<u>MNIST</u> Handwritten digit classification using neural network

Innovative Term Paper

The documentation for the fulfillment of Innovative term paper

By

**Soni Sumit Raghukant
(22MCA063)**

**3MCA202 Software Engineering
Department of Computer Science & Engineering,
Institute of Technology,
Nirma University,
Ahmedabad 382 481
February 2023**

Abstract:

Deep learning is the most powerful machine learning technology in the current generation. This article presents the results of handwriting analysis using deep learning. Every hand is unique. So hands vary from person to person. can be written in two ways. One is online written confirmation and the other is offline written confirmation. Online handwriting information system, get ideas while operating , offline handwriting knowledge, working on scanned images. Offline writing is the hardest writing to find. The MNIST dataset is used for code numbers, which includes code numbers 70000. Many machine learning algorithms have been developed that can be used for classification. This article examines the accuracy and performance evaluation of the algorithm convolutional neural network (CNN). The proposed method identified with an overall accuracy of 93%.

To develop a holistic device for handwritten English individual reputation for manually crammed forms by systematically synthesizing a sturdy handwritten textual person dataset for proper representation of handwriting.

As a part of this observation, 572 copies of a form had been filled through over 200 exceptional individuals to introduce demographic variants. The bureaucracy had been then scanned and every handwritten man or woman inside the forms was labeled and extracted the usage of standard photo processing strategies. The dataset of eighty four, 712 person images created by using this method (HW-dataset) made out of each alphabetical and numerical characters. Three hybrid datasets (h-EH) had been then fashioned by means of MNIST stands for Mixed National Institute of Standards and Technology, which created the handwritten digit dataset. It is one of the most The MNIST database (Modified National Institute of Standards and Technology database)

combining EMNIST datasets and the HW-dataset based on Digits (h-EHd – 329, 668 man or woman pics), Alphabets (h-EHa – 163, half individual snapshots) and a combination of Digits and Alphabets (h-EHm – 189, 586 character pics). An anchor based photograph extraction approach was used along with a Multi-Channel CNN (MCCANN) version which became educated on three versions of h-EH, to automate the technique of digitization of handwritten forms.

The classification accuracies of the MCCNN for h-EHa, h-EHd and h-EHm are ninety three%, ninety six% and ninety three% respectively for check records. models educated on only the MNIST dataset perform poorly on taking a look at facts. An anchor based object detection technique used at the side of MCCANN educated on h-EH produces fantastic outcomes in digitizing hand crammed paperwork. touch loose solutions will advantage incidence due to the emergence of danger of fomites in the international.

In this sort of space, manual coping with bureaucracy for the cause of statistics access, digitization and records handling can be considered as capacity fitness and protection dangers. The solution supplied inside the present day work uses an aggregate of models that's skilled on a hybrid handwritten statistics set with excessive demographic variability. The version evolved as part of this take a look at is well acceptable for enabling contact loose dealing with files

Introduction:

OCR (Optical Character Recognition) is the use of technology to distinguish printed or handwritten text characters within digital images of physical documents, such as a scanned paper document. The basic process of OCR involves

examining the text of a document and converting the characters into a code that can be used for data processing. OCR is sometimes also called text recognition.

MNIST is a large database of handwritten digits commonly used for training various image processing systems. The database is also widely used for machine learning training and testing. It was created by "remixing" samples from the original NIST data sets.

The developers felt that since the NIST training dataset was taken from American Census Bureau employees, while the test dataset was taken from American high school students, it was not very suitable for machine learning experiments. In addition, the black and white images from NIST were normalized to fit within a 28x28 pixel bounding box and smoothed to introduce gray levels.

The MNIST database contains 60,000 training images and 10,000 test images. Half of the training set and half of the test set were taken from the NIST training dataset, while the other half of the training set and the other half of the test set were taken from the NIST test dataset. The original creators of the database keep a list of some of the methods tested on it. In their original paper, they use a support vector machine to obtain an error rate of 0.8%.

Literature Review:

HANDWRITTEN RECOGNITION BY USING MACHINE LEARNING APPROACH by P.Thangamariappan and Dr.J.C.Miraclin Joyce Pamila where they discussed Handwritten Recognition, MNIST Dataset, Deep Learning, Convolutional Neural Networks (CNN) and The result of their work shows that the maximum accuracy 98.5% was obtained in MNIST dataset using the multilayer perceptron neural network technique.

In Handwritten Data Digitization Using an Anchor based Multi-Channel CNN (MCCNN) Trained on a Hybrid Dataset (h-EH) by Abhinandan Chiney and team, they enlighted the terms Hand Written Data Digitization; Anchor based Multi-Channel CNN, OCR, Image Processing, NLP, Language Science and proposed a holistic Anchor based text isolation method followed by Multi-Channel CNN model is developed to digitize hand-filled forms. A representative hybrid dataset of character images is created for the purpose of training the models. This dataset is found to be more representative of varied demographics as compared to a standard dataset like EMNIST.

In Assessing Four Neural Networks on Handwritten Digit Recognition Dataset (MNIST) by Feiyang Chen, Nan Chen, Hanyang Mao, and Hanlin Hu, they made use of Neural Network, CNN, CapsNet, DenseNet, ResNet, MNIST and proposed which models perform better across divided MNIST datasets. We compared four models on the MNIST dataset with different divisions and showed that CapsNet performs best across datasets. Additionally, we also observe surprisingly that CapsNet requires only a small amount of data to achieve excellent performance. Finally, we will apply CapsNet's ability to generalize in other image recognition fields in the future.

In AN ENSEMBLE OF SIMPLE CONVOLUTIONAL NEURAL NETWORK MODELS FOR MNIST DIGIT RECOGNITION by Sanghyeon An Minjun Lee Sanglee Park Heerin Yang Jungmin So, they mentioned The MNIST handwritten digit data set is often used as an entry-level data set for training and testing neural networks. While achieving 99% accuracy on the test set is rather easy, correctly classifying the last 1% of the images is challenging. People have tried many different network models and techniques to increase test accuracy, and the best accuracy reported reaches approximately 99.8%.

A Survey of Handwritten Character Recognition with MNIST and EMNIST by Alejandro Baldominos and team, they had mentioned that there has been no significant changes in the artificial intelligence techniques used for computer vision. Indeed, most of the merit of the good performance can be attributed instead to hardware improvements, which has eventually led to more complex architectures and the ability to deal with higher amounts of data.

Tools & Techniques used:

Machine learning model: A machine learning model is defined as a mathematical representation of the output of the training process. Machine learning is the study of different algorithms that can improve automatically through experience & old data and build the model. A machine learning model is similar to computer software designed to recognize patterns or behaviors based on previous experience or data. The learning algorithm discovers patterns within the training data, and it outputs an ML model which captures these patterns and makes predictions on new data.

The implementation of the proposed model is done using python.

The libraries used are as follows:

calendar for EPOCH, dataclasses for data class, numpy, matplotlib.pyplot, seaborn, cv2, PIL for Image, tensorflow, tensorflow for keras, keras.datasets for mnist, tensorflow.math for confusion_matrix.
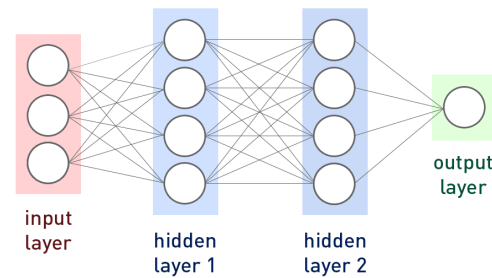
TensorFlow platform and it's uses:

TensorFlow is an end-to-end open source platform for machine learning. TensorFlow is a rich system for managing all aspects of a machine learning system; however, this class focuses on using a particular TensorFlow API to develop and train machine learning models.

Keras library:

Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Up until version 2.3, Keras supported multiple backends, including TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML. Keras is a high-level API to build and train neural network models. Keras does not depend on TensorFlow, and vice versa . Keras can use TensorFlow as its backend.

PIL library:

Python Imaging Library is a free and open-source additional library for the Python programming language that adds support for opening, manipulating, and saving many different image file formats. It is available for Windows, Mac OS X and Linux.

Seaborn library:

Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

OpenCV/ CV2 library:

OpenCV is a great tool for image processing and performing computer vision tasks. It is an open-source library that can be used to perform tasks like face detection, objection tracking, landmark detection, and much more. It supports multiple languages including python, java C++.

Activation Function in model training:

An Activation Function decides whether a neuron should be activated or not. This means that it will decide whether the neuron's input to the network is important or not in the process of prediction using simpler mathematical operations.

RELU activation function:

A rectified linear activation function, or ReLU for short, is a piecewise linear function that will output directly if it is positive, otherwise it will output zero. It has become the default activation function for many types of neural networks because a model that uses it is easier to train and often performs better.

Sigmoid activation function:

Loading MNIST data from keras .dataset

(X_train, Y_train), (X_test, Y_test) = mnist.load_data()

Inorder to reduce load over machine in order to process the images which are having color in range 0-255, we reduce it into range of 0-1

X_train = X_train/255
X_test = X_test/255

Building the Neural Network
Flatten - we can't feed data as a matrix so we convert it into a single line using it.
The input shape is the size of the image.

Dense - all the layers comes into previous or next layer, (no of neurons, activation function = relu)
relu = rectified linear unit
sigmoid
2 layers are used The final layer is output layer so that the neurons of the previous layers are connected to this layer. The final layer should have total 10 neurons as we have a total of (0-9) 10 values.

```
# initializing model variable
model = tf.keras.Sequential(
[keras.layers.Flatten(input_shape=(28,28,1)),
keras.layers.Dense(50, activation = 'relu'),
keras.layers.Dense(50, activation = 'relu'),
keras.layers.Dense(10, activation = 'sigmoid')])
```

optimizer is used to determine the most optimium model parameter
ex. linear regression w b
loss is used for label removing
Compliing the model

metrics =
accuracy = no of correct prediction/ total number of data passed

```
model.compile(
    optimizer='adam',
    loss = 'sparse_categorical_crossentropy',
    metrics = ['accuracy']
)
```

```
# Train neural network
model.fit(X_train, Y_train, epochs=10)
```

Character Recognition System:

Character recognition is the process that allows computers to recognize written or printed characters, such as numbers or letters, and convert them into a form that the computer can use.

MNIST Dataset:

The set of images in the MNIST database was created in 1998 as a combination of two of NIST's databases: Special Database 1 and Special Database 3. Special Database 1 and Special Database 3 consist of digits written by high school students and employees of the United States Census Bureau, respectively.

The MNIST database contains 60,000 training images and 10,000 testing images. Half of the training set and half of the test set were taken from NIST's training dataset, while the other half of the training set and the other half of the test set were taken from NIST's testing dataset. The original creators of the database keep a list of some of the methods tested on it.

Other relevant datasets:

- 3D MNIST
  Voxels instead of pixels
- Multiple MNIST
  Multiple MNIST digits with 80% overlap. No official implementation available but this might do the job.
- multiMNIST
  MNIST with multiple digits next to each other(not just one character)
- MultiDigitMNIST
  Multiple digits in 2D plane (not just directly next to each other)

Deployment/ Output Screenshot:
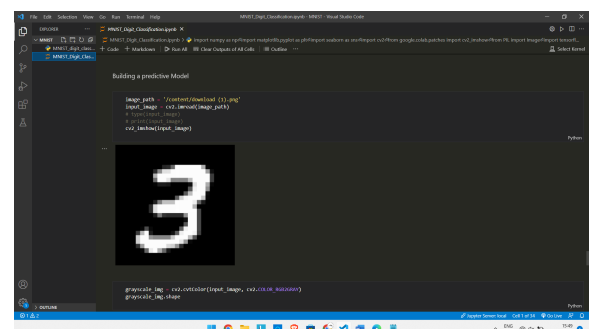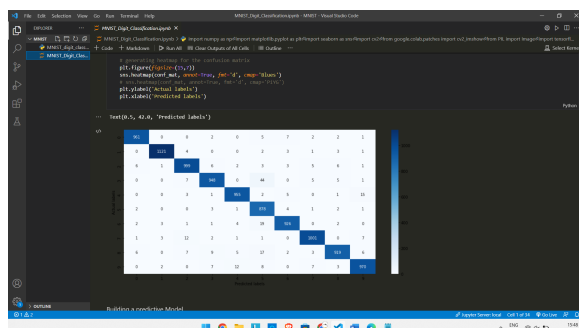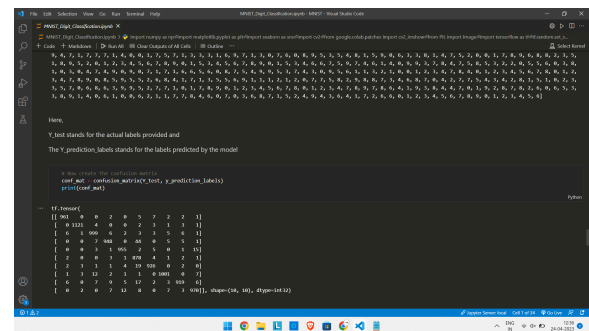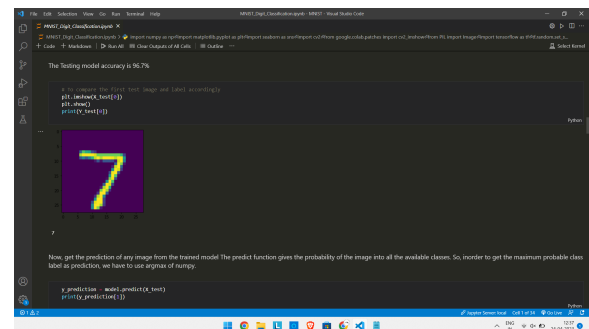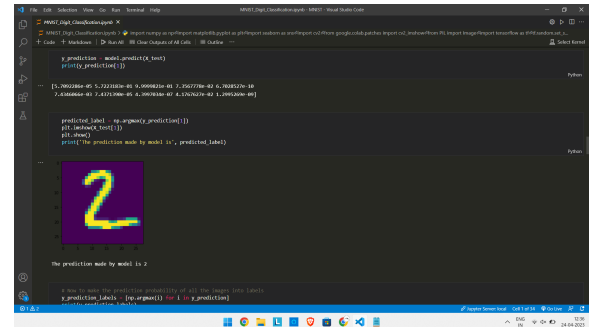
References:

[1] Handwritten Data Digitization Using an Anchor based Multi-Channel CNN (MCCNN) Trained on a Hybrid Dataset (h-EH)
By Abhinandan Chineya,b, Anwesh Reddy Paduria,* , Narayana Darapanenia,c, Santosh Kulkarnia,c, Manish Kadama,b, Ishan Kohlia,d, Malarvizhi Subramaniyan
in Procedia Computer Science 189 (2021) 175–182

[2] HANDWRITTEN RECOGNITION BY USING MACHINE LEARNING APPROACH
By P.Thangamariappan and Dr.J.C.Miraclin Joyce Pamila
in Online March 2020 in IJEAST

[3] Assessing Four Neural Networks on Handwritten Digit Recognition Dataset (MNIST)
By Feiyang Chen, Nan Chen, Hanyang Mao, Hanlin Hu
In arXiv:1811.08278v2 [cs.CV] 20 Jul 2019

[4] AN ENSEMBLE OF SIMPLE CONVOLUTIONAL NEURAL NETWORK MODELS FOR MNIST DIGIT RECOGNITION
By Sanghyeon An Minjun Lee Sanglee Park Heerin Yang Jungmin So
In arXiv:2008.10400v2 [cs.CV] 5 Oct 2020

[5] A Survey of Handwritten Character Recognition with MNIST and EMNIST
By Alejandro Baldominos * , Yago Saez and Pedro Isasi

Published: 4 August 2019

[6] HANDWRITTEN CHARACTER RECOGNITION USING CNN WITH EXTENDED MNIST DATASET
By Zaiban Kaladgi*1, Sarvesh Gupta*2, Moiz Jesawada*3, Vishal Kandale*4
In e-ISSN: 2582-5208 International Research Journal of Modernization in Engineering Technology and Science Volume:04/Issue:05/May-2022

[7] Analytical Study Of Handwritten Character Recognition: A Deep Learning Way
By Dr. Pooja Raundale and Hadi Maredia

[8] Handwritten Digit Recognition and Text Conversion using MNIST Dataset
By 1K. Deepthi Krishna, 2Sathvika. R, 3K. Kavya Sree, 4Raga Deepika K, 5Satya Priyanka
In 2020 IJCRT | Volume 8, Issue 5 May 2020

[9] Handwritten Digits, Alphabet & Symbol Recognition Using Machine Learning
By Prof. K. S. Chandwani1 , Akanksha Nimje2 , Avanti Zade3 , Divya Surjuse4 , Sakshi Ingole5 , Simran Sheikh6
In June 2021| IJIRT | Volume 8 Issue 1

[10] Recognition of Handwritten Digits and English Texts using MNIST and EMNIST Datasets
By Asha B. Shetty1 , Navya N. Ail2 , M. Sahana3 , Sushmitha4 , Varsha P. Bhat5*
In International Journal of Research in Engineering, Science and Management Volume 4, Issue 7, July 2021