Handwritten Character Recognition using Neural Networks forBanking Applications

Shreya Mhalgi
BE,Computer Engineering Student
MMCOE, Karvenagar
Pune, India

Ketki Ganu
BE, Computer Engineering Student
MMCOE, Karvenagar
Pune, India

Prajakta Marne
BE, Computer Engineering Student
MMCOE, Karvenagar
Pune, India

Radhika Phadke
BE, Computer Engineering Student
MMCOE, Karvenagar
Pune, India

Swati Shekapure
Asst Professor, Computer Engineering
MMCOE, Karvenagar
Pune, India

ABSTRACT

Banks often accept handwritten forms for various purposes like application for creating or closure of accounts, loans, net banking, etc. The form takes a lot of user information consisting of sensitive data viz. Aadhar card number, pan card number. This information is usually taken in pen-paper format and needs entry to the bank database to document the particulars in the system or the bank requires to store a physical copy of the form for future reference. Manual entry of these details into the bank database is a tedious process and might be erroneous at times. Also, maintaining the original copy of the form or like document generate stockpiles of paper. In an attempt to overcome these discrepancies, the proposed problem statement provides a solution by making use of Handwritten Character Recognition which will input data in the form of an image to store and maintain it in a digital library

Keywords

Artificial Neural Networks, Deep learning, Convolutional Neural Networks(CNN), Handwritten Character Recognition(HCR)

1. INTRODUCTION

Despite the availability of digital writing tools, many tasks such as filling the form for Bank related work are still preferred as pen paper. This cannot be considered as a limitation of technology but the convenience of users over keyboard, mouse, and touchscreens. Loss of information from physically stored records, difficulty in accessing the records, erroneous manual entry to the database are the common problems faced especially in the Indian banking sector. Taking credentials from customers is pen-paper based, but at the same time, it is difficult to store and access physical documents efficiently. Also manual entry of user credentials to the database, introduces the risk of sensitive data not being digitalized or erroneous manual data entry to the system. Thus, storage, access, and retrieval could be made digital, without any need to store the physical hard copy of the document. This process can find efficient solutions by making use of Deep learning technology.[1] Handwritten character recognition is the process of recognizing the handwritten text, then feeding it to the character recognition model as an electronic translation of images.[2] The process would happen by making use of the Convolutional Neural Networks (CNN)-image processing and then storing the contents of the image into the database in machine-editable text format, using the python libraries. Convolutional Neural Networks are primarily used in the field of pattern of pattern recognition and are analogous to ANN. CNN comprise of neurons that have the ability to optimize themselves through learning that is these take raw input vector and perform operations to get a final output of the class score.[3] In this proposed system, the system would take blocks of handwritten characters as an input vector and the same would be recognised and stored in a database.

1.1 Why use CNN?

Architecture of CNN is well suited for 2D data and it uses 2D convolutional layers to process data such as images. Convolutional layer is the key building block of CNN. According to Alexander Del Toro Barba, the ConvNets are computationally efficient than the machine learning algorithms and have the ability to handle unstructured data, thus have successfully evolved in areas such as image recognition.[5]
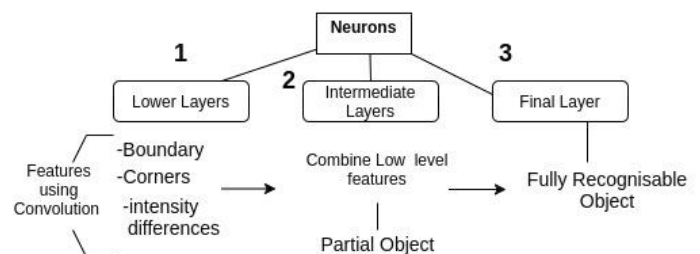

Fig. 1. Feature Extraction in Neurons

Matrix multiplication with different sections using randomly generated matrices (Kernels) is done to find a good feature. If good feature is not found, the algorithm backpropagates for improvement.[6] Using this improving algorithm, the machine self-learns.

1.2 CNN over LSTM

LSTM and RNN and derivatives make sequential predictions of given data, that is to reach the present processing cell, the long-term information has to travel through each cell. Although LSTM can is able to learn and remeber and learn a lot of information, the switch gates in the path make it rather more complicated.

In contrast CNN makes spatial-correlations. According to [6], LSTM is a general class of the Recurrent Neural Networks(RNN). RNN is mainly used for problems, which require long range information to be remembered.

1.3 Problems with RNN [7]

1. Vanishing gradient-
When updating gradient in back-propagation is small the algorithm does not learn anything,
2. Exploding gradient-
Model cannot handle the large values generated due to the same matrix multiplication, thus can crash.

The purpose of this problem statement concerns the bank employees convenience for managing customer particulars at the same time providing an efficient service to the customer.
The proposed system is a technological solution without substituting the pen-paper input from the user in the banking sector. The aim is to digitalize the handwritten text by scanning the document and storing it in digital format. Thus, the data can be stored into tables that can be retrieved easily as and when required. In this project, the challenging task would be selecting particulars from the form and putting it in the desired table. The approach to this problem would use offline handwritten character recognition using Convolutional Neural Networks.

2. LITERATURE SURVEY

Introduction of the optical character recognition and knowing the advantages of this technology gave rise to the conduction of experiments with OCR. The second generation of IBM OCR was used to recognize numerals back then in the 1960s.[10]
Upgrading the generations, further research was carried out in 1975-1980, for OCR to operate on large hand-written data sets with poor print quality than before. Online character recognition used to put some restrictions on the orders and number of strokes for Japanese characters. This disadvantage was overcome in the offline methodology for Japanese character recognition, which used nonlinear shape normalization.[11] In the year 1999, recognition was performed as a combination of online and offline character recognition where bitmap image was used as an input for offline method thus achieving a 73 percent recognition rate. Further, the use of the Hidden Markov Model (HMM) in the recognition systems was encouraged to upgrade the recognition rate for offline character recognition.[12] Algorithms such as image segmentation and recognition in union with the HMM suggested optimization of word recognition problems.

In the year 2005, Liu et al proposed a system of gradient-based feature extraction. The moto being to improve the efficiency of character recognition using the demarcation schemes. Further, in 2006, the system was improved making use of low character resolution.
Use of glyph image quality estimation and glyph recognition was used for the same.[13] The proposed system by Kannan et al (2008) was used for recognition of preprocessing cursive Tamil characters made use of the HMM(Hidden Markov Model) was claimed robust and flexible at the same time.[14] Use of neural networks for character recognition purposes started gaining momentum then. In 2009, Prasad et al proposed a technique for recognition of characters from the Gujarati language, making use of neural networks. Image enhancement and prepossessing were carried out before image recognition.[15] Use of CNN peaked especially after 2011 when handwriting analyzed by Ciresan achieved a tiny 0.27 percent error rate.[16] But the document recognition work using the neural networks was instigated way back in 1998.[17] In the hybrid system based on Generalized Regression Neural Network which used wavelet transform (2012), gave an insight that classification algorithms had more robustness than the algorithms used till then.[18]

CNN is a computationally expensive algorithm like any other neural network algorithm. The reason behind this is CNN requires a huge amount of training data and cannot properly handle input transformations.[19] Although these intensive tasks are handled much better with the advancements in the hardware like GPUs, the errors in handwritten text recognition differ to an extent and as a result, the outputs are averaged.[20] However, the NYtimes has already harnessed this method to store the old print in their archives. Thus, the banking application of CNN of scanning forms and storing them into databases foresees a compelling solution for the current bank operations. CNN has embarked in the applications by overcoming many approaches in image processing, yet persists with some drawbacks. For instance, they are not robust to affine transformation and do not take the spatial relationships within the image into considerations. A fairly new trend in research has come up with a robust solution for the drawbacks CNN which would even help to get rid of the 0.27% error in the handwriting analysed by Ciresan. Capsule networks providing a solution for CNN by training the kernels to activate specific parts of the object thus setting an equivariance among the kernel equations. CapsNets give an improved performance with the combined use of kernels and dynamic routing to collect opinions from individual capsules.[21] Neural networks have now the ability to learn with a mere 200 training samples per class,[22] making the conversion of handwritten text to digitized text more efficient and accurate.

2.1 Advantages of the System

—A handwritten form can be digitalized and stored in database easily.
—Easy to store and access information.
—A lot of time can be saved.
—Helping bank employees to reduce the inaccuracy caused due to manual entries.

2.2 Limitations of the system

—Can only detect English characters.
—Can only detect English characters.
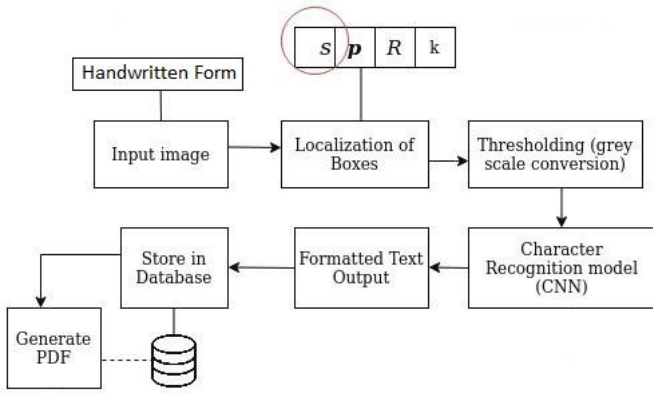—Quality of the final image depends on quality of the initial image.
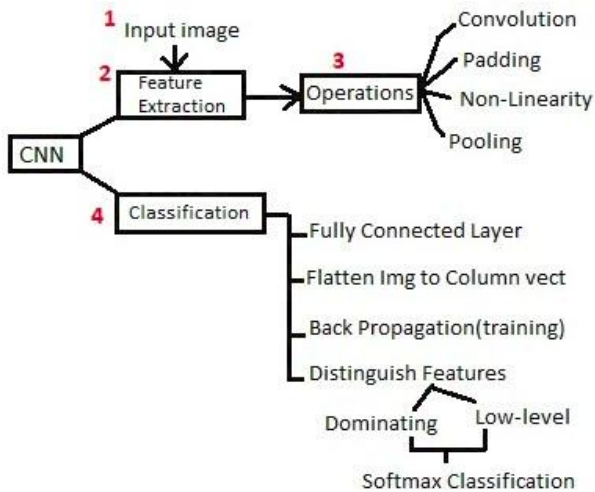
Fig. 2.   Proposed System Architecture.

Fig. 3.   Convolutional Neural Networks.

Fig. 4.   Original Form

After thresholding (replacing each pixel with a black pixel if the image intensity is than some fixed constant or a white pixel if image intensity is greater than the constant) the image, box detection would begin.

3. IMPLEMENTATION

CNN will be trained with the NIST dataset[24]

3.1 Input

The input would be fed to the system via a scanner or any optical device. It would contain the customer details which have to be stored in the bank system database.

3.2 Localization of Boxes

Boxes would be detected on the inverted image by detecting the vertical lines, horizontal lines and summing them up.

Fig. 5.   After thresholding the image

Fig.-. shows the steps in box detection

3.3   Preprocessing

Before the image is given to the CNN model for recognition, we have cleaned it in three steps as follows-

1. Gray scaling:
It is the process of converting an image from other color spaces to shades of gray. Gray scaling ensures dimensionality reduction-color spaces(RGB, CMYK, etc) having different directions are reduced to one, thus reducing the complexity of the model.[25]

Fig. 6.   Gray Scaling

2. Black and white (binary) conversion:
Often a gray scale image is converted to a binary image to define the part of image that needs to be recognised. Such pixels could be set to 1 for more accurate recognition.

Fig. 7.   Black and white conversion

3. Erosion:
It is called as the morphological operation in which the pixels from the object boundary are removed so that the substantive object remains and small white noises, islands and objects are removed.[26]

Fig. 8.   Erosion

3.4 CNN Model

The CNN model developed for this bank system consists of 10 convolutional layers, 5 max-pooling layers, and 3 dense layers, a total of 13 layers. Images of size 64 x 64 are fed to the model for training and are passed through a stack of convolutions equipped with ReLU[28], followed by max pooling, which is done in a 2 x 2-pixel window similar to[29]. Upon training a subpart of the NIST dataset consisting of Block letters and numbers a training accuracy of 95% was achieved with 10 epochs. Further, training the entire NIST dataset which comprised Block letters, small case letters as well as numbers a training accuracy of 86% was reached, keeping the number of epochs constant.
Further testing the model with entire NIST dataset produced the following results (Table 1.)-

Table 1. Accuracy of the model

No.	Handwritten Form Type	Accuracy
1	Block letters + numbers	74.04
2	Block letters + small case letters + numbers	86.0

3.5 Keras for implementation

Kears is especially preferred for the implementation of neural networks due its high level implementation modules.With neural networks getting deeper, it might become difficult for the user to tune with all the parameters. Keras makes this tasks easy. It abstracts low level libraries like TensorFlow[30] and enables the user to not to focus on the low level details of the implementation.[31]

3.6 output of predict.py to text file to do corrections(if any)

After training the dataset with CNN and testing the input against the output, the result generated from the predict.py file is then put to a text file that is opened for the bank employee to check for any discrepancies in the credentials that were scanned. The bank employee gets the access to do corrections (if any) to the scanned form. Normally, the output on the console(for the python file) can be written as it is to a text file with the terminal command(Ubuntu) $python3 file.py > textfile.txt$. The same command was included in the implementation for the automatic generation of the text file from the predicted output by importing the os module.[32]

3.7 Textfile to a table in Database

The content of the textfile is read line by line making use of the file handling operations in python.[33] The value name: XYZ is split using the splitline function and only the value XYZ is appended to the particular column in the table of the MySQL database.

### 3.8 Save form as pdf with time stamp and date, text file is deleted

The textfile is now converted to a pdf file using the FPDF library [34] in python. The data from textfile is now stored in a pdf file and this pdf is stored on the local machine with the current date and time stamp. This pdf will hold all the details on the form which was filled by the customer. This file will act as an e-form and serve as the backup for the database.

4. CONCLUSION AND FUTURE WORK

This paper proposes a system to encourage the use of Handwritten Character Recognition methods in the Indian banks which would ease out the data management, access, maintenance, and storing facilities. This would certainly contribute to making banking transactions more efficient and more autonomous. As already automated passbook printing machines have helped to dequeue the lines in front of the desk, implementing this system. Convolutional Neural Networks have been the center-forward in image recognition algorithms, yet the system seeks a future with optimal training sets with capsule networks.

5. REFERENCES

[1] Xin Jia. Image recognition method based on deep learning. 28-30 May 2017.

[2] Palak Patel Chirag I Patel, Ripal Patel. Handwritten character recognition using neural networks. *International Journal of Scientific and Engineering Research*, 2(5):1–6, May 2011.

[3] Ryan Nash Keiron O'Shea. An introduction to convolutional neural networks. *CoRR*, abs/1511.08458, 2015.

[4] Herbert F. Schantz. The history of ocr optical character recognition, 1982.

[5] Alexander Del Toro Barba(Machine Learning Specialist(Lead) at Google). How artificial intelligence is revolutionizing finance, January 2017. linkedin.com.

[6] Eugenio Culurciello. The fall of lstm/rnn, April 2018. Towards Data Science.

[7] Vladlen Koltun Shaojie Bai, J. Zico Kolter. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *CoRR*, abs/1803.01271, 2018.

[8] Geoffrey So. Should we abandon lstm for cnn, March 2019 2019. Published at medium.com.

[9] Yann LeCun Xiang Zhang, Junbo Jake Zhao. Character-level convolutional networks for text classification. *CoRR*, abs/1509.01626, 2015.

[10] Shunji Mori. Historical review of ocr research and development. July 1992. Invited Paper,PROCEEDINGS OF THE IEEE. VOL 80. NO 7.

[11] S S Kumar Deepa Berchmans. Optical character recognotion: An overview and an insight. 2014. International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT).

[12] Ishiqaki K Akiyama K Nakagawa M Tanaka H, Nakajima K. Hybrid pen-input character recognition system based on integration of online-offline recognition. pages 209–212. Proceedings of the fifth International Conference on Document Analysis and Recognition (ICDAR-1999) .

[13] Ruwei Dai Chunmei Liu, Chunheng Wang. Low resolution character recognition by image quality evaluation. pages 864–867. Proceedings of the Eighteenth International Conference on Pattern Recognition (ICPR 2006).

[14] Suresh RM Kannan R J, Prabhakar R. Off-line cursive handwritten tamil character recognition. pages 159–164, 2008. International Conference on Security technology.

[15] Prasad R S Prasad J R, Kulkarni U V. Offline handwritten character recognition of gujarati script using pattern matching. pages 611–614, 2009. 3 Rd International Conference on Anti-counterfeiting Security and Identification in communication.

[16] Luca Maria Gambardella Jurgen Schmidhuber Dan Claudiu Ciersan, Ueli Meier. Convolutional neural network committees for handwritten character classification. 2011. 2011 International Conference on Document Analysis and Recognition, 10.1109/ICDAR.2011.229.

[17] Yoshua Bengio Yann LeCun, Leon Bottou and Patrick Haner. Gradient-based learning applied to document recognition. pages 611–614, 1998. PROC. OF THE IEEE, NOVEMBER 1998.

[18] Tian Ruixia Yang Gang Wang Yutao, Qin Tingting. Recognition of license plate character based on wavelet transform and generalized regression neural network. pages 1881–1885, 2012. Control and Decision Conference (CCDC), 2012 24th Chinese.

[19] Alejandro Baldominos, Yago Saez, Pedro Isasi. A survey of handwritten character recognition with mnist and emnist. MDPI, Appl. Sci.2019, (15), 3169.

[20] Rachel Wiles. Have we solved the problem of handwriting recognition?, June 2019. towardsdatasicence.com.

[21] Swarnendu Ghosh Ritesh Sarkhel Bodhisatwa Mandal, Suvam Dubey and Nibaran Das. Handwritten indic character recognition using capsule networks. *CoRR*, abs/1901.00166, 2019.

[22] Hirunima Jayasekara Jathushan Rajasegaran Suranga Seneviratne Vinoj Jayasundara, Sandaru Jayasekara and Ranga Rodrigo. Handwritten character recognition with very small datasets. *CoRR*, abs/1904.08095, 2019.

[23] Raymond Kurzweil. Introduces the first omni-font optical character recognition system, 1974. HistoryofInformation.com.

[24] Nist handprinted forms and characters database. nist.gov.

[25] Sourabh Sinha. Python — gray scaling of images using opencv. GeeksforGeeks-Acomputer science portal for geeks.

[26] Types of morphological operations. mathworks.com.

[27] Image processing in idl. northstar-dartmouth.edu.

[28] Geoffrey E. Hinton Alex krizhevsky, Ilya Sutskever. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems 25 (NIPS)*, 2012.

[29] Andrew Zisserman Karen Simonyan. Very deep convolutional networks for large-scale image classification. 2015. ICLR conference.

[30] Tensorflow. https://www.tensorflow.org/.

[31] Why use keras. Keras documentation. keras.io.

[32] Miscellaneous operating system interfaces. Documentation the Python standard library.

[33] Chinmoy Lenka. File handling in python. GeeksforGeeks-A computer science portal for geeks.

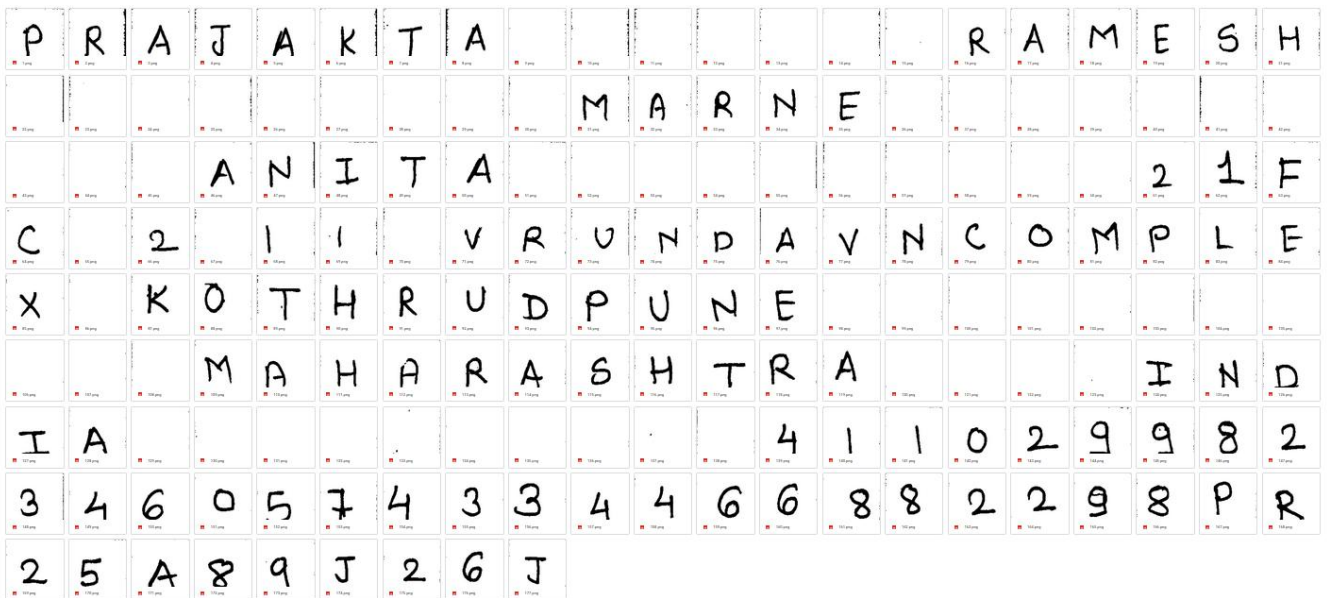[34] Tutorial, pyfpdf. https://pyfpdf.readthedocs.io/en/latest/Tutorial/index.ht
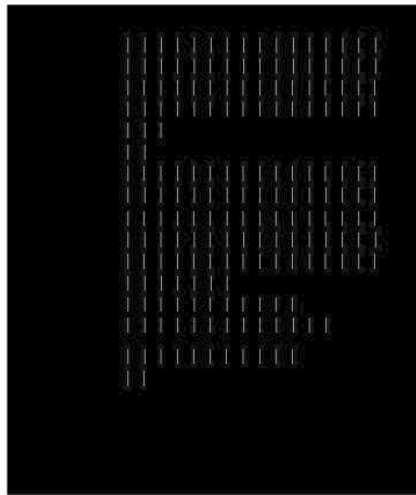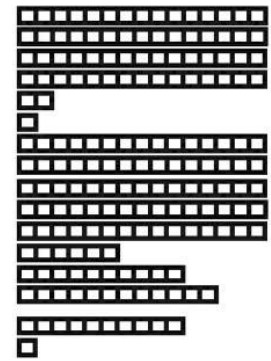
Fig. 9. Box Detection

Fig. 10. Result: Box Detection