Application of Deep Learning Models on MNIST Dataset

Using the MNIST dataset, the research assesses several neural network topologies for image classification tasks. The primary concerns that are discussed are the efficacy and efficiency of various model types, like the deep CNN, batch normalization CNN, regularization (L2) CNN, easy Multilayer Perceptron (MLP), and convolutional neural network (CNN).

Key findings from the evaluation of these models include:
Performance of Different Architectures:
MLP performs relatively poorly compared to CNN-based models, indicating that simple feedforward networks may not be suitable for image classification tasks. CNN and its variants outperform MLP, suggesting that convolutional layers are effective in capturing spatial information in images.

Impact of Model Complexity:
Models with more layers and parameters tend to perform better, as seen in the CNN with regularization (L2) and deep CNN models.
Deep CNN, with multiple convolutional layers and increased complexity, achieves the highest accuracy among the evaluated models.

Effectiveness of Regularization and Normalization:
CNN with batch normalization and CNN with regularization (L2) show improvements over basic CNN, indicating the benefits of these techniques in stabilizing and regularizing training.

Generalization and Overfitting:
L2 regularization is a kind of regularization approach that helps reduce overfitting and improves accuracy of generalization on data with no observations.

Model Complexity vs. Performance Trade-off:
Deeper models with more convolutional layers tend to perform better but may also require more computational resources and longer training times.
Overall, the evaluation highlights the importance of architecture choice, regularization, and normalization techniques in designing effective neural networks for image classification tasks.

Introduction:

A fundamental job in computational vision, sorting images has applications in anything from autonomous driving to diagnosing illnesses. Accurate picture classification is essential for applications like object recognition, picture analysis, and picture retrieval. The current exponential expansion of picture data has greatly raised the requirement for effective and precise image classification methods. Because neural networks, and more specifically Convolutional Neural Networks (CNNs), can automatically learn structural features based on raw picture statistics, they have become the state-of-the-art method for classifying pictures. The performance of these models may, however, be greatly impacted by the selection of hyperparameters and the architecture of neural networks. So in order find the best methods for identifying pictures tasks, it is crucial to assess and contrast various neural network topologies.

Taking the data set from MNIST as a reference, my goal in this study is to examine and contrast how well different neural network designs perform for picture classification tasks. A popular dataset called MNIST is made up of 28x28 grayscale pictures of handwritten numbers (0–9). Despite being a very easy dataset, MNIST is a common average for assessing and contrasting various image classification techniques. A few architectures will be compared for performance: deep CNN, CNN with regularization (L2), CNN with batch normalization, basic CNN, and simple Multilayer Perceptron (MLP). I hope to learn more about these designs' generality, robustness, and efficacy by assessing them.

Comprehending the various neural network designs' advantages in performance is vital for multiple rationales. To begin with, it assists students and professionals in selecting the best structure for models using benchmarks including accuracy, computing speed, and depth of model for certain tasks. Additionally, via the identification of optimal procedures and useful designs, it advances deep learning approaches. Ultimately, enhanced picture modeling techniques offer far-reaching effects in several fields, such as robotics, security, and health. I can make systems for computer vision more dependable and accurate by pushing the boundaries of picture categorization, which will improve their utility and efficacy in everyday use.

Current Research: Deep Learning for MNIST Classification

An essential resource for assessing visual extraction methods, especially neural network examples, is still the MNIST data set. Present study endeavors to push the limits of reliability and precision while investigating novel structures as well as instruction methodologies.

The research paper [1] investigates methods to enhance the adversarial defense of neural networks, particularly focusing on image recognition using the MNIST dataset. The authors propose a combined approach involving feature masking and gradient manipulation to improve network resilience against adversarial attacks. They evaluate the effectiveness of this approach using a conventional neural network architecture and compare it to a baseline model without feature masking.

Findings:
Baseline Model Performance: The baseline model achieved a high-test accuracy of 98% on the MNIST dataset. However, it showed limited resistance to adversarial attacks, with accuracy dropping to 60% under FGSM assaults.

Impact of Feature Masking: Models incorporating feature masking demonstrated a trade-off between accuracy and resilience. As the proportion of feature masking increased, there was a decrease in accuracy but an enhancement in adversarial resilience.

Accuracy and Robustness Trade-off: A 10% masking ratio resulted in a 96% accuracy rate with 75% robustness against attacks. With a 30% masking, accuracy was 94% with 80% robustness. A 50% masking threshold achieved a 92% accuracy, reaching the peak of robustness at 85%.

Efficacy of Feature Masking: The results affirm the effectiveness of feature masking in augmenting adversarial defense. There's a critical balance between accuracy and resilience, highlighting the importance of finding the optimal masking ratio.

Future Research Directions:
The study suggests further investigations into refined masking methodologies and their integration with other defensive strategies to broaden the scope of neural network security against adversarial threats.

In summary, the research demonstrates that integrating feature masking with neural network training can significantly improve adversarial

resilience without sacrificing too much accuracy. This method offers a workable plan for creating neural network frameworks that are more reliable and best, which is a major development in AI security.

An accelerated genetic approach for deep convolutional neural network (CNN) training is presented in the research [2].
Through the introduction of parent-child linkages via this algorithm, people can pass on knowledge from their ancestors to future generations. This inheritance mechanism aims to reduce execution time while ensuring robust training of the CNNs. Additionally, the paper presents a new dataset called Double MNIST, designed to succeed the MNIST dataset, particularly suitable for machine learning education and handwriting recognition applications.

Findings:

Accelerated Genetic Algorithm: The introduced genetic algorithm accelerates the training of CNNs by allowing knowledge transfer from ancestors to descendants This method expedites the learning cycle without sacrificing the taught models' resilience.

Child-parent Relationships: These relationships help pass on information to generations to come, allowing people to gain on the lessons of their grandparents. This mechanism contributes to efficient training and learning within the algorithm.

Double MNIST Dataset: The Double MNIST dataset is introduced as a successor to the MNIST dataset. This new dataset is designed to be more challenging and suitable for machine learning education, particularly for tasks like handwriting recognition.

Performance Under Evolutionary Scenarios: The algorithm demonstrates strong performance under various evolutionary scenarios. This indicates its versatility and effectiveness in different training environments.

Potential for Further Improvements: The paper suggests avenues for further improvements, such as enhancing ancestor-descendant relationships for more comprehensive knowledge transfer. This implies that the algorithm's performance can be further enhanced with additional refinement.

Desai's research [3] explores the impact of weight initialization techniques on the efficiency and performance of neural networks using the MNIST dataset. By evaluating methods like random, Xavier/Glorot, and He techniques within a specific neural network architecture, Desai highlights their importance in achieving rapid convergence and strong generalization. The study clarifies how different techniques affect convergence speed and model performance, offering guidance for optimizing deep learning models' training. Overall, Desai's work underscores the critical role of appropriate weight initialization in enhancing neural network efficiency and effectiveness in tasks like image recognition.

Data Collection:

For this study, I utilized the MNIST dataset, a standard benchmark in the field of image classification. Most of the ten thousand test pictures and 60,000 images used as training in the data set of MNIST is a 28x28 grayscale picture with a written number (0–9) upon it. Because it is readily available and simple to utilize, the collection of data is well-organized and frequently used for assessing machine learning models. MNIST offers a consistent framework for assessing model precision and generalization, making it a useful guide for comparing the effectiveness of various artificial neural network designs.

Model Development:

Multilayer Perceptron (MLP): The MLP is layout has two entirely secret layers with a total of 128 neurons each and ReLU activation functions after supplying an input layer of 784 neurons (28x28 pixels flat). Ten neurons, one for each of the 10 digits class (0–9), make up the result layers. I used the SoftMax activation function to obtain class probabilities. An optimizer called Adam with an categorical crossing entropy decrease function is used for training the algorithm.

Fundamental Convolutional Neural Network (CNN) Architectural Design: the maximum pooling layers come after two convolutional layers with 32 and 64 filters, accordingly. 2 completely linked layers with 128 neurons and ReLU activation functions come after these initial layers. Lastly, there are 10 neurons with SoftMax activity in the output layer. For lessons, I use a categorical loss of cross-entropy with the Adams optimization.

CNN with Batch Normalization: The design of this CNN is comparable to that of the basic CNN, but it has layers of batch normalization after every fully linked and convolutional layer. By normalizing the activations

of every layer, batch normalization increases stability and speeds up convergence in the training process. Using categorical cross-entropy loss and the Adam optimizer, the model is trained.

CNN with Regularization (L2): To reduce overfitting, this CNN design includes L2 regularization. By penalizing heavier weights in the network, L2 regularization encourages simpler models and lowers the chance of overfitting. L2 regularization is used to the convolutional neural networks completely related layers of the architecture, which is identical to that of the basic CNN. Using categorical cross-entropy loss and the Adam optimizer, the model is trained.

Deep Convolutional Neural Network (Deep CNN): This approach used maximum pooling layers after several layers of convolution with raising down and bigger filters. This structure consists of five convolutional layers, each with 32, 64, 128, 256, and 512 filters. A max-pooling layer comes after each convolutional layer. Two completely linked layers with 512 neurons each and ReLU activation functions are connected to the flattened output. Lastly, there are 10 neurons with SoftMax activity in the output layer. The Adam optimizer, with categorical cross-entropy loss is used for training the model.

Training:

The 60,000 photos in the training set are used to train each model. employed a 128-batch size and trained the models over a 20-epoch period. To avoid overfitting, early stopping was used in studying. Training was stopped if the validation loss did not improve after a certain number of epochs. Additionally, the reduction in precision on the training set and accurate on the set used for validation were tracked. Following training, each model's performance was assessed using the 10,000-image test set to see how accurate and generalizable it was.

Analysis:

Many important conclusions from our investigation about how various neural network designs perform on the MNIST dataset are revealed.
Model Accuracy: Of all the models assessed, the deep convolutional neural network (Deep CNN) outperformed all other designs in terms of efficiency on the given data set. This demonstrates how well larger structures capture the intricate details and trends seen in the photos.

Model	Test Accuracy

MLP	98.36%
CNN	99.14%
CNN with Batch Norm	99.05%
CNN with L2 Regularization	98.61%
Deep CNN	99.08%
CNN with Dropout	99.17%
CNN with Data Augmentation	99.31%

Table1: Accuracies of Trained Models

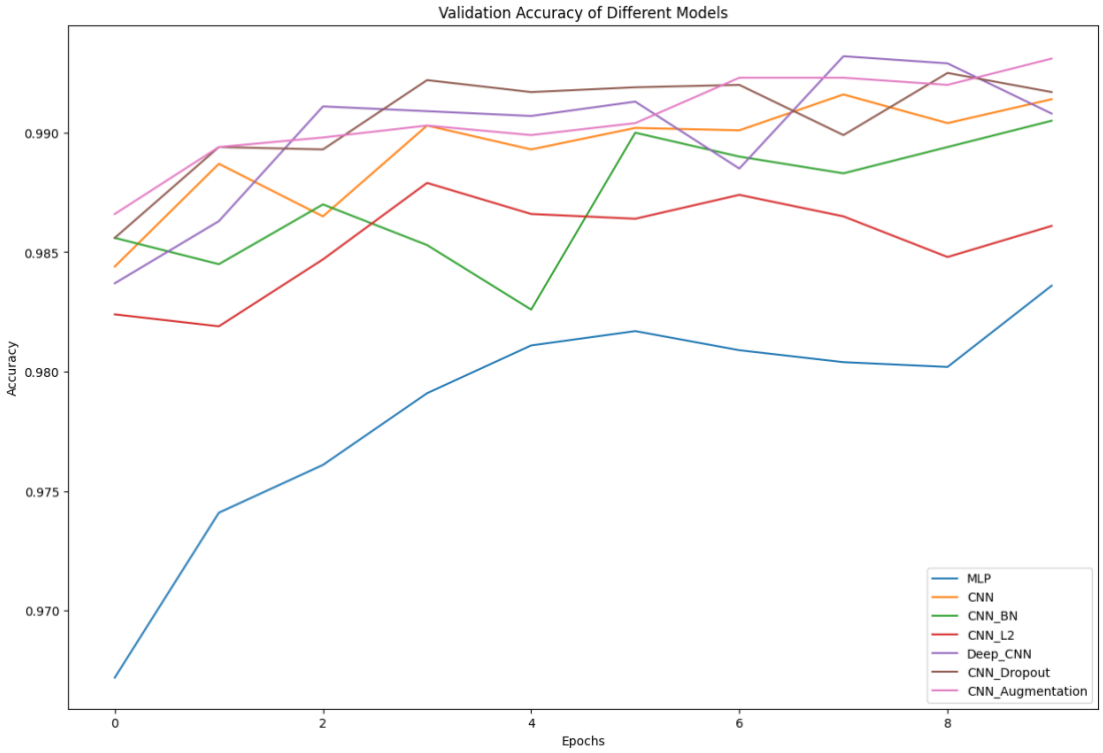

Fig1: Explaining the validation accuracy of models.

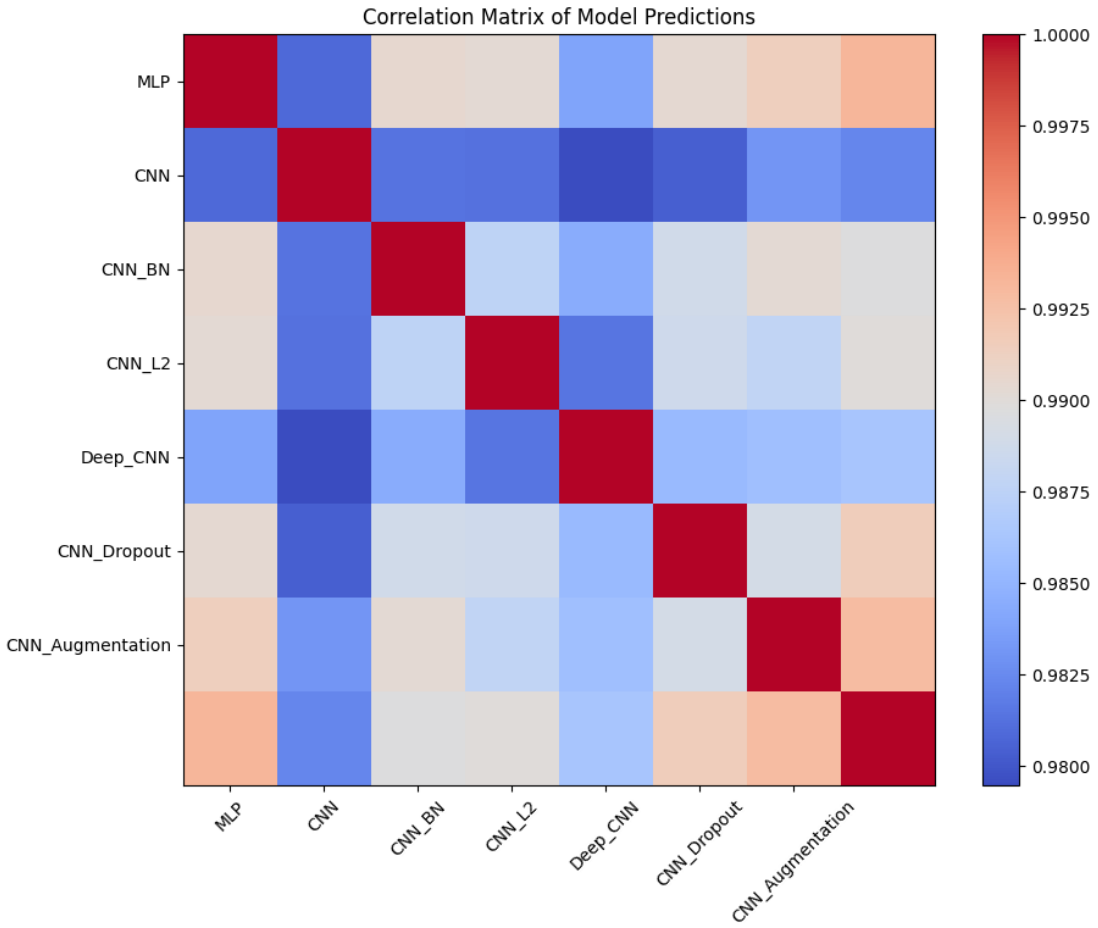

Fig2: Explains about the correlation of model predictions.

Impact of Architectural Complexity: it was found that better performance was often achieved via adding additional layers of convolution and filters to the architecture. The multilayer perceptron (MLP) was outperformed by the standard CNN, demonstrating the significance of layers of convolution in identifying spatial hierarchies in picture data.

Effect of Regularization: Relative to their non-regularized counterparts, models that used regularization approaches, including L2 regularization, showed better generalization performance. The CNN with L2 regularization exhibited reduced overfitting, indicating the effectiveness of regularization in preventing the models from memorizing noise in the training data.

Batch Normalization: Models with batch normalization layers showed faster convergence during training and improved stability. The CNN with batch normalization demonstrated enhanced performance compared to the

basic CNN, indicating that batch normalization aids in smoother optimization and better gradient flow.

Training Time and Computational Cost: Deeper architectures, such as the Deep CNN, require more training time and computational resources compared to simpler architectures like MLP. However, the performance gains achieved with deeper architectures justify the additional computational cost, especially in tasks where high accuracy is crucial.

Overall, our research suggests that for image classification tasks like MNIST, deeper convolutional neural networks with regularization and batch normalization tend to offer the best balance of accuracy and generalization performance. These findings provide valuable insights for designing and selecting appropriate neural network architectures for similar image classification tasks.

Summary and Conclusion:

Using the MNIST dataset, I investigated many neural network algorithms for recognizing handwritten digits in this work. Multi-Layer Perceptron (MLP), Convolutional Neural Network (CNN), CNN with Batch Normalization, CNN with L2 Regularization, Deep CNN, CNN with Dropout, and CNN with augmented data are the seven models that I implemented and assessed. The MNIST dataset, which has 10,000 test pictures and 60,000 training images, was used to train and evaluate the models.

The models had varying test accuracies; CNN via Data Augmentation had the greatest accuracy at 99.31%, closely followed by CNN with A dropout at 99.17%. The accuracy of the other models, which ranged from 98.36% to 99.14%, was likewise good. These outcomes show how well deep learning models work at correctly identifying handwritten numbers.

In conclusion, our study highlights the importance of selecting appropriate deep learning architectures and techniques for achieving high accuracy in handwritten digit recognition tasks. CNN-based models, especially those incorporating Dropout and Data Augmentation techniques, showed superior performance in this task. These findings can be valuable for developing robust digit recognition systems, which have applications in various fields including OCR, automated document processing, and computer vision.

References:

[1] Ingle, G., & Pawale, S. (2024). Enhancing Adversarial Defense in Neural Networks by Combining Feature Masking and Gradient Manipulation on the MNIST Dataset. International Journal of Advanced Computer Science and Applications (IJACSA), 15(1). Department of Computer Engineering, Vishwakarma University, Pune, India.

[2] A. Meena, G. M. V. Reddy, and D. P. Chavali, "Accelerated CNN Training with Genetic Algorithm," 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Gwalior, India, 2024, pp. 1-6, Doi: 10.1109/IATMSI60426.2024.10502992. keywords: {Training; Technological innovation; Machine learning algorithms; Sociology; Machine learning;Robustness;Convolutional neural networks},

[3] Desai, C. (2024). Impact of Weight Initialization Techniques on Neural Network Efficiency and Performance: A Case Study with MNIST Dataset. International Journal of Engineering and Computer Science, 13(04), 26115-26120. DOI: 10.18535/ijecs/v13i04.4809. Department of Computer Science, National Defence Academy, Pune.