## SUMMER TRANING PROJECT REPORT

## ON

**“Image recognition using CNN on CIFAR 10 Dataset”**

## Submitted in the partial fulfillment for the award of Degree of Bachelor in

## BBA (CAM): 2021-24

**Under the Guidance: Submitted by:**

Ms. Aditi Kaushik Shivam Chaurasia

Assistant Professor, CPJCHS 00324201921

Batch: 2021-24



## CHANDERPRABHU JAIN COLLEGE OF HIGHER STUDIES & SCHOOL OF LAW

**An ISO 90012015 Certified Institute (Approved by the Govt. Of NCT of Delhi Affiliated to**

**Guru Gobind Singh Indraprastha University, Delhi and Approved by Bar Council of India)**

**Plot No OCF Sector A8, Narela, New Delhi 110040**

# CERTIFICATE

This is to certify that Summer Training Project Report entitled “Image recognition using CNN on CIFAR10 Dataset” which is submitted by Shivam Chaurasia in partial fulfillment of the requirement for the award of degree BBA (CAM) to GGSIP University, Dwarka, Delhi is a record of the candidate own work carried out by him under my supervision.

**Date: Supervisor Signature**

# ACKNOWLEDGEMENT

I offer my sincere thanks and humble regards to Chander Prabhu Jain College of Higher Studies & School of Law, GGSIP University, New Delhi for imparting us very valuable professional learning in summer training Project Report of BBA (CAM).

I pay my gratitude and sincere regards to Ms. Aditi Kaushik, my project Guide, for imparting her knowledge. I am thankful to her as she has been a constant source of advice, motivation and inspiration. I am also thankful to her for giving her suggestions and encouragement throughout the project work.

I take the opportunity to express my gratitude and thanks to our computer Lab staff and library staff for providing me with the opportunity to utilize their resources for the completion of the project. I am also thankful to my family and friends for constantly motivating me to complete the project and providing me with an environment which enhanced my knowledge.

**Student’s Signature**

Executive Summary

This study delves into the realm of image recognition using Convolutional Neural Networks (CNNs) with a specific focus on the CIFAR10 dataset. CIFAR10 is a dataset containing 60,000 32x32 color images distributed across ten diverse classes. The dataset's complexity and variety make it a valuable tested for image classification tasks, motivating the need for robust and reliable CNN models.

The research commences by designing and training a CNN model, optimizing it for the recognition of objects and scenes depicted in the CIFAR10 images. Despite meticulous model development and training procedures, an intriguing discovery was made during initial testing: the model provided incorrect classifications for some images, only to subsequently deliver correct answers when the same tests were repeated.

This observation raises pertinent questions regarding the stability and reproducibility of deep learning models. The phenomenon underscores the need for thorough model evaluation, parameter tuning, and consideration of factors that may impact model convergence. The study implies that machine learning practitioners should exercise caution and diligence when interpreting model performance results, especially in cases where model accuracy is on the borderline between correct and incorrect classifications. Further research is warranted to unravel the underlying causes of such variability and to develop strategies for ensuring consistent and dependable image recognition using CNNs.

# TABLE OF CONTENT

|  |  |  |
| --- | --- | --- |
| **S. NO.** | **CONTENT** | **PAGE NO.** |
| **1.** | **CHAPTER 1: INTRODUCTION**   * 1. Introduction of Company “AcmeGrade”   2. Image Recognition using CNN on CIFAR   3. Convolutional Neural Network (CNN)   4. How CNNs work   5. CIFAR10 | 1  1-3  3  3-6  7 |
| **2.** | **CHAPTER 2: INDUSTRY PROFILE** | 8-10 |
| **3.** | **CHAPTER 3: RESEARCH METHODOLOGY**   * 1. Introduction   2. Problem and Objectives   3. Literature Review   4. Data Collection   5. Preprocessing   6. Model Selection   7. Model Training   8. Model Evaluation | 11  11  12  12  13  14  14  15 |
| **4.** | **CHAPTER 4: ANALYSIS AND INTERPRETATION** | 16-21 |
| **5.** | **CHAPTER 5: SUMMARY AND CONCLUSION**   * 1. Summary   2. Conclusion | 22  22-23 |
| **6.** | **REFERENCES** | 24 |

CHAPTER: 1

Introduction

* 1. Introduction of “AcmeGrade”

AcmeGrade is a private limited company located in Bangalore, Karnataka, India. It was incorporated on August 5, 2021, and has an authorized share capital of INR 1.00 lac and a total paid up capital of INR 50,000.00.

AcmeGrade offers supervised internships and creative industry relevant projects to help young people learn and up skill themselves. They have internship programs in various fields such as computer science engineering, information technology, mechanical engineering, civil engineering, digital marketing, finance, and human resources.

* 1. “Image Recognition Using CNN”

Image recognition using Convolutional Neural Networks (CNN) is a deep learning technique for identifying and classifying objects or patterns within digital images. CNNs have revolutionized the field of computer vision and are widely used in applications such as facial recognition, autonomous vehicles, medical image analysis, and more. In this explanation, I'll provide a detailed overview of how CNNs work for image recognition.

1. Convolutional Layer: CNNs start with convolutional layers. These layers apply filters or kernels to the input image. Filters are small grids of numbers that slide over the image to detect patterns or features like edges, corners, or textures. Each filter creates a feature map, which highlights a particular feature in the input image. Multiple filters can be applied in parallel to capture different features.
2. ReLU Activation: After each convolution operation, a rectified linear unit (ReLU) activation function is applied element wise to the feature maps. ReLU introduces nonlinearity, allowing the network to learn complex patterns.
3. Pooling Layer: Pooling layers are used to down sample the feature maps. They reduce the spatial dimensions while retaining important information. Max pooling is a common technique where the maximum value in a local region is retained, and the rest are discarded. This helps in reducing computational complexity and improving translation invariance.
4. Multiple Convolution and Pooling Layers: CNNs consist of multiple convolution and pooling layers stacked together. As the network progresses through these layers, it captures increasingly abstract and complex features.
5. Flattening: The output of the convolutional and pooling layers is a three dimensional tensor. Before feeding it into a fully connected neural network, this tensor is flattened into a one dimensional vector.
6. Fully Connected Layers: These layers are similar to those in a traditional feed forward neural network. They take the flattened vector as input and perform classification or regression tasks. Fully connected layers combine the information learned from previous layers to make predictions.
7. Softmax Activation: For image classification, the output layer typically uses a softmax activation function. Softmax converts the network's raw scores into class probabilities. The class with the highest probability is considered the network's prediction.
8. Loss Function: CNNs use a loss function, such as cross entropy, to quantify the difference between the predicted probabilities and the true labels. The goal during training is to minimize this loss function, typically using gradient descent or its variants.
9. Back propagation and Training: The network's weights are updated using back propagation, which computes gradients of the loss with respect to the network's parameters. Training involves iteratively adjusting the weights using an optimization algorithm like stochastic gradient descent.
10. Regularization Techniques: To prevent over fitting, various techniques like dropout, batch normalization, and weight decay are applied to the network.
11. Data Augmentation: Data augmentation techniques, such as rotation, scaling, and flipping, are used to increase the diversity of the training data and improve the network's generalization.
12. Transfer Learning: In practice, pertained CNN models on large datasets (e.g., Image Net) are often used as a starting point. These models can be fine-tuned for specific image recognition tasks, saving significant training time and resources.
13. Inference: Once the CNN is trained, it can be used for image recognition tasks by feeding it new images, and it will produce predictions based on the learned features.

CNNs have been highly successful in image recognition because they automatically learn hierarchical features from images, making them robust and effective for a wide range of computer vision tasks. The architecture and hyper parameters of CNNs can be tailored to specific image recognition tasks, making them versatile tools in the field of deep learning and computer vision.

* 1. **Convolutional Neural Network (CNN)**

A Convolutional Neural Network (CNN) is a type of artificial neural network designed for tasks involving visual data, such as image and video recognition. It is particularly powerful in capturing spatial hierarchies of features through the application of convolutional and pooling layers.

* **Definition:**

A Convolutional Neural Network is a class of deep neural networks, most commonly applied to analyzing visual data. It is specifically designed to automatically and adaptively learn spatial hierarchies of features from input images through the use of convolutional and pooling layers.

* 1. **How CNNs Work:**

CNNs are effective in capturing hierarchical representations of features in images, making them well suited for tasks such as image classification, object detection, and segmentation. The convolutional layers enable the network to automatically learn relevant filters, reducing the need for manual feature engineering. This adaptability makes CNNs a powerful tool for various computer vision applications.

Convolutional Neural Networks (CNNs) are a type of deep neural network designed for processing structured grid data, such as images. They have proven to be highly effective in tasks like image recognition, object detection, and image segmentation. Here's a detailed overview of how CNNs work:

1. **Input Layer:**

* Input Data

CNNs take as input multidimensional data, typically in the form of images represented as grids of pixel values.

The dimensions of the input are typically height, width, and channels (e.g., red, green, and blue in the case of a

Color image).

1. **Convolutional Layer:**

* Convolution Operation

The core building block of a CNN is the convolutional layer. Convolution involves sliding a small filter (also called a kernel) over the input image and computing the dot product of the filter weights and the pixel values in the receptive field. This operation is performed at multiple positions across the entire image, producing a feature map.

* Activation Function:

Each element in the feature map is then passed through a nonlinear activation function (commonly ReLU) to introduce nonlinearity.

1. **Pooling Layer:**

* Pooling Operation

Pooling layers are used to reduce the spatial dimensions of the input volume. Max pooling is a common technique where, for each region in the feature map, only the maximum value is retained, discarding the rest. This reduces computation and helps make the network more robust to variations in scale and orientation.

1. **Flattening:**

* Flattening the Output

After several convolutional and pooling layers, the spatial dimensions are reduced, and the network produces a 3D volume of activations. To connect this to a fully connected layer, these activations are flattened into a vector.

1. **Fully Connected (Dense) Layer:**

* Neural Network Layers

The flattened vector is fed into one or more fully connected layers. These layers operate as a traditional neural network, connecting every input neuron to every output neuron with associated weights. Nonlinear activation functions (e.g., ReLU) are often applied to the outputs of these fully connected layers.

1. **Output Layer:**

* Final Predictions

The final fully connected layer typically has as many neurons as there are classes in the classification task. The output is often passed through a softmax activation function to convert the raw scores into probabilities. The class with the highest probability is the predicted class.

1. **Training:**

* Loss Function

During training, the network's predictions are compared to the actual labels using a loss function (e.g., cross entropy loss for classification tasks).

* Back propagation

The gradient of the loss with respect to the network parameters is computed using back propagation.

* Optimization

An optimization algorithm (e.g., stochastic gradient descent) is used to update the weights of the network in the direction that minimizes the loss.

1. **Repeat:**

* Stacking Layers

The convolutional, pooling, and fully connected layers can be stacked to form a deep architecture. Deeper networks can capture more complex features and hierarchical representations.

1. **Parameters:**

* Learnable Parameters

The parameters of the network, including the weights and biases, are learned during training through the optimization process.

1. **Testing/Prediction:**

* Forward Pass

During testing or prediction, new data is fed through the trained network using a forward pass to obtain predictions.

1. **Evaluation:**

* Performance Metrics

The predictions are evaluated using appropriate performance metrics for the specific task, such as accuracy, precision, recall, or F1 score.

In summary, CNNs use convolutional and pooling layers to automatically and adaptively learn spatial hierarchies of features from input data, enabling them to capture intricate patterns and representations in images. The combination of convolutional operations, nonlinear activation functions, and optimization during training allows CNNs to excel in tasks related to image processing and pattern recognition.

* 1. **CIFAR10 Dataset**

The CIFAR10 dataset is a well-known benchmark dataset in the field of computer vision and machine learning. It consists of 60,000 32x32 color images in 10 different classes, with 6,000 images per class. The dataset is split into 50,000 training images and 10,000 testing images.

The 10 classes in the CIFAR10 dataset are as follows:

1. Airplane
2. Automobile
3. Bird
4. Cat
5. Deer
6. Dog
7. Frog
8. Horse
9. Ship
10. Truck

Each image in the dataset is a low-resolution (32x32 pixels) color image, and the goal is typically to train a machine learning model, often a Convolutional Neural Network (CNN), to correctly classify these images into one of the 10 classes.

The CIFAR10 dataset is commonly used for educational purposes, benchmarking new machine learning models and algorithms, and as a standard dataset for testing the performance of image classification algorithms. Its relatively small size and simplicity make it a good starting point for experimenting with and understanding various computer vision techniques.

If you are working with machine learning or deep learning projects, CIFAR10 is a great dataset to use for practice and experimentation. Many deep learning frameworks and libraries provide easy access to this dataset for training and evaluation purposes.

CHAPTER: 2

Industry Profile

2.1 AcmeGrade

Acmegrade Private Limited is a Private incorporated on 05 August 2021. It is classified as Nongovt Company and is registered at Registrar of Companies, Bangalore. Its authorized share capital is Rs. 100,000 and its paid up capital is Rs. 50,000. It is involved in other computer related activities [for example maintenance of websites of other firms/ creation of multimedia presentations for other firms etc.]

Acmegrade Private Limited's Annual General Meeting (AGM) was last held on N/A and as per records from Ministry of Corporate Affairs (MCA), its balance sheet was last filed on N/A.

Directors of Acmegrade Private Limited are Rahul Kumar and Vybhav.

Acmegrade Private Limited's Corporate Identification Number is (CIN) U72900KA2021PTC150439 and its registration number is 150439.Its Email address is aurangabadrahul5@gmail.com and its registered address is 480/A 18TH A MAIN HSR LAYOUT 3RD SECTOR BANGALORE Bangalore KA 560102 IN.

Current status of Acmegrade Private Limited is Active.

2.2 Director Details

|  |  |  |
| --- | --- | --- |
| DIN | Director Name | Designation |
| 09270897 | Rahul Kumar | Director |
| 03270898 | Vybhav | Director |

2.3 Company Details

|  |  |
| --- | --- |
| CIN | U72900KA2021PTC150439 |
| Company Name | ACMEGRADE PRIVATE LIMITED |
| Company Status | Active |
| RoC | RoCBangalore |
| Registration Number | 150439 |
| Company Category | Company limited by Shares |
| Company Sub Category | Nongovt Company |
| Class of Company | Private |
| Date of Incorporation | 05 August 2021 |
| Age of Company | 2 years, 3 month, 8 days |
| Activity | Other computer related activities [for example maintenance of websites of other firms/ creation of multimedia presentations for other firms etc.] |
| Number of Members | 0 |

2.4 Financial Details

|  |  |
| --- | --- |
| Authorized Capital | 100,000 |
| Paid-up Capital | 50,000 |

2.5 Contact Details

|  |  |
| --- | --- |
| State | Karnataka |
| PIN Code | 560102 |
| Country | INDIA |
| Address | 480/A 18TH A MAIN HSR LAYOUT 3RD SECTOR BANGALORE BANGALORE KA 560102 IN |
| Email | aurangabadrahul5@gamil.com |
| Website | https://www.acmegrade.com/ |

CHAPTER: 3

Research Methodology

3.1 Introduction

Research Methodology refers to the systematic approach and procedures used in the study, analysis, and development of IT related solutions, systems, or knowledge. Research in IT can cover a broad range of topics, including software development, system design, network analysis, cyber security, artificial intelligence, and more. The methodology adopted in IT research guides the researchers through the process of investigation, experimentation, or problem-solving. Research methodology in the context of developing an image recognition model using Convolutional Neural Networks (CNN) involves outlining the systematic approach and procedures used to conduct the research.

Here's a general outline of the research methodology for creating an AI image recognition model:

3.2 Problem Definition and Objectives:

* Problems:

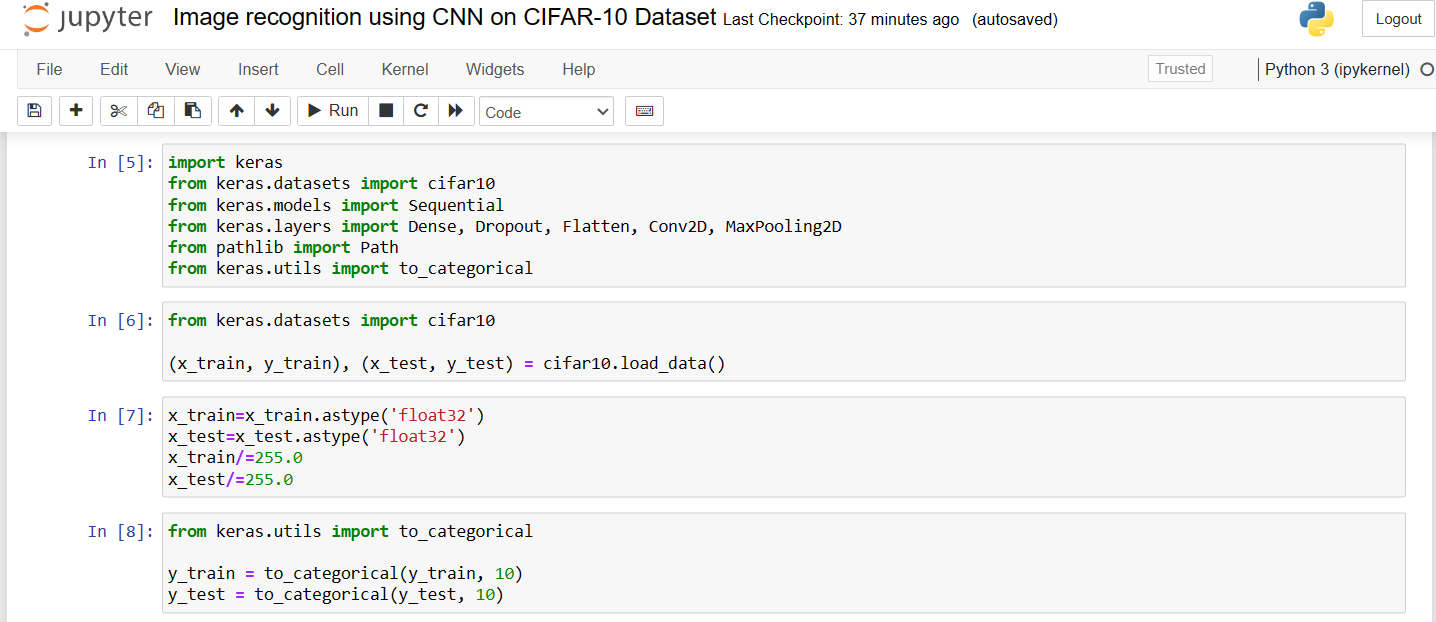
1. Unaware of Library files
2. Unaware of software (Anaconda, Jupyter Notebook)
3. Unaware of CIFAR10 Dataset
4. Unable to access zip files

* Objectives:

1. To learn how to make an Image recognition model using CNNs.
2. To learn and explore coding platform.

3.3 Literature Review:

In this Image Recognition Model, I am using libraries such as "TensorFlow," "Keras," "CIFAR10," etc., and the scripting language employed is "Python." I work on "Jupyter Notebook" with the support of the "Anaconda" software. Additionally, I use the "Command Prompt" for updating Python and installing libraries. However, I am encountering some issues while installing the TensorFlow library and extracting zipped files.



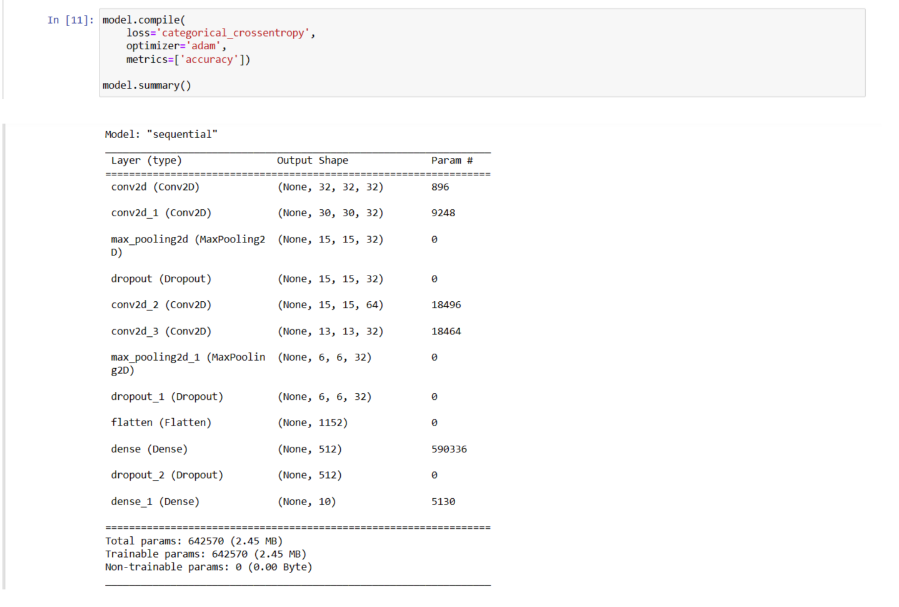
3.4 Data Collection:

I have downloaded CIFAR10 dataset online, and it is zipped file so that through command prompt, it is imported in the model training. Gather a diverse and representative dataset for training and testing your CNN model. Ensure that the dataset covers a wide range of scenarios and variations that your model might encounter in real-world applications. Properly annotate the data with labels for supervised learning.

3.5 Preprocessing:

Prepare and preprocess the dataset for training. This may involve tasks such as resizing images, normalizing pixel values, data augmentation to increase the diversity of the dataset, and handling any missing or corrupted data.



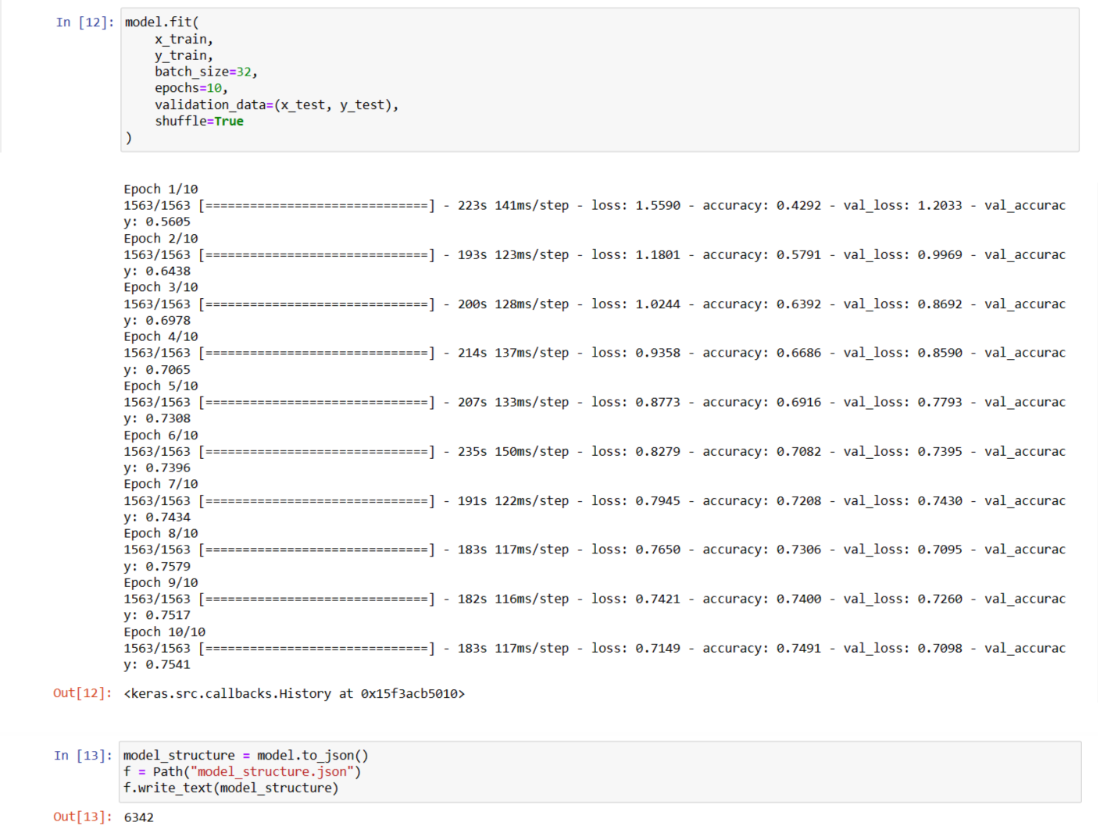


3.6 Model Selection:

Choose an appropriate CNN architecture for your image recognition task. This could involve selecting a pertained model or designing a custom architecture based on the complexity of your problem and available resources.

3.7 Model Training:

Train the selected CNN model on your prepared dataset. This involves feeding the model with labeled images and adjusting its weights through back propagation to minimize the prediction errors. Monitor the training process for convergence and avoid over fitting.



3.8 Model Evaluation:

Evaluate the performance of your trained model on a separate test dataset. Use metrics such as accuracy, precision, recall, and F1 score to assess how well the model generalizes to new, unseen data.



CHAPTER: 4

Analysis and

Interpretation

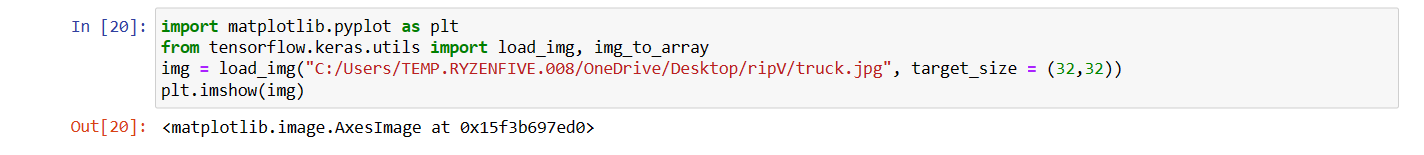
**4.1 Analysis and Interpretation**

Analysis refers to the process of examining and breaking down a complex entity or set of information into its individual components or parts. In various fields such as science, business, literature, and data science, analysis involves a systematic examination, investigation, or evaluation to understand the structure, function, or nature of the subject under study. The goal of analysis is often to gain insights, identify patterns, relationships, or trends, and make sense of the information in a more comprehensive and organized manner.

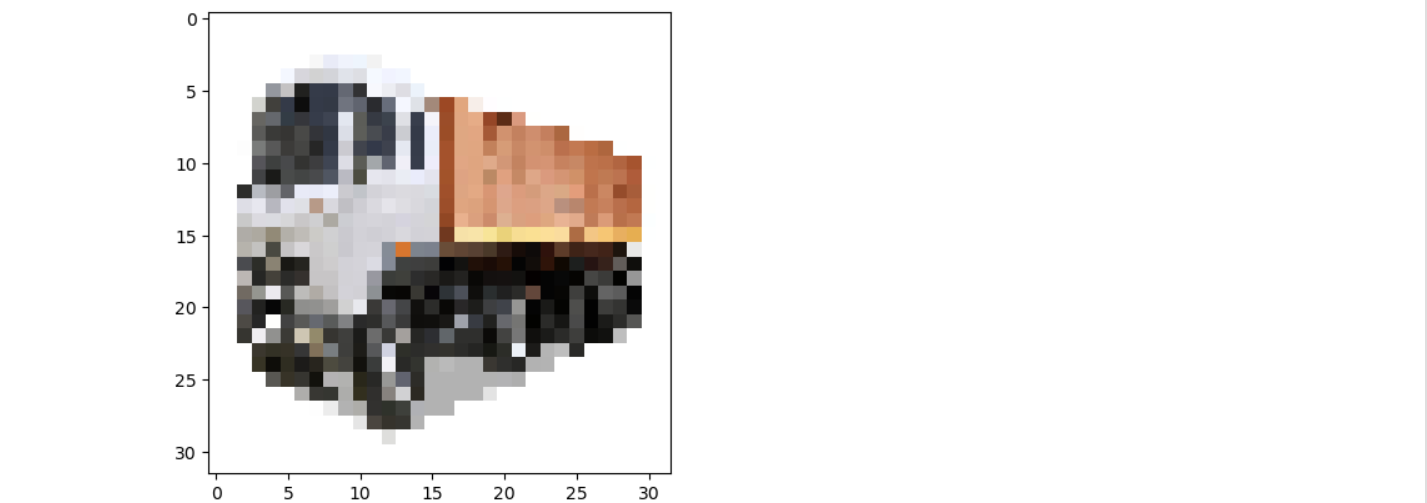
Interpretation involves explaining or assigning meaning to the results or findings obtained through analysis. It is the process of deriving significance, understanding implications, and providing context to the information that has been analyzed. Interpretation goes beyond the raw data or facts and involves making connections, drawing conclusions, and offering explanations. In various contexts, interpretation is subjective and relies on the expertise, background, or perspective of the interpreter.

After making model, Here I am calling an image for testing propose and let’s check it is working or not and working then analyzing its accuracy.

* **Test Number 1, Image of Truck**
* Here I am calling the image from device through its path and giving to the model.

****

* Here image is fetched from the Device to the model, and this is image of “truck” which is present in CIFAR10 dataset.

****

* Here it is analyzing the image’s characteristics and predicting the image that what is it. And It is predicted and print “This is an image is a truck likelihood” and Yes, it is true, it is a truck’s image.

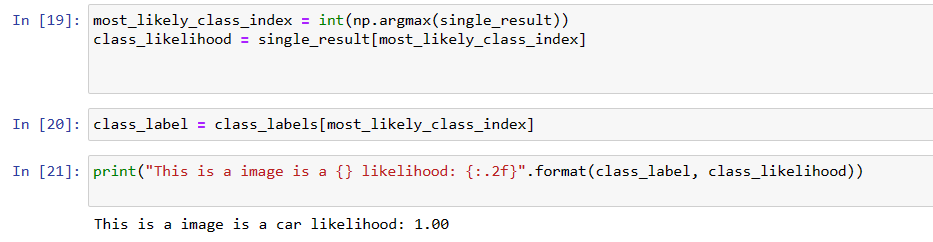
****

**Interpretation:** The correct prediction of a truck in the image by the CNN on the CIFAR10 dataset demonstrates the model's ability to effectively recognize and classify objects. This success highlights the CNN's capacity to learn and generalize from training data, identifying key features that enable accurate predictions. The application of such image recognition has broad implications, from autonomous vehicles to object detection in diverse industries.

* **Test Number 2, Image of Bird**

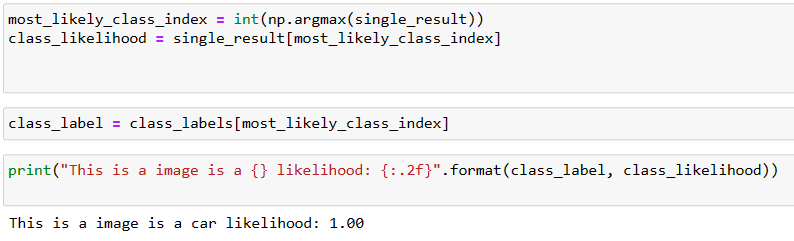
****

****

****

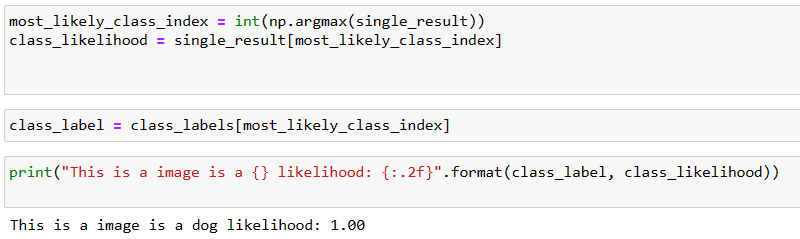
**Interpretation:** The incorrect prediction of a bird as a car in the image by the CNN on the CIFAR10 dataset indicates a misclassification. This could be due to insufficient learning of bird features during training or visual similarities between birds and cars. Addressing such misclassifications is important for improving the model's accuracy in applications requiring precise object recognition, such as wildlife monitoring.

* **Test Number 3, Image of Car**

****

**Interpretation:** When a car image is given to a CNN for image recognition on the CIFAR10 dataset, a correct prediction indicates that the model has effectively learned and recognized key features of cars during training. This success validates the model's ability to generalize and classify unseen images, with potential applications in areas like autonomous vehicles and surveillance systems.

* **Test Number 4, Image of Dog**

****

**Interpretation:** a correct prediction of a dog in your image by the CNN demonstrates the model's ability to learn and generalize meaningful features, making it a reliable tool for image recognition tasks, particularly within the scope of the CIFAR10 dataset.

**4.2 Model’s Interpretation:**

Image recognition using Convolutional Neural Networks (CNNs) on the CIFAR10 dataset with an accuracy of 75.41% indicates a reasonably effective model, given the complexity of the dataset. CIFAR10 consists of 60,000 32x32 color images in 10 different classes, making it a challenging task for the model to correctly identify objects.

The 75.41% accuracy suggests that the model is performing well on average but may still struggle with certain classes or specific instances. In your case, the misclassification of a bird as a car in the testing phase indicates a limitation of the model. Misclassifications can occur due to various reasons, including the complexity of distinguishing between similar looking objects or insufficient training data for certain classes.

Possible reasons for the misclassification of the bird image as a car could include:

1. Similar Features: Birds and cars may share certain visual features, such as round shapes or specific colors, making it challenging for the model to differentiate between them.
2. Limited Representations: The model might not have encountered enough diverse examples of birds in various poses, lighting conditions, or backgrounds during training, making it less adept at recognizing birds in unfamiliar contexts.
3. Over fitting or under fitting: The model might be over fitting to the training data or under fitting, leading to poor generalization on new, unseen examples.
4. Model Complexity: The architecture of the CNN may not be complex enough to capture the intricate details that differentiate birds from cars, especially if the dataset is highly diverse and contains subtle variations.

To improve the model's performance, you could consider the following steps:

* Data Augmentation: Increase the diversity of the training set by applying random transformations (e.g., rotations, flips, zoom) to the images. This helps the model generalize better to variations in the test set.
* Fine-tuning the Model: Adjust the hyper parameters, try different architectures, or fine-tune the existing one to better capture the nuances of the dataset.
* Confusion Matrix Analysis: Examine the confusion matrix to identify which classes are frequently confused. This can guide adjustments to the model or additional training data for specific classes.
* Class specific Analysis: Investigate the specific characteristics of misclassified classes (e.g., birds and cars) to understand where the model is struggling and address those challenges.

Improving model accuracy and robustness is an iterative process that involves refining the model architecture, training procedures, and dataset quality.

CHAPTER: 5

Summary and

Conclusion

Summary:

The project focused on implementing image recognition using a Convolutional Neural Network (CNN) on the CIFAR10 dataset. CIFAR10 is a challenging dataset with 60,000 32x32 color images spread across ten different classes. The CNN model was designed to learn and recognize features in these images and classify them into their respective categories.

The process involved importing the necessary libraries, loading and preprocessing the dataset, defining a CNN architecture, training the model, and evaluating its performance. Various components like convolutional layers, pooling layers, and fully connected layers were used to construct the CNN. The model was trained with the training set and evaluated on a separate test set. The model's accuracy on the test set was used to gauge its performance.

Conclusion:

It's not unusual for a deep learning model, such as a CNN, to provide incorrect answers during its initial predictions and then improve its accuracy over time. This phenomenon can be attributed to several factors:

1. Initialization: The model's initial weights are random, and its first predictions are based on these random weights. It may take time for the model to adjust its weights through training and start making more accurate predictions.

2. Stochasticity: Deep learning models often rely on stochastic optimization algorithms like stochastic gradient descent. These algorithms involve randomness, and the model might find different local minima in the loss landscape during different training runs.

3. Training Time: The model learns from the training data over time. The more epochs it undergoes and the more data it sees, the better its performance becomes. It refines its feature extraction and decision-making capabilities as it trains.

4. Hyper parameter Tuning: The performance of the model can be influenced by various hyper parameters like learning rate, batch size, and network architecture. Fine-tuning these hyper parameters can significantly impact model accuracy.

It's essential to consider that deep learning models, including CNNs, require careful tuning and multiple training iterations to reach their optimal performance. If the model initially provides incorrect answers but later gives correct answers, it indicates that the model is learning and adapting, which is a positive sign. This also highlights the importance of training for an adequate number of epochs and using techniques like learning rate schedules and early stopping to achieve the best results.

In conclusion, the success of an image recognition model using CNN on the CIFAR10 dataset is not solely determined by the initial predictions. The model's performance should be evaluated over multiple training runs and with rigorous testing to ensure that it consistently provides accurate and reliable results.

REFERENCE

Websites

* [https://www.companydetails.in/company/acmegradeprivatelimited](https://www.companydetails.in/company/acmegrade-private-limited)
* <https://www.acmegrade.com/>
* <https://chat.openai.com/>
* [https://www.zaubacorp.com/company/ACMEGRADEPRIVATELIMITED/U72900KA2021PTC150439](https://www.zaubacorp.com/company/ACMEGRADE-PRIVATE-LIMITED/U72900KA2021PTC150439)