CE712 MINI PROJECT REPORT

Project Title: Classification of satellite image patches via Convolutional Neural Networks using Transfer learning

CE-712: Digital Image Processing of Remotely Sensed Data

Name:-	Vaidehi Patil	Roll No:-	170110012
Department	Department of Electrical Engineering	Program:-	Dual Degree
Name:-	Rishabh Ramteke	Roll No:-	170070046
Department	Department of Electrical Engineering	Program:-	B.Tech.

1. Introduction

CNN model is a deep neural networks hierarchy that has hidden layers among the general layers to help the weights of the system learn more about the features found in the input image. The convolutional layer applies an array of weights to all of the input sections from the image and creates the output feature map. The pooling layers simplifies the information that is found in the output from the convolutional layer. The last layer is the fully connected layer that oversees the gathering of the findings from former layers and provides an N-dimensional vector, where N stands for the total number of classes.

Transfer learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. In the context of this project's pipeline transfer learning implies inter-layer transfer of information and the fact that the information extraction happens at the layer before the last fully connected layer. This has been preferred in order to understand and analyze whether the influence of reflectance and shading information affects the accuracy of the feature extractor and outline if these two factors can be considered two main characteristics that help to decide the category the input image belongs to.

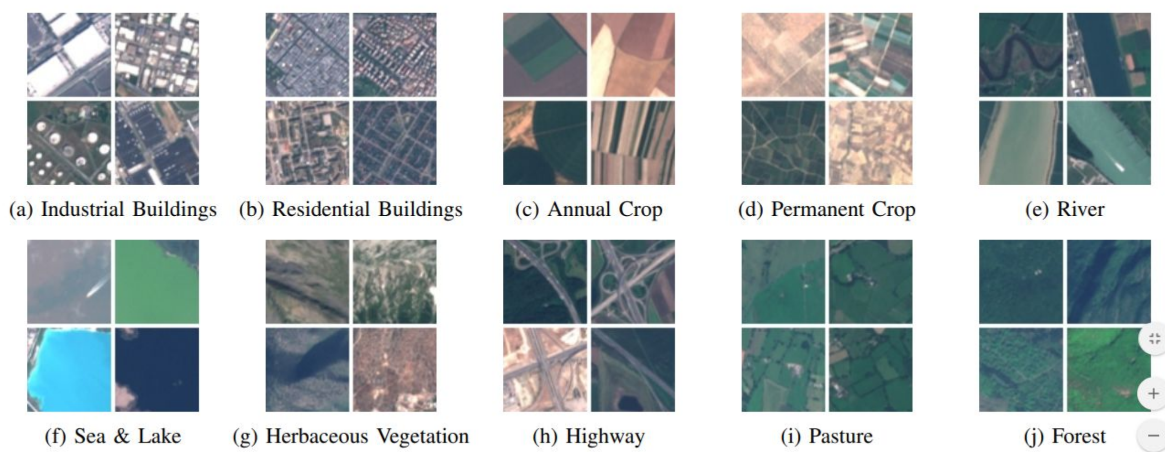
The aim of land use and land cover classification is to automatically provide labels describing the represented physical land type or how a land area is used (e.g., residential, industrial). The classification model can be used for detecting land use or land cover changes and how it can assist in improving geographical maps. It has applications in a wide variety of domains like agriculture, disaster recovery, climate change, urban development, or environmental monitoring.

2. Data sets used

EUROSAT

<u>Basic information:</u>

The EuroSAT dataset consists of 27,000 georeferenced and labelled image patches with 10 different land use and land cover classes. Each class contains 2,000 to 3,000 images. The image patches measure 64x64 pixels. We used the RGB version of the dataset. Dataset is based on Sentinel-2 satellite images.



<u>Reason for choosing this data :</u>

The cities covered in the dataset are distributed over the 34 European countries which ensure the high intra-class variance inherent to remotely sensed images. It consists of images recorded all over the year which gives it a variance as high as possible inherent in the covered land use and land cover classes.

Within one class of the EuroSAT dataset, different land types of this class are represented such as different types of forests in the forest class or different types of industrial buildings in the industrial building class. Between the classes, there is a low positive correlation which makes the dataset diverse and closest to real-world application.

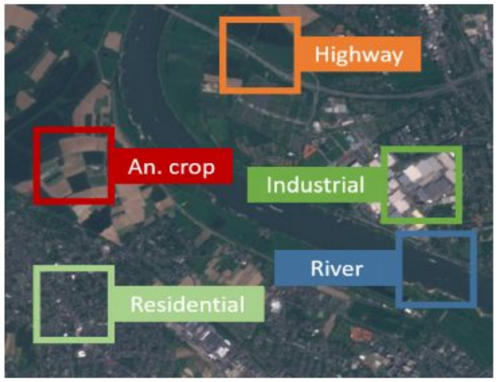


Fig. 1: Land use and land cover classification based on Sentinel-2 satellite images. Patches are extracted with the purpose to identify the shown class. This visualization highlights the classes annual crop, river, highway, industrial buildings and residential buildings.

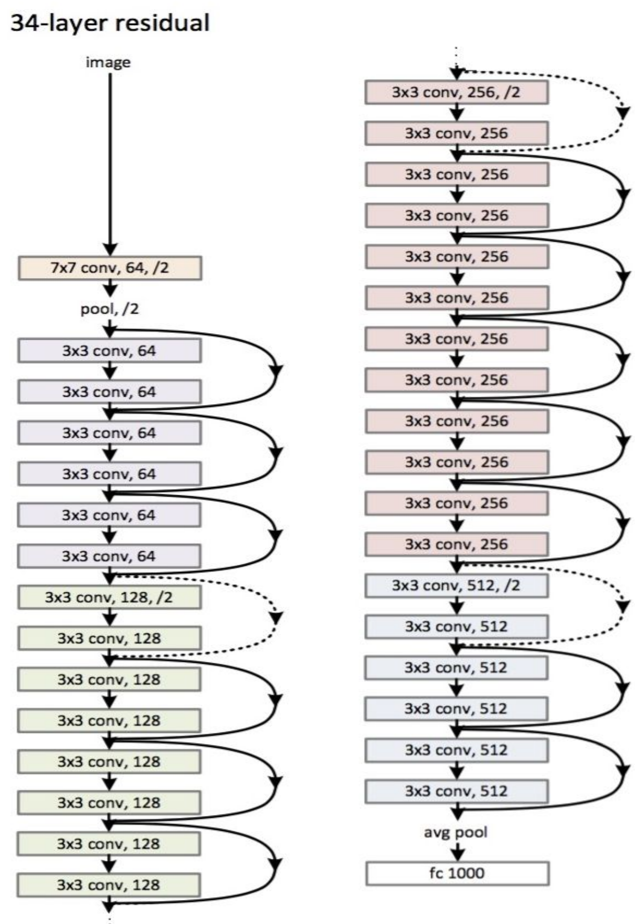
3. Methodology Implemented in python

Transfer learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem.

We used two standard convolutional neural network(CNN) architectures Resnet-50 and Mobilenet pre-trained on a standard image classification dataset Imagenet.

We took the CNNs pre-trained on ImageNet, removed the last fully-connected layer (this layer's outputs are the 1000 class scores for a different task like ImageNet), then treated the rest of the CNN as a fixed feature extractor for the new dataset. This is motivated by the observation that the earlier features of a ConvNet contain more generic features (e.g. edge detectors or colour blob detectors) that should be useful to many tasks, but later layers of the ConvNet becomes progressively more specific to the details of the classes contained in the original dataset.

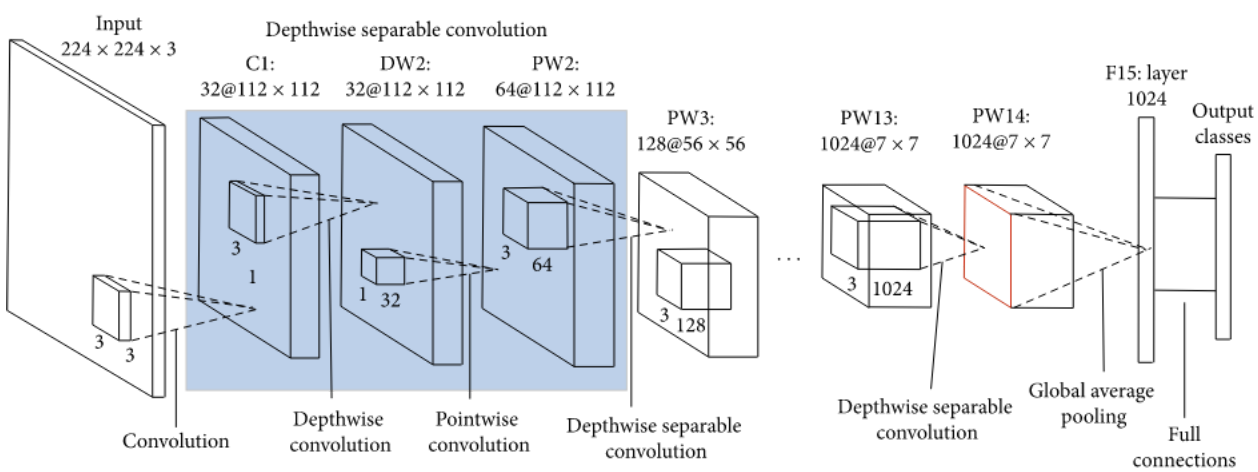
Resnet-50



The above image shows the Resnet-50 network. ResNet, short for Residual Networks, is a classic neural network used as a backbone for many computer vision tasks. It uses skip connection to add the output from an earlier layer to a later layer. This helps it mitigate the vanishing gradient problem. We have used this for the classification of remote sensing data.

We added one dense layer on top of Resnet-50 with a softmax activation function which gives a 10-dimensional output which can be interpreted as the probability that it belongs to each of the 10 classes. The corresponding ground truth is a 1 hot bit 10-dimensional vector which has 1 for the class it belongs to and 0 for the rest. The loss function used is categorical cross-entropy and the optimizer used is adam optimizer to train the CNN.

Mobilenet :



The above image shows the Mobilenet network. Mobilenet is lightweight in its architecture compared to Resnet. It uses depthwise separable convolutions which basically means it performs a single convolution on each colour channel rather than combining all three and flattening it. This has the effect of filtering the input channels.

We added one dense layer on top of Mobilenet with a softmax activation function which gives a 10-dimensional output which can be interpreted as the probability that it belongs to each of the 10 classes.

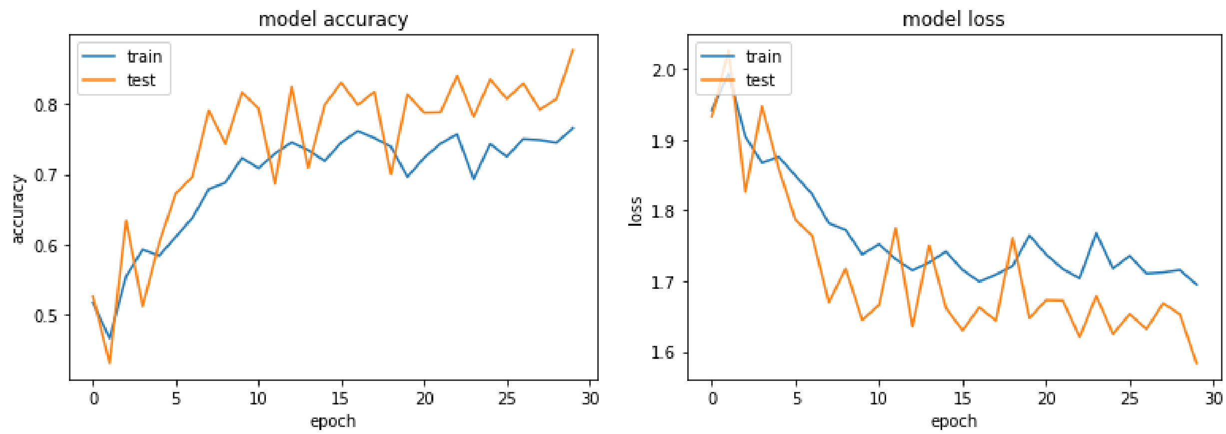
The corresponding ground truth is a 1 hot bit 10-dimensional vector which has 1 for the class it belongs to and 0 for the rest. The loss function used is categorical cross-entropy and the optimizer used is adam optimizer to train the CNN.

4. Results and Discussion

We calculated the overall classification accuracy to evaluate the performance of the two models on the considered datasets.

MobileNet

Link to code :
https://colab.research.google.com/drive/1p-AAs0W9BhEGGFTsloC6atrUk70jOLtc?usp=sharing



For Mobilenet pretrained network, the model achieved a validation accuracy of 83.7% at 30th epoch. The above plots show how loss and accuracy varies with epoch during the training process. We used Adam optimizer with a learning rate of 0.001.

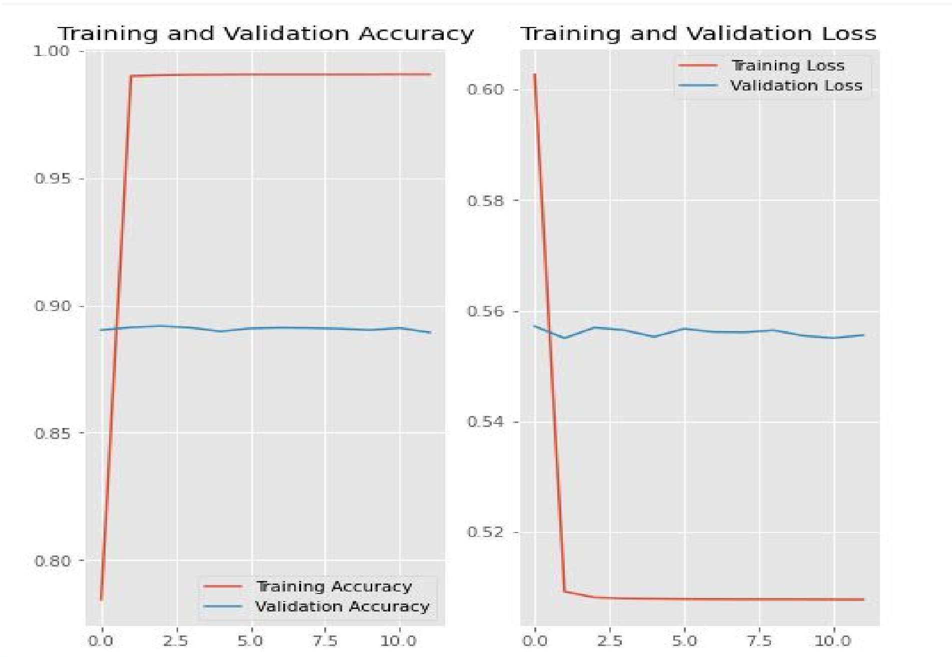
Hyperparameter Tuning in Resnet-50 architecture to improve accuracy:

Link to the code:
https://colab.research.google.com/drive/1ZTmPCtgNv-ecBrlzXrklW8haXHjK-R6P?usp=sharing

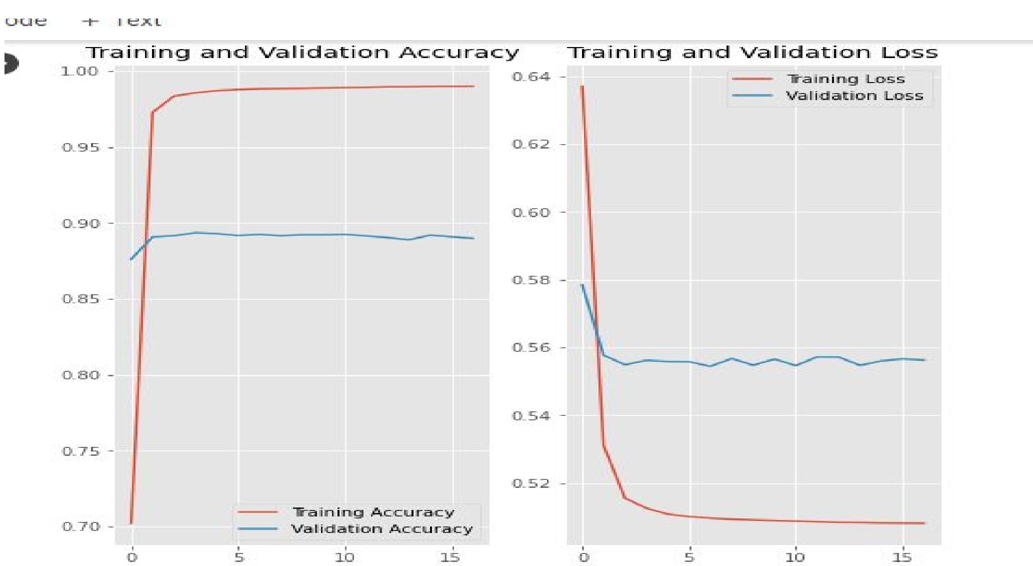
For the Resnet-50 pretrained network, the model achieved a validation accuracy of 92.83% and training accuracy of 97.68% at 48th epoch. The below plots show how loss and accuracy varies with epoch during the training process.

Experiments with Learning rate :

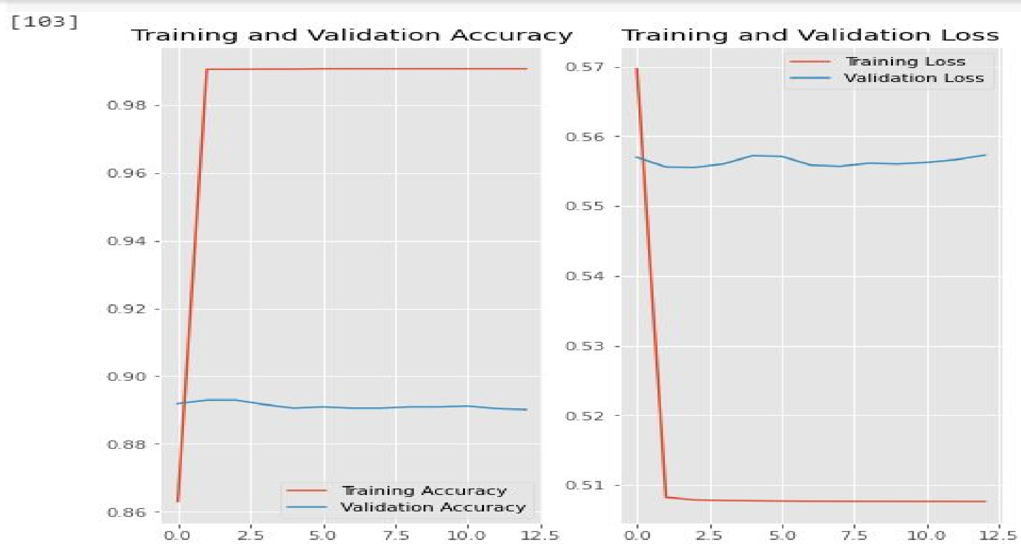
We used Adam optimizer for this particular experiment.



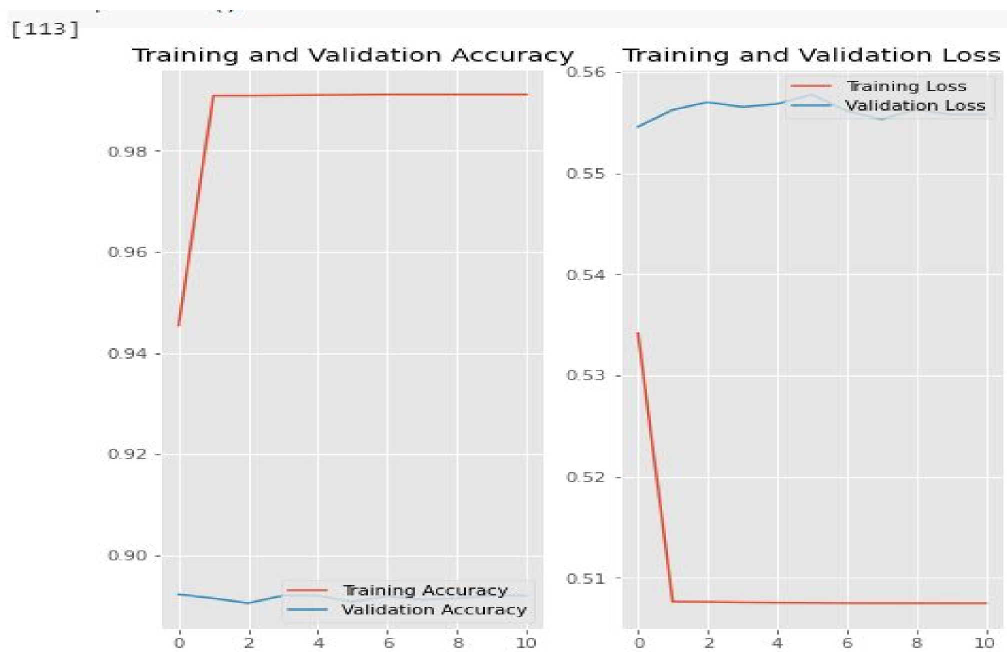
For Learning rate = 0.001 =1e-4. We get Validation accuracy of 88.2



For Learning rate = 0.0001 = 1e-3. We get Validation accuracy of 89.6



For Learning rate = 0.01 = 1e-2. We get Validation accuracy of 89.1

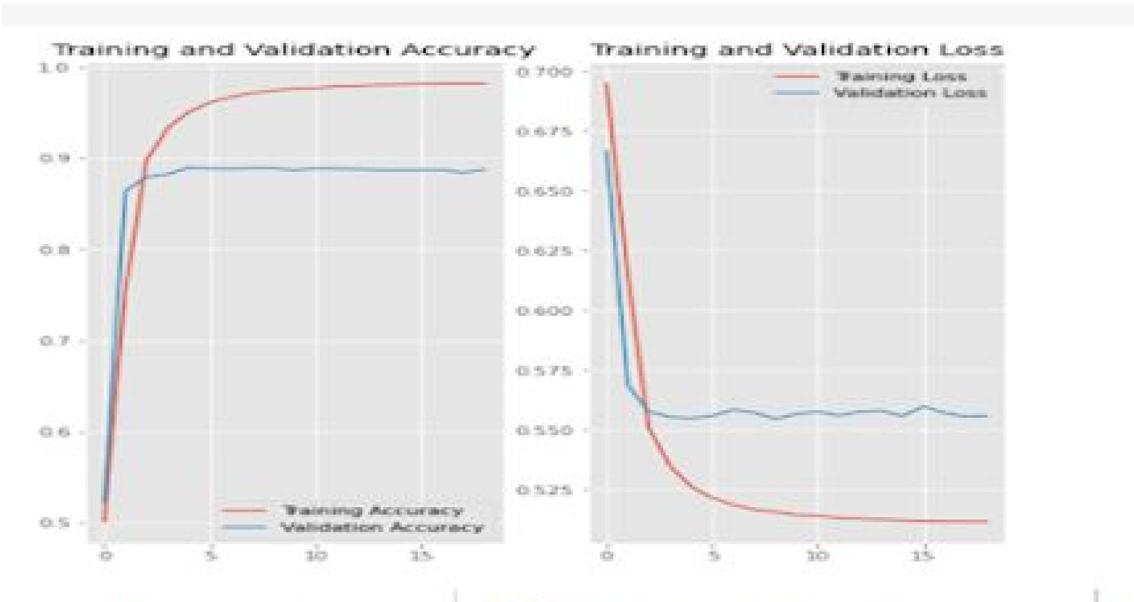


For Learning rate = 0.1 = 1e-1. We get Validation accuracy of 88.7

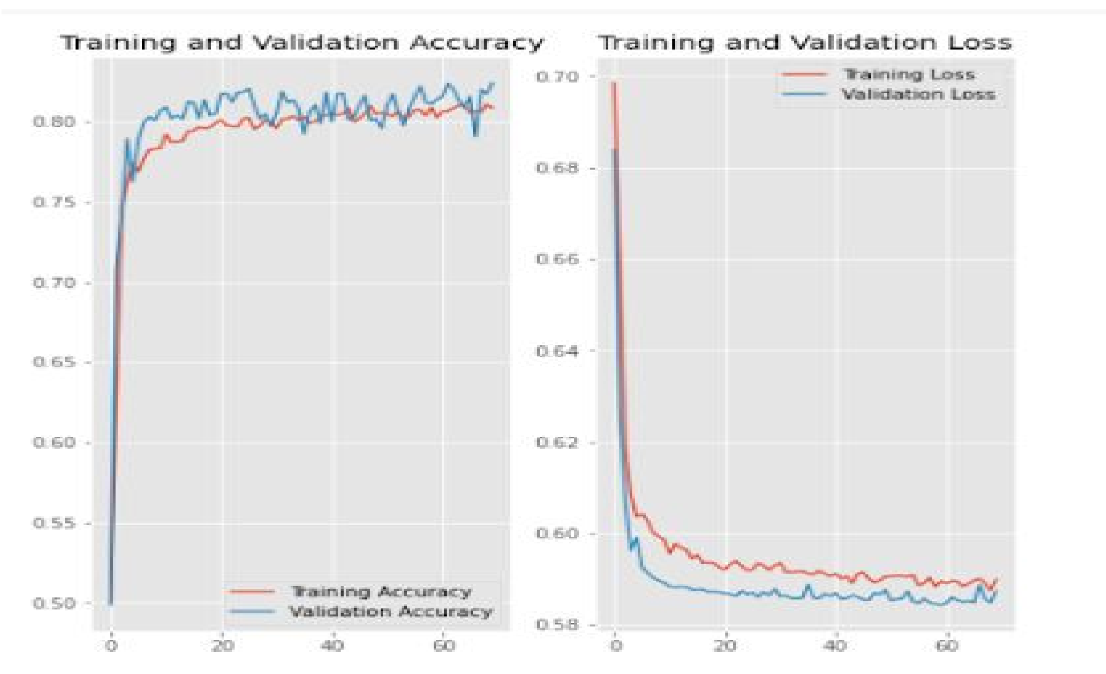
Clearly, Learning rate of 0.001 performs the best as the loss converges faster.

Experiments with Various optimizers :

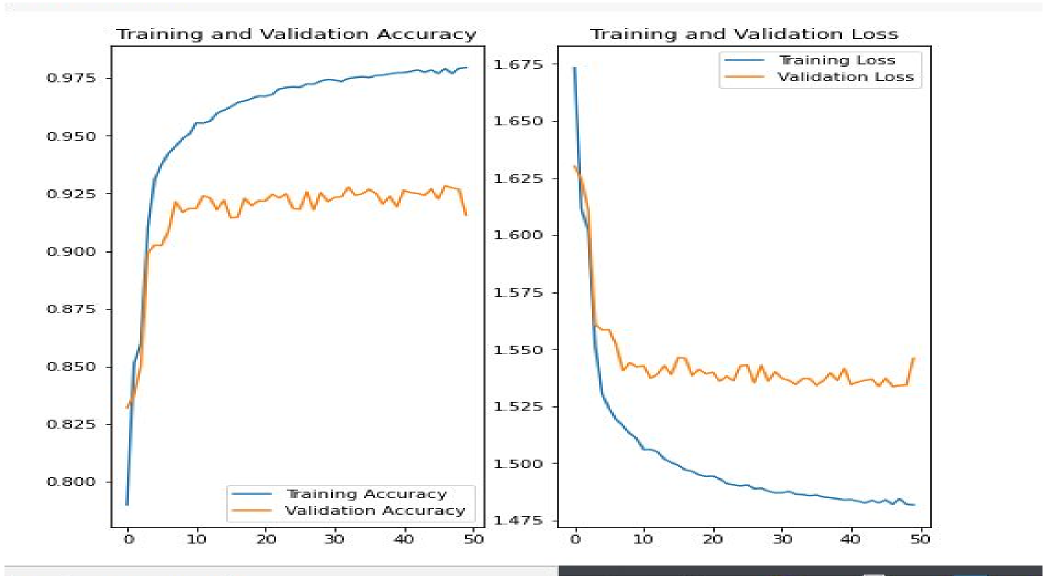
We used the same learning rate (0.001) for this particular experiment.



Stochastic Gradient Descent Optimizer : 89.17% validation accuracy with default learning rate 0.001



Gradient Descent Optimizer : 81.05% validation accuracy with default learning rate 0.001



Adam Optimizer: Best performance with 92.76% validation accuracy with default learning rate 0.001

Thus from these two experiments, we find that Resnet architecture combined with Adam optimizer & 0.001 learning rate gives the best accuracy.

Moreover, we can see that loss decreases smoothly in case of Resnet-50 compared to MobileNet. MobileNet also achieves a reasonable accuracy even though Resnet-50 outperforms it. Resnet-50 has more parameters while mobilenet has a lightweight architecture. Thus, in case of a complex training data with 3 channels such as EuroSat, it obviously performs better which is also proven from the above experiments.

5. Summary

Convolutional Neural Networks play a crucial role in the field of remote sensing for supervised classification of land cover features. It can be used to study various features of the earth using satellite data. CNN with transfer learning gives us reasonably good accuracy which can be clearly observed from the above results. The Resnet-50 model performed better than Mobilenet on the Eurosat dataset. ResNet-50 has more number of parameters to be used so it is expected that it will show better performance as compared to the MobileNet as more is the number of parameters higher is the expressivity of the model. Hyperparameters like learning rate and optimizers can be tuned to enhance the performance of deep learning models. Adam optimiser with a learning rate of 0.01 performs the best as it strikes a balance between faster convergence and better generalization.