| | | | | |
|---|---|---|---|---|
| | **Bangladesh Army University of Science and Technology (BAUST)** | | | |
| | ***Department of Computer Science and Engineering*** | | | |
| | Assignment #1, Winter 2023 | | Level-4 | Term-I |
| | Course Code: CSE 4131 | Course Title: Artificial Neural Networks and Fuzzy Systems | | |
| | Submission Date: 11/05/23 | CO Number: CO2 | Full Marks: 15 | |
| ID: 200101056 | | Name: Md Ali Arman | | |

# Handwritten Digit Recognition System using Machine Learning in Python

## 1. Introduction

Applications of handwritten digit recognition range from scanning handwritten documents and reading postal codes to reading digits on bank checks and serving as a fundamental problem in computer vision. Convolutional Neural Networks (CNNs) have become the preferred method for recognizing handwritten digits due to their impressive performance in this area. Convolutional neural networks (CNNs) are neural networks that were developed for the purpose of performing image recognition. They consist of many interconnected "layers" that carry out tasks like convolution, pooling, and non-linear activations. Because CNNs can automatically learn hierarchical representations of the input image data, they are ideally suited for identifying recurring design elements and other visual characteristics. A convolutional neural network (CNN) is trained with images of handwritten digits in order to recognize them. The network is trained to recognize individual numbers and classify them appropriately. Once the CNN has been trained, it can be used to identify previously unseen handwritten digits. The MNIST dataset, which includes 70,000 grayscale images of handwritten digits from 0 to 9, is the most popular dataset used for training and evaluating CNNs for handwritten digit recognition. The data set contains 60,000 images that will be used for training and 10,000 images that will be used for testing. When it comes to the challenging task of recognizing handwritten digits, CNNs have proven to be a game-changer in the field of computer vision. They are still being studied because of the potential impact they could have in many different fields.

## 2. Literature Review

### 2.1 Improved Handwritten Digit Recognition Using Convolutional Neural Networks (CNN)

The ability to read handwriting has become increasingly useful in the modern era of digitization. License plate readers, postal letter sorters, Cheque truncation system (CTS) scanners, historical document preservationists, library and bank archivists, etc. are just some of the main users of this technology. Because of their superiority over shallow neural architectures in automatic feature detection, which requires human intervention in most cases, deep neural architectures are increasingly attractive. Image classification, object recognition, recommendation systems, signal processing, natural language processing, computer vision, and face recognition are just some of the many domains in which convolutional neural networks (CNNs) find use. A CNN requires little in the way of pre-processing or feature extraction because it incorporates these processes into its classification step.

**Table 1.** Shallow neural network vs deep neural network.

| Factors | Shallow Neural Network (SNN) | Deep Neural Network (DNN) |
|---|---|---|
| **Number of hidden layers** | - single hidden layer (need to be fully connected). | - multiple hidden layers (not necessarily fully connected). |
| **Feature Engineering** | - requires a separate feature extraction process.<br>- some of the famous features used in the literature include local binary patterns (LBPs), histogram of oriented gradients (HOGs), speeded up robust features (SURFs), and scale-invariant feature transform (SIFT). | - supersedes the handcrafted features and works directly on the whole image.<br>- useful in computing complex pattern recognition problems.<br>- can capture complexities inherent in the data. |
| **Requirements** | - emphasizes the quality of features and their extraction process.<br>- networks are more dependent on the expert skills of researchers. | - able to automatically detect the important features of an object (here an object can be an image, a handwritten character, a face, etc.) without any human supervision or intervention. |
| **Dependency on data volume** | - requires small amount of data. | - requires large amount of data. |

**Figure 01: Table of Shallow neural network vs Deep Neural Network**

It automatically extracts rich, related features from images and provides high recognition accuracy, even with limited training data. Accuracy as high as 98% or 99% has been achieved using a CNN model for handwritten digit recognition from the MNIST benchmark database. An ensemble model was developed by combining several different CNN models, and its reported accuracy was 99.73 percent. When the SVM (support vector machine) capability of minimizing the structural risk is combined with the CNN model capability of extracting the deep features, Niu and Suen report an extraordinary recognition accuracy of 99.81%. Building diverse ensembles of deep neural networks (DNN), Alvear-Sandoval et al. recently achieved an error rate of 0.19% on MNIST.
Extensive testing on the MNIST dataset is used to determine the importance of the hyper-parameters proposed here. They also confirmed that hyper-parameter fine-tuning is critical for optimizing CNN performance. Using the Adam optimizer, they improved upon all previous reported results for the MNIST database, achieving a recognition rate of 99.89%.
This is the first study to exhaustively explore all the parameters of CNN architecture that maximize recognition accuracy on the MNIST dataset, which is the main novelty of this work. The accuracy of a pure CNN model was not able to be matched by other researchers. Similar recognition accuracy as achieved in this work was achieved by other researchers using ensemble CNN network architectures for the same dataset, at the expense of increased computational cost and high testing complexity.

### 2.2 Recognition of Handwritten Numerals of various Indian Regional Languages using Deep Learning

The use of computer vision and pattern recognition to decipher handwritten digits is an exciting area of study. It's a crucial part of the postal especially useful automation services in India, where so many different languages and scripts are in use. Because of this, the recognition system faces a number of obstacles, such as the fact that handwriting comes in both cursive and print forms.

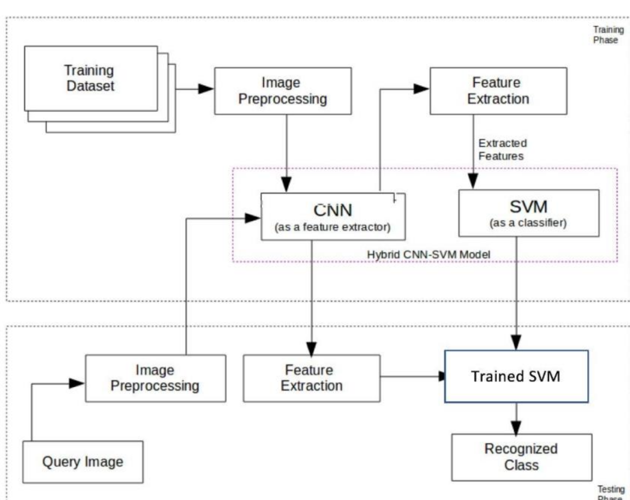


**Figure 02: Proposed work**

Using Convolutional Neural Network (CNN) and Support Vector Machine (SVM), this paper suggests a method for accurate handwritten digit recognition. The goal of the proposed work is to create a system that can read digits written in regional languages like Bangla, Devanagari, Oriya, and Telugu. Single-digit-resolution input images are first normalized by the proposed system, after which CNN serves as a feature extractor and SVM acts as a classifier. Jadavpur University's CMaterdb (which contains Telugu numerals) and ISI Kolkata's benchmark database (which contains Bangla, Devanagari, and Oriya numerals) have been used in experiments.

| Language | CNN | CNN-SVM [Linear] | CNN-SVM [RBF] | State-of-art results |
|---|---|---|---|---|
| Bangla | 95.84% | 99.14% | 99.14% | 98.78% [29] |
| Devanagari | 95.30% | 98.82% | 99.41% | 96% [15] |
| Oriya | 89.92% | 91.68% | 94.54% | 97.64% [30] |
| Telugu | 94.37% | 98.66% | 99.16% | 94% [23] |

**Figure 03: Result Table**

The results show that their model achieves an accuracy of 99.41% when translating from the Devanagari script, 99.14% when translating from Bangla, 99.16% when translating from Telugu, and 94.54% when translating from Oriya.

For deep learning models to be trained and tested using the proposed method, a substantial number of computational resources are needed. Poorly written, noisy, or otherwise illegible handwritten numbers may cause the model to underperform. Handwritten digits that vary greatly from those in the training dataset may also prove difficult for the model to decipher.

### 2.3 Exploring Deep Learning Techniques for Kannada Handwritten Character Recognition: A Boon for Digitization

All official government and medical records in Karnataka State were written by hand in the Kannada language. Reproducing the contents of these old documents by typing them up is a time-consuming chore because of how obscure the writing is. So, a computer-based system is required to bridge the gap between machines and humans. This paper explores the use of deep learning techniques for feature extraction and proposes a method for improving the quality of images used in a Kannada handwritten character recognition system.

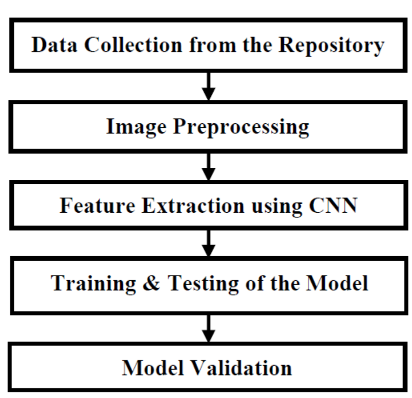


**Figure 04: Block diagram of the proposal Model**

The proposed method's layout is straightforward and simple to facilitate user comprehension. The work was tested using the publicly available Chars74K dataset. Twenty-five different handwritten Kannada vowels and consonants were tested across 657 different classrooms. Model accuracy was evaluated over 15 iterations using the Categorical Cross-entropy loss function.

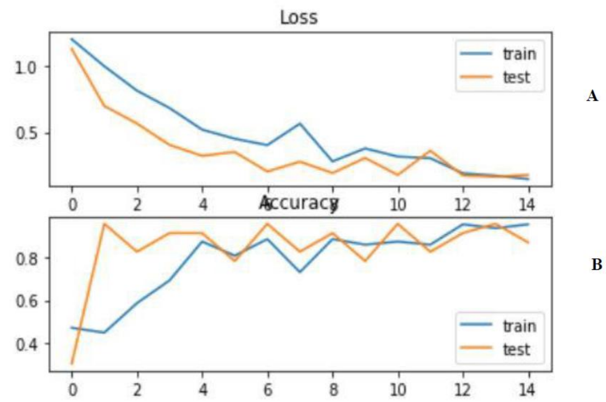


**Figure 05: Line Plots of (A) Categorical Cross-Entropy Loss and (B) Accuracy over Training Epochs**

The model's predictive performance, on the training set, was 95.11 percent, and on the testing set, it was 86.1 percent. The proposed model would be very helpful for record-keeping in public administration.
Different people's handwriting styles can cause letters to overlap, which can result in lower accuracy values for handwritten data captured in real time. Kannada handwritten reports were the only form of documentation available in government offices and healthcare departments in Karnataka; typing up the contents of these old documents would be a laborious task that could lead to mistakes.

### 2.4 Handwritten Digit String Recognition using Convolutional Neural Network

Recognizing strings is a crucial part of many computer vision applications. Combinations of CNN and RNN have seen widespread use recently for tackling the problem of string recognition. On the other hand, training RNNs is not only challenging but also time-consuming.

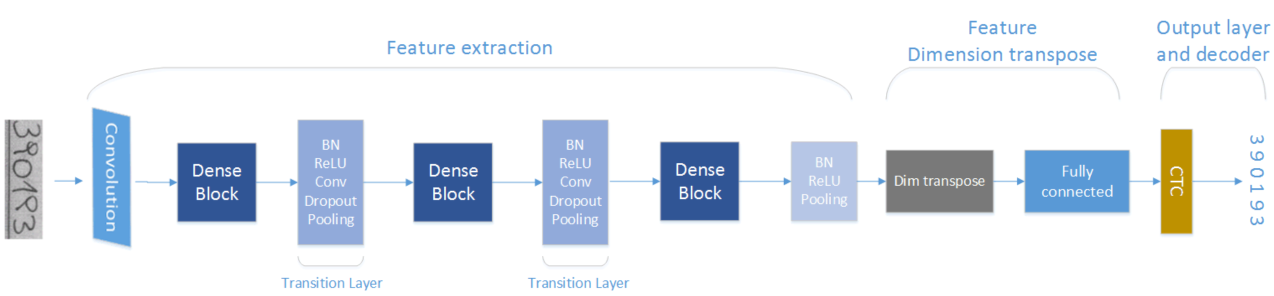


**Figure 06: The proposed architecture with three dense blocks.**

They proposed CNN-only architecture to the problem of handwritten digit string recognition (HDSR) in this paper. Each layer of this network serves a specific purpose: first, feature extraction; second, feature dimension transposition; and third, output. Inspired by DenseNet's astounding results, we make use of dense blocks to carry out feature extraction. To decode the feature sequence and compute the loss, a CTC (connectionist temporal classification) output layer is placed at the very top of the network. In between the feature extraction layer and the output layer are several feature dimension transposition layers.

| Methods | CAR-A | CAR-B | CVLHDS |
|---|---|---|---|
| Tebessa I [1] | 0.3705 | 0.2662 | 0.5930 - |
| Tebessa II [1] | 0.3972 | 0.2772 | 0.6123 |
| Singapore [1] | 0.5230 | 0.5930 | 0.5040 |
| Pernambuco [1] | 0.7830 | 0.7543 | 0.5860 |
| BeiJing [1] | 0.8073 | 0.7013 | **0.8529** |
| FSPP [14] | 0.8261 | 0.8332 | 0.7923 |
| CRNN [12] | 0.8801 | 0.8979 | 0.2601 |
| Saabni [3] | 0.8580 | | - |
| our previous work [13] | 0.8975 | 0.9114 | 0.2704 |
| Proposed | **0.9220** | **0.9402** | 0.4269 |

**Figure 07: Recognition rates of different models on the datasets CVL and CAR.**

Tests on the ORAND-CAR-A and ORAND-CAR-B datasets show that the proposed method significantly outperforms the state-of-the-art, achieving recognition rates of 92.2% and 94.02%, respectively.

There are some restrictions to the proposed method, despite the fact that it performed very well on the test set. There is the possibility of poor performance on images that divert significantly from the training data, such as those with radically different handwriting styles or backgrounds. Another restriction is that input images with noise or other artifacts may cause the model's performance to suffer.

## 3. Result and Analysis

"Improved Handwritten Digit Recognition Using Convolutional Neural Networks (CNN)":

When the proposed method's SVM (support vector machine) capability of minimizing structural risk is combined with the CNN model's capability of extracting deep features, Niu and Suen report an extraordinary recognition accuracy of 99.81%.

"Recognition of Handwritten Numerals of various Indian Regional Languages using Deep Learning":

The proposed method achieved an accuracy of 99.41% when translating from the Devanagari script, 99.14% when translating from Bangla, 99.16% when translating from Telugu, and 94.54% when translating from Oriya.

"Exploring Deep Learning Techniques for Kannada Handwritten Character Recognition: A Boon for Digitization":

The proposed method achieved an accuracy on the training set of 95.11 percent and on the testing set of 86.1 percent.

"Handwritten Digit String Recognition using Convolutional Neural Network":

The proposed method achieved recognition rates of 92.2% and 94.02%, respectively.

According to the results presented above, all four thesis papers utilize deep learning techniques to successfully recognize handwritten digits or characters. There have been numerous papers written on the topic of handwritten digit recognition, but they have all taken a slightly different tack, such as identifying digits written in a variety of Indian regional languages, delving into deep learning techniques for Kannada handwritten character recognition, etc.

These papers show that convolutional neural networks (CNNs) and other deep learning methods can accurately recognize handwritten digits and characters.

## 4. Conclusion

Convolutional Neural Networks (CNNs) and other deep learning techniques have been shown to be effective in recognizing handwritten digits and characters in the papers.

The papers show how CNNs can adapt to different scripts and languages without losing accuracy when recognizing handwritten numbers and characters. Pre-processing techniques such as normalization and resizing, as well as optimization methods such as Adam optimizer and cross-entropy loss function, are emphasized in these papers as crucial to achieving high accuracy in handwritten digit and character recognition.

The analysis of these papers as a whole demonstrates the utility of deep learning techniques, and in particular CNNs, for the recognition of handwritten digits and characters across a variety of languages and scripts. The papers show that high accuracy in recognizing handwritten digits and characters is possible with careful optimization and appropriate architecture design. Handwritten digit and character recognition using deep learning techniques is an exciting and rapidly growing field that promises to enable numerous digital, automated, and machine learning applications.

## References

[1] Ahlawat, Savita, et al. "Improved handwritten digit recognition using convolutional neural networks (CNN)." Sensors 20.12 (2020): 3344.
[2] Chaurasia, Saumya, and Suneeta Agarwal. "Recognition of handwritten numerals of various indian regional languages using deep learning." 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON). IEEE, 2018.
[3] Rao, Abhishek S., et al. "Exploring deep learning techniques for Kannada handwritten character recognition: a boon for digitization." International Journal of Advanced Science and Technology 29.5 (2020): 11078-11093.
[4] Zhan, Hongjian, Shujing Lyu, and Yue Lu. "Handwritten digit string recognition using convolutional neural network." 2018 24th International Conference on Pattern Recognition (ICPR). IEEE, 2018.