Handwriting Digit Recognition on Mnist Using CNN

Ms. G. Chandraleka[1], Mrs. K. Dhamayanthi[2]
[1, 2] Dept of MCA
[2]Assistant Professor, Dept of MCA
[1, 2] Francis Xavier Engineering College, Tamilnadu, India

Abstract- *The capacity of pc applications to understand numbers written via way of means of human beings is known as handwritten digit recognition. This is hard to do with a device due to the fact handwritten numbers aren't usually best and may be special shapes and sizes. A transcribed wide variety manipulate body is a way to resolve this trouble that makes use of more than a few photograph and acknowledges the wide variety withinside the photograph. Using the MNIST dataset and the PyTorch library, a convolutional neural community version turned into evolved for handwritten digit recognition. The capacity of a pc to understand human handwritten numbers from diverse sources, which include images, papers, contact paths, etc., and classify them into ten predefined categories (0-9) is known as handwritten digit recognition. It has been the situation of countless studies withinside the discipline of calligraphy. Checking the wide variety has diverse tasks, as an instance confirming the wide variety plate, organizing correspondence, the financial institution genuinely tests the processing, etc. We face many limitations with regards to spotting handwritten numbers. due to the special writing forms of special businesses of people, due to the fact it's far whatever however an optical ID. For handwritten digit recognition, this look at compares numerous device analyzing and deep analyzing algorithms. We used aid for this. Convolutional Neural Network, Multilayer Perceptron and Vector Machine The correlation among those calculations is primarily based totally on their sensitivity, violations, and take a look at training time, supported via way of means of drawings and guides evolved for viewing matplotlib.*

Keywords- handwritten digit recognition, CNN, deep learning, MNIST dataset, epochs, hidden levels, stochastic gradient descent and again propagation.

I. INTRODUCTION

Over time, the number of fields studied in depth increases. Deep research uses a convolutional neural network (CNN) to analyze the visible image [1, 2]. Object recognition, face recognition, robotics, video analysis, segmentation, pattern recognition, natural language processing, spam detection, topic classification, regression analysis, speech recognition, an example of image classification that can be done using convolutional neural networks. Accuracy in these areas, including the popularity of handwritten numbers, has been brought to human-level perfection through the use of deep convolutional neural networks (CNNs). A biological model of the mammalian visual system is the only one that stimulates the CNN structure. Cells in the visual cortex of the cat are sensitized to a small point in the visual field, which is diagnosed by the receptive field [3]. This was observed by D.H. Hubel et al. in 1062. Neocognitron [], D.H. Hubelin et al. sample version stimulated by paintings. [5, 6] became the primary vision of the laptop. In 1980 it was changed to supply through Fukushima. In 1998, the CNN framework of LeCun et al. [7] which had seven layers of a convolutional neural network. It was cleverly converted to handwritten numbers directly from the pixel values of the images [8]. The training version uses the gradient descent and forward regularization [9]. In handwritten popularity numbers, characters are entered using the enter key. The version can be diagnosed by the system. A simple synthetic neural network (ANN) has an input layer, an output layer, and some hidden layers between the input and output layers. CNN has a completely comparable structure to ANN. Each layer of ANN has many neurons. The weighted sum of all neurons in a layer becomes the input to the neuron in the next layer, including unbiased costs. In CNN, a layer has 3 dimensions. Here, not all neurons are fully connected. Instead, each neuron in the layer is connected to a nearby receptive field. A value trait creates community as an effort to teach. It compares the output of the community with the primary output. The token propagates back to the system again and again to replace the common weights and biases in all receiving fields to lower the cost of the value function, thus increasing the community efficiency [10]. Y. LeCun et al., The generalizability of learning networks can be significantly improved by providing constraints on the task domain. This article shows how such constraints can be integrated back into the distribution network through the network architecture. This approach has been successfully applied to identify numbers in handwritten zip codes provided by the US Postal Service. A single network learns the entire recognition process, moving from the normalized image of the character to the final classification.

II. RELATED WORK

K. Gaurav, Bhatia P. K. [1] Et al, This article covers different character recognition preprocessing techniques with different types of images, ranging from simple documents based on a handwritten shape to documents with colorful and complex backgrounds of varying intensity. This article discusses various pre-processing techniques such as error detection and correction, image enhancement techniques, contrast stretching, binarization, denoising techniques, normalization and segmentation, morphological processing of the image. But even with all of the above techniques, it may not be possible to achieve the highest accuracy in a pretreatment system.

Salvador Espana-Boquera et al [2], inThis hybrid Hidden Markov Paper (HMM) model is designed to recognize unlimited offline handwritten text. In this case the structural part of the optical model was modeled using Markov chains and a multilayer perceptron is used to estimate the emission probability.

Pradeep et al. [3] proposed Classification based on the neural network of the handwriting recognition system. Each individual character is scaled to 30 x 20 pixels for processing. They used binary functionality to train a neural network. However, these features are not robust. In post-processing, recognized characters are converted to ASCII format.The input layer has 600 neurons equal to the number of pixels. The output layer has 26 neurons.

Salvador Espana-Boquera et al. [4] in this paper hybrid Hidden Markov Model (HMM) model is proposed for recognizing unconstrained offline handwritten texts. In this, the structural part of the optical model has been modeled with Markov chains, and a Multilayer Perceptron is used to estimate the emission probabilities.

In [5], Diagonal feature extraction has been proposed for offline character recognition. It is based on the SSN model. Two approaches with 54 and 69 functions were chosen to build this neural network recognition system. To compare the recognition efficiency of the proposed diagonal feature extraction method, the neural network recognition system is trained using the horizontal and vertical feature extraction methods. The diagonal feature extraction method was found to give a recognition accuracy of out of 97.8% for 54 articles and 98.5% for 69 articles.

A. Brakensiek, J. Rottland, A. Kosmala, J. Rigoll [6] et al, This article describes an offline handwriting recognition system based on Hidden Markov Models (HMM) using discrete and hybrid modeling techniques. Compare handwriting recognition experiments using a discrete approach and two different hybrid approaches consisting of a discrete structure and a semi-continuous structure. To develop the system, an approach without segmentation is considered

B.

R. Bajaj, L. Dey, S. Chaudhari et al [7], employed three different kinds of features, namely, the density features, moment features and descriptive component features for classification of Numerals. They proposed multi classifier connectionist architecture for increasing the recognition reliability and they obtained 89.6% accuracy for handwritten numerals.

Sandhya Arora in [8], used four feature extraction techniques namely, intersection, shadow feature, chain code histogram and straight line fitting features. Shadow features are computed globally for character image while intersection features, chain code histogram features and line fitting features are computed by dividing the character image into different segments. On experimentation with a dataset of 4900 samples the overall recognition rate observed was 92.80% for Devanagari characters.

Mohammed Z. Khedher, Gheith A. Abandah, and Ahmed M. Al Khawaldeh [9] et al, this paper describes that Recognition of characters greatly depends upon the features used. An off-line recognition system based on the selected features was built. Evaluation of the importance and accuracy of the selected features is made. The recognition based on the selected features give average accuracies of 88% and 70% for the numbers and letters, respectively.

Pritpal singh et al. [10] Mentioned wavelet transforms based handwritten character and numeral recognition for Gurumukhi script. Color images are converted in gray scale and median filter is applied to remove noise. Binarized image is then normalized to 32 X 64 pixels size. The decomposition of the image into different frequency bands is obtained by successive low pass and high-pass filtering of the signal and down-sampling the coefficients after each filtering.

III. THEORY

Proximal guide Vector gadget (PSVM),MultilayerPer ceptron, guide Vector machine (SVM), Random woodland, Bayes net, Naive Bayes, J48, and Random Tree area number of the various algorithms used in handwritten digit reputation systems.Theaccuracythat these algorithms pro vide is, according with preceding studies, withinthe variety of:

- Proximal SVM - 98%
- Multi-facet Perceptron - 90%
- SVM - 87%
- Bayes Net - 84%
- J48 - 79%
- Irregular Tree - 75%

Despite the fact that these calculations can prove valuable in some applications for this innovation, many different applications, such as banking applications, require better results that can be achieved with different calculations compared to the previously referenced applications.

A convolutional neural network (CNN) can be used in handwritten digit recognition systems to reduce errors and increase overall performance. Our proposed system uses a CNN with multiple pooling and convolutional layers and a 3x3 kernel to achieve this. Our model uses 60,000 28x28 grayscales during the training. Compared with traditional algorithms such as SVM, Multilayer Perceptron, Bayes Net, Random Forest and so on, our model achieves about 99.16% accuracy. used to set up handwritten systems that recognize numbers. The goal of the proposed work is to recognize user-defined handwritten numbers and identify the integer that the user enters - by default as a decimal - before converting it to binary, octal or hexadecimal. For this purpose, a graphical user interface (GUI) is developed where the user is presented with a canvas widget that allows them to draw handwritten strings for conversion and detection.

A1. *Research Methodology*

The input layer consists of 28 x 28 pixel images, which means that the network contains 78 neurons as input data. The input pixels are shades of gray with a value of 0 for white pixels and 1 for black pixels. This CNN model has five hidden layers. The first hidden layer is convolutional layer 1, which is responsible for extracting features from the input data. This layer performs a convolution operation on small local areas by combining the filter with the previous layer..

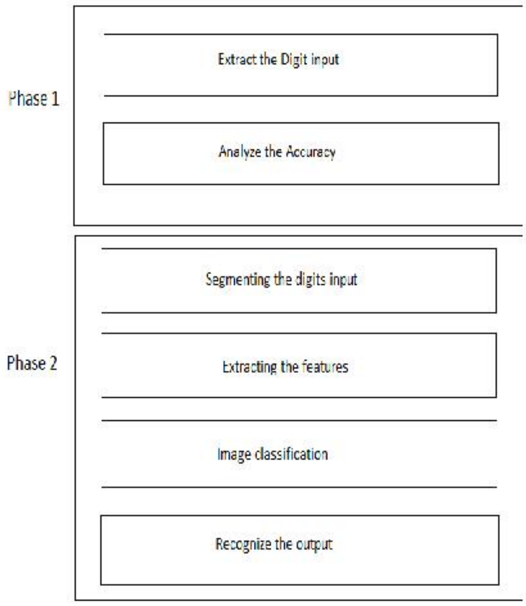


Fig 1 Researh Methodology

A2.*Algorithm Implementation*

A seven-layered convolutional neural network with one input layer, five hidden layers, and one output layer is designed and demonstrated in order to recognize the digits written by hand.

Step1: The network's input data consists of 784 neurons because the images in the input layer are 28 by 28 pixels. The grayscale input pixels have values of 0 for white pixels and 1 for black pixels. This CNN model has five hidden layers here. Convolution layer 1 is the first hidden layer, and it is in charge of extracting features from input data. By combining a filter from the layer before it, this layer performs a convolution operation on small, localized areas. Moreover, it comprises of various component maps with learnable portions and corrected direct units (ReLU).

Step2: The location of the filters is determined by the kernel size. To improve the model's performance, ReLU is used as an activation function at the end of each convolution layer and as a fully connected layer. The pooling layer 1 is the next hidden layer. It reduces the model's number of parameters, computational complexity, and output information from the convolution layer. Max pooling, min pooling, average pooling, and L2 pooling are the various pooling methods. Max pooling is used to subsample each feature map's dimension in this case. Pooling layer 2 and convolution layer 2, which perform the same tasks as convolution layer 1 and pooling layer 1, with the exception of their differing feature maps and kernel sizes.

Step3: After the pooling layer, a Flatten layer is used to turn the 2D featured map matrix into a 1D feature vector, letting the fully connected layers handle the output. Another hidden layer, also known as the dense layer, is a fully connected layer. Similar to the hidden layer of Artificial Neural Networks (ANNs), this layer is fully interconnected and links each neuron from the previous layer to the next. To diminish overfitting, dropout regularization strategy is utilized at completely associated layer 1. During training, it randomly kills some neurons to boost the network's performance and make it stronger.

Step4: As a result, the network is less likely to overfit the training data and is better able to generalize. The digits from 0 to 9 are determined by ten neurons in the network's output layer. The activation function softmax, which is used in the output layer to improve the model's performance, classifies the output digit from 0 to 9, which has the highest activation value.

IV. EXPERIMENT AND RESULT

It is a collection of thousands of handwritten images that have been used to train classification models using machine learning methods. As a supporter of this number statement, we prepare a multifaceted perceptron that uses Tensorflow - v2 to recognize transcribed numbers.

In practice, one of the central problems of pattern recognition applications is the recognition of handwritten characters. The advantages of a digital receipt remember to organize the mail, the bank really looks at the processing, structures the flow of information, etc. The process of enabling machines to recognize human handwritten text is called handwriting recognition. Handwritten numbers are imperfect, vary from person to person, and can be made to different tastes, making it difficult for a machine to accomplish this task.

A1. Simulation Environment

Live code, equations, visualizations, and text documents can all be created and shared with Jupyter Notebook, an open-source web application. Project Jupyter staff is in charge of maintaining Jupyter Notebooks. This is a random project from his Python project, which included its own Python notebook project. The core programming languages it supports are what gave it its name, Jupyter: Julia, Python, and R. Jupytercomes with an IPython kernel that can be used to compose Python programs, however more than 100 different parts areavailable. Welldone. As a scientific labnotebook, Jupyter notebooks are especially useful for computational physics and extensive data analysis with computer programs.

Google Colab, also known as Colaboratory, is a cloud-based, free Jupyter notebook environment that doesn't need any configuration. Support for users' free GPU and TPU You can write and run code, store and share your analysis, and use powerful computing tools right from your browser with Colaboratory, all for free. The product guarantees collaboration, as the name suggests. AJupyter notebook that makes use of Google Docs' linking feature. Additionally, nothing needs to be updated because it runs on Google servers. Your Google Drive account is where notebooks are stored. It offers a platform that makes it possible for anyone to create deep learning applications by utilizing well-known libraries like PyTorch, TensorFlow, and Keras. It offers a computer-friendly means of avoiding the expense of extensive ML operations training.

A1. Architecture diagram

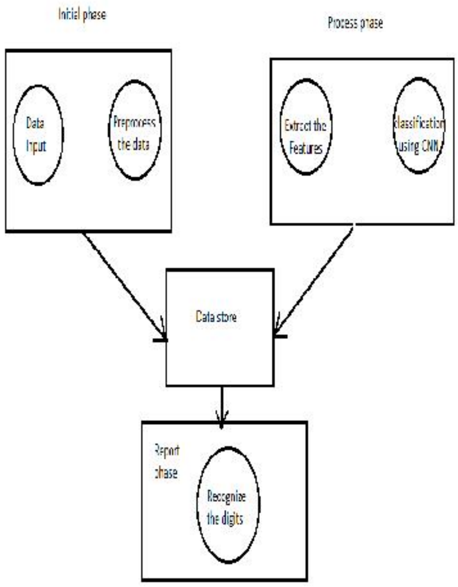


Fig 2 Architecture diagram

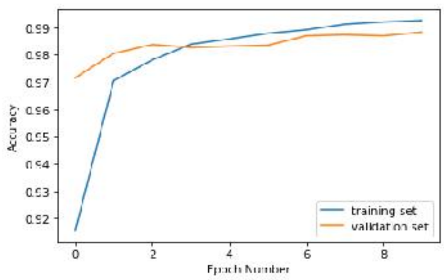


Fig 3: Accuracay plot curve for training and validation

Fig 4: Visualise validation predicted data

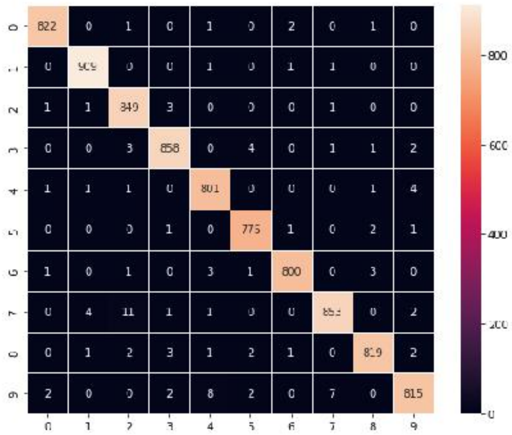
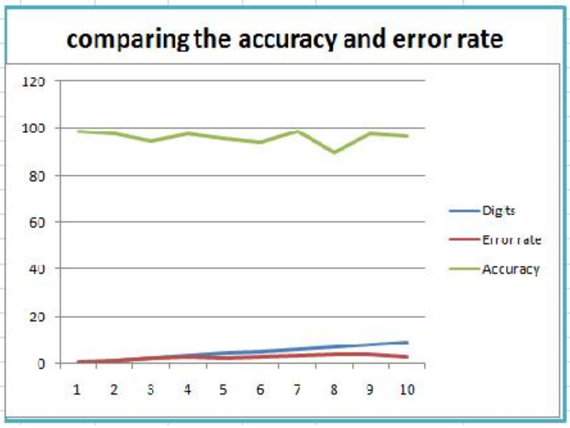


Fig 5:Confusion matrix of validation dataset

Fig 6: Visualize and test prediction

A2. Performance Metrics

Digits	Error rate	Accuracy
0	0.4	99
1	1.2	98
2	1.9	95
3	2.6	98
4	2.4	96
5	3	94
6	3.4	99
7	3.9	90
8	3.7	98
9	2.7	97



V. DISCUSSION AND CONCLUSION

Handwritten Digit Recognition employing deep learning techniques is demonstrated in this paper. In order to compare the classifiers, the most widely used machine learning algorithms—KNN, SVM, RFC, and CNN—have all been trained and tested on the same data. High levels of precision can be achieved with these deep learning methods. This method improves the accuracy of classification models by more than 99%, focusing on which classifier performs better than others. A CNN model can achieve an accuracy of approximately 98.72 percent by employing Keras as its backend and Tensorflow as its software. In this initial test, CNN has an accuracy of 98.72 percent, KNN has an accuracy of 96.67 percent, and RFC and SVM aren't very good.

VI. FUTURE SCOPE

Applications based on deep and machine learning algorithms have practically no limit to their future development. We can work on a denser or hybrid algorithm in the future to solve many problems with more diverse data than the current set of algorithms. As a result of the above differentiation and future development, we can attain high-level functioning applications that can be used in classified or government agencies as well as for the common people. We can use these algorithms in hospitals for detailed medical diagnosis, treatment, and monitoring of patients. We can use it in surveillance systems to keep track of suspicious activity under the system. We can use it in fingerprint and retinal scanners. Database filtering applications. Equipment checking for national forces. Utilizing these algorithms in both day-to-day and high-level applications (i.e., at the corporate or government level) can assist us in creating an atmosphere of safety, awareness, and comfort. Because of their absolute accuracy and advantages over numerous major issues, application-based AI and deep learning are the technological future.

REFERENCES

[1] M. Wu and Z. Zhang, Handwritten Digit Classification using the MNIST Data Set, 2010.

[2] A. Dutta and A. Dutta, Handwritten digit recognition using deep learning, International Journal of Advanced Research in Computer Engineering and Technology (IJARCET), vol. 6, no. July 7, 2017.

[3] Al Maadeed, Somaya and Abdelaal Hassaine, Automatic prediction of age, gender and nationality using offline handwriting. EURASIP Journal of Image and Video Processing, no. 1 2014.

[4] Gaurav Jain, Jason Ko, Handwritten Digits Recognition, Project Report, University of Toronto, 2008-11-21.

[5] Hamid, Norhidayu Abdul and NilamNur Amir Sjarif, Manwritten Recognition Using SVM, KNN and Neural Network, arXiv preprintarXiv:1702.00723 (2017).

[6] R.G.Mihalyi, Classification of Handwritten Numbers with Support Vector Machines, 2011.

[7] Z. Dan, C. Xu, The Recognition of Handwritten Digits Based on BP Neural Networks and the Implementation on Android, julkaisussa: 3-a Internacia Konferenco pri Intelligent System Design and Engineering Applications, s. 198-1509, 2013.

[8] Y. Liu and Q. Liu, "Convolutional Neural Networks with High Margin Softmax Loss Function for Cognitive Load Recognition"2017 36th Chinese Control Conference (CCC), 2017.

[9] Y. LeCun et al., "Backpropagation Applied to handwritten zipcode discovery", Neural computing, vol. 1, no., page 51-551, 1989.

[10] A. Kriĵevski, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks", julikussa Advances in neural information processing systems, 2012, s. 1097-1105.

[11] D. Hubel and T. Wiesel, “Aberrant visual projections in the Siamese cat,” The Journal of physiology, vol. 218, no. 1, s. 33-62, 1971.

[12] J. Jin, K. Fu ja C. Zhang, "Traffic Sign Recognition with Hinge Loss Trained Convolutional Neural Networks", IEEE Transactions intelligentsete transpordisüsteemide kohta, vol. 15, nr. 5, pp. 1991-2000.

[13] M. Wu and Z. Zhang, "Classification of handwritten numbers using the mnist dataset", Course Project CSE802: Pattern Classification and Analysis, 2010.

[14] Y. LeCun, Y. Bengio and G. Hinton, "Profunda lernado", naturo, vol. 521, no. 7553, p. 36, 2015.

[15] D. Cireşan, U. Meier and J. Chem. Schmidhuber, "Multi-Column Deep Neural Networks for Image Classification," arXiv preprint arXiv:1202.275, 2012.

[16] K. Fukushima and S. Miyake, "Neocognitron: SelfOrganizing Neural Network Model for the Mechanism of Visual Pattern recognition," Competition and Cooper neural networks: Springer, 1982, pp. 267-285.

[17] Y. LeCun et al., "Handwritten digit detection with a backpropagation network" in Advances in neuroinformation Processing Systems, 1990, pp. 396.

[18] Y. LeCun, L . Bottou, Y. Bengio, Kaj P. Haffner, "Applying Gradientbased Learning to Document Recognition," Proceedings IEEE, vol. 86, no. 11, p. 2278-232, 1998.

[19] R. Hecht-Nielsen, "Theory of the backpropagation neuronet", in Neural networks for perception: Elsevier, 1992, p. 65-93

[20] Y.LeCun, "LeNet-5, a convolutional hermover network", URL: http://yann. lecun. com/exdb/lenet, voi. 20, 2015.

[21] S.J. Russell kaj P. Norvig, The technical intellect: the contemporary name. malajzianoj; Pearson Education Limited, 2016.

[22] S.Haykin, "Neural Review: A Comprehensive Foundation: MacMillan College," New York.

[23] K.B. Lee, S. Cheon, kaj C.O. Kim, “Convolutional neural network for semiconductor defect classification and diagnosismanufacturing processes", IEEE Transactions on Semiconductor Manufacturing, vol. 30, no. 2, pp. 135-12, 2017.

[24] K. G. Pasi and S. R. Naik, "Effect of parameter variations on the accuracy of a convolutional neural network", in 2016 International Conference on Computing, Analytics and Security Trends (CAST), 2016, pp. 398-03: IEEE.

[25] D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella and J. Schmidhuber, "A flexible, high-performance convolutional neural networknetworks for image classification" Twenty-second International Joint Conference on Artificial Intelligence, 2011

[26] K. Isogawa, T. Ida, T. Shiodera kaj T. Takeguchi, Deep Shrinkage Convolutional Neural Network for Adaptive Noise Reduction, "IEEE Signal Processing Letters, kd. 25, nr 2, lk

[27] C. C. Park, Y. Kim, and G. Kim, "Sentence search in image stream using recurrent coherence convolutional networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 0, no. , s. 95-957, 2018.
[28] Y. Yin, J. Wu ja H. Zheng, "Ncfm: Accurate Handwritten Digit Recognition Using Convolutional Neural Networks," 2016. International Joint Conference on Neural Networks (IJCNN), 2016, pp. 525-531: IEEE.
[29] L. Xie, J. Wang, Z . Wei, M. Wang, and Q. Tian, "Disturblabel: Regulating cnn on the loss layer", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 753-762.
[30] A. Tavanaei and A. S. Maida, "Multilayer unsupervised learning in a spiking convolutional neural network", 2017 International Joint Conference on Neural Networks (IJCNN), 2017, lk 2023-2030: IEEE 22-228, 2018.