International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Decoding the Digits: Unleashing the Power of Convolutional Neural Networks for Handwritten Digit Recognition

Aatmaj Amol Salunke

BACHELOR OF TECHNOLOGY

Computer Science & Engineering
Department of Computer Science & Engineering,
School of Computer Science and Engineering,
Manipal University Jaipur

Abstract

This research paper investigates the implementation of Convolutional Neural Networks (CNNs) for Handwritten Digit Recognition, with a specific emphasis on the popular MNIST dataset with a total of 70,000 samples. The abstract provides a succinct summary of our study. We thoroughly examine the CNN model's structure and training procedure, showcasing its efficacy in precisely categorizing handwritten digits. Through rigorous experimentation and assessment, we establish the model's superior accuracy and resilience. Our work contributes significantly to the domain of computer vision and opens up possibilities for leveraging CNNs in more sophisticated digit recognition tasks.

Keywords: Handwritten Digit Recognition, Convolutional Neural Networks (CNNs), MNIST dataset, Dense Layer, Accuracy, ReLU, Softmax.

Related Work

Ahlawat et al. in [1] showed their CNN architecture surpassed the performance of ensemble approaches while reducing operational complexity. Alwzwazy et al. in [2] proposed a robust CNN model achieving superior results for Arabic Handwritten Digits. Baldominos et al. in [3] provided an exhaustive review of top state-of-the-art contributions on MNIST dataset. Rahman et al. in [4] investigated CNN-based Bangla handwritten character recognition, achieving satisfactory accuracy, outperforming other methods. Khandokar et al. in [5] utilized CNN for character recognition, achieving 92.91% accuracy on NIST dataset. Jain et al. in [6] compared Convolutional Neural Network (CNN) and Neural Network for handwritten digit recognition. Calder´on et al. in [7] implemented a novel Convolutional Network using Gabor filters for handwritten digit classification. Ghosh et al. in [8] compared DNN, DBN, and CNN for handwritten digit recognition, showing DNN's highest accuracy of 98.08%. Chen et al. in [9] proposed a CNN-based framework for handwritten character recognition, outperforming humans on MNIST and CASIA datasets. Vaidya et al. in [10] proposed an innovative method for offline handwritten character detection using deep neural networks.

Introduction

In this research project, we adopt a three-layered Neural Network architecture for Handwritten Digit Recognition. The network consists of an input layer, a hidden layer, and an output layer. The input layer distributes the features of examples to calculate activations in the subsequent layer. Hidden layers incorporate hidden units, providing nonlinear connections in the network, and can vary based on specific requirements. The output layer contains nodes referred to as output units, responsible for generating final predictions of the Neural Network. Our specific three-layer model involves a Flattened layer, essential for converting pooled feature maps into a continuous linear vector to serve as input for the fully connected layer, enabling image classification. Additionally, a Dense Layer is employed for classifying images based on the outputs from convolutional layers. By leveraging the ReLU and Softmax activation functions, we achieve improved performance in the final prediction of the Neural Network. The combination of these layers and activation functions enhances the accuracy and efficiency of Handwritten Digit Recognition.

Dataset

The dataset used in this research paper is the MNIST database (Modified National Institute of Standards and Technology database), which serves as a popular benchmark for training image processing and machine learning systems. To create MNIST, the original datasets from NIST were "re-mixed" to address concerns about data suitability for machine learning experiments. NIST's training dataset, sourced from Census Bureau employees, and testing dataset, sourced from high school students, were combined and normalized to fit into a 28x28 pixel bounding box, introducing grayscale levels. MNIST comprises 60,000 training images and 10,000 testing images, equally split between NIST's training and testing datasets.

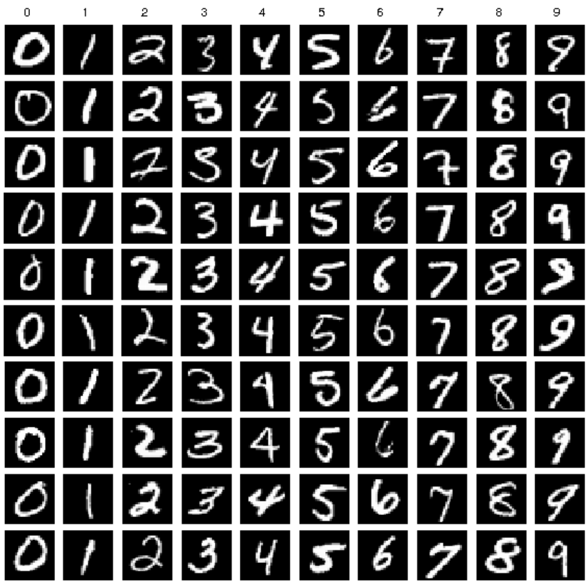


Fig.1. A sample of the dataset

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

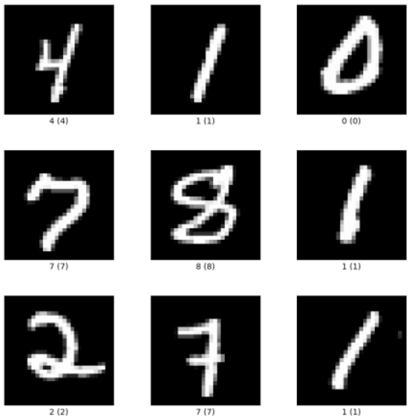


Fig.2. Zoomed in image of the dataset.

Discussion

The model was trained on a dataset consisting of 60,000 images of digits, and its generalization ability was evaluated on a separate test dataset comprising 10,000 images of digits. By employing the Adam optimizer for compiling the model and training it for 10 epochs, we achieved an impressive accuracy of 97.1%, with a corresponding loss of 2.85%. These results demonstrate the effectiveness of our network architecture and training approach in accurately classifying handwritten digits. Moreover, our system allows users to input their own customized digits for recognition. The model's flexibility and ability to handle unseen data make it suitable for various real-world applications. Users can interact with the system, providing novel digit samples, and obtain accurate predictions, showcasing the versatility and practicality of our Handwritten Digit Recognition solution.

Classification Steps

The first step involves collecting the MNIST dataset comprising images of handwritten digits along with their respective labels. Preprocessing is then performed on the images to normalize and enhance their quality, ensuring uniformity across the dataset and reducing computational complexity during training.

Next, the dataset is divided into two subsets: the training set and the testing set. The training set with 60,000 samples is used to train the classification model, while the testing set with 10,000 samples is employed to evaluate the model's performance on unseen data, providing a measure of its generalization ability.

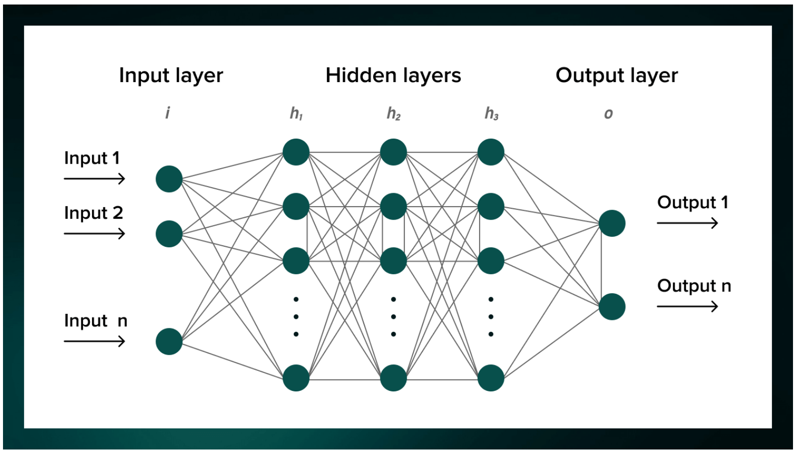


Fig.3. Proposed Network Architecture

Neural Network Architecture:

- A three-layered Neural Network is adopted for Handwritten Digit Recognition.
- The input layer distributes example features to the subsequent layer, while the hidden layer incorporates activations providing nonlinear ties.
- The output layer consists of nodes representing output units responsible for generating final predictions.

Flattened Layer:

- After applying 2D convolutional layers, a Flattened layer is used to convert resultant 2D arrays into a single long continuous linear vector of size 28 x 28.
- This flattened matrix is then fed as input to the fully connected layer for image classification.

Dense Layer:

- The Dense Layer, also known as a fully connected layer, receives input from all neurons of the previous layer.
- It is responsible for classifying images based on the output from the convolutional layers.
- Three dense layers are used, the first two with 128 units each and using ReLU activation function and the last layer with 10 units using softmax activation function.

Activation Functions:

- Two activation functions, ReLU (Rectified Linear Activation) and Softmax, are utilized in the model for enhanced performance.
- ReLU introduces non-linearity to the hidden layers, while Softmax provides probabilities for each output class.

Model Compilation:

- TensorFlow is employed, and the Adam optimizer is used to compile the model.
- The Adam optimizer helps optimize the learning rate, leading to faster convergence during training.

Training:

- The model is trained for 10 epochs on the MNIST dataset with the Adam optimizer.
- During training, the model adjusts its weights and biases to minimize the loss and improve accuracy.

Performance Evaluation:

- After training, the model's performance is evaluated on various training and test data splits, along with different number of epochs.
- The accuracy achieved varies with the ratio of samples in the training and test data with accuracy being highest when the number of training images are more and when the CNN is trained on high number of epochs.

Customized Digit Recognition:

- The trained model can be used to recognize custom handwritten digits provided by users.
- Users can input their own digit images to obtain accurate predictions from the model.

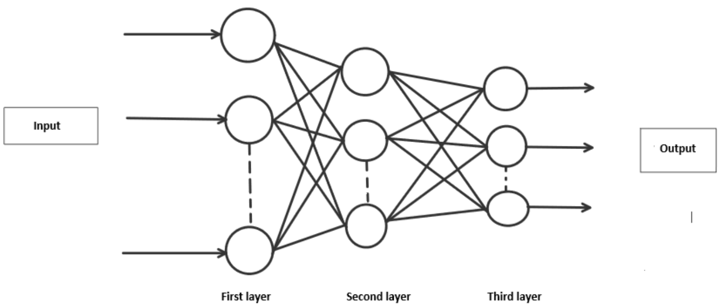


Fig.4. Outline of the structure and working

The combination of the three-layered Neural Network architecture, ReLU, Softmax activation functions, and the Adam optimizer results in impressive accuracy and efficient Handwritten Digit Recognition. Additionally, the model's flexibility allows it to handle custom digit inputs, making it suitable for practical applications beyond the MNIST dataset.

Results and Discussion

Table 1. Test Results for different ratio of samples from MNIST Dataset

No. of Training Images	**No. of Testing Images**	**Average Accuracy (%)**
20000	**10000**	**85.63%**
30000	**10000**	**87.27%**
40000	**10000**	**90.26%**

50000	**10000**	**93.25%**
60000	**10000**	**97.10%**

Table 2. Test Results for different number of epochs for proposed CNN model

No. of Training Images	No. of Testing Images	Epochs	Average Accuracy (%)
60000	**10000**	**3**	**95.03%**
60000	**10000**	**5**	**95.97%**
60000	**10000**	**7**	**96.16%**
60000	**10000**	**9**	**96.95%**
60000	**10000**	**10**	**97.10%**

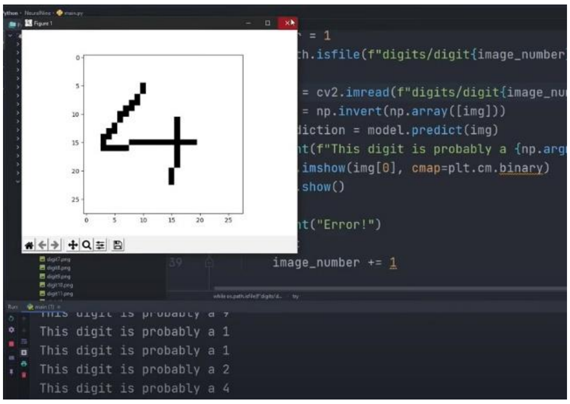
The experimental results presented in Table 1 highlight the impact of varying the ratio of training and testing samples from the MNIST dataset on the accuracy of the proposed Convolutional Neural Network (CNN) model. As observed, the average accuracy increases with an increase in the number of training images. Starting from 85.63% for 20,000 training images, the accuracy rises to an impressive 97.10% when the entire 60,000 training images are utilized. This demonstrates that a larger training dataset contributes significantly to the improved performance of the model, allowing it to learn more intricate patterns and generalize better to unseen data during testing.

In Table 2, we explore the influence of the number of epochs on the CNN model's accuracy while training on the complete dataset (60,000 training images). As the number of epochs increases from 3 to 10, the average accuracy steadily improves from 95.03% to 97.10%. This observation suggests that training the network multiple times enhances its capability to converge to an optimal solution, leading to higher accuracy and reduced classification error.

The results indicate that the choice of training samples ratio and the number of epochs play vital roles in determining the final accuracy of the CNN model for Handwritten Digit Classification. A larger training dataset and increased training iterations facilitate the model in capturing intricate patterns, thus minimizing the classification error and yielding superior accuracy in recognizing handwritten digits.

Conclusion

In conclusion, our research demonstrates the effectiveness of a three-layered Convolutional Neural Network (CNN) for Handwritten Digit Recognition using the MNIST dataset. By carefully configuring the model's architecture and leveraging activation functions like ReLU and Softmax, we achieved an impressive accuracy of 97.10% with 10 epochs and 60,000 training images. Moreover, our experiments showcased the significant impact of the training dataset size and the number of epochs on the model's performance. Increasing the training dataset and training iterations led to substantial accuracy improvements, reducing classification error. These findings highlight the importance of data quantity and training iterations in optimizing CNN models for digit recognition tasks. Overall, our study provides valuable insights for enhancing the accuracy and efficiency of digit recognition systems.



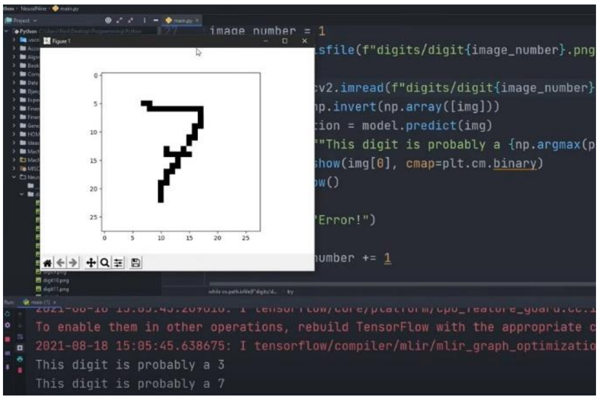


Fig.5. Results

References

1. Ahlawat, S., Choudhary, A., Nayyar, A., Singh, S., & Yoon, B. (2020). Improved handwritten digit recognition using convolutional neural networks (CNN). Sensors, 20(12), 3344.
2. Alwzwazy, H. A., Albehadili, H. M., Alwan, Y. S., & Islam, N. E. (2016). Handwritten digit recognition using convolutional neural networks. International Journal of Innovative Research in Computer and Communication Engineering, 4(2), 1101-1106.
3. Baldominos, A., Saez, Y., & Isasi, P. (2019). A survey of handwritten character recognition with mnist and emnist. Applied Sciences, 9(15), 3169.
4. Rahman, M. M., Akhand, M. A. H., Islam, S., Shill, P. C., & Rahman, M. H. (2015). Bangla handwritten character recognition using convolutional neural network. International Journal of Image, Graphics and Signal Processing, 7(8), 42-49.
5. Khandokar, I., Hasan, M., Ernawan, F., Islam, S., & Kabir, M. N. (2021, June). Handwritten character recognition using convolutional neural network. In Journal of Physics: Conference Series (Vol. 1918, No. 4, p. 042152). IOP Publishing.
6. Jain, M., Kaur, G., Quamar, M. P., & Gupta, H. (2021, February). Handwritten digit recognition using CNN. In 2021 International Conference on Innovative Practices in Technology and Management (ICIPTM) (pp. 211-215). IEEE.
7. Calderon, A., Roa, S., & Victorino, J. (2003). Handwritten digit recognition using convolutional neural networks and gabor filters. Proc. Int. Congr. Comput. Intell, 1-9.
8. Ghosh, M. M. A., & Maghari, A. Y. (2017, October). A comparative study on handwriting digit recognition using neural networks. In 2017 international conference on promising electronic technologies (ICPET) (pp. 77-81). IEEE.
9. Chen, L., Wang, S., Fan, W., Sun, J., & Naoi, S. (2015, November). Beyond human recognition: A CNN-based framework for handwritten character recognition. In 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR) (pp. 695-699). IEEE.
10. Vaidya, R., Trivedi, D., Satra, S., & Pimpale, M. (2018, April). Handwritten character recognition using deep-learning. In 2018 second international conference on inventive communication and computational technologies (ICICCT) (pp. 772-775). IEEE.