Journal of Intelligent Decision Making and Information Science

www.jidmis.org.org
eISSN: 3079-0875

Comparative Analysis of Neural Network Architectures for Image Classification

Muhammad Rizwan[1], Hamza Naveed[1], Rana Muhammad Zulqarnain[2,*]

[1] Department of Software Engineering, Faculty of Computer Science and Information Technology, University of Central Punjab, Lahore, Punjab, Pakistan. E-mail: muhammad.rizwan02@ucp.edu.pk; hamzanaveed255@gmail.com
[2] School of Mathematical Sciences, Zhejiang Normal University, Jinhua, Zhejiang, China. E-mail: ranazulqarnain7777@gmail.com

ARTICLE INFO

Article history:
Received 23 November 2024
Received in revised form 1 March 2025
Accepted 9 May 2025
Available online 14 May 2025

Keywords:
Neural Networks; MLP; CNN; Image Classification

ABSTRACT

An important challenge in the growing domains of computer vision and machine learning is the precise categorisation of images, which drives the growth of innovative methodologies and methods. The proposed study will examine three distinct picture classification frameworks. The structures include a Single Layer Network (SLN), a Multi-Layer Perceptron (MLP), and a Convolutional Neural Network (CNN) based on LeCun's technique. The study will employ the MNIST digit identification dataset, the Fashion-MNIST dataset, and the CIFAR-10 dataset. Precision, recall, and F1-score are the measures used for model evaluation. Compared to the SLN and CNN models, the MLP model exhibits superior precision and recall.

1. Introduction

Image detection is a significant difficulty that promotes the advancement of advanced techniques and models in data science and computer vision, fields that are perpetually evolving and remain subjects of ongoing research. Various methodological techniques are employed to analyse and evaluate the intricate graphical information used in this process. The practical implementation of this project distinguishes it from others. This research aims to examine and discriminate among three distinct approaches to photo categorisation. This discussion encompasses three types of structures: a Single Layer Network (SLN), a Multi-Layer Perceptron (MLP), and a Convolutional Neural Network (CNN) based on Le Cun's methodology. The MNIST digit identification dataset, the Fashion-MNIST dataset, and the CIFAR-10 dataset will be standards for the studies conducted throughout this project.

The MNIST dataset, consisting of greyscale images of written numbers, is frequently utilised as an essential illustration in neural networks for image classification. An alternate, albeit more complex, method for evaluating clothing and accessories involves utilising the Fashion-MNIST dataset. This dataset presents more details with data organised similarly. The CIFAR-10 dataset poses

[*] *Corresponding author.*
E-mail address: ranazulqarnain7777@gmail.com

https://doi.org/10.59543/jidmis.v2i.11481

a greater dispute, comprising colour images across ten different types, including autos and animals. This renders a task more arduous to achieve. A comprehensive grasp of the visual attributes depicted in the photographs is necessary to categorise the images. The selection of this information enables a thorough evaluation of the performance and robustness of the proposed categorisation methods. This step will enable an assessment to be conducted. One of the neural networks that is the most fundamental in terms of classifying issues is the single-layer network. The reason for this is that its structure is straightforward, and it has linear decision limits.

The Multi-Layer Perceptron is difficult to comprehend since it incorporates one or more different levels concealed from view. This not only makes it easier to recognize nonlinear correlations in the data, but it also has the potential to deliver a greater level of accuracy compared to more fundamental models. The groundbreaking research LeCun conducted in deep learning was the impetus for developing the Convolutional Neural Network. This network is designed to extract features and categorize data by utilizing spatial hierarchies in the data; this study inspired it. Although three other models have been established, this model is believed to be the most advanced and efficient one for processing complex visual data. To determine whether these models are appropriate for a variety of photo classification tasks, this research aims to investigate the performance of these models on three distinct datasets to identify their advantages and limits. This will determine whether these models are suitable for the tasks.

The outcomes of this study not only contribute to a deeper academic and practical comprehension of the challenges involved with the classification of images, but they also offer support in selecting appropriate models based on the specific characteristics of the data. This endeavour aims to elucidate the circumstances in which every model delivers the highest possible degree of performance, with the ultimate objective of enhancing the area of machine learning. In addition to being of tremendous help in directing future research and applications in the field of photo classification, this will also be of great use in other significant domains.

2. Literature review

Desai and Shah [1] highlighted the advantages of employing neural networks for breast cancer detection, such as reduced examination duration and improved accuracy relative to traditional techniques. Kurani et al. [2] presented a thorough review of the application of Support Vector Machines (SVM) and Artificial Neural Networks (ANN) for forecasting stock market trends. LeCun et al. [3] extensively contributed to image identification by focussing on the application of replication systems for handwritten digit identification tasks using little processing. The specialised network structure for this challenge produced significant results, achieving a 1% mistake ratio and a 9% rejection rate for written zip code digits. Findings indicate that large replication models are proficient in everyday image recognition, necessitating intricate preparation methods. Data sets that serve have significantly impacted the implementation of machine learning in detecting images, aiding in the design and assessment of algorithms. The MNIST database structure, established by Deng [4], is a prominent dataset fundamental to recognising handwriting digits. This allows scholars to explore various algorithms, promoting progress in this field. Denker et al. [5] emphasised the capability of neural networks to recognise handwritten digits, especially for the interpretation of zip code numerals. Krizhevsky et al. [6] broadened the domain of image classification tasks and developed intricate visual patterns; CIFAR-10 includes several classes, providing a solid foundation for investigating deeper convolutional neural network architectures. Xiao et al. [7] presented an updated MNIST database, a more intricate substitute for MNIST.

Ullah et al. [8] focused on arrhythmia identification in cardiac surveillance utilising sophisticated deep neural network methodology. Yang et al. [9] investigated permanent dwell-time-controlled devices with stochastic data expulsions, offering pragmatic ways to improve system resilience and reliability in real-life situations. Duan et al. [10] introduced an observer-based failure identification structure for continuous-discrete networks utilising the TS fuzzy method, representing a notable advancement in enhancing diagnostic accuracy in nonlinear systems. Diao and Zhang [11] presented a decision tree-based approach to improve managerial effectiveness, illustrating the capabilities of neural network algorithms in decision-making procedures. Yang et al. [12] devised a synchronous resilient controller process employing H∞ sliding mode management for uncertain leaping singularly perturbation structures. Xing et al. [13] examined stability and Hopf bifurcation in (n + m)-neuron double-ring networks characterised by numerous time lapses. Jia et al. [14] examined finite-time synchronisation in fractional-order delayed memristive neural networks. Yu et al. [15] introduced an adaptive finite-time control paradigm for these structures, tackling issues related to uncertain control gains and fulfilling predetermined performance targets. Zhang et al. [16] formulated an observer-based sliding mode control for fuzzy stochastic switching systems subjected to deceptive threats. Xu et al. [17] proposed an innovative optimal tracking control strategy for switched linear systems, utilising data-driven techniques to improve system efficiency independent of detailed system formulations.

3. Model description

We developed three architectural designs, which will be addressed individually in this section.

3.1. Single-layer neural network

The single-layer neural network consists of a solitary layer, utilising picture dimensions of 28 by 28 for MNIST digits and MNIST fashion, but the image size is 64 by 64 for Cifar10. A classification layer is, after that utilised, comprising ten neurones. The mechanism of activation utilised in the detection layer is the sinusoidal function.

3.2. Multi-layer neural network

The multi-layer model consists of two hidden layers, with an initial input size of 28 by 28 in the case of MNIST digit and MNIST fashion and an image size of 64 x 64 in the case of Cifar10. The first hidden layer has a size of 24, and the second has a size of 16. The final layer serves as the classification layer, comprising 10 neurons. ReLU activation functions were applied to the hidden layers, while the sigmoid activation function was utilized in the last layer.

3.3. Convolution neural network

The “Handwritten Digit Recognition with a Back Propagation Network” paper initially introduced the CNN layer network. Our model adopts the same proposed architecture, consisting of two CNN layers. Four filters of size five by 5 comprise the first layer, and then there is average pooling with filters of size two by 2. Following average pooling with a filter size of 2 by 2, the second layer consists of 12 filters of size 5 by 5. Subsequently, a flattened layer is applied. The final layer is the classification layer, replacing SVM with a single network for classification. All CNN layers utilize ReLU activation, and the classification layer employs the sigmoid activation function.

4. Performance measure

We utilized commonly used performance measures to check the models’ ability to classify the images.

4.1. Accuracy

Accuracy measures the overall correctness of predictions, indicating the ratio of correct predictions to the total number of predictions.

$$\text{Accuracy} = (TP+TN)/(\text{Total Predications}) \quad (1)$$

Accuracy is a general metric for assessing overall model performance. However, in imbalanced datasets, where one class significantly outnumbers others, accuracy may not be the most informative metric, as the majority class could dominate it.

4.2. Precision

Precision is the ratio of correctly predicted positive observations to the total predicted positives.

Precision = TP/(TP+ FP) (2)

Precision is valuable when the cost of false positives is high. In the context of digit recognition, precision would indicate how many of the predicted positive instances (correctly identified digits) are indeed correct.

4.3. Recall

Recall measures the ratio of correctly predicted positive observations to all actual positives.

Recall = TP/(TP+ FN) (3)

Recall is crucial when the cost of false negatives is high. For instance, in digit recognition, recall would indicate the model's ability to identify and capture all cases of a specific digit correctly.

4.4. F1-Score

F1 Score is the harmonic mean of Precision and Recall, providing a balance between the two metrics.

F1 Score =2 × (Precision×Recall)/(Precision+Recall) (4)

F1-Score is beneficial when there is an uneven class distribution or when false positives and negatives must be minimized. It is a good overall metric for evaluating classification models on balanced datasets.

5. Experimental setup

The section will walk through the experimental setting and dataset details used in the comparative study.

5.1. Experimental Settings

All the results and experiments are conducted at Google Collab Intel Xeon CPU with two vCPUs (virtual CPUs), 13 GB of RAM, and NVIDIA Tesla K80 with 12 GB of VRAM (Video Random-Access Memory). We used PyTorch for all model training, torchvision for datasets, and transformations to prepare the dataset for model training.

5.2. Datasets

Discussing the MNIST digit recognition dataset, the Fashion-MNIST dataset and the CIFAR-10 dataset thoroughly examines standard datasets utilized in machine learning and computer vision. Benchmark datasets are essential for testing and assessing the performance of machine learning algorithms. They provide standardised tasks to evaluate methodologies and monitor development over the years. The MNIST digit recognition dataset, Fashion-MNIST, and CIFAR-10 collections are prominent sources for assessing algorithms in recognising image tasks. This is a detailed presentation encompassing the following datasets:

5.2.1. MNIST Digit

The dataset [1] is renowned in the machine learning industry. The dataset contains 70,000 grayscale photographs of handwritten numbers ranging from 0 to 9. It is split into a training set of 60,000 photos and a test set of 10,000 images. The dimensions of each image are 28 by 28 pixels. The MNIST digit dataset was initially developed by Yann LeCun, Corinna Cortes, and Christopher J.C. Burges. It is widely used to assess the effectiveness of image processing systems, especially in recognizing handwritten digits.

5.2.2. MNIST Fashion

The Fashion-MNIST dataset [7] by Zalando Research is a more difficult version of the original MNIST dataset. It consists of 70,000 black and white photos representing ten different fashion item categories, including T-shirts, trousers, and gowns. The dataset, like MNIST, is divided into a training set containing 60,000 images and a test set including 10,000 images, each image being 28x28 pixels in size. Fashion-MNIST is designed to be a straight substitute for the original dataset, providing a more challenging classification task due to the increased intricacy of fashion item photos compared to handwritten digits.

5.2.3. Cifar-10

The CIFAR-10 dataset [4] is a commonly used standard for assessing algorithms related to object recognition in machine learning. The dataset comprises 60,000 colour photographs divided into ten classifications, including aeroplanes, cars, birds, and cats, with 6,000 images in each class. The dataset comprises 50,000 training and 10,000 test images, each measuring 32x32 pixels. Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton developed CIFAR-10, which is more challenging than MNIST and Fashion-MNIST since it contains colour photos and a wider variety of objects.

5.3. Models Training

We use the cross-entropy loss function for all models, with a batch size of 255 for all datasets. The Adam optimizer is employed to optimize the weights. We utilize Reduce LROn- Plateau to adjust the learning rate dynamically, setting the mode to ’min,’ the factor rate to 0.8, the patience factor to 2, and an initial learning rate of 0.01. We run 50 epochs for all models across MNIST Digit and MNIST Fashion dataset cases and 100 epochs for all models across the Cifar-10 dataset.

5.4. Experimental results

5.4.1. Single layer neural network

The performance of the single-layer neural network across three diverse datasets—MNIST Digit, MNIST Fashion, and Cifar-10—reveals varying degrees of success. In the MNIST Digit dataset, the model exhibits robust performance with high precision, recall, and F1-score across most digits, indicating its proficiency in digit classification. The overall accuracy of 92% suggests a reliable performance on this well-established dataset.

On the MNIST Fashion dataset, the model demonstrates satisfactory precision, recall, and F1-score for some classes, such as class 1 (Trousers) and class 9 (Ankle boot). However, its performance is notably lower for classes like 6 (Shirt), where precision, recall, and F1-score are comparatively lower. The overall accuracy of 83% indicates moderate success, but the model struggles with specific fashion categories.

The Cifar-10 dataset poses a more significant challenge for the single-layer neural network. Precision, recall, and F1-score for most classes are notably lower than the MNIST datasets. The model’s accuracy on Cifar-10 is only 30%. The single-layer model is trained on all datasets, and the results can be seen in Table 1.

Table 1: Single-layer neural network

Dataset	Label	Precision	Recall	F1-Score	#Samples
	0	0.95	0.98	0.96	980
	1	0.97	0.97	0.97	1135
	2	0.93	0.89	0.91	1032
		0.88	0.92	0.90	1010
MNIST Digit	3	0.91	0.91	0.91	982
	4	0.91	0.82	0.86	892
	5	0.94	0.94	0.94	958
		0.93	0.93	0.93	1028
	6	0.85	0.89	0.87	974
Accuracy	7	0.89		0.89	1009
			0.90	0.92	10000

	8				
	9				
	0			0.80	1000
	1	0.74	0.87	0.96	1000
MNIST Fashion	2	0.98	0.95	0.72	1000
	3	0.69	0.77	0.84	1000
	4	0.82	0.85	0.73	1000
	5	0.71	0.74	0.92	1000
	6	0.94	0.90	0.51	1000
	7	0.63	0.43	0.91	1000
Accuracy	8	0.89	0.93	0.94	1000
	9	0.94	0.93	0.94	1000
		0.94	0.94	0.83	10000
	0	0.31	0.68	0.43	1000
	1	0.42	0.49	0.45	1000
	2	0.21	0.43	0.28	1000
	3	0.22	0.37	0.28	1000
Cifar-10	4	0.23	0.17	0.20	1000
	5	0.27	0.08	0.13	1000
	6	0.29	0.10	0.15	1000
Accuracy	7	0.46	0.22	0.29	1000
	8	0.44	0.24	0.31	1000
	9	0.42	0.20	0.27	1000
				0.30	10000

The MNIST digit dataset shows a good performance as compared to the other dataset. Overall, on MNIST digit, this model is good, but labels 4 and 8 cannot perform better. The loss plots are also shown for all datasets in Fig. 1 to visualize how the models converge.

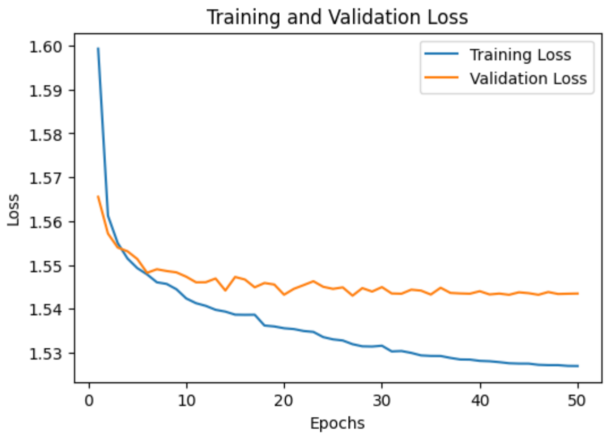

Figure 1 (a): MNIST Digit Dataset Losses

Figure 2 (b): MNIST Fashion Dataset Losses

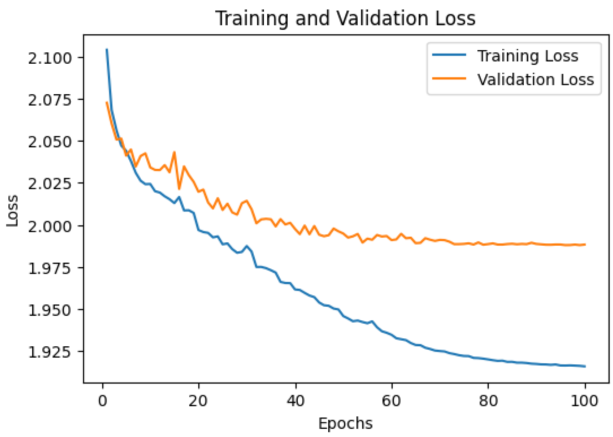

Figure 3 (c): Cifar-10 Dataset Losses

Fig. 1: Training and Validation dataset losses for model Single Layer across all the datasets

5.4.2. Multi-layer neural network

The multi-layer neural network demonstrates a notable performance improvement compared to its single-layer counterpart across the three datasets. In the MNIST Digit dataset, the model achieves higher precision, recall, and F1-score for all digits, reaching an impressive overall accuracy of 95%. The enhanced ability to capture intricate patterns and features in the data is evident, leading to superior classification results. Similarly, on the MNIST Fashion dataset, the multi-layer model outperforms the single-layer one. It exhibits increased precision, recall, and F1 scores for several fashion classes. Despite challenges with class 6 (Shirt), the overall accuracy improves to 84%, indicating the more profound architecture's effectiveness in capturing nuanced features in fashion images.

The most significant advancement is observed in the Cifar-10 dataset. The multi-layer model substantially enhances precision, recall, and F1-score across most classes compared to the single-layer network. Although the accuracy remains modest at 33%, the deeper architecture demonstrates a better capacity to discern complex patterns within diverse and challenging images, showcasing its suitability for more intricate classification tasks. The overall results of these three datasets are shown in Table 2

Table 2: multi-layer neural network

Dataset	Label	Precision	Recall	F1-Score	#Samples
MNIST Digit	0	0.96	0.99	0.97	980
	1	0.98	0.98	0.98	1135
	2	0.95	0.95	0.95	1032
	3	0.95	0.94	0.95	1010
	4	0.93	0.95	0.94	982
	5	0.96	0.95	0.95	892
	6	0.97	0.96	0.96	958
	7	0.96	0.95	0.95	1028
	8	0.93	0.95	0.94	974
	9	0.96	0.92	0.94	1009
Accuracy				0.95	10000
MNIST Fashion	0	0.70	0.90	0.79	1000
	1	0.98	0.97	0.97	1000
	2	0.64	0.89	0.75	1000
	3	0.85	0.89	0.87	1000
	4	0.75	0.66	0.70	1000
	5	0.95	0.93	0.94	1000
	6	0.81	0.33	0.47	1000
	7	0.85	0.97	0.90	1000
	8	0.97	0.95	0.96	1000
	9	0.97	0.87	0.91	1000
Accuracy				0.84	10000
Cifar-10	0	0.37	0.69	0.48	1000
	1	0.41	0.71	0.52	1000
	2	0.23	0.51	0.32	1000
	3	0.25	0.52	0.34	1000
	4	0.27	0.14	0.18	1000
	5	0.27	0.03	0.05	1000
	6	0.43	0.05	0.09	1000
	7	0.50	0.32	0.39	1000
	8	0.48	0.22	0.30	1000
	9	0.42	0.11	0.18	1000
Accuracy				0.33	10000

The converging loss function plots in Fig. 2.

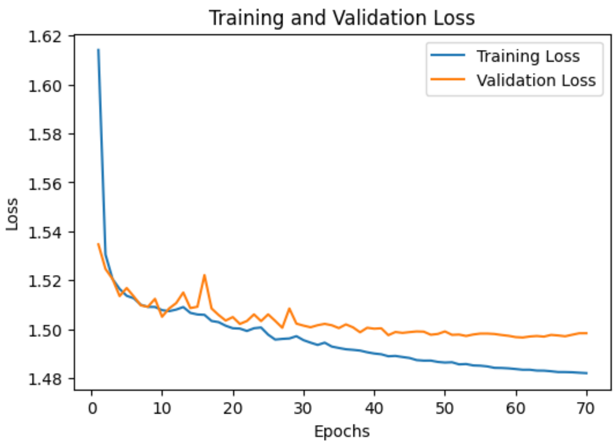

Figure 2 (a): MNIST Digit Dataset Losses

Figure 2 (b): MNIST Fashion Data Set Losses

Figure 2 (c): Cifar-10 Dataset Losses

Fig. 2: Training and Validation dataset losses for model multi-layers across all the datasets

5.4.3. Convolution neural network

The Convolutional Neural Network (CNN) outperforms single-layer and multi-layer neural networks across all three datasets, showcasing its efficacy in image classification tasks. In the MNIST Digit dataset, the CNN achieves near-perfect precision, recall, and F1-score for most digits, resulting in an overall accuracy of 99%. CNN's ability to capture hierarchical features and patterns in images contributes to its superior performance in digital recognition.

Similarly, on the MNIST Fashion dataset, the CNN exhibits notable improvements over the previous architecture. It achieves higher precision, recall, and F1-score for various fashion classes, leading to % overall accuracy of 86%. CNN's capability to automatically learn relevant features and hierarchies in image data enhances its performance in classifying diverse fashion items.

The most substantial advancement is observed in the Cifar-10 dataset. CNN significantly outperforms single-layer and multi-layer models, achieving higher precision, recall, and F1 scores across most classes. Despite the challenging nature of the dataset, CNN demonstrates an overall accuracy of 40%, highlighting its ability to handle complex and diverse images better than its predecessors. The overall results of these three datasets are shown in Table 3.

Table 3: Convolution Neural Network

Dataset	Label	Precision	Recall	F1-Score	#Samples
MNIST Digit	0	0.98	1.00	0.99	980
	1	0.99	0.99	0.99	1135
	2	0.98	0.99	0.99	1032
	3	0.99	1.00	0.99	1010
	4	0.99	0.99	0.99	982
	5	0.98	0.98	0.98	892
	6	1.00	0.98	0.99	958
	7	0.99	0.98	0.99	1028
	8	0.99	0.99	0.99	974
	9	0.99	0.97	0.98	1009
Accuracy				0.99	10000
MNIST Fashion	0	0.70	0.91	0.79	1000
	1	0.99	0.97	0.98	1000
	2	0.80	0.78	0.79	1000
	3	0.86	0.89	0.87	1000
	4	0.77	0.74	0.75	1000
	5	0.97	0.97	0.97	1000
	6	0.66	0.49	0.56	1000
	7	0.93	0.96	0.95	1000
	8	0.98	0.95	0.97	1000
	9	0.97	0.95	0.96	1000
Accuracy			0.87	0.86	10000
Cifar-10	0	0.41	0.71	0.52	1000
	1	0.50	0.74	0.60	1000
	2	0.29	0.50	0.36	1000
	3	0.29	0.52	0.37	1000
	4	0.31	0.24	0.27	1000
	5	0.23	0.07	0.10	1000
	6	0.56	0.36	0.43	1000
	7	0.55	0.33	0.41	1000
	8	0.53	0.28	0.37	1000
	9	0.49	0.21	0.29	1000
Accuracy				0.40	10000

The converging loss function plots in Fig. 3.

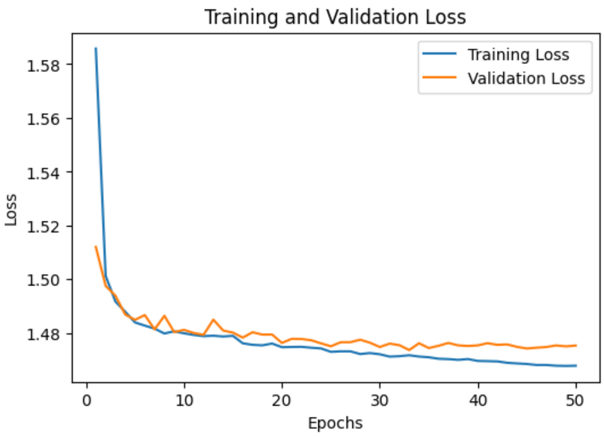

Figure 3 (a): MNIST Digit Dataset Losses

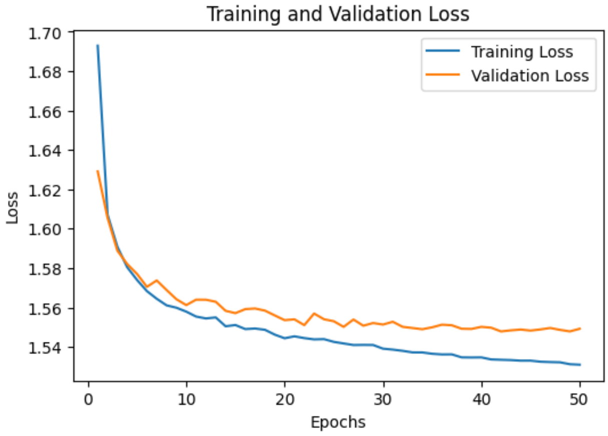

Figure 3 (b): MNIST Fashion Dataset Losses

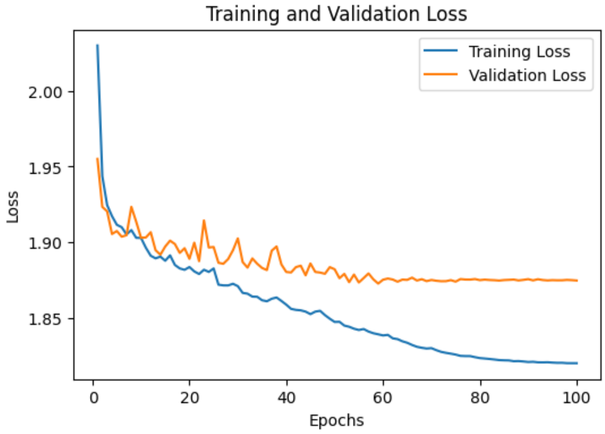

Figure 3 (c): Cifar-10 Dataset Losses

Fig.3: Training and Validation dataset losses for model CNN across all the datasets

5.5. Analysis

The misclassified images across the single-layer, multi-layer, and convolutional neural network (CNN) models provide valuable insights into their limitations. The single layer network struggles with intricate patterns, particularly in Cifar-10, where its simplicity hampers its ability to discern complex features, and the misclassified across MNIST Fashion and Digit are shown in Fig. 4.

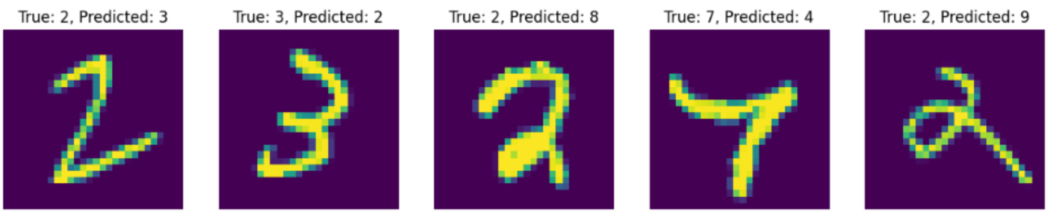

Figure 4 (a): MNIST Digit miss-classifications

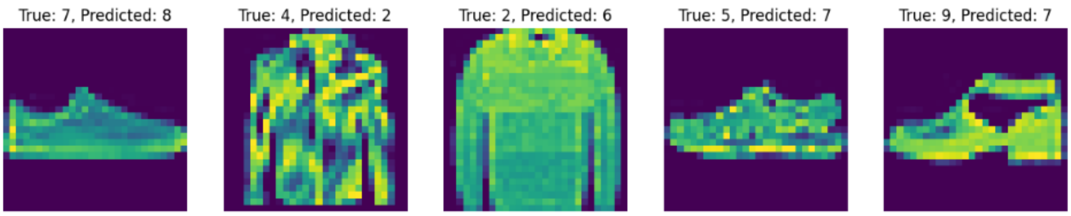

Figure 4 (b): MNIST Fashion miss-classifications

Fig.4: Single Layer Neural Network miss-classifications images

The model confused Number 2 with 3 and 3 with 2 because of the similarity of both words and generalization of the model; that is because it misclassified 2 as 3 and 3 as 2.

The multi-layer network showcases improved performance but challenges distinguishing fine details, leading to misclassifications, especially in MNIST Fashion, where class boundaries are subtle, and the misclassified across MNIST Fashion and Digit are shown in Fig. 5.

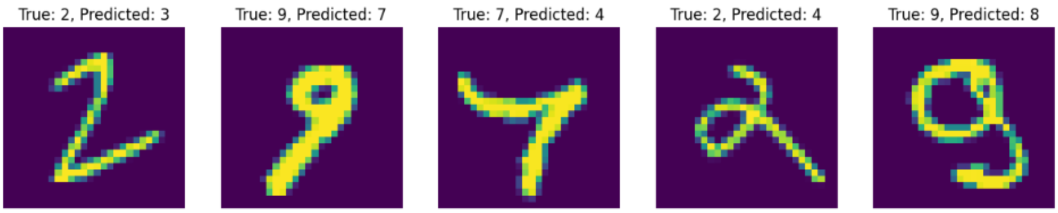

Figure 5 (a): MNIST Digit miss-classifications

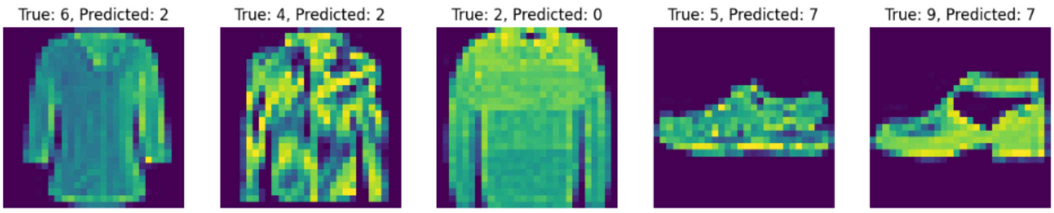

Figure 5 (b): MNIST Fashion miss-classifications

Fig. 5: Erroneous categorization of images is produced by multi-layer neural networks.

Because of the similarities between the two terms and the model's generalisation, the model misidentified the number 9 as both 8 and 7. This is because it misclassified 9 as both eighth and seventh.

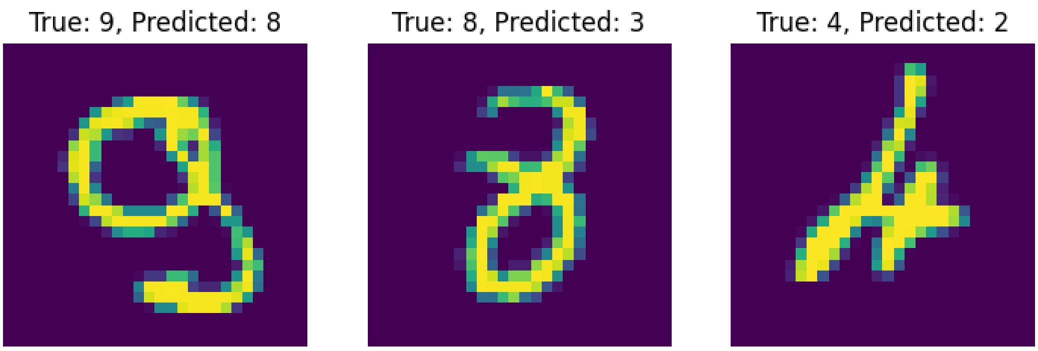

Figure 6 (a): MNIST Digit miss-classifications

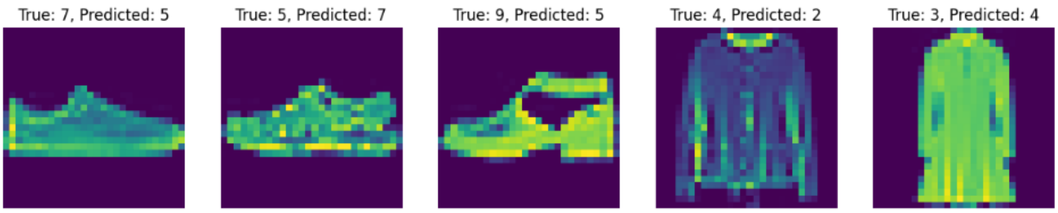

Figure 6 (b): MNIST Fashion miss-classifications

In contrast to the previous models, the Convolutional Neural Network (CNN) does not have any misclassifications in fig. 6 because it has a lower generalization of pictures. Yet, it does incorrectly classify 9 as 8 and 8 as 3. Its simplicity limits its capacity to capture intricate features present in the diverse images of Cifar-10 and the varied fashion items in MNIST Fashion. The limitations become more evident in lower accuracy scores, especially on Cifar-10, where the model struggles to discern patterns effectively.

While the most robust, CNN encounters misclassifications due to its sensitivity to lighting, orientation, and background in specific images. Despite its superiority, ambiguous features can still mislead CNN, especially in the diverse and complex Cifar-10 dataset. These misclassifications underscore the importance of model architecture and dataset characteristics, and the misclassification across MNIST Fashion and Digit are shown in Fig. 6. While CNN excels overall, addressing its sensitivity to nuanced features and further refining architectures could contribute to even more accurate classifications.

The MLN enhances the structure's depth, leading to better outcomes on all datasets relative to the SLN. This underscores the significance of ordering and rationalization in acquiring complicated aspects. Nevertheless, the MLN remains inadequate, especially on Cifar-10, suggesting that a deeper infrastructure alone may not address the dataset's diversification and intricacy. CNN, specifically engineered for image-related tasks, demonstrates considerable superiority over SLN and MLN. Its capacity to autonomously acquire hierarchical features and identify spatial patterns renders it appropriate for image categorization. CNN demonstrates exceptional accuracy on MNIST Digit and MNIST Fashion, surpassing other architectures. Nonetheless, its performance on Cifar-10 is rather small while being the most effective of the three models, highlighting the dataset's intrinsic complexity.

Sun et al. [18], Duan et al. [19], and Wang et al. [20] have contributed to the advancement of control systems by emphasizing fuzzy models, sliding mode control, and event-triggered techniques. Further studies in this domain may focus on integrating sliding mode control with sophisticated prediction algorithms to improve stability in highly nonlinear systems influenced by uncertainty and outside factors. Furthermore, creating flexible or self-learning control systems would benefit dynamic contexts. Huang et al. [21] proposed a new multi-objective grey target decision model for supplier selection, offering a structure with the potential for further expansion. Future studies may investigate adaptive decision-making methods, the integration of machine learning methods for robotics and improved accuracy, multi-criteria implementations across several domains, and the quantification of uncertainty to improve the resilience of decision-making. Ali et al. [22] and Marwan et al. [23] establish a foundation for the continued investigation of fractional-order systems in complex dynamic systems. Future research priorities encompass the simulation of cancer dynamics to integrate supplementary biological factors, fractal-based diagnostics for enhanced clinical imaging techniques, optimization through fractional-order methodologies in engineering systems, and applying these strategies to fluid dynamics. Abidin et al. [24] elucidated well-posedness in variable-exponent function spaces for micropolar fluid equations, paving the stage for further exploration. Future research may focus on developing comprehensive fluid models for multiphase and turbulent flows, and sophisticated computational models to address practical fluid dynamics challenges.

6. Conclusions

Our investigation of single-layer (SLN), multi-layer (MLN), and convolutional neural network (CNN) architectures in the MNIST Digit, MNIST Fashion, and CIFAR-10 datasets has produced significant insights into their ability to perform. The single-layer model displayed efficacy in fundamental tasks like MNIST digit identification but encountered difficulties addressing the complex structure of CIFAR-10 and the variety of MNIST Fashion. The depth incorporation in the multi-layer model showed enhancement, particularly on Cifar-10; however, it remained inferior to the more sophisticated CNN. The convolutional neural networks distinguished themselves as the top performing, performing in all scenarios and showcasing their ability to capture complex characteristics in visual data. Challenges included the resource-intensive nature of the CNN and its vulnerability to excess fitting. To tackle these issues, possible enhancements encompass data augmentation, transfer learning, architectural refinement, and hyperparameter optimization. This thorough examination underscored the advantages and disadvantages of various designs and stressed the importance of strategic model training methodologies. The selection of architecture and meticulous hyper-parameter tweaking and training techniques are crucial for attaining optimum efficiency across various datasets.

Author Contributions

For research articles with several authors, a short paragraph specifying their individual contributions must be provided. The following statements should be used “Conceptualization, M. R. and H. N.; methodology, H. N.; software, M. R.; validation, M. R., H. N. and R. M. Z.; formal analysis, R. M. Z.; investigation, H. N.; resources, M. R.; data curation, H. N.; writing—original draft preparation, M. R., H. N. and R. M. Z.; writing—review and editing, R. M.Z.; visualization, H. N.; supervision, R. M. Z.; project administration, R. M. Z. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was not funded by any grant.

References

1. Desai, M., & Shah, M. (2021). An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). Clinical eHealth, 4, 1-11. https://doi.org/10.1016/j.ceh.2020.11.002.
2. Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. Annals of Data Science, 10(1), 183-208. https://doi.org/10.1007/s40745-021-00344-x.
3. LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten digit recognition with a back-propagation network. Advances in neural information processing systems, 2.

https://proceedings.neurips.cc/paper_files/paper/1989.
4. 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. https://doi.org/10.1109/MSP.2012.2211477.
5. Denker, J., Gardner, W., Graf, H., Henderson, D., Howard, R., Hubbard, W., ... & Guyon, I. (1988). Neural network recognizer for handwritten zip code digits. Advances in neural information processing systems, 1. (1988). https://proceedings.neurips.cc/paper_files/paper/1988.
6. Krizhevsky, A., Nair, V., & Hinton, G. (2010). 5(4), 1. Cifar-10 (canadian institute for advanced research). URL http://www. cs. toronto. edu/kriz/cifar. Html
7. Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms. https://doi.org/10.48550/arXiv.1708.07747.
8. Ullah, W., Siddique, I., Zulqarnain, R. M., Alam, M. M., Ahmad, I., & Raza, U. A. (2021). Classification of arrhythmia in heartbeat detection using deep learning. Computational Intelligence and Neuroscience, 2021(1), 2195922. https://doi.org/10.1155/2021/2195922.
9. Yang, Y., Chen, F., Lang, J., Chen, X., & Wang, J. (2021). Sliding mode control of persistent dwell-time switched systems with random data dropouts. Applied Mathematics and Computation, 400, 126087. https://doi.org/10.1016/j.amc.2021.126087.
10. Duan, Z., Ding, F., Liang, J., & Xiang, Z. (2023). Observer-based fault detection for continuous–discrete systems in TS fuzzy model. Nonlinear Analysis: Hybrid Systems, 50, 101379. https://doi.org/10.1016/j.nahs.2023.101379.
11. Diao, Y., & Zhang, Q. (2021). Optimization of Management Mode of Small-and Medium-Sized Enterprises Based on Decision Tree Model. Journal of Mathematics, 2021(1), 2815086. https://doi.org/10.1155/2021/2815086.
12. Yang, C., Li, F., Kong, Q., Chen, X., & Wang, J. (2021). Asynchronous fault-tolerant control for stochastic jumping singularly perturbed systems: An H∞ sliding mode control scheme. Applied Mathematics and Computation, 389, 125562. https://doi.org/10.1016/j.amc.2020.125562.
13. Xing, R., Xiao, M., Zhang, Y., & Qiu, J. (2022). Stability and Hopf bifurcation analysis of an (n+ m)-neuron double-ring neural network model with multiple time delays. Journal of Systems Science and Complexity, 35(1), 159-178. https://doi.org/10.1007/s11424-021-0108-2.
14. Jia, T., Chen, X., He, L., Zhao, F., & Qiu, J. (2022). Finite-time synchronization of uncertain fractional-order delayed memristive neural networks via adaptive sliding mode control and its application. Fractal and Fractional, 6(9), 502. https://doi.org/10.3390/fractalfract6090502.
15. Yu, Z., Zhao, F., Ding, S., & Chen, X. (2022). Adaptive pre-assigned finite-time control of uncertain nonlinear systems with unknown control gains. Applied Mathematics and Computation, 417, 126784. https://doi.org/10.1016/j.amc.2021.126784.
16. Zhang, N., Qi, W., Pang, G., Cheng, J., & Shi, K. (2022). Observer-based sliding mode control for fuzzy stochastic switching systems with deception attacks. Applied Mathematics and Computation, 427, 127153. https://doi.org/10.1016/j.amc.2022.127153.
17. Xu, Y., Liu, Y., Ruan, Q., & Lou, J. (2023). Data-driven optimal tracking control of switched linear systems. Nonlinear Analysis: Hybrid Systems, 49, 101355. https://doi.org/10.1016/j.nahs.2023.101355.
18. Sun, Q., Ren, J., & Zhao, F. (2022). Sliding mode control of discrete-time interval type-2 fuzzy Markov jump systems with the preview target signal. Applied Mathematics and Computation, 435, 127479. https://doi.org/10.1016/j.amc.2022.127479.
19. Duan, Z. X., Liang, J. L., & Xiang, Z. R. (2022). H∞ control for continuous-discrete systems in TS fuzzy model with finite frequency specifications. Discrete Contin Dyn Syst S, 64(1), 1-18. http://dx.doi.org/10.3934/dcdss.2022064.
20. Wang, H., Chen, X., & Wang, J. (2022). H∞ sliding mode control for PDT-switched nonlinear systems under the dynamic event-triggered mechanism. Applied Mathematics and Computation, 412, 126474. https://doi.org/10.1016/j.amc.2021.126474.
21. Huang, B., Miao, J., & Li, Q. (2022). A Vetoed Multi-objective Grey Target Decision Model with Application in Supplier Choice. Journal of Grey System, 34(4). https://openurl.ebsco.com/EPDB%3Agcd%3A3%3A15617539/detailv2?sid=ebsco%3Aocu%3Arecord&id=ebsco%3Agcd%3A161945001&bquery=IS%200957-3720%20AND%20VI%2034%20AND%20IP%204%20AND%20DT%202022&page=1&link_origin=scholar.google.com

22. Ali, G., Marwan, M., Rahman, U. U., & Hleili, M. (2024). Investigation Of Fractional-Ordered Tumor-Immune Interaction Model Via Fractional-Order Derivative. FRACTALS (fractals), 32(06), 1-10. 10.1142/S0218348X24501196.
23. Abidin, M. Z., Marwan, M., Ullah, N., & Mohamed Zidan, A. (2023). Well-Posedness in Variable-Exponent Function Spaces for the Three-Dimensional Micropolar Fluid Equations. Journal of Mathematics, 2023(1), 4083997. https://doi.org/10.1155/2023/4083997.
24. Marwan, M., Han, M., Dai, Y., & Cai, M. (2024). The Impact Of Global Dynamics On The Fractals Of A Quadrotor Unmanned Aerial Vehicle (Quav) Chaotic System. Fractals, 32(02), 2450043. https://doi.org/10.1142/S0218348X24500439.