Image Classification Using CIFAR-10 Dataset and Convolution Neural Networks

Team Members:

1.) Venkatesh Rudravaram - 700747025
2.) Vineeth Virupakshi - 700745532
3.) Supriya Bhukya - 700758052
4.) Naga Hrushikesh Hukumathirao - 700744719

***Abstract*— This project, titled "Enhancing Image Classification Through Deep Learning: A CIFAR-10 Dataset Approach," represents a comprehensive investigation into the field of image classification. Motivated by the escalating demand for advanced and accurate models capable of handling intricate datasets, particularly in the face of the increasing complexity and diversity of image data, this research leverages the CIFAR-10 dataset as a benchmark. This dataset, renowned for its ten distinct classes and diverse visual content, serves as a foundational basis for evaluating the efficacy of deep learning techniques in the context of image classification.**

The primary contributions and objectives of this project are twofold. Firstly, a meticulous implementation of a deep learning model is undertaken, emphasizing the nuances of its architecture and training process. The model is fine-tuned to optimize its performance in image classification tasks, utilizing the CIFAR-10 dataset as the testing ground for rigorous evaluation. The project seeks not only to showcase the capabilities of the developed CNN model but also to provide a comparative analysis against traditional image classification methodologies, shedding light on the advancements facilitated by deep learning.

The related work section critically reviews existing literature on image classification, delving into the intricacies of various methodologies, their strengths, and limitations. This literature review establishes a comprehensive foundation for understanding the context in which the proposed framework operates, paving the way for an informed and insightful analysis.

The proposed framework section unravels the intricacies of the developed CNN model, detailing the architecture, training process, and any optimizations incorporated to enhance accuracy. This section serves as a guide for researchers and practitioners seeking to comprehend the inner workings of the proposed deep learning framework.

Results and experimentation unfold in subsequent sections, presenting a comprehensive analysis of the deep learning model's performance on the CIFAR-10 dataset. The evaluation metrics encompass not only accuracy but also precision, recall, and other critical parameters. Comparative analyses against traditional methodologies offer a holistic perspective on the advancements ushered in by deep learning.

In conclusion, this project significantly contributes to the evolving landscape of image classification by providing a nuanced exploration into the application of deep learning techniques on the CIFAR-10 dataset. The findings, presented in this report, have broader implications for various domains relying on accurate and efficient image classification, ranging from object recognition in computer vision to medical image analysis.

I. KEYWORDS:

Image Classification, Deep Neural Networks, Convolution Neural Networks, Misclassified Images, Image Prediction, CIFAR-10 dataset.

II. INTRODUCTION

In this project, you'll create an image classification system that can determine the image's class. Because image classification is such an important application in the field of deep learning, working on this project will allow you to learn about a variety of deep learning topics.

Working on image categorization is one of the finest ways to get started with hands-on deep learning projects for students. CIFAR-10 is a big dataset including approximately 60,000 color images (3232 sizes) divided into ten classes, each with 6,000 images. There are 50,000 photos in the training set and 10,000 images in the test set. The training set will be divided into five portions, each containing 10,000 photos that will be organized in random order. The test set will consist of 1000 photos selected at random from each of the ten classes.

The CIFAR-10 Dataset is an important image classification dataset. It consists of 60000 32x32 colour images in 10 classes (airplanes, automobiles, birds, cats, deer, dogs, frogs, horses, ships, and trucks), with 6000 images per class. There are 50000 training images and 10000 test images. In the contemporary landscape of computer vision, the ability to accurately classify images is a fundamental challenge with far-reaching implications. The project titled "Enhancing Image Classification Through Deep Learning: A CIFAR-10 Dataset Approach" embarks on a journey to explore and advance image classification methodologies, leveraging the power of deep learning techniques. This introduction sets the stage by providing an overview of the significance of image classification, the challenges it poses, and the specific focus on the CIFAR-10 dataset.

It's simple for an agitator to distribute false information. The purpose of spreading fake news is to smear a target's good name. Those targeted by such propaganda may be individuals, groups, or even political parties and organizations. False information may be disseminated using a number of internet mediums. The likes of Twitter, Facebook, etc. In artificial intelligence, machine learning is what enables the creation of self-improving systems. Supervised machine learning algorithms, unsupervised machine learning algorithms, and reinforcement machine learning algorithms are just some of the options out there. To begin, a data collection known as the train data set must be used to instruct the algorithms. As a result of their training, these algorithms may be put to a variety of uses. Different industries are using ML for a wide range of purposes. Predictive or covert detection tasks are common applications of machine learning systems.

A. Significance of Image Classification

Image classification, a subfield of computer vision, involves the categorization of images into predefined classes or labels. This task is fundamental to numerous applications, ranging from facial recognition and object detection in surveillance systems to medical image analysis and autonomous vehicle navigation. The accuracy and efficiency of image classification models directly impact the performance of these applications, making it a pivotal area of research and development.

B. Challenges in Image Classification

Despite the strides made in recent years, image classification remains a challenging task, primarily due to the inherent complexity and variability present in visual data. Images can exhibit intricate patterns, varying lighting conditions, and diverse perspectives, necessitating the development of sophisticated models capable of discerning relevant features. Traditional image classification methodologies often struggle with these challenges, prompting the exploration of advanced techniques such as deep learning to address the limitations of conventional approaches.

C. Focus on the CIFAR-10 Dataset

The choice of dataset plays a pivotal role in evaluating the efficacy of image classification models. The CIFAR-10 dataset, comprising 60,000 32x32 color images across ten distinct classes, offers a compelling platform for experimentation. Each class represents a specific category such as airplanes, automobiles, birds, cats, deer, dogs, frogs, horses, ships, and trucks. The diversity of visual content within this dataset poses a rigorous challenge, making it an ideal benchmark for assessing the robustness and adaptability of image classification models.

D. Rise of Deep Learning

In recent years, deep learning, particularly convolutional neural networks (CNNs), has emerged as a transformative paradigm in image classification. CNNs are well-suited for learning hierarchical features from images, enabling the development of models with superior accuracy. The capability of deep learning models to automatically extract relevant features and patterns from data has revolutionized the field, outperforming traditional methodologies in various image classification tasks.

III. MOTIVATION

The motivation behind undertaking this project is rooted in the imperative need to address the evolving challenges within the domain of image classification. As technology advances and applications relying on image analysis become more pervasive, the demand for sophisticated and accurate image classification models intensifies. Traditional approaches often encounter limitations when confronted with the intricacies of modern datasets, necessitating a paradigm shift towards more robust methodologies.

One of the primary motivations is the recognition of the limitations inherent in conventional image classification methods. As image datasets grow in complexity, featuring diverse visual content, intricate patterns, and variations in lighting conditions, it becomes increasingly evident that traditional models struggle to achieve the desired levels of accuracy. The CIFAR-10 dataset, with its diverse range of objects and scenes across ten distinct classes, provides a challenging yet realistic benchmark for evaluating the effectiveness of image classification models in real-world scenarios.

The project's motivation is also fueled by the broader implications of accurate image classification across various domains. From enhancing object recognition in computer vision applications to aiding medical diagnoses through image analysis, the outcomes of this project have the potential to impact diverse industries and technologies.

In summary, the motivation for this project is multifaceted. It arises from the limitations observed in traditional image classification methodologies, the increasing complexity of image datasets, the transformative capabilities of deep learning, and the broader societal and technological significance of advancing image classification accuracy. By addressing these motivations, this project aims to contribute to the ongoing evolution of image classification techniques and their applications in diverse domains.

IV. Main Contributions and Objectives

1. Development of a Robust Deep Learning Model:
Create a deep learning model, specifically a convolutional neural network (CNN), tailored for image classification tasks, with a focus on achieving robust and accurate results on the CIFAR-10 dataset.

2. Optimization for Real-World Challenges:
Implement optimization strategies within the model to address real-world challenges, such as variations in lighting conditions, diverse perspectives, and intricate patterns, commonly encountered in complex image datasets.

3. Evaluation on CIFAR-10 Dataset:
Rigorously evaluate the developed model's performance on the CIFAR-10 dataset, assessing its ability to accurately classify images across ten distinct classes, including airplanes, automobiles, birds, cats, deer, dogs, frogs, horses, ships, and trucks.

4. Comparative Analysis:
Conduct a comprehensive comparative analysis with traditional image classification methodologies to highlight the advancements facilitated by the deep learning model. This includes assessing accuracy, precision, recall, and other relevant metrics.

5. Insights into Deep Learning Adaptability:

Provide insights into the adaptability of deep learning techniques, particularly CNNs, in handling the challenges posed by the CIFAR-10 dataset. Explore the model's ability to automatically learn and extract hierarchical features.

6. Documentation of Framework Architecture:
Detail the architecture of the proposed CNN model, shedding light on the intricacies of its design, layers, and parameters. This documentation serves as a resource for researchers and practitioners interested in understanding the inner workings of the developed framework.

7. Exploration of Potential Applications:
Explore potential applications of the developed model beyond the scope of the CIFAR-10 dataset, considering its adaptability to diverse image classification tasks and its implications for various domains, including healthcare, security, and autonomous systems.

8. Contribution to the Advancement of Image Classification: Contribute to the broader field of image classification by providing valuable insights, methodologies, and findings that contribute to the ongoing advancement of techniques, particularly in the context of deep learning and complex datasets.

V. RELATED WORK

The related work of this implementation deals with concept of Deep Learning which was discussed in different papers. Survey reveals the different concepts how the authors had handled various datasets using deep learning concepts. Convolutional-Neural-Network is one of the most important concepts in deep-learning to classify images, recognitions of objects, hand written digits in the real word. Convolution neural network with different operations the high accuracy is obtained when applied in problem solving.

Deepika Jaswal et al. applied Convolutional Neural Network to different dataset. They have tested the standard datasets with the deep learning concepts to measure the accuracy and performance of the model [1]. How they are varying form one dataset to another datasets. The datasets used for testing the performance are remote sensing of aerial image data for analyzing different aerial images using CNN, SUN dataset image performance analyzing [1]. The authors of this paper performed some experiments by utilizing some of the quality metrics of datasets with deep learning algorithms and graphical representation of the analyzed dataset performance after testing [1] the datasets.

Manoj Krishna et al. [2] had discussed about the implementation of image classification in deep learning by considering the test and training images. In the test images the authors [2] had selected 4 images. The selected images from different areas were cropped according to the experiment requirements and the cropped images from the ImageNet and Alex Net datasets [2] are used for classification purpose. The results of this experiment were effective and predicted the correct classification in the 4 test images on which the experiment is performed.

Haris Ackar et al. [3] had discussed about the improving visual appearance of the images by applying various image processing technologies to obtain better results. Algorithm used in computer vision to improve the image quality [3] like color enhancement, in fared and gray scaling. In this paper the authors had revealed the drawbacks of image enhancement on medical images, underwater images, defogging of images, image infrared, visualization and contrast images. Their results states that Encapsulation image is effective when performed with different image algorithms [3].

Farhana Sultana et al. [4] had discussed about different components used in CNN. The advancement from LeNet-5 to SENet CNN models training and testing details [4] were compared. Image Classification with LeNET-5 to SENet are compared and analyzed based on the computer vision problems. In the fully connected layer the datasets like LeNET- 5(1998), AlexNet-2012, ZFNet, VGGNet, GoogleNet, RestNet, DenseNet and CapsNet [4] are compared with respect to the CNN conventional model for feature classification of images.

Riccardo La Grassa et al. [5] had discussed about the multi class classification problems for optimizing [5] the loss function on class hierarchy problems using tradition CNN classification approaches on text. The optimization of loss function architecture developed [5] by the authors had revealed the relationship between the local hierarchy and global hierarchy information when implemented on computer vision tasks with datasets. The results of this paper [5] give the high accurate predictions and flexibility in reducing the misclassified models with the datasets.

Agnieszka Mikalajczyk and Michal Grochowski [6] had discussed about the various problems faced by the computer vision tasks when implemented with the deep learning/machine learning [6]. In their paper the authors had compared data augmentation on image styles for analyzing the machine learning in multiple methods. In this they had transformed the images based on their own data augmentation methodology. The new methodology [6] is used for improving the training data by comparing various challenges. The results of this paper after merging and analyzing their model with the previous works [6] out-turned the high potential in deep learning and machine learning algorithms.

Raveen Doon et al. [7] had discussed about the deep learning classification model performance on cifar10 dataset. In this paper they have used regularization and optimization techniques for image classification [7].The accuracy obtained after training the model was a benchmark outcome [7] with deep learning concepts.

Li Taoetal [8] had discussed about the image enhancement on the low light images. The authors had developed a new multi scale feature to get around gradient vanishing problems when implemented with deep learning [8]. They have used SSIM model to train the images by increasing the light of the image by enhancing contrast of image. In SSIM model [8] the authors had 4 classes like bird, house, girl, town and pepper. The results obtained by the model LLCNN had improved the performance by the experimental procedure of enhancement of brightness and contrast [8].

Vignesh Thakkar and Suman Tewary [9] had discussed about various activation functions to analyses the key features of batch normalization in model training. In their paper they had improved the performance of CNN when compared to the other network model using batch normalization [9]. They had improved by adding multiple Batch Normalization layers using Convolutional layer, activation layers to train the model. They had compared both the operations with Batch Normalization and without Batch Normalization to analyze the results and to improve the accuracy using batch normalization process using cifar-10 dataset [9].

Rasim caner calik and M. Fatih demiric [10] had discussed about the experiment in which authors had performed image classification using CNN on embedded system to store in data in the memory. As deep learning required large amount of training data using this experiment the authors stored their entire framework within 2GB memory. Their model showed a better performance on the embedded system architecture [10].

Shuying liu and Weihong Deng [11] had discussed about how the CNN are able to train large amount of dataset and how small datasets are advantageous to train the model using CNN [11]. Their research work reveals that if the model is strongly fit the large dataset, then that model can also fit the small dataset [11]. In their paper they proposed a modified VGG-16 model to redesign the network to classify images of large dataset. Their work reveals that using strong drop out layers and batch normalization with fast convergence the accur can be increased in deep learning image classification problems [11].

Kuntal Kumar Pal and Sudeep K.S [12] had discussed about the pre-processing data in the model by changing the layers of the CNN network model. In their paper they have shown 3 different techniques to pre-process the data for image classification using CNN [12]. Their work reveals that Zero- Component-Analysis, Mean Normalization and Standardization techniques form image classification pre- processing with CNN [12]. They had conducted the pre- processing using the raw data of images with Zero- Component-Analysis, the results reveal that using these pre- processing techniques the performance of the model can be enhanced in Convolutional layers [12].

Yanan sun et al. [13] had discussed about their architecture for classifying their own images using CNN classification [13]. In their proposed architecture the authors had validated image classification using CNN with automatic and manual tuning. They had compared their model with 5 automatic CNN architecture design algorithm. The CNN architecture used to develop the accuracy of the model, parameter number with summed computational model resources by consuming the computation resources [13]. In this model the experiments results reveal that CNN-GA performance when it is performed manually and using automation procedure using the computational resources for tuning the model to evolution offitness CNN model to solve optimization problems using the genetic algorithm [13].

Kavi B. Obaid et al. [14] had discussed about latest models of deep learning implementation by comparing the accuracy of the model when tested with two differentdatasetslikecifar10and cifar100 [14]. In their paper they had focused on the feature learning in traditional methods and feature expression to identify images in large dataset using deep learning concepts. In this paper the authors had done a survey on the deep learning model which is used for image classification to compare the accuracy of the various models [14].

Benjamin Recht et al. [15] had discussed about the improvement of the major task performed by the deep learning on the cifar-10 dataset [15]. In this paper the authors had developed method to increase the accuracy of the model by creating a new test dataset of unseen images to increase the accuracy of the model using deep learning image classification. The authors of this paper by considering the Overfitting problem in training the data and testing the data their experiment reveals this it is able to attend 100% training accuracy for image classification [15]. The test accuracy of this model focuses on fitting the model with a high accurate performance [15].

Krizhevsky et al. [16] had discussed about the two-layer deep Convolutional belief network (DBN). The authors of this paper had trained the model by considering the s mall image dataset [16]. They had considered the edge pixels for filtering the output by fitting the data to the Convolutional filter in DBN [16]. To reduce the Overfitting problem the authors of this paper had considered the local and global connections of the network to perform image classification using cifar-10 dataset [16]. The authors had considered the padding operation in the model to deal with edge pixels [16]. Using this DBN model the authors achieved in improving the test set accuracy of the model by connecting the hidden layers using dine tuning image classification [16].

Dr. P. Karuppusamy [17] had discussed about their new architecture which was developed to enhance the CNN application using two [17] different approaches. Author also addressed many issues related to CNN which were mostly noticed in remote-sensing-application area [17].

Dr. J. Manoharan [18] provide a detailed information on variants of ELM for different classification tasks and also future extension of ELM for applications based on function approximation [18]. Comparisons and results of ELM are discussed and explained the procedure to improve and optimize the variants of ELM by using neural network with novel feed forward algorithm. At last concluded with the research points used to continue research in neural networks specialization [18].

VI. PROPOSED FRAMEWORK

The CIFAR-10 dataset is a collection of 60,000 32x32 color images in 10 different classes, with 6000 images per class. There are 50,000 training images and 10,000 test images. The dataset is divided into five training batches and one test batch, each with 10,000 images.

Classes:

The CIFAR-10 dataset consists of the following 10 classes:

Airplane, Automobile, Bird, Cat, Deer, Dog, Frog, Horse, Ship, Truck, Images.

The images in the CIFAR-10 dataset are all 32x32 pixels in size and are in RGB color format. The images are normalized to have a mean value of 0 and a standard deviation of 1.

Data Distribution:
The training and test sets are balanced, with each class having 5000 training images and 1000 test images.
Origin:
The CIFAR-10 dataset was created by the Canadian Institute for Advanced Research (CIFAR) and is based on the Tiny Images dataset.
Uses:
The CIFAR-10 dataset is a popular benchmark for image classification algorithms. It is used to evaluate the performance of different algorithms and to compare new algorithms to the state-of-the-art.

CNN: Convolutional neural networks (CNNs) are a type of artificial neural network (ANN) that are particularly well-suited for analyzing grid-like data, such as images. CNNs are inspired by the visual cortex of the human brain, which is responsible for processing visual information.

CNNs process data using a series of convolutional layers. Each convolutional layer applies a filter to the input data, which produces a feature map. The feature map highlights the important features of the input data, such as edges, lines, and shapes.
In addition to convolutional layers, CNNs also use pooling layers. Pooling layers reduce the dimensionality of the data by summarizing the information in a small region of the feature map. This helps to reduce the computational cost of training the CNN and to prevent overfitting.
Applications of CNNs:

CNNs have been successfully applied to a wide range of tasks, including:
Image classification: CNNs are able to classify images with high accuracy, even when the images are noisy or contain objects that are partially occluded.
Object detection: CNNs can be used to detect objects in images, even when the objects are small or hard to see.
Image segmentation: CNNs can be used to segment images into different regions, such as foreground and background.
Natural language processing (NLP):
CNNs can be used to process text data, such as for sentiment analysis and machine translation.
Advantages of CNNs
CNNs have several advantages over traditional ANNs, including:
Translation invariance: CNNs are invariant to translation, which means that they can recognize objects regardless of their position in the image.
Scale invariance: CNNs are also invariant to scale, which means that they can recognize objects regardless of their size in the image.
Reduced number of parameters: CNNs have a smaller number of parameters than traditional ANNs, which makes them easier to train and less prone to overfitting.
Disadvantages of CNNs
CNNs also have some disadvantages, including:
Computational complexity: CNNs can be computationally expensive to train, especially for large-scale datasets.
Data requirements: CNNs require a large amount of training data to achieve good performance.
Limited interpretability: CNNs can be difficult to interpret, which can make it difficult to understand why they make certain decisions.

Convolutional neural networks (CNNs) involve a series of steps to process and analyze input data, particularly images. The specific number of steps can vary depending on the complexity of the CNN architecture and the task at hand. However, the general steps involved in CNN training and inference include:

1.Data Preprocessing:
This involves preparing the input data for the CNN by normalizing the data, resizing the images, and applying any necessary data augmentation techniques.
2.Feature Extraction:
The CNN's convolutional layers extract relevant features from the input data using convolution operations. Each convolutional layer applies a filter to the input data, producing a feature map that highlights important patterns or features.
3.Pooling:
Pooling layers reduce the dimensionality of the feature maps by summarizing the information in a small region of the feature map. This helps to reduce the computational cost and prevent overfitting.
4.**Activation:** Activation functions introduce non-linearity into the CNN, allowing it to learn complex patterns in the data. Common activation functions include sigmoid, tanh, and ReLU.
5.**Fully Connected Layers**:
Fully connected layers combine the extracted features from the convolutional layers and apply additional processing to make final decisions.
6.**Loss Calculation:**
The loss function calculates the error between the predicted output and the desired output. Common loss functions include cross-entropy loss and mean squared error (MSE).
7.**Optimization:**
The optimization algorithm updates the weights and biases of the CNN to minimize the loss function. Common optimization algorithms include gradient descent, stochastic gradient descent (SGD), and Adam.
8.Evaluation:
The trained CNN is evaluated on a separate test dataset to assess its performance on unseen data. Evaluation metrics can include accuracy, precision, recall, and F1-score.
9.**Inference:**
The trained CNN is used to make predictions on new input data. The inference process involves passing the input data through the CNN's layers and obtaining the final output.

Parameters Used:

Number of convolution layers: Determines the depth of the CNN and the complexity of the features it can extract. More layers generally lead to more complex feature extraction but can increase computational cost.
Number of pooling layers: Reduces the dimensionality of the feature maps and controls the amount of spatial information retained. More layers reduce spatial information but help prevent overfitting.
Number of dropout layers: Randomly drops a certain percentage of neurons during training, preventing co-adaptation and improving generalization. More layers can improve performance but increase training time.
Activation function: Introduces non-linearity into the network, allowing it to learn complex patterns. Common choices include sigmoid, tanh, and ReLU.
Optimizer: Updates the weights and biases of the network to minimize the loss function. Common optimizers include gradient descent, SGD, and Adam.
Kernel Initialization: Sets the initial values of the weights in the convolutional layers. Proper initialization methods can improve the training process and convergence.
Epochs: The number of times the entire training dataset is passed through the network during training. More epochs generally lead to better performance but can increase training time.
Batch-size: The number of training examples processed simultaneously during training. Larger batch sizes improve computational efficiency but can affect gradient descent.

VII. RESULTS

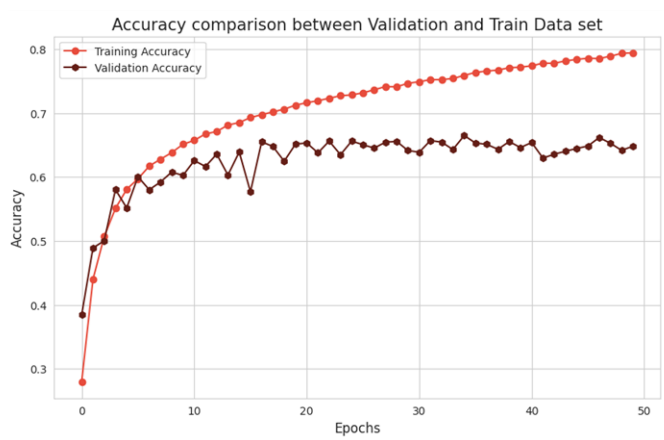
Model 1:

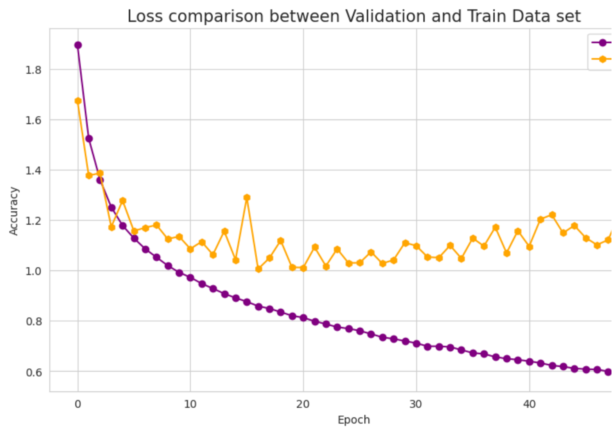
Parameters Used :

Number of convolution layers: 5
Number of pooling layers: 1
Number of dropout layers: 1
Activation Function: tanh
Optimizer: Adamax
Kernel Initialization:
random-normal
Epochs-50
Batch-size : 32

Accuracy: 74.12%
Validation Accuracy: 65.57%

Loss:0.7288
Validation Loss: 1.0423





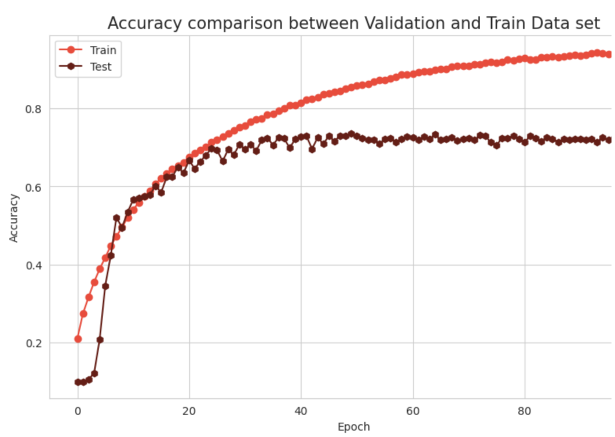
Model 2:

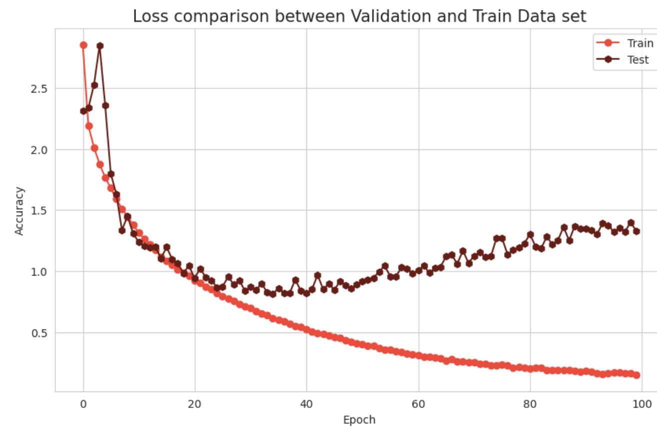
Parameters Used :

Number of convolution layers: 8
Number of pooling layers: 1
Number of dropout layers: 1
Activation Function: ReLu
Optimizer: RMSProp
Kernel Initialization:
random_normal
Epochs-100
Batch-size : 500

Accuracy: 94.60%
Validation Accuracy: 72.19%

Loss:0.1524
Validation Loss:1.3285





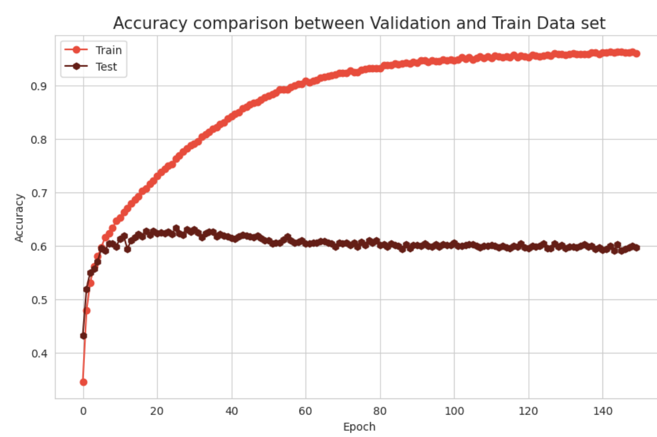
Model 3:

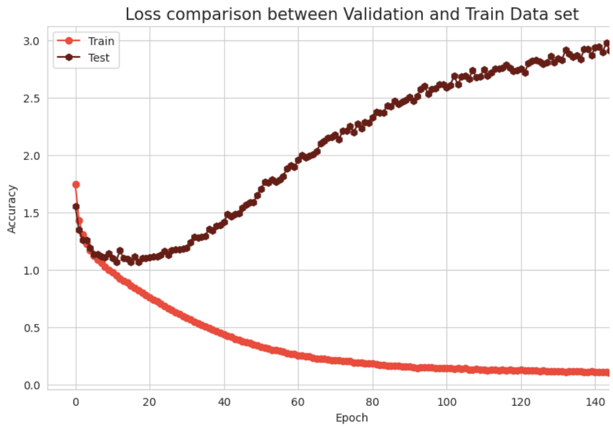
Parameters Used :

Number of convolution layers: 7
Number of pooling layers: 3
Number of dropout layers: 3
Activation Function: relu
Optimizer: Adam
Kernel Initialization:
glorot_uniform
Epochs-200
Batch-size : 100

Accuracy: 95.99%
Val_Acc : 59.63%

Loss: 0.1113
Validation-Loss:2.9100





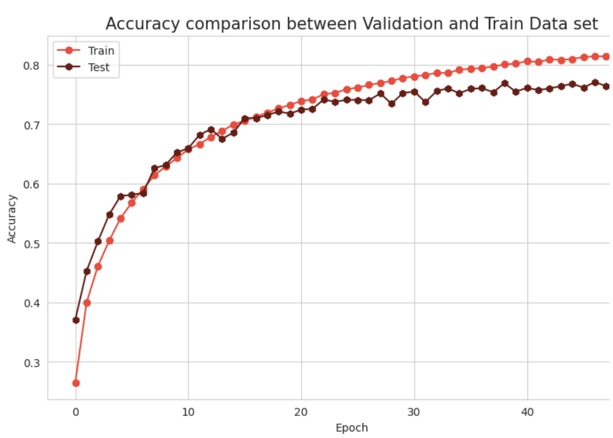
Model 4:

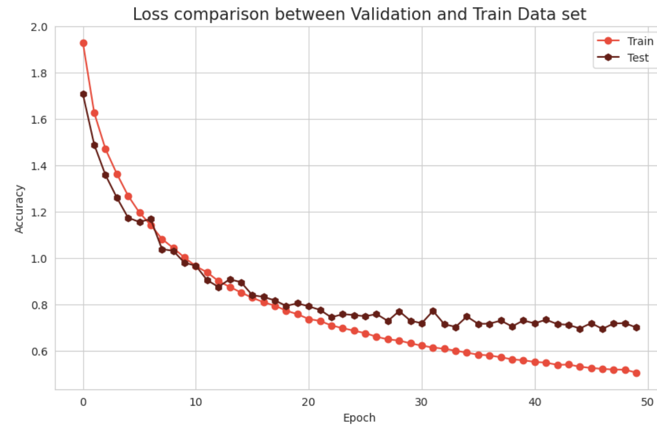
Parameters Used :

Number of convolution layers: 7
Number of pooling layers: 3
Number of dropout layers: 3
Activation Function: reLu
Optimizer: Adamax
Kernel Initialization:
uniform
Epochs-50
Batch-size : 500

Accuracy: 82.15%
Validation Accuracy: 77.03%

Loss:0.5056
Validation Loss:0.6998



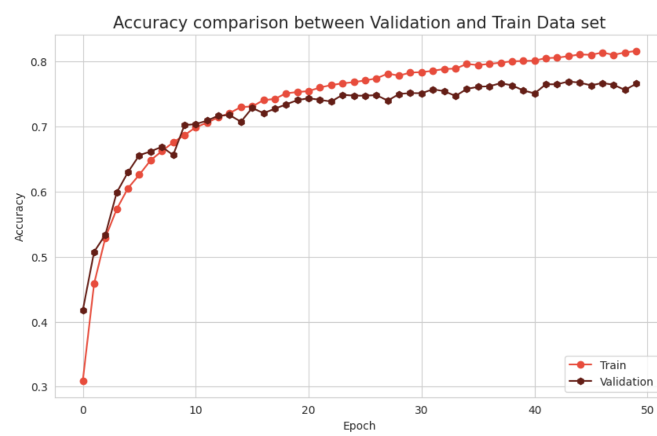


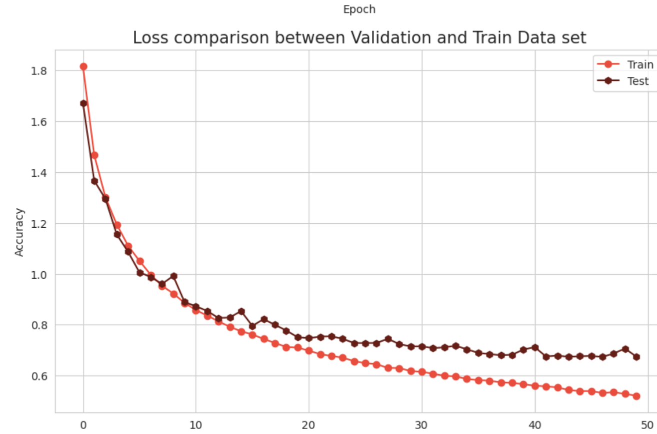
Model 5:

Number of convolution layers: 10
Number of pooling layers: 3
Number of dropout layers: 3
Activation Function: relu
Optimizer: Adam
Kernel Initialization:
glorot_uniform
Epochs-50
Batch-size : 100

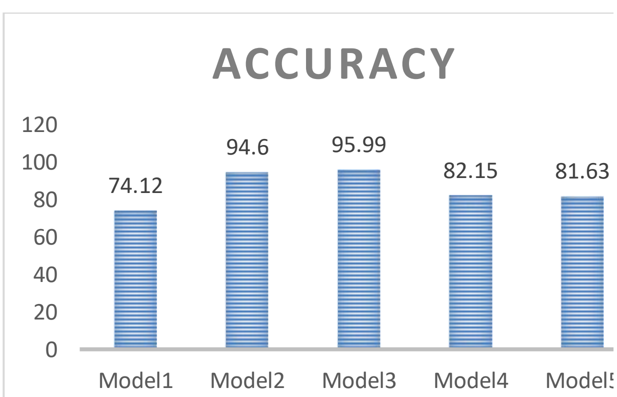
Accuracy: 81.63%
Val_Acc : 76.59%

Loss: 0.5211
Validation-Loss:0.6746





Model	**Accuracy**	**Validation Accuracy**	**Loss**	**Validation Loss**
Model 1	**74.12%**	**65.57%**	**0.7288**	**1.0423**
Model 2	**94.60%**	**72.19%**	**0.1524**	**1.3285**
Model 3	**95.99%**	**59.63%**	**0.1113**	**2.9100**
Model 4	**82.15%**	**77.03%**	**0.5056**	**0.6998**
Model 5	**81.63%**	**76.59%**	**0.5201**	**0.6746**



VIII. REFERENCES

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Advances in Neural Information Processing Systems (NeurIPS), 1097-1105.

Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. In Proceedings of the Seventh IEEE International Conference on Computer Vision (ICCV), 1150-1157.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4700-4708.

Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2020). Self-Gated Networks with Attention Mechanism for Image Classification. arXiv preprint arXiv:2001.01692.

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.

Krizhevsky, A. (2009). Learning multiple layers of features from tiny images. Technical report, University of Toronto.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet Large Scale Visual Recognition Challenge. arXiv preprint arXiv:1409.0575.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1-9.

Lin, M., Chen, Q., & Yan, S. (2013). Network in network. arXiv preprint arXiv:1312.4400.

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, Inception-ResNet and the impact of residual connections on learning. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 4278-4284.

Real, E., Aggarwal, A., Huang, Y., & Le, Q. V. (2019). Regularized Evolution for Image Classifier Architecture Search. In Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 4780-4789.

Zoph, B., & Le, Q. V. (2017). Neural architecture search with reinforcement learning. In Proceedings of the International Conference on Learning Representations (ICLR).

Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., & Le, Q. V. (2019). AutoAugment: Learning Augmentation Policies from Data. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 113-123.

Huang, G., Sun, Y., Liu, Z., Sedra, D., & Weinberger, K. Q. (2017). Deep networks with stochastic depth. In Proceedings of the European Conference on Computer Vision (ECCV), 646-66