Farsi Digit Recognition Using GAN-Generated Data and Convolutional Neural Networks

Farah Jawad Al-Ghanim
Department of Computer Science, College of Computer Science and Information Technology
University of Al-Qadisiyah
Al-Diwaniyah, Iraq
farah.jawad@qu.edu.iq
Orcid.org/0009-0004-6754-662X

Nisreen Ryadh Hamza
Department of Computer Science, College of Computer Science and Information Technology
University of Al-Qadisiyah
Al-Diwaniyah, Iraq
nesreen.readh@qu.edu.iq

DOI: http://dx.doi.org/10.31642/JoKMC/2018/120102
Received Jun. 29, 2025. Accepted for publication Mar.8, 2025

Abstract—*Handwritten digit recognition is one of the most active study areas in computer vision because to its numerous applications such as automatically identifying the digits in bank checks and car numbers. Handwritten Latin digits have been the subject of extensive research over the last three decades, whereas Persian handwritten digits have received far less attention. For this reason, we will concentrate on the problem of recognizing Persian (Farsi) handwritten numerals. The main challenge in the recognition of Persian handwritten digits is the presence of different patterns in Persian digit writing, which complicates the feature extraction process. An appropriate approach for automated feature extraction has been the focus of most earlier investigations since handcrafted feature extraction methods are complex and have unstable performance levels. This paper studies the use of a dataset of Persian handwritten digits generated by a Generative Adversarial Network (GAN) to develop a highly accurate Convolutional Neural Network (CNN) model for digit recognition. The proposed CNN architecture achieved a test accuracy of 99.7%, demonstrating its effectiveness. This study highlights the viability of GAN-generated datasets for machine learning applications, especially in resource-constrained scenarios.*

Keywords— Handwritten Digit Recognition; Convolutional Neural Network; Persian digits; CNN; Farsi.

I. INTRODUCTION

Attempting to make systems intelligent, machine learning has emerged as a new multidisciplinary area in recent years. Among the many applications of artificial intelligence, machine learning modifies and investigates techniques and algorithms that enable computers and systems to learn and train. Machine learning is the theory that a system can learn and recognize things much like a person.[1] Deep learning is one of the most used machine learning techniques. The basic idea of a perceptron introduced the concept of deep learning in the 1950s. In the 1980s, a multi-layer perceptron structure was then established. Nevertheless, before 2000, perceptrons' capacity for learning remained restricted. The algorithms then veered in the direction of deep learning and may catch scientists' interest. Machine learning includes the subfields of deep learning and hierarchical learning.[2]

In light of the great development witnessed by artificial intelligence and the great transformation witnessed by systems, it has become necessary to work on developing methods to distinguish between different data, and the use of artificial intelligence methods has become indispensable at the present time due to the scientific development witnessed in intelligence and has become included in all fields of science and even in areas of life where the challenges have become great and the methods are advanced. In the domains of optical character recognition (OCR) and machine vision, handwritten number recognition is a crucial research topic. There is no doubt that the use of modern technologies has contributed greatly to the detection and discrimination between different data sets, including handwritten Persian numbers. Given the beauty of the Persian script, the abundance of curves and lines in writing numbers, and the variation of handwriting, this has become a clear challenge. Because the use of handwritten Persian numbers requires special skills [3]

One kind of artificial neural network (ANN) that has acquired popularity in various computer vision applications is the convolutional neural network (CNN). Convolution, pooling, and fully connected are several layers of the construction blocks that CNN uses to automatically and adaptively learn spatial hierarchies of information via backpropagation. As it learns the features through its multiple neural layers (neurons), which contain many neurons that respond to a specific type of input rather than other inputs, it is a collection of interconnected neurons with numerical weights that are changed during the training process so that it responds when a pattern or image is entered and recognizes it. The initial patterns are found by the first layer, which is followed by the patterns of patterns by the second layer, the patterns of those patterns by the third layer, and so on.[4] A key task in pattern recognition and machine learning, handwritten digit recognition has applications in data digitalization, banking, and postal services. Large-scale datasets for Farsi handwritten digits are scarce, despite the general usage of datasets such as an MNIST_dataset. This paper addresses this gap by using a GAN-generated dataset to train a CNN model. The leveraging of artificial data to get the greatest accuracy for Farsi digit recognition is what makes this study innovative. The proposed CNN model's properties are also explained in detail in the study, along with how the presented approach achieves high performance.

The remainder of the study document is structured as follows: The literature review is presented in Section 2. The proposed CNN model, with its dataset, system design, and CNN model architecture, is in Section 3. Section 4 discusses the comparative study and the results of the proposed CNN model. Section 5 concludes the paper and highlights the topic for future study.

II. RELATED WORKS

A detailed overview of the relevant literature on the issue of handwritten digit recognition is in this section. The review covers a variety of techniques, such as those that use deep neural networks, classical method features, machine learning classifiers, transfer learning models, and combinations of all these.

In the classical methods, Sadri, Javad et al [5], achieved a 98.57% accuracy rate in digit identification by combining a multilayer perceptron neural network with the gradient histogram approach. In Parseh, Mohammad Javad et al [6], By using principle component analysis (PCA), the authors sought to decrease the dimension of the feature vector. An SVM was then used to classify the acquired feature vector, yielding a 99.07% classification accuracy. In Hajihashemi et al [7], Using the HODA dataset, a system based on Holography graph neurons achieved an 88.9% recognition rate using vector symbolic structures. In Sethy, Abhisek et al [8], presented a subsequential technique that integrated RF and LS-SVM classifiers, acquiring a 99% accuracy on an Odia dataset of hand-written.

In the deep learning methods (DL), This methods outperformed traditional methods, and indicated encouraging results.In Latif, Ghazanfar et al [9], proposed a DNN model for identifying handwritten characters in several Eastern languages, including Western Arabic, Urdu, Farsi, Devanagari, and the Eastern Arabic. In Siddique, Fathma et al [10], Study how different numbers of layers and epochs influence a CNN's model accuracy in classifying handwritten digits. for CNN's performance assessment used MNIST dataset to experiment.

In Akhlaghi, Maryam et al [11], Deep neural networks are used to propose a dependable technique for recognizing Farsi handwritten phone numbers. To identify a Farsi handwritten digit string, the suggested segmentation approach is used to first break the string down into single digits, and then a single Farsi handwritten digit recognition algorithm is used to classify each segment. The digit string of the Farsi phone number picture is ultimately produced by classifying. The suggested method has a 94.6% accuracy rate. In et al [12], CNNs and bagging weighted majority ensemble learning were combined to build the CBWME model. To detect Farsi handwritten digits, the VGG16, ResNet18, and Xception CNN models were employed as basis classifiers in this model. The bagging weighted majority ensemble learning was investigated for integrating the base classifiers' outputs. Three databases—HODA, CENPARMI, and IFHCDB—were used to evaluate the model's performance, and the outcomes were contrasted with those of the CNN models. The CBWME model was found to have the highest average recognition accuracy (97.65%) in the HODA dataset, followed by the Xception model (95.9%), ResNet18 model (93.75%), and VGG16 model (90.26%). The accuracy ordering were identical to those found in the CENPARMI and IFHCDB datasets.

In Miri, A., et al [13], In order to classify digits in the HODA dataset, a novel algorithm based on probabilistic neural networks (PNN) and MLP neural networks is proposed. This paper presents a novel CNN architecture that incorporates dilated convolutions; the proposed model efficiently expands receptive fields without increasing model depth, resulting in improved accuracy and decreased computational overhead; additionally, a combined loss function is used to improve overall performance. In Fateh, Amirreza et al [14], This study uses the Multi-Resolution Attention (MRA) module and transfer learning to propose a unique framework for handwritten number recognition in 12 languages. It delivers excellent precision, outperforming earlier methods with an accuracy boost of up to 2% for several languages. Chinese handwritten digits with the highest precision. The technique uses MRA-enhanced UNet for language identification, image enhancement, and transfer learning-based language-specific digit recognition. A dataset of 100,000 digitally created pictures was employed, and models were adjusted for multilingual recognition. In Noori, et al [15], In contrast to existing models, this research proposes a unique convolutional neural network (CNN) architecture for Persian digit identification that is shallower. reduced the number of trainable parameters by using the HODA dataset for training. We suggest a mixed loss function to increase accuracy and add dilated convolution layers to capture more features without raising parameters. The suggested network performs quicker and with more precision.

In contrast to previous studies that relied on classical feature extraction or machine and deep learning models, the proposed CNN model balances accuracy and computational cost.

III. METHODOLOGY

The proposed system combines a strong CNN architecture with GAN-generated data to recognize Persian handwritten digits with high accuracy. Preparing the dataset, preprocessing it, designing the CNN model, training it, and evaluating its performance are the key phases. The phases used in this study are illustrated in Figure 1 below.

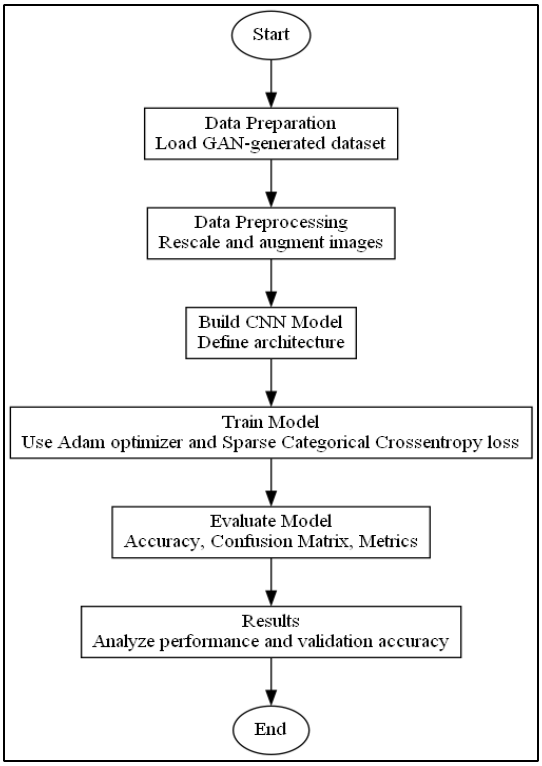

Fig. 1. The proposed system diagram

A. Dataset Description

The study's dataset consists of 150,000 Persian handwritten digit images.[16] the images in this dataset are generated using state-of-the-art Generative Adversarial Networks GANs.[17] There are 15,000 pictures in each class (digits 0-9) as Figure 2, to provide a balanced representation. The images are suitable for training convolutional neural networks due to their resolution (28x28). GANs ensure variety and realism in the dataset because they mimic handwritten digits in the real world. Additionally, the homogeneity of the dataset allows the model to learn digit attributes without bias, providing a controlled environment for training and testing.

This dataset was downloaded from the website Kaggle (www.kaggle.com). This site is a leading platform for data science and machine learning, offering access to high-quality datasets, cloud-based computing resources, and collaborative tools. Kaggle’s machine learning competitions foster innovation, while its datasets serve as benchmarks for AI research. For projects like signature verification, iris/fingerprint recognition, and disease classification, Kaggle provides invaluable resources to improve models and validate findings.

Fig. 2. illustrates an example of the appearance of a (0-9) digits generated by GANs.

B. The Proposed Model Architecture

The CNN architecture has been suitably built to reproduce the specific features of handwritten digits. This includes several layers:

1- Input: This layer for processing the input images and uses 3 channels.

2- Convolutional: Uses three Conv2D layers with a different filter size and uses ReLU activation.

3- Flattening: This layer constructs a 1D-feature vector from the feature maps.

4- Fully Connected: Finally, used Dense layer with ReLU activation (for feature extraction), This layer assists the model in learning more complex data representations. Then the output layer with softmax activation (for classification).

5- The Training Model: used Sparse Categorical Crossentropy as Loss Function. In the optimizer step, the Adam optimizer is used which dynamically modifies learning rates to reach quicker convergence. The dataset is split into a training set (70%), a validation set (20%), and a test set (10%). Finally, trained up to 20 epochs. 32 batch size.

CNNs do well in hierarchical feature extraction. The first convolutional layers of the model recognized simple edges and textures. Intermediate layers captured patterns peculiar and distinct to Farsi digits, such as characteristic loops or strokes. To classify each digit, fully connected layers combined these features. Table (1) displays the summary proposed CNN Architecture:

TABLE I. THE PRPOSED MODEL ARCHITECTURE SUMMARY TABLE

Layer Type	Parameter	Shape	The Goal
Rescaling	None	(28, 28, 3)	Normalize: pixel values to [0, 1]
Conv2D (32 filters)	Kernel: 3x3 ReLU	(26, 26, 32)	Extract features
MaxPooling 2D	Pool size: 2x2	(13, 13, 32)	Down- sample spatial dimensions
Conv2D (64 filters)	Kernel: 3x3 ReLU	(11, 11, 64)	Learn more complex features
MaxPooling 2D	Pool size: 2x2	(5, 5, 64)	Further reduce spatial dimensions
Conv2D (64 filters)	Kernel: 3x3 ReLU	(3, 3, 64)	Learn high- level abstract features
Flatten	None	(576,)	Flatten the feature maps
Dense	Units: 128 ReLU	(128,)	Combine extracted features
Dense	Units: 10 Softmax	(10,)	Output: probabilities for each digit class

IV. RESULTS AND DISCUSSION

The results indicate the advantage of using the GAN-generated datasets and the efficacy of the proposed CNN model. However, there are issues with generalization when artificial data is used. Validating the model's applicability needs testing on a dataset of Farsi handwritten digits from the actual world. Also, in situations when real-world data is limited, the study demonstrates how the GANs may produce a variety of high-quality datasets for training machine learning models.

The proposed CNN model has a test accuracy of 99.7%. All classes had nearly ideal accuracy, recall, and F1-scores, according to the classification report. A few misclassifications are seen in the confusion matrix, especially between visually similar digits. Consistent performance across every digit class is one of the main study conclusions. The majority of mistakes happen between digits that have similar forms, such as "2", "3", and "6" The matrix analysis supports the CNN model's good - generalization on the validation set.

Model performance was evaluated using four evaluation Metrics. The accuracy for the percentage of correctly classified samples. Precision, Recall, F1-Score to assess per-class performance. Finally, Figure 3 show the Confusion Matrix is used to visualize misclassification patterns and identify challenging digit pairs. Below is an illustration of these metrics:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP + FN}} \qquad (2)$$

$$\text{F} - \text{score} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \qquad (3)$$

$$\text{Accuracy} = \frac{\text{TP + TN}}{\text{TP + TN + FP + FN}} \qquad (4)$$

The entire result that obtained by using the suggested model is shown in Table (2).

TABLE II. CLASSIFICATION REPORT

Classes	Precision	Recall	F1-score
class_ 0	1.00	1.00	1.00
class_ 1	0.99	1.00	1.00
class_ 2	0.98	1.00	0.99
class_ 3	1.00	0.99	0.99
class_ 4	1.00	1.00	1.00
class_ 5	1.00	1.00	1.00
class_ 6	1.00	1.00	1.00
class_ 7	1.00	1.00	1.00
class_ 8	1.00	0.99	1.00
class_ 9	1.00	1.00	1.00

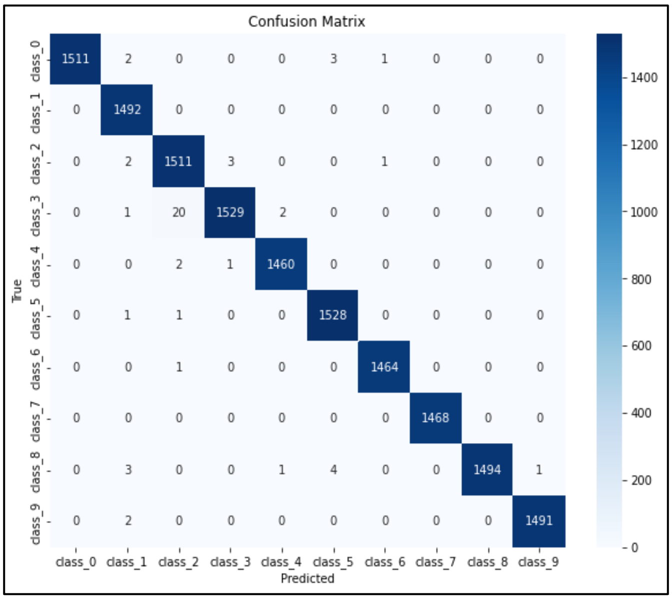

Fig. 3. Confusion matrix for GANs-dataset

The curves for accuracy and loss are depicted in the two Figures (4 and 5), which present information about how the proposed model learns.

The accuracy gradually enhances as the numeral of epochs increases as in Figure 4, displaying the learning process. The effectual generalization of the proposed CNN models to unseen data and the effectiveness in recognizing Farsi-handwritten digits are shown by the close alignment of the accuracy training and validation.

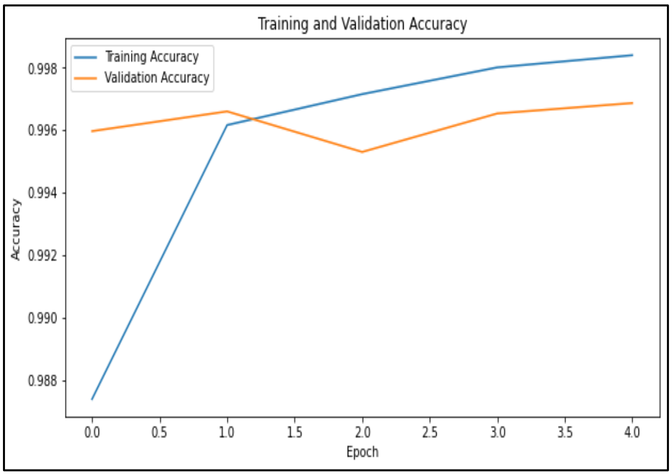

Fig. 4. Training and Validation Accuracy for GANs-dataset

Training loss decreases in the early epochs, indicating the assimilation of the proposed CNN model for the data and improvement in predictions. When the validation loss stays constant at a certain level, the model has reached a stable learning stage. The small difference between the two losses assures that overfitting has not taken place, strengthening the model's generalization strength. As shown in Figure 5.

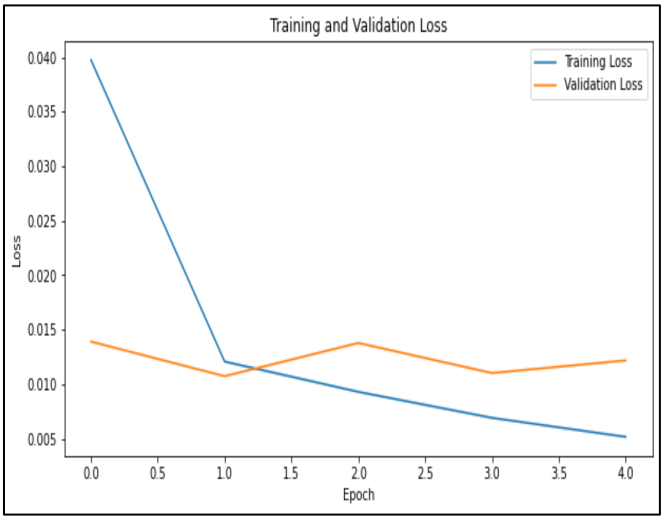

Fig. 5. Training and Validation Loss for GANs-dataset

These Figures indicate that the CNN model is properly trained, minimizes loss, and continuously improves accuracy. The lack of a significant difference between training and validation accuracy/loss validates the model's robustness and suitability for real-world applications.

Through random image testing, the model's dependability in correctly identifying Farsi digits was proven. The prediction model is illustrated in Figure 6.

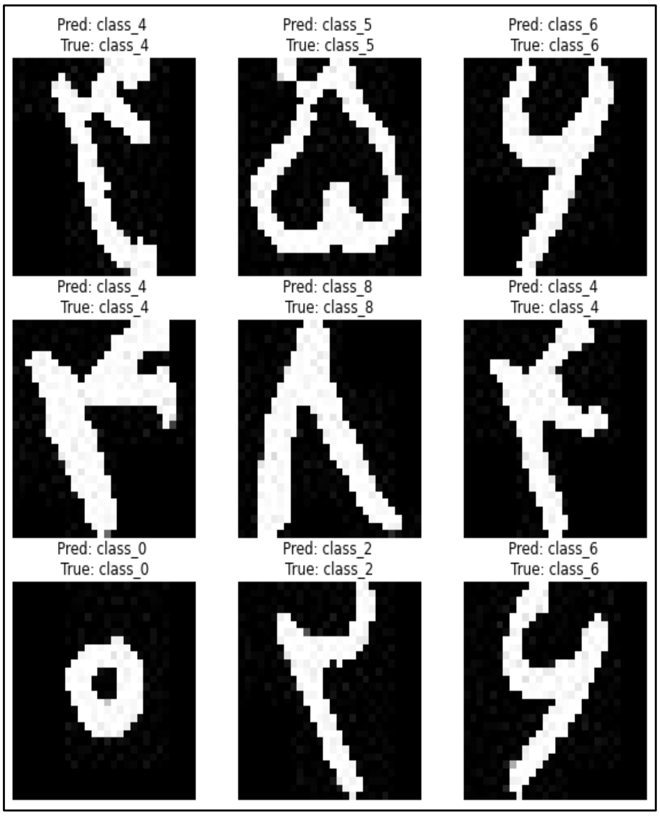

Fig. 6. Some Test Image (Predicted and True Labels)

To compare with the earlier works in terms of using this dataset. Up to the conclusion of this research study, no other published studies have used this dataset, which was obtained from the Kaggle website and was recently released at the end of 2023.

To prove the accuracy and efficiency of the proposed system, we trained the proposed model on well-known real dataset, which is (HODA Farsi Digit Dataset). It is the first database of handwritten Farsi digits was created as part of the Recognizing Farsi Digits and Characters in SANJESH Registration Forms MSc project at Tarbiat Modarres University. Completed in 2005, this project has been finished in collaboration with Hoda System Corporation. The Figure 7, show sample digit in this dataset.

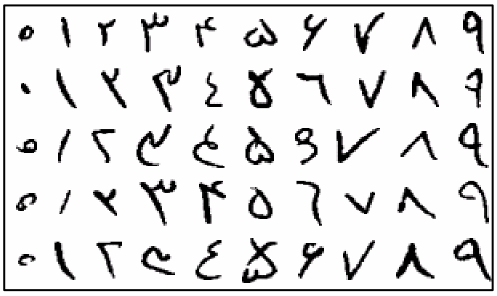

Fig. 7. Samples of different handwriting styles [18]

The dataset's samples are handwritten characters taken from over 12,000 Iranian university entrance examination registration forms. There are 102,352 total picture samples in the dataset, which are divided into three sets: 60,000 training samples, 20,000 test samples, and 22,352 remaining samples. [19] In this study used the remaining samples as validation set. the model achieved an overall test accuracy of 99.34%, confirming its strong generalization capability when dealing with unseen data. Table (3) show entire result as:

TABLE III. CLASSIFICATION REPORT FOR HODA DATASET

Classes	Precision	Recall	F1-score
class_ 0	1.00	0.99	0.99
class_ 1	1.00	1.00	1.00
class_ 2	0.98	1.00	0.99
class_ 3	0.99	0.98	0.99
class_ 4	0.99	0.99	0.99
class_ 5	0.99	1.00	0.99
class_ 6	0.99	1.00	0.99
class_ 7	1.00	1.00	1.00
class_ 8	1.00	1.00	1.00
class_ 9	1.00	0.99	0.99

The classification report illustrates that the CNN performs consistently across every digit and is not affected by the imbalance class.

The confusion matrix proves the model can distinguish between various classes and has strong performance in Farsi handwritten digit recognition. Figure 8 illustrates the confusion matrix:

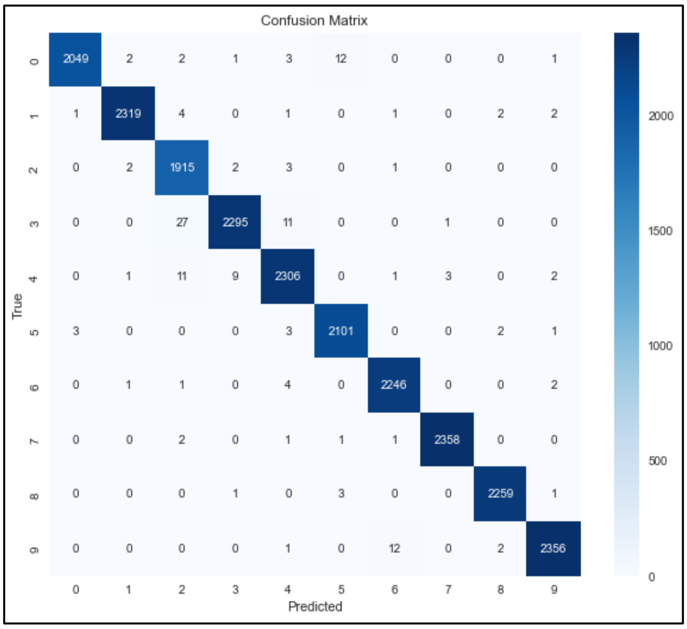

Fig. 8. confusion matrix of hoda dataset

The accuracy curve shows a steady increase and a gradual improvement in both training and validation accuracy over the epochs without signs of overfitting, indicating that the model generalizes well to unseen data.

The loss decreased steadily, the final validation loss is still low and aligns closely with the training loss, the model stays good-regularized and does not suffer from overfitting.

The following two figures (9 and 10) reflect stable learning behavior.

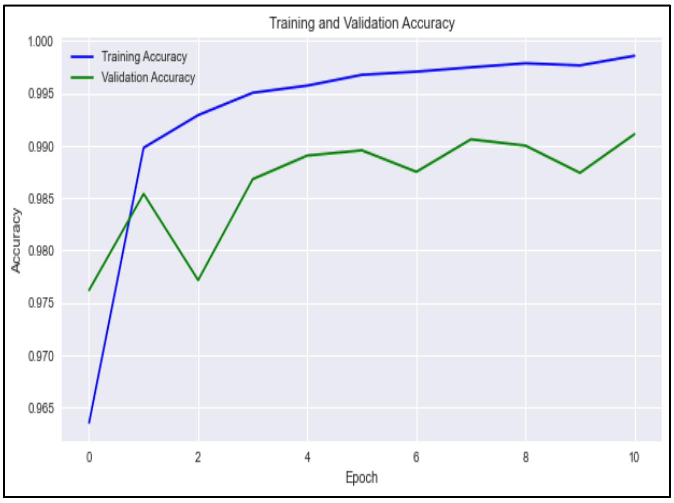

Fig. 9. Training and Validation Accuracy of huda dataet

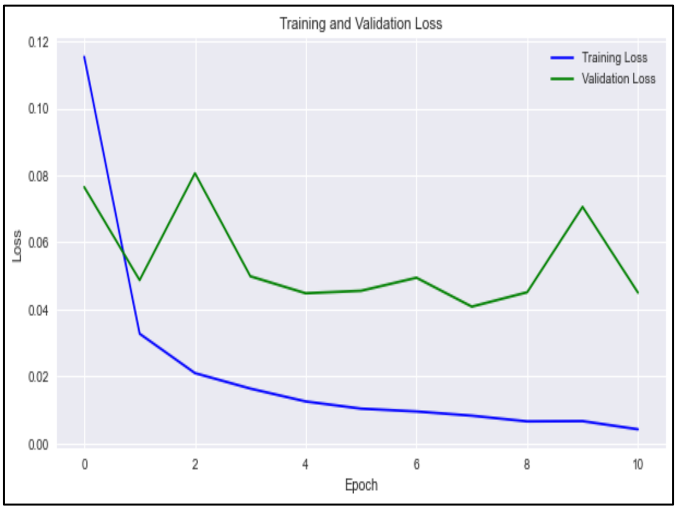

Fig. 10. Training and Validation Loss of hoda dataset

The model effectively learns the characteristics of the handwritten Persian digits while retaining good generalization skills, as demonstrated by the training and validation curves.

Finally, with all of the above and according to the final results, the model is accurate, which makes it appropriate for real-world applications such as handwritten document digit recognition.

V. Conclusion

Farsi (Persian) handwritten digit recognition is a branch of pattern recognition. The most significant obstacle in the recognition of Persian handwritten digits is the presence of different patterns in Persian digit writing, which complicates the feature extraction process. Using both visual aids and life experience, the human brain can simultaneously recognize the written digits. As a result, more effective systems are created for pattern recognition or handwritten digit identification employing human learning models. This paper shows that deep learning models can be trained in resource-constrained environments utilizing datasets produced by the GANs model. The proposed CNN model in this paper recognized Persian handwritten digits with exceptional precision. Future research will concentrate on using advanced regularization techniques, expanding the methodology to more scripts or languages, and using different real-world datasets.

References

[1] Sun, Tianming, Bin Feng, Jinpeng Huo, Yu Xiao, Wengan Wang, Jin Peng, Zehua Li et al. "Artificial intelligence meets flexible sensors: emerging smart flexible sensing systems driven by machine learning and artificial synapses." Nano-Micro Letters 16, no. 1 (2024): 14.

[2] Schmidhuber, Jürgen. "Deep learning in neural networks: An overview." Neural networks 61 (2015): 85-117.

[3] J. M. Alghazo, G. Latif, A. Elhassan, L. Alzubaidi, A. Al-Hmouz, and R. Al-Hmouz, “An online numeral recognition system using improved structural features–a unified method for handwritten Arabic and Persian numerals,” Journal of Telecommunication, Electronic and Computer Engineering (JTEC), vol. 9, no. 2–10, pp. 33–40, 2017.

[4] Yamashita, Rikiya, et al. "Convolutional neural networks: an overview and application in radiology." Insights into imaging 9 (2018): 611-629

[5] Sadri, Javad, Mohammad Reza Yeganehzad, and Javad Saghi. "A novel comprehensive database for offline Persian handwriting recognition." Pattern Recognition 60 (2016): 378-393.

[6] Parseh, Mohammad Javad, and Mojtaba Meftahi. "A new combined feature extraction method for Persian handwritten digit recognition." International Journal of Image and Graphics 17.02 (2017): 1750012.

[7] Hajihashemi, Vahid, et al. "A pattern recognition based Holographic Graph Neuron for Persian alphabet recognition." 2020 International conference on machine vision and image processing (MVIP). IEEE, 2020.

[8] Sethy, Abhisek, Prashanta Kumar Patra, and Soumya Ranjan Nayak. "A hybrid system for handwritten character recognition with high robustness." Traitement du Signal 39.2 (2022): 567.

[9] Latif, Ghazanfar, et al. "Deep convolutional neural network for recognition of unified multi-language handwritten numerals." 2018 IEEE 2nd International workshop on Arabic and derived script analysis and recognition (ASAR). IEEE, 2018.

[10] Siddique, Fathma, Shadman Sakib, and Md Abu Bakr Siddique. "Recognition of handwritten digit using convolutional neural network in python with tensorflow and comparison of performance for various hidden layers." 2019 5th international conference on advances in electrical engineering (ICAEE). IEEE, 2019.

[11] Akhlaghi, Maryam, and Vahid Ghods. "Farsi handwritten phone number recognition using deep learning." SN Applied Sciences 2.3 (2020): 408.

[12] Nanehkaran, Yaser Ahangari, et al. "A pragmatic convolutional bagging ensemble learning for recognition of Farsi handwritten digits." The Journal of Supercomputing 77.11 (2021): 13474-13493.

[13] Miri, A., and M. Khedmati. "DEVELOPMENT OF AN ALGORITHM FOR PERSIAN HANDWRITING DIGITS RECOGNITION BASED ON MLP AND PNN CLASSIFIERS AND USING CLUSTER CENTERS." Sharif Journal of Industrial Engineering & Management 39.1 (2023): 23-34.

[14] Fateh, Amirreza, et al. "Advancing Multilingual Handwritten Numeral Recognition With Attention-Driven Transfer Learning." IEEE Access 12 (2024): 41381-41395.

[15] Noori, Hossein. "Improving Persian Handwritten Digit Recognition using Convolutional Neural Network." Journal of Algorithms and Computation 56.1 (2024): 55-72.

[16] Amirmahdi Aboutalebi, 2024, Persian HandWritten Digits dataset, https://www.kaggle.com/datasets/amirmahdiabbootalebi/persian-handwritten-digits

[17] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative Adversarial Networks.

[18] https://farsiocr.ir/مجموعه-داده/مجموعه-ارقام-دستنویس-هدی/.

[19] [19] Khosravi, H., Kabir, E.: Introducing a very large dataset of handwritten Farsi digits and a study on their varieties. Pattern Recognit. Lett. 28(10), 1133–1141 (2007)