Vol.78 (3), pp.10889–10931, September 2018.

[9] L.G.Hafemann, R.Sabourin, and L.S.Oliveira, "Learning features for offline handwritten signature verification using deep convolutional neural networks", *Pattern Recognition*, vol.70, pp. 163–176, 2017.

[10] B.S. Prachitee and S.D. Sheetal, "A classification technique using associative classification learning", *International Journal of Computer Applications*, vol. 20, no.5, pp. 20-28, 2011.

[11] M.Elleuch and M. Kherallah, "Boosting of deep convolutional architectures for Arabic handwriting recognition", *International Journal of Multimedia Data Engineering and Management (IJMDEM)*, Vol. 10,No.4, pp. 26-45, October 2019.

[12] A.T. Kabakuş, "A Comparison of the State-of-the-Art Deep Learning Platforms: An Experimental Study",*Sakarya University Journal of Computer and Information Sciences*, vol. 3, no. 3, pp.169- 182, 2020.

[13] L. Vasquez, A. G. Ravelo-García, J. B. Alonso, M. K. Dutta and C. M. Travieso, "Writer identification approach by holistic graph metric features using off-line handwritten words", *Neural Computer Application*, vol. 32, pp. 1–14, 2018.

[14] S.He and L.Schomaker, "Global-Context Residual Recurrent Neural Networks for Writer Identification", *Pattern Recognition*, Vol.117, pp.107975, September 2021.

[15] H. T. Nguyen, C. T. Nguyen, T. Ino, and M. Nakagawa, "Text-independent writer identification using convolutional neural network," *Pattern Recognition Letters*, Vol. 121, pp. 104-112, 2019.

[16] N.Altwaijry and I.Al.Turaiki "Arabic handwriting recognition system using convolutional neural network", *Neural Computing and Applications*,Vol. 33 ,No.7,pp.2249–2261,April 2021

[17] G.Tong,Y. Li, H.Gao, H.Chen, H.Wang and X.Yang, "A multi-scale attention CRNN for Chinese text line recognition in natural scenes", *International Journal on Document Analysis and Recognition (IJDAR)*, Vol.23, pp.103–114, 2020.

[18] L.Dawei,M.Wu and Y.Hu, "Offline Writer Identification Using Convolutional Neural Network and VLAD Descriptors", In:*Proc.of 7th International Conf. On Artificial Intelligence and Security(ICAIS)*,Dublin,lreland,pp.253-264,July 2021

[19] Vidushi Garg, "Handwritten Text Classification using Deep Learning", *International Research Journal of Engineering and Technology (IRJET)* ,Vol.7,No.05,pp.3485-3490,May 2020.
[20] N.Zouhaira, M.Anis and K.Monji, "Densely Connected layer to improve VG Gnet-based CRNN for Arabic handwriting text line Recognition", International Journal of Hybrid Intelligent Systems, Vol. 17, No. 3-4, pp 113–127,2021