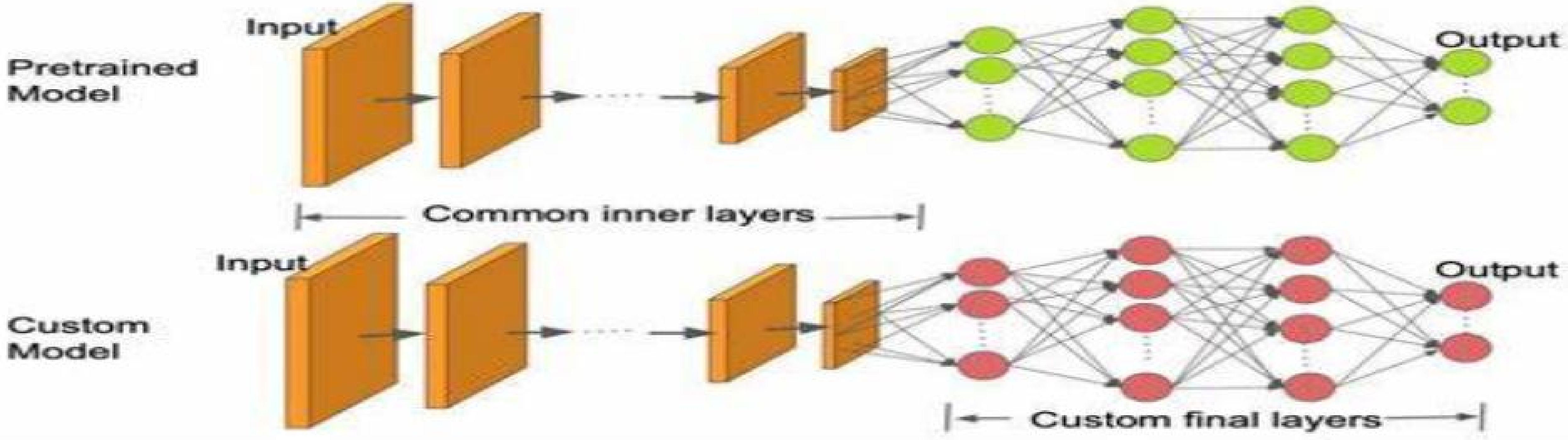
Transfer Learning for Image Classification

BY

NAME: DANIEL NDUBISI EZE

STUDENT_ID: 22009281



Introduction to CIFAR10-Choosen dataset

- Cifar10 Dataset which was named after the Canadian Institute for Advanced Research that funded the project is the choice for this task. The dataset has managable size: 60,000 images with 32*32 pixels. And relatively good number of classess (10) making it ideal when computational resources in taken into consideration. Though it has some limitations in form of low image resolution, it is suitable for beginners as it has became a benchmark for evaluating performances of various image classifcation tasks. Cifar10 dataset comes with a predefined training and testing split that consists of 50000 training images and 10000 testing images.

Introduction to CIFAR10-Choosen dataset

- The set back in using the CIFAR10 dataset is the low image resolutions.

Therefore, the dataset would need extensive augumentation for an improved model performance. Fig below shows the original randomly selected images of the dataset before preprocessing.

Overview of Transfer Learning

- Tranfer learning is a machine learning technique already developed model for a task is being used as a starting point for delevoping another model on a related task. This method is used on task where a large dataset like CIFAR10 (used is this task) and extensive computational resources is required to develop a model. It leverages the concept of pretrained model on a large dataset like Imagnet to achieve reasonable gain in model performance less data and computational resources.
- The concept of transfer learning is important in this task of image classification as it leads to improved accuracy and efficiency. It has a wide range of applications in real-time object recognision in Automobile (mostly found in Tesla automobiles) and in medical imaging.
- Keras, as an open sources high level nueral network API was used in this task. This is because of its abilities to create and modify deep learning models. And uses simple and intiutive interface which helps in experimentation and prototyping of the model.

Pretrained model selection

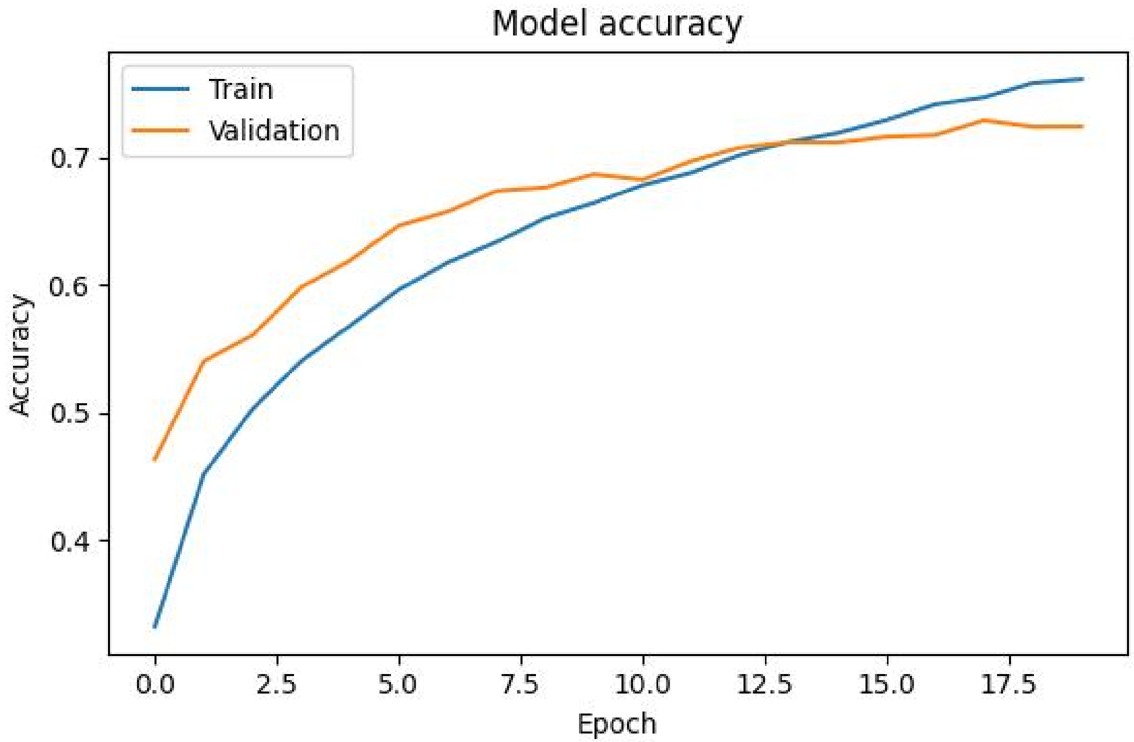
- MobileNetV2: Introduced by google researchers in 2018 which is an evolution of original mobileNet model is aimed at improving the efficiency and performance of nueral networks on devices with limited computational power.
- The original purpose of MobileNetV2 was to enable the deployment of deep neural networks on devices with constrained resources without compromising too much on model accuracy. Its design choices, such as depthwise separable convolutions, linear bottlenecks, and efficient architecture, reflect a commitment to addressing the challenges associated with running complex models in resource-constrained environments.

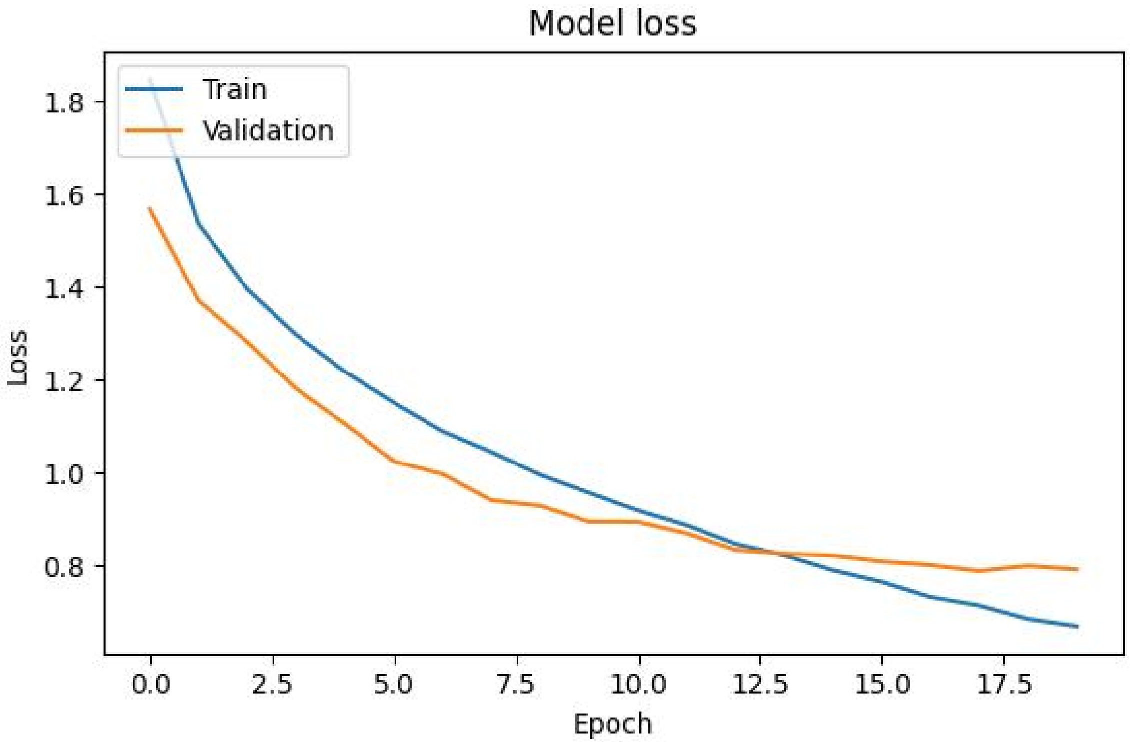
Methodology

- **Data preprocessing:** Dataset from CIFAR10 was loaded together with its labels and subsequent preprocessing of data done. These include normalizing the image pixel between 0 and 1, applying OneHot Encoder to the labels to represent categorical labels in a numeric form, Data Augumentation and image shape resizing. These steps were expeident to make sure the dataset takes input shape for pretrained model and for better efficiency/generalisation of the model.
- **Fine-tuning:** The base_model layers were fine-tuned by freezing the first 10 layers so that the weigths of these layers will not be updated during training while the remaining layers were allowed to adopt the characteristics of the Cifar10 dataset while training the model.

Trained model from scratch

As can be shown below from the plot of the performance of the model trained from scratch, The model performance is balanced evidenced by the gap between the training and validation loss. The model correctly classified approximatley 72% of the class.

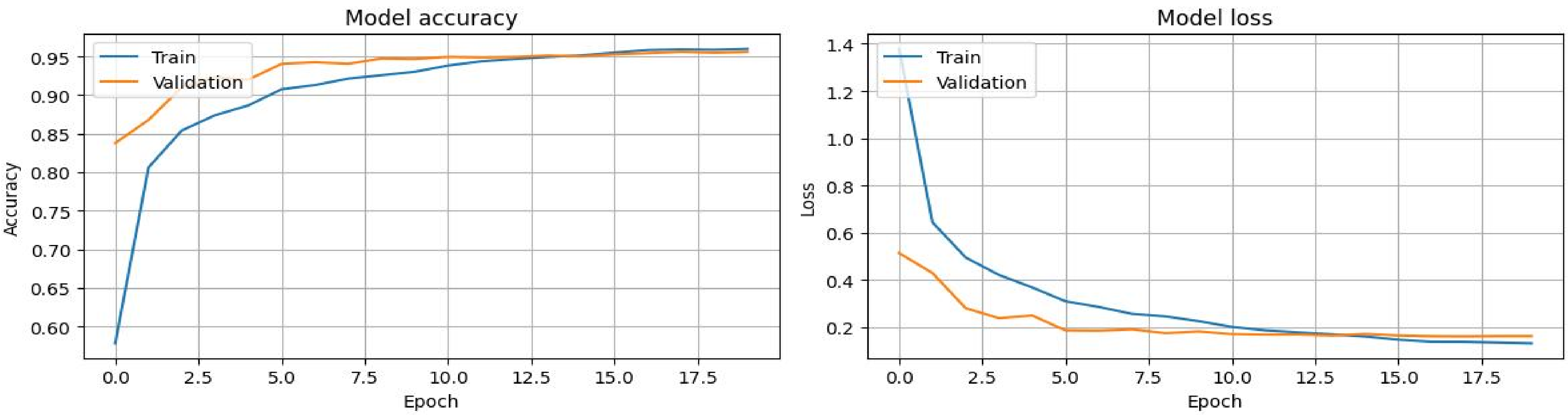




Pretrained model using MobileNetV2

- Initialy, there were signs of overfitting and the model is not generalizing well when the base_model was frozen. This prompted the freezing of the first 10 layers of the model to retain the knowledge learned from the original task. The pretrained model was tested on the test_image and it could be inferred that it generalize well to even an unseen data. Thus the plot of the performance of the model.

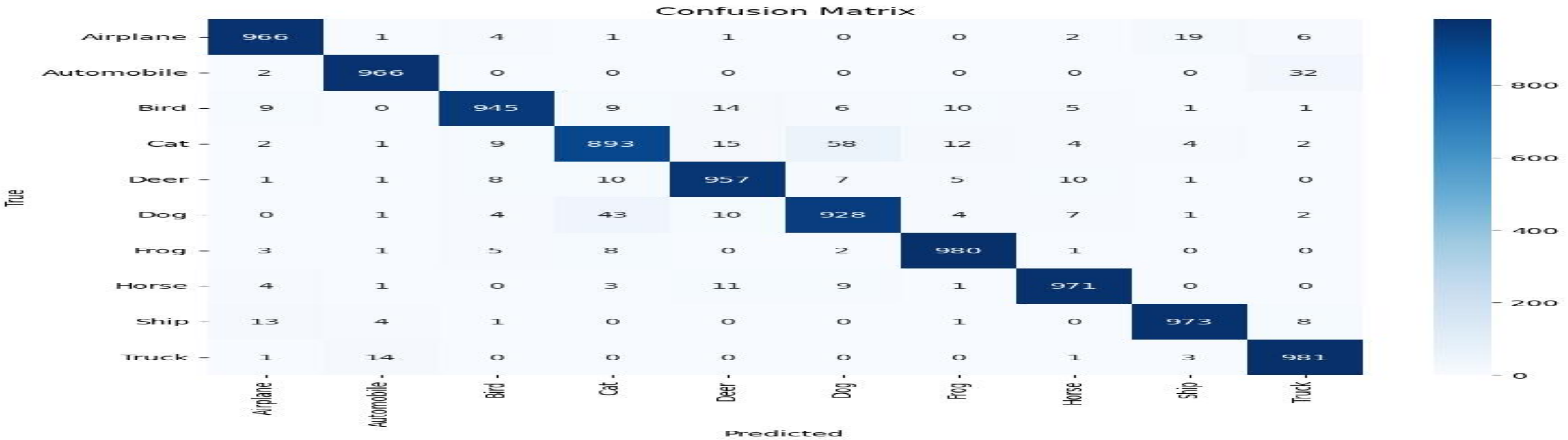
pretrained model



Comparison between the two models

Pretrained model	Trained model from scratch
The model converges faster	Take more time for the model to converge
Acheived higher accuracy and better generalisation. recorded accuracy of over 95%	Able to acheive 72% accuracy
Minimal loss recorded: Training; 20%, val; 25%	high loss recorded: Training: 67%, Val; 79%
Effective with limited dataset like Cifar10	May require a larger dataset to achieve comparable result

The confusion matrix shows that the pretrained model is best fitted for this task.



Limitations and potential areas of improvement

- **Model size**. MobileNetV2 used for this task has a large model size. It required computational resources to fine-tune it.
- **Image size resizing**. The dataset used comes with image size of (32,32,3) while that of the pretrained model is (224,224,3). resizing my dataset to even (96,96,3: as the least shape MobileNetV2 image can take) required upgraded version of GPU.
- **fine-tuning steps:** It was time consuming to fine-tune the model for optimal performance. Serveral steps used before getting the model perform as desired.

Potential areas of improvement: Development of more efficient and adaptive way of fine-tuning to prevent overfitting, improved model performance and generalization.

Bias Mitigation: Biases from the pretrained model can be transferred to the new task. Mitigating this transferable biases will help in the optimal performance and generalization of the model.

Conclusion: The chioce between pretrained model and training from scratch depends on the size of the dataset, computational resources and similarity of task at hand.