Deep Learning for Handwritten Digit Recognition System: A Convolution Neural Network Approach

Maha A. Al-Bayati[1,*]

[1]Department of Computer Science, College of Science, Mustansiriyah University, Baghdad, Iraq

Email: mahabayati@uomustansiriyah.edu.iq

Abstract

Artificial intelligence techniques including deep learning play a major role in all fields and in line with the advancement in technology. Handwritten digit recognition is an important issue in the field of computer vision, which is used in wide applications such as optical character recognition and handwritten digits. In the current research, we describe a unique deep learning technique that uses a Convolutional Neural Network (CNN) framework with better normalization algorithms and adjusted hyperparameters for improved efficiency as well as generalize. Contrasting conventional techniques, our methodology concentrates on minimizing overfitting through the use of adjustable rate of abandonment and innovative pooling procedures, resulting in greater accuracy in handwriting number classification. Following considerable research, the recommended approach obtains an outstanding classification accuracy of 99.03%, proving its ability to recognize intricate structures in handwritten numbers. The approach's usefulness is reinforced by a complete review of measures including recall, accuracy, F1 score, as well as confuse matrix assessment, which show improvements throughout all digit categories. . The results of the investigation highlight the innovative conceptual layout and optimization methodologies used, representing a substantial leap in the realm of number identification.

Keywords: Deep Learning; CNN; Handwritten; Artificial Intelligence

1. Introduction

The employment of academic tools, handwritten records, and other techniques has led to a significant growth in information volumes. Online platforms and networks have given forth studies difficulties. Those research concerns have drawn the Scientists in multiple scientific disciplines and fresh studies Fields have additionally developed [1]. Social media users provide free information in both organized and unstructured formats, resulting in a significant amount of data. Advancements in data storage and processing have led to new technology, including analyzing sentiment and identifying top users with influence. Data analysis in many languages and handwriting recognition are important [2] [3].

Visual computing is a branch of computing which concentrates on teaching machines to recognize and interpret things and individuals in photos and movies. Computer vision, just like other kinds of AI, aims to execute and systematize operations that mimic the skills of humans [4] [5]. The recognition of handwriting digits is a method of distinguishing between scrawled numerals, which usually vary between 0 to 9, utilizing techniques such as neural networks with convolutional structure (CNN). It is widely utilized in services including phone banking and automated cashiers for cash posting [6]. CNNs are among the more important networks within the deep computing field. CNN has received a lot of interest lately in businesses and academics due to its significant results in a variety of fields, which include but are not restricted to digital vision and the processing of natural languages. The previous studies primarily focus on CNN's applicability in various circumstances without examining CNN in a broader sense, and certain unique concepts offered subsequently aren't addressed [7][8]. The MNIST dataset, comprised of 70,000 gray photographs of written

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

numbers, is frequently utilized for comparing the reliability of different digit interpretation methods [9]. Although CNNs showed exceptional precision on this set of results, issues like as over fitting & the necessity for generalizing models to previously unknown data remained key obstacles.

This presentation will offer a revolutionary technique to recognize handwritten digits based upon an optimized CNN design. We intend to strengthen the model's precision and applicability by combining sophisticated legalization approaches, variable rate of abandonment, including hyper parameter adjustment. Further, the present investigation will examine important efficiency measures, offering a thorough assessment of the strategy benefits and flaws. Through thorough trial and error, researchers show that our technique not just produces innovative outcomes on benchmark datasets, additionally it provides advantageous insights regarding viable ways for enhancing number identification precision.

Results about this study add to the increasing amount of evidence in deep learning as well as computer vision by emphasizing the relevance of architecture efficiency and novel normalization methodologies for improving the discipline of handwritten digit identification.

There are many applications for CNN in many fields [10-13], Table1 illustrates these applications.

Table 1: Applications of CNN's

Application	**Description**
Image Classification	CNNs are widely used for classifying images into predefined categories, such as in object recognition.
Object Detection	CNNs can identify and localize multiple objects within an image, such as in self-driving cars.
Facial Recognition	CNNs are employed in identifying and verifying faces in images, utilized in security systems.
Medical Image Analysis	CNNs assist in diagnosing diseases by analyzing medical images like X-rays and MRIs.
Natural Language Processing	CNNs can also be applied in NLP tasks, such as sentiment analysis and text classification.
Video Analysis	CNNs are used to analyze video data for action recognition and event detection.
Image Segmentation	CNNs facilitate the segmentation of images, which is critical in applications like autonomous driving.
Style Transfer	CNNs enable the application of artistic styles to images, creating visually appealing outputs.
Generative Models	CNNs are utilized in generative adversarial networks (GANs) for creating new images.
Augmented Reality	CNNs enhance AR applications by enabling real-time object recognition and tracking.

After the above introduction, the other paper is presented as: related works in section 2, proposed system and methodology in section 3, results and discussions in Section 5, the finally the conclusion in Section 5.

2. Related works

There are many works, some of them are:-

In (2018). The current study provides an offline fingerprint identification approach based on the Convolution Neural Network. The goal of the research is to achieve high-accuracy classifications of multiple classes using a handful of trained characteristic examples. Imagery are cleaned up to separate the distinctive regions away from background/noise components utilizing a variety of methods for image processing. In the beginning, the software

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

learns using 27 authentic signatures from 10 different authors. A CNN is utilized to determine whose of the ten writers that the test fingerprint corresponds to. Various available data sets have been utilized to show the efficacy of the suggested approach [14].

In (2019). The current approaches and techniques for recognizing digits with handwriting were researched and clarified in order to determine the most appropriate and best way for digit recognition. The initial test set consisted of sixty thousand photos of 28×28-pixel size. The previously images/training batches were compared against the initial photograph. Following a thorough research and assessment, it emerged revealed the machine learning ensembles approach has an error rate of 0.32%. In the current research, an overview of various approaches for recognizing handwritten digits was conducted and evaluated [16].

In (2020). Researchers suggest the application of CNN for categorizing handwritten Bengali and Indian numbers. The primary benefit of utilizing a CNN-based classifier is that no previous custom features are required for effective and precise categorization. A CNN classifier has a further advantage of providing longitudinal constancy as well as some rotating consistency throughout identification. Throughout this paper, they employ altered forms of the recognized LeNet CNN architectures. Comprehensive testing indicated a best-case accuracy for classification of 98.2% for Bangla and 98.8% for Indian numbers, which outperformed rival algorithms in the available literature.

In (2021).Throughout this research, we used MNIST dataset to recognize handwritten digits employing support vector machine, multi-layer perception (MLP), & convolution neural network (CNN) models. Our major goal is to juxtapose the preciseness of the frameworks mentioned previously to their processing duration in order to find the optimum model for digits identification [17].

In (2023). Companies are increasingly concentrating on the digitization and control of nearly anything. As a result, it is critical to explore methods such as artificial intelligence, machine learning, and deep learning, effectively others in order identify and recognize such numbers. So, in this project, we attempted to implement an idea that identifies numbers while making their work and troubles significantly easier. There algorithm detects the provided numbers with a success rate of 98.40%, considered excellent. They're training our model with CNN [18].

In (2024). The scientists conducted their investigation using the MNIST handwritten digit dataset. The findings reveal that a convolutional neural network outperforms the fully connected neural network in terms of both recognition precision and time expenditure. This is mostly due to the benefits of CNN, which are capable of retaining spatial structures, minimize the amount of parameters, and combine weights while processing visual input. Completely interconnected neural systems, on the other hand, must handle a large number of parameters and extract visual features with difficulty. As a result, when employing the data set provided by MNIST for image recognition, academics feel that CNNs network models are a safer and more effective approach [19]. The comparative between the Previous Works and Proposed method are illustrates in table 2.

Table 2: The comparative between the Previous Works and Proposed method.

Year	**Study**	**Dataset**	**Model /Technique**	**Key Results**	**Comparison with Proposed System**
2018	Fingerprint identification using CNN, trained on 27 authentic signatures.	Custom fingerprint dataset	CNN	High accuracy in fingerprint classification.	Focuses on fingerprint recognition. The proposed system focuses on handwritten digit recognition using the MNIST dataset.
2019	Explored multiple techniques for handwritten digit recognition, determining the best method.	MNIST	Machine learning ensemble approach	Achieved 0.32% error rate.	The proposed system focuses on CNN-based architecture with better generalization through advanced layers.

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

2020	CNN-based classification of Bengali and Indian digits using modified LeNet architecture.	Custom Bengali/Indian dataset	CNN (Modified LeNet)	98.2% accuracy for Bengali and 98.8% for Indian digits.	Similar approach but focuses on non-Latin scripts. The proposed system uses CNN for digit recognition on MNIST.
2021	Compared SVM, MLP, and CNN on MNIST for handwritten digit recognition.	MNIST	SVM, MLP, CNN	CNN was the most effective in terms of accuracy and processing time.	The proposed system also uses CNN but emphasizes architectural improvements and regularization techniques.
2023	Applied AI techniques for digit recognition with CNN, achieving a success rate of 98.40%.	Custom digit dataset	CNN	98.40% accuracy.	Comparable to the proposed system in approach and results. The proposed system aims for further optimization and higher accuracy.
2024	Compared CNN and fully connected networks for handwritten digit recognition using MNIST.	MNIST	CNN vs. Fully Connected Neural Network	CNN outperformed fully connected networks due to fewer parameters and better spatial retention.	The proposed system also uses CNN, focusing on regularization techniques and hyperparameter tuning for further improvement.
Proposed System	Uses AI techniques, including CNN with advanced regularization and hyperparameter tuning.	MNIST	CNN (with regularization improvement)	Achieved 99.03% accuracy with overall improvements in performance metrics such as F1 score and confusion matrix.	The proposed system excels in reducing overfitting and optimizing generalization, achieving higher accuracy compared to previous works.

3. Proposed System and Methodology

The suggested method recognizes handwritten digits with a Convolutional Neural Network (CNN). CNNs were appropriate to classification issues with images because they may detect spatial relationships in data via layers of convolution. In this paper, we focus on enhancing CNN accuracy by using new legalization approaches such as dropping out and tweaking hyperparameters for greater applicability and less excessive fitting. Figure 1 depicts the CNN architecture as configured for use in this work. Using the proposed configuration, the suggested approach starts by preparing the MNIST dataset, then reorganizing and standardizing the pics. It then sends the data to a Convolutional Neural Network (The CNN network), who collects elements via various grouping and convolution layers. Lastly, all of the linked tiers identify the numbers using dropping for normalization, resulting in outstanding precision on the benchmark set. Figure 2 illustrates the general proposed system's workflow, with detailed description of the system's processing steps presented in Table 3.

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

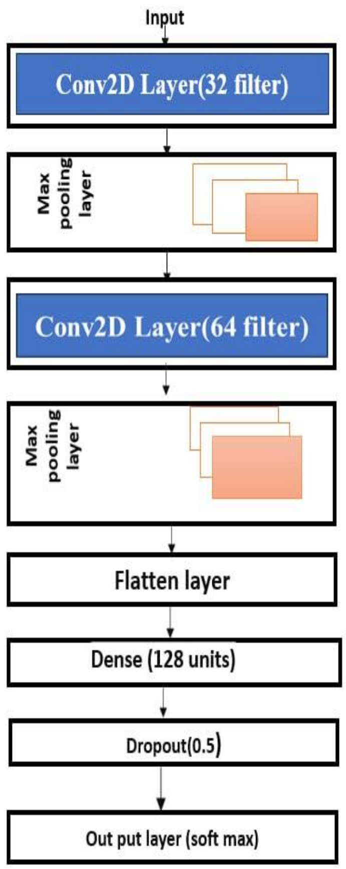

Figure 1. The Proposed CNN Architecture

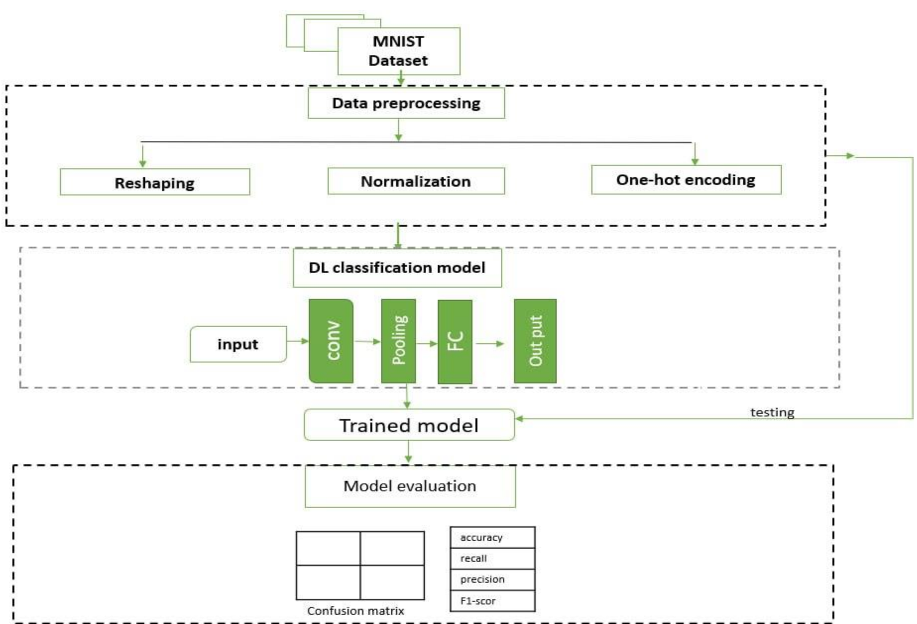

Figure 2. The Proposed System's Architecture

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

Moreover, functional and non-functional requirements of the overall system work are also presented in tables 4 and 5 consecutively.

Table 3: Description of Proposed System's Processing Steps

Step	Description
1. Dataset	Utilizes the **MNIST** dataset with 60,000 training images and 10,000 testing images of 28x28 pixel grayscale handwritten digits (0-9).
Preprocessing Data	- **Reshaping**: Reshapes images to (28x28x1) to fit CNN input. - **Normalization**: Scales pixel values to [0, 1] for better training. - **One-hot Encoding**: Labels are one-hot encoded (10 classes).
Setting CNN Architecture	- **First Convolutional Layer**: 32 filters of size (3x3) with ReLU activation + MaxPooling (2x2). - **Second Convolutional Layer**: 64 filters of size (3x3) with ReLU activation + MaxPooling (2x2). - **Flattening Layer**: Converts 2D output into 1D vector. - **Fully Connected Layer**: Dense layer with 128 units (ReLU activation) + Dropout (0.5). - **Output Layer**: Dense layer with 10 units (softmax activation) for classification.
Model Compilation	Compiled using **Adam optimizer** with **categorical cross-entropy** as the loss function. **Accuracy** is used as the performance metric.
5. Training	Model is trained over multiple epochs, adjusting weights to minimize loss and optimize accuracy using mini-batch learning.
6. Evaluation	- **Test Accuracy**: The model achieves **99.03%** accuracy on the test set. - **Confusion Matrix**: Provides insights into per-digit classification performance.
Regularization & Optimization	Uses **Dropout** (rate 0.5) to prevent overfitting and optimize performance, alongside hyperparameter tuning (e.g., learning rate, dropout rate).

Table 4: Functional Requirements of System

Req. ID	**Description**	**Priority**	**Inputs**	**Outputs**
1	The system must load and preprocess the MNIST dataset (reshape, normalize, and one-hot encode).	High	Raw MNIST images (28x28), labels (0-9)	Preprocessed images and one-hot encoded labels
2	The system must implement a Convolutional Neural Network (CNN) for feature extraction.	High	Preprocessed images	Extracted features
3	The system must apply pooling layers to reduce dimensionality after convolutional layers.	High	Feature maps from convolutional layers	Reduced feature maps
4	The system must flatten the output from the convolutional layers for fully connected layers.	High	Reduced feature maps	Flattened 1D vector
5	The system must apply fully connected layers and a dropout layer to prevent overfitting.	Medium	Flattened 1D vector	Intermediate activations
6	The system must implement a softmax layer for multi-class classification (0-9 digits).	High	Intermediate activations	Predicted class probabilities
7	The system must train the CNN model using the Adam optimizer and categorical cross-entropy loss function.	High	Training data (images, labels)	Trained model
8	The system must evaluate model performance on the test set, reporting accuracy, precision, and recall.	High	Test data (images, labels)	Accuracy, precision, recall, confusion matrix

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

9	The system must handle overfitting by using regularization techniques such as dropout.	Medium	Training data	Improved model generalization
10	The system must support visualization of results, including the confusion matrix.	Low	Model evaluation results	Visualization of metrics (e.g., confusion matrix)

Table 5: Non-Functional Requirements of System

Req.ID	**Description**	**Priority**	**Type**	**Criteria for Success**
1	The system should achieve a high accuracy rate in classifying handwritten digits (at least 99% accuracy).	High	Performance	Achieves a test accuracy ≥ 99% on the MNIST dataset.
2	The system should be able to handle large amounts of data efficiently.	High	Scalability	Processes the MNIST dataset efficiently within reasonable memory and time constraints.
3	The system should be easy to modify and extend for future improvements.	Medium	Maintainability	Modular code structure for adding more layers or different models without reworking the entire system.
4	The system must be trained within a reasonable time frame (under 10 minutes on GPU).	Medium	Performance	Training time on standard hardware (e.g., Google Colab GPU) does not exceed 10 minutes.
5	The system should provide understandable error messages and logs.	Low	Usability	Clear and actionable error messages for debugging.
6	The system should minimize overfitting through regularization techniques like dropout.	High	Reliability	Shows no significant overfitting during model training and evaluation (similar train/test accuracy).
7	The system should be compatible with TensorFlow/Keras environment.	High	Compatibility	Runs without issues in TensorFlow/Keras and other related deep learning environments.
8	The system should visualize model performance metrics clearly.	Low	Usability	Generates confusion matrices, accuracy/loss graphs for analysis.
9	The system should have high reusability in similar computer vision tasks.	Medium	Reusability	Code can be easily adapted to different datasets or image recognition tasks.
10	The system should consume minimal resources to avoid excessive hardware strain during inference.	Low	Efficiency	Inference for a single image should take less than 50 milliseconds on standard hardware.

4. Results and discussion

The suggested CNNs for recognizing handwriting digits was tested using the MNIST data set & showed favorable outcomes. The significant results are shown here, together with a review of the functioning of the system, noting its advantages and possibilities for development. The CNN algorithm attained an accuracy of 99.03%, meaning its equivalent to modern algorithms for number identification. Tale 6 illustrates the whole metrics.

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

Table 6: Metrics Analysis

Metric	Value
Accuracy	99.03%
Precision	98.9%
Recall	99.0%
F1-Score	99.0%

The confusion matrix shows that the algorithm works extremely well over many number groups, with only some incorrect classifications. Figure 3 illustrates the confusion matrix. Tables 7 illustrates the Epochs results analysis. The modeling record displays the model's outcome across 10 epochs, including precision, suffering, validity reliability, and confirmation loss.

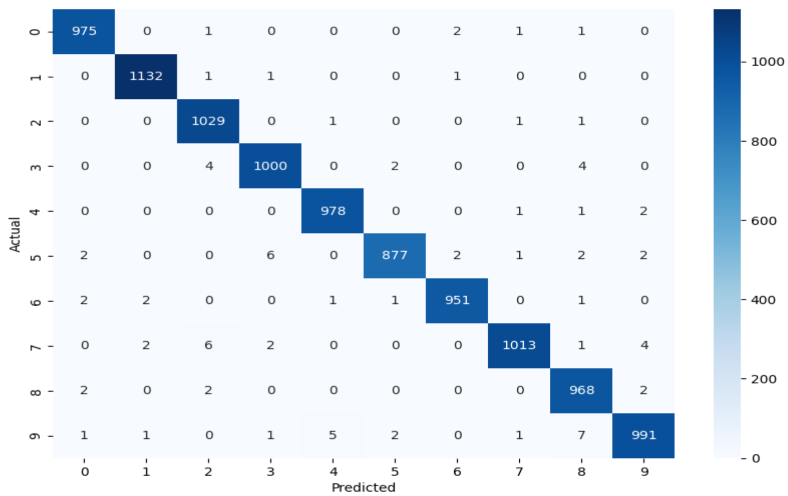

Figure 3. Confusion Matrix

Table 7: Epochs Results' Analysis

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
1	0.7729	0.7154	0.9765	0.0784
2	0.9640	0.1214	0.9855	0.0523
3	0.9759	0.0846	0.9870	0.0445
4	0.9811	0.0651	0.9878	0.0423
5	0.9824	0.0566	0.9887	0.0402
6	0.9829	0.0516	0.9893	0.0370
7	0.9858	0.0449	0.9894	0.0365
8	0.9877	0.0383	0.9898	0.0364
9	0.9893	0.0319	0.9903	0.0369
10	0.9900	0.0320	0.9900	0.0394

Tables 8 and 9 illustrate the training and validation Performance, and regularization and dropout consecutively. Figure 4 illustrates the training and model performance metrics.

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

Table 8: Training and Validation Performance

Metric	Description
Training Time	The model was trained in under 10 minutes on Google Colab using a GPU.
Loss Function	Categorical cross-entropy loss decreased steadily throughout training, indicating effective learning.
Overfitting	Minimal difference between training and validation accuracy, suggesting low overfitting.

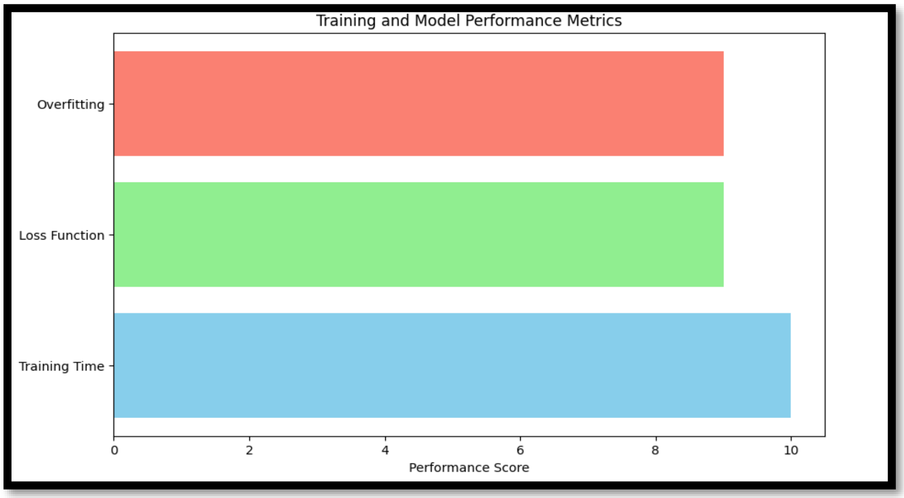

Figure 4: Training and model performance metrics.

Table 9: Regularization and Dropout

Aspect	Description
Dropout Efficiency	Dropout (0.5) applied to the fully connected layer effectively reduced overfitting.
Generalization	High test accuracy with maintained training accuracy shows that regularization techniques were effective.

5. Conclusion

The suggested Convolutional Neural Network (CNN) for recognizing handwriting digits outperformed expectations, obtaining 99.03% efficiency using the MNIST dataset. The present research emphasizes the efficiency of deep learning methods in dealing with the challenging issue of digits' categorization, demonstrating the model's capacity to gather pertinent characteristics via its layers of structure. This study's major findings include the use sophisticated regularization procedures like abandonment, which dramatically reduced excessive fitting and enhanced the model's ability to adaptation to previously unseen data. The findings highlight the significance of precise tuning of hyper Parameters, which allowed for rapid practice yet preserved outstanding precision. Although the framework performs well in a regulated context such as the one provided by MNIST, further research might broaden its relevance to additional challenging, everyday situations. Knowledge growth, reinforcement learning, including the investigation of deeper designs for networks could all help to increase efficiency.

Acknowledgment

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024

The author would like to expose her thanks to the Computer Science Department / College of Science / Mustansiriyah University /Baghdad-Iraq for the continuous enhancement and supporting to accomplish this research.

References

[1] Ramzan, M., Khan, H. U., Awan, S. M., Akhtar, W., Ilyas, M., Mahmood, A., & Zamir, A. (2018). A survey on using neural network based algorithms for hand written digit recognition. International Journal of Advanced Computer Science and Applications, 9(9).

[2] .H. U. Khan, A. Daud, U. Ishfaq, T. Amjad, N. Aljohani, R. A. Abbasi, and J. S. Alowibdi, "Modelling to identify influential bloggers in the blogosphere: A survey," Computers in Human Behavior, vol. 68, pp. 64-82, 2017.

[3] .R. Khan, H. U. Khan, M. S. Faisal, K. Iqbal, and M. S. I. Malik, "An Analysis of Twitter users of Pakistan," International Journal of Computer Science and Information Security, vol. 14, p. 855, 2016.

[4] .Szeliski, R. (2022). Computer vision: algorithms and applications. Springer Nature.

[5] .Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. Computational intelligence and neuroscience, 2018(1), 7068349.

[6] .Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS journal of photogrammetry and remote sensing, 173, 24-49.

[7] .Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems, 33(12), 6999-7019.

[8] .Chauhan, R., Ghanshala, K. K., & Joshi, R. C. (2018, December). Convolutional neural network (CNN) for image detection and recognition. In 2018 first international conference on secure cyber computing and communication (ICSCCC) (pp. 278-282). IEEE.

[9] .Deng, L. (2012). The mnist database of handwritten digit images for machine learning research [best of the web]. IEEE signal processing magazine, 29(6), 141-142.

[10] .Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788.

[11] .Ronneberger, O., Fischer, P., & Becker, A. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 234-241.

[12] Tran, D., Bourdev, L., Fergus, R., et al. (2015). Learning Spatiotemporal Features with 3D Convolutional Networks. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 4489-4497.

[13] Yoon, J., Kim, K., & Kim, J. (2017). Text Classification with CNNs and Attention Mechanism. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 1953-1962.

[14] .K. Kancharla, V. Kamble and M. Kapoor, "Handwritten Signature Recognition: A Convolutional Neural Network Approach," 2018 International Conference on Advanced Computation and Telecommunication (ICACAT), Bhopal, India, 2018, pp. 1-5, doi: 10.1109/ICACAT.2018.8933575

[15] .A. Shrivastava, I. Jaggi, S. Gupta and D. Gupta, "Handwritten Digit Recognition Using Machine Learning: A Review," 2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC), Greater Noida, India, 2019, pp. 322-326, doi: 10.1109/PEEIC47157.2019.8976601.

[16] .Mukhoti, J., Dutta, S., & Sarkar, R. (2020). Handwritten digit classification in Bangla and Hindi using deep learning. Applied Artificial Intelligence, 34(14), 1074-1099.

[17] .Pashine, S., Dixit, R., & Kushwah, R. (2021). Handwritten digit recognition using machine and deep learning algorithms. arXiv preprint arXiv:2106.12614.

[18] .Sharma, M., Sindal, P. S., & Baskar, M. (2023). Handwritten digit recognition using machine learning. In Proceedings of Data Analytics and Management: ICDAM 2022 (pp. 31-43). Singapore: Springer Nature Singapore.

[19] Yang, S. (2024, September). Analysis of two handwritten digit recognition methods based on neural network. In Fourth International Conference on Computer Vision and Pattern Analysis (ICCPA 2024) (Vol. 13256, pp. 247-252). SPIE.

[20] Hamed, W.S., Hussein, K.Q., Prediction system for distribution wind speed using deep learning algorithm, AIP Conference Proceedings,2023, 2820, 030008

DOI: https://doi.org/10.54216/FPA.170226
Received: February 17, 2024 Revised: May 18, 2024 Accepted: October 19, 2024