iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: VI Month of publication: June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35616

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Various Models for the Conversion of Handwritten Text to Digital Text

Bhavyasri Maddineni[1], Sushma Sheshita Dulam[2], Champabai Chowhan[3], V. Sudha Rani[4]
[1, 2, 3]Student, [4]Associate Professor, ECE, Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad

Abstract: Handwritten Text Recognition (HTR) also known as Handwriting Recognition (HWR) is the detection and interpretation of handwritten text images by the computer. Handwritten text from various sources such as notebooks, documents, forms, photographs, and other devices can be given to the computer to predict and convert into the Computerized Text/Digital Text. Humans find easier to write on a piece of paper rather than typing, but now-a-days everything is being digitalized. So, HTR/HWR has an increasing use these days. There are various techniques used in recognizing the handwriting. Some of the traditional techniques are Character extraction, Character recognition, and Feature extraction, while the modern techniques are segmenting the lines for recognition, machine learning techniques, convolution neural networks, and recurrent neural networks. There are various applications for the HTR/HWR such as the Online recognition, Offline Recognition, Signature verification, Postal address interpretation, Bank-Cheque processing, Writer recognition and these are considered to be the active areas of research. An effective HTR/HWR is therefore needed for the above stated applications. During this project our objective is to find and develop various models of the purpose.
Keywords: Optical Character Recognition (OCR), Artificial Intelligence, Machine Learning, Deep Learning, Neural Networks (NN), Artificial Neural Networks (ANN), Convolution Neural Networks (CNN), Recurrent Neural Networks (RNN).

I. INTRODUCTION

Handwritten Text Recognition (HTR) also known as Handwriting Recognition (HWR) is the detection and interpretation of handwritten text images by the computer. Handwritten text from various sources such as notebooks, documents, forms, photographs, and other devices can be given to the computer to predict and convert into the Computerized Text/Digital Text.

The recognition can be of two types:

1) *Offline Recognition:* In this technique, there is an automatic conversion of text from the images into the digital text. The data in this technique is said to be the static representation of handwriting. This is of two types: Traditional Techniques, Modern Techniques.
2) *Online Recognition:* In this technique, there is an automatic conversion of text as it is written. Here, sensors are used to identify the pen-tip movements. The data in this technique is said to be the digital representation of handwriting. In this technique, the signals are converted into the corresponding digital characters. The online handwriting recognition interface includes elements such as a pen for the writing purpose, touch-surface integrated with the output display, application to determine the pen-tip movements and to convert them to the digital text. The steps involved in this technique are: Preprocessing, Feature extraction, Classification.

II. LITERATURE REVIEW

A. Handwritten Text Recognition System Based on Neural Network

In this paper, an efficient approach towards the development of handwritten text recognition system is being proposed. This system contains a 3-layer Artificial Neural Network (ANN) using supervised learning. The choice of optimal feature vectors greatly varies the accuracy of any system. In this system bit map representation of input samples is used as the feature vector. The feature vector is pre-processed and then along with the target vector it is applied to the ANN model.

Two algorithms are used in this system: Resilient Back Propagation, Scaled Conjugate Gradient.

This system is being tested on different handwritten images with different styles. The proposed system first performs pre-processing on the scanned imaged to remove the noise; bit map image representation is used for the feature extraction.

The accuracy obtained in this system is about 95%. The system proposed gave good results for images with different styles, sizes, alignment, and background. The system classified most of the characters correctly even if the image has noise such as containing text in the background.

B. Improving CNN – RNN Hybrid Networks for Handwriting Recognition

In this, paper two factors, namely, architectures and the availability of large-scale annotated data is explored systematically for improving handwritten recognition for scanned images of text.

The architecture proposed is the modified CNN – RNN hybrid architecture.

The work focuses on:

1) Effective initialization of network using synthetic data for pre-training
2) Image normalization for slant correction
3) Domain specific data transformation and distortion

C. Handwritten Text Recognition using Deep Learning

This paper classifies the handwritten words present in the document so that the handwritten text can be converted to digital format.

Two main approaches are implemented to achieve this task: classifying words and character segmentation.

For the first approach, Convolution Neural Network (CNN) model is used with different architectures to train the model that can accurately classify words.

For the second approach, Long Term Memory networks (LSTM) with convolution for constructing bounding boxes for each character. Later, the segmented characters are passed to a CNN model for classification. Then each word is reconstructed according to the classification and segmentation results.

D. Handwritten Chinese Text Recognition using Separable Multi – Dimensional Recurrent Neural Network

In this paper, the recognition model is designed using Separable MDLSTM (SMDLSTM). When compared to the traditional MDLSTM, this method extracts contextual information, consume less computational efforts, resources.

III. EXISTING MODEL

A. Optical Character Recognition (OCR)

OCR recognizes character and they are differentiated based on the shape. It is used as an input device for pre - printed documents to read alpha-numeric characters. For the recognition to happen OCR first creates a bitmap for the images and the corresponding ASCII values, when the recognition is performed the character which is matched in the bitmap is read, otherwise rejected. It produces a high degree of accuracy, and also has a fast processing. But, it performs the task well for the printed text rather than the handwritten text.

IV. PROPOSED MODEL

OCR technology achieves the accuracy greater than 99% for the typed characters in the input images with high quality. But it provides less accuracy when there are different types of handwritings, differences in spaces and irregularities of handwriting. Thus, the accuracy that is provided by the OCR systems for handwritten characters is less when compared to the accuracy that is provided by the OCR systems for typed characters. OCR system does not contain many tools to handle handwriting recognition. As the handwritings differ from person to person, traditional OCR system cannot recognize everyone's handwriting. To recognize all the different handwritings successfully complex deep learning algorithms are to be used.

The recognition of handwritten text highly depends on Neural Networks. By using these algorithms the performance of handwriting recognition tools can be increased.

V. DATASETS

A. Scikit Learn's Digit Dataset

The dataset contains 8X8 image of a digit.

Classes	10
Samples per class	~180
Total samples	1797
Dimensionality	64
Features	Integers 0-16

Table 1. Scikit's Learn Digit Dataset

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

B. Keras's MNIST Digit Dataset

The dataset contains 28X28 image of a digit.

Classes	10
x_train	(60000, 28, 28)
y_train	(60000,)
x_test	(10000, 28, 28)
y_test	(10000,)

Table 2. Keras's MNIST Dataset

C. Extra Keras's EMNIST Balanced Type Dataset

The dataset contains digits-, upper- and lower- case handwritten letters.

Train	112800
Test	18800
Total	131600
Classes	47

Table 3. Extra Keras's EMNIST Dataset

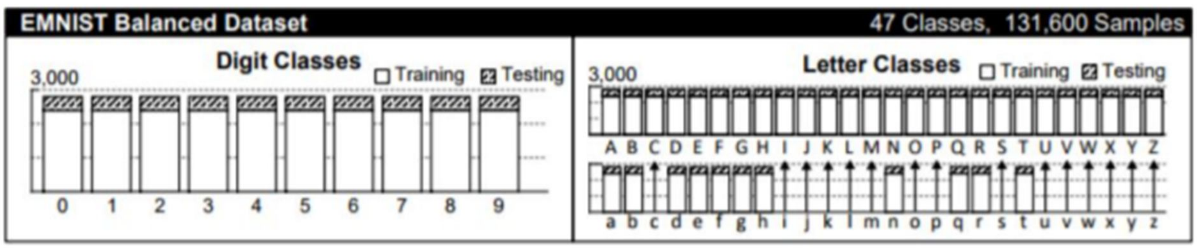


Fig 1. EMNIST Balanced Dataset

VI. DEVELOPED MODELS

Dataset	Model	Description	Accuracy
Sklearn's Digit	Multinomial Logistic Regression	Default parameters	97.55 %
	Support Vector Machine	C=1.1	99.16 %
	Decision Tree	Default parameters	100 %
	Random Forest	n_estimators=35 criterion=entropy	98.6 %
Keras's MNIST	Neural Networks	Layer 1: type=dense, activation=relu Layer 2: type=Dense, activation=sigmoid Optimizer=adam Loss=sparse_categorical_crossentropy Epochs=7	98.08%
	Convolution Neural Networks	Layer 1: type=Conv2D, activation=relu Layer 2: type=MaxPooling2D Layer 3: type=Dropout Layer 4: type=Conv2D, activation=relu Layer 5: type=MaxPooling2D Layer 6: type=Dropout Layer 7: type=Flatten	99.27 %

		Layer 8: type=Dense, activation=sigmoid Layer 9: type=Dense, activation=softmax Optimizer=adam Loss=sparse_categorical_crossentropy Epochs=10	
	Recurrent Neural Networks	Layer 1: type=LSTM, activation=relu Layer 2: type=Dropout Layer 3: type=LSTM, activation=sigmoid Layer 4: type=Dense, activation=relu Layer 5: type=Dropout Layer 6: type=Dense, activation=softmax Optimizer=Adam, Learning rate=1e-3, Decay=1e-5 Loss=sparse_categorical_crossentropy Epochs=5	98.18%
Extra Keras's EMNIST	Neural Networks	Layer 1: type=Dense, activation=relu Layer 2: type=Dense, activation=sigmoid Layer 3: type=Dense, activation=softmax Optimizer=Adam Loss=sparse_categorical_crossentropy Epochs=20	83.04 %
	Convolution Neural Networks	Layer 1: type=Conv2D, activation=relu Layer 2: type=Conv2D, activation=relu Layer 3: type=Conv2D, activation=relu Layer 4: type=Flatten Layer 5: type=Dense, activation=relu Layer 6: type=Dense, activation=relu Layer 7: type=Dense, activation=softmax Optimizer=Adam Loss=sparse_categorical_crossentropy Epochs=7	78.26 %
	Recurrent Neural Networks	Layer 1: type=LSTM, activation=relu Layer 2: type=Dropout Layer 3: type=LSTM, activation=relu Layer 4: type=Dropout Layer 5: type=Dense, activation=relu Layer 6: type=Dropout Layer 7: type=Dense, activation=softmax Optimizer=Adam, Learning rate=1e-3, Decay=1e-5 Loss=sparse_categorical_crossentropy Epochs=3	98.22 %

Table 4. Accuracy for various models and datasets

VII. CONCLUSION

A. Multiclass / Multinomial Logistic Regression

1) This model helps to understand the relationships among the variables present in the dataset.
2) When parameters are estimated, simultaneous models give small standard errors than when compared to logistic regression.
3) For other categories, choice of reference class does not show any effect on the parameter estimation.
4) This algorithm cannot solve nonlinear problems as it has linear decision surface.

B. Support Vector Machine
1) It is very effective for high dimensional data.
2) It can perform well when the number of rows is more than the number of features of data.
3) When classes in data and points are separated it works well. Both Regression and classification use this algorithm.
4) It works well with the image dataset.
5) If classes in the dataset are overlapped it does not perform well.
6) It is very difficult to choose optimal kernel.
7) Training of large dataset takes more time.
8) It is more complex and difficult to understand. We cannot explain the classification as it is not a probabilistic model.

C. Decision Tree
1) This classifier processes same like a human brain while making a decision.
2) This algorithm is simple and easy to understand.
3) It is used to solve problems related to decision-making.
4) It is complex as it contains many layers.
5) Over fitting is major issue in this algorithm.
6) As the number of classes increases the complexity of computation also increase.

D. Random Forest
1) This algorithm reduces the over fitting problem in decision trees and it also reduces the variance and improves accuracy.
2) It can solve the problems of both classification and regression.
3) This works well with categorical and continuous variables.
4) Missing values can be automatically handled by this algorithm.
5) Rule based approach is used in this algorithm instead of distance calculation.
6) So, feature scaling is not required. Complexity is more in this algorithm.
7) Training period is more for this algorithm.

E. Neural Networks
1) Once the model is trained, it can work on incomplete information.
2) Provides fault tolerance.
3) It does parallel processing.
4) It can also be used for non-linear and complex relationships.
5) Good knowledge of applied maths, algorithms, probability and statistics, distributed computing is required.

F. Convolution Neural Networks
1) CNN does feature extraction for every location in an image, which is a saving in parameters.
2) It can be used for 1 dimensional problems like time series, 3 dimensional problems like image classification.
3) This model is less flexible.
4) This model is not used for total unstructured data.

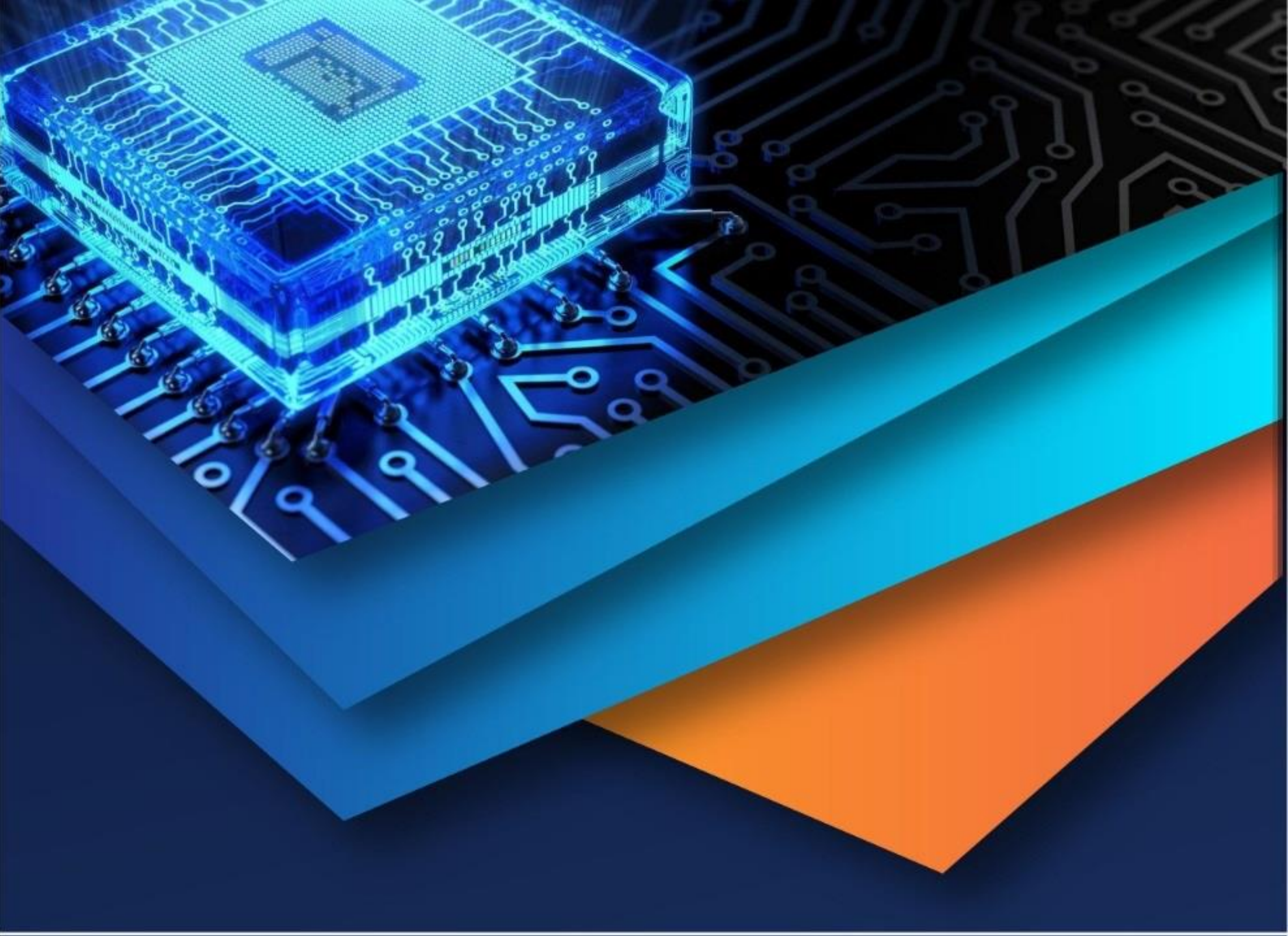
G. Recurrent Neural Networks
1) This model is used with convolution layers to extend the powerful pixel neighbourhood.
2) It is useful in time series prediction because of the feature to remember previous input gradient exploding and vanishing problems.
3) It is very difficult to do training in RNN.
4) It cannot compute very lengthy sequences if the usage is Relu as activation feature.

VIII. FUTURE SCOPE

The models can be extended to recognize and convert handwritten words and sentences to corresponding digital format using RCNN.

REFERENCES

[1] Handwritten Text Recognition System based on Neural Networks by Ahmed Mahi Obaid, Hazem M. El Bakry, M. A. Eldosuky, A. I. Shehab
[2] Handwritten Text Recognition using Deep Learning by Batuhan Balci, Dan Saadati, Dan Shiferaw
[3] Handwritten Chinese Text Recognition using Separable Multi – Dimensional Recurrent Neural NETWORKS by Yi-Chao Wu, Fei Yin, Zhuo Chen, Cheng-Lin Liu
[4] Improving CNN – RNN Hybrid Networks for Handwriting Recognition by Karthik Dutta, Praveen Krishnan, Minesh Mathew and C. V. Jawahar
[5] keras.io
[6] kaggle.com



10.22214/IJRASET

INDEX COPERNICUS
45.98

ISRA JIF
IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY