Deep Learning Fundamentals
Assessment 2

Menuja Minu Mabotuwana Jagodage Don

November 10, 2024

1 Introduction

Image classification based on machine learning, and more specifically convolutional neural network, is gaining more and more focus due to the high success rate. With real-world applications for object recognition, such as robotic controls or autonomous vehicles, it is one of the most researched topics in the field of computer vision (Hussain, Bird, and Faria 2019).

In 2009, researchers Alex Krizhevsky, Vinod Nair, and Geoffrey at the Toronto University created dataset CIFAR-10 as a technical report. It was designed to evaluate deep neural networks for image classification applications that are widely used. Because of the simplicity of each 32x32 pixel image and the huge size of the dataset, CIFAR-10 is a perfect tool for testing and improving image classification techniques (Krizhevsky 2009). Nowadays, CIFAR-10 is one of the most fundamental and broadly accepted datasets, comprising ten categories such as airplanes, birds, vehicles, dogs, cats, deer, frogs, horses, ships, and trucks.

Figure 1: Sample Pictures per Class

Separated into training and test set, the dataset sums up to 60000 colored images equally distributed among each label, enabling a balanced training process across all machine learning appliances.
The primary objective of this report is to start with a unweighted ResNet-18 implementation

and iteratively perform fine tuning to increase the overall performance. In a final conclusion, the resulting predictive accuracy is compared to other models.

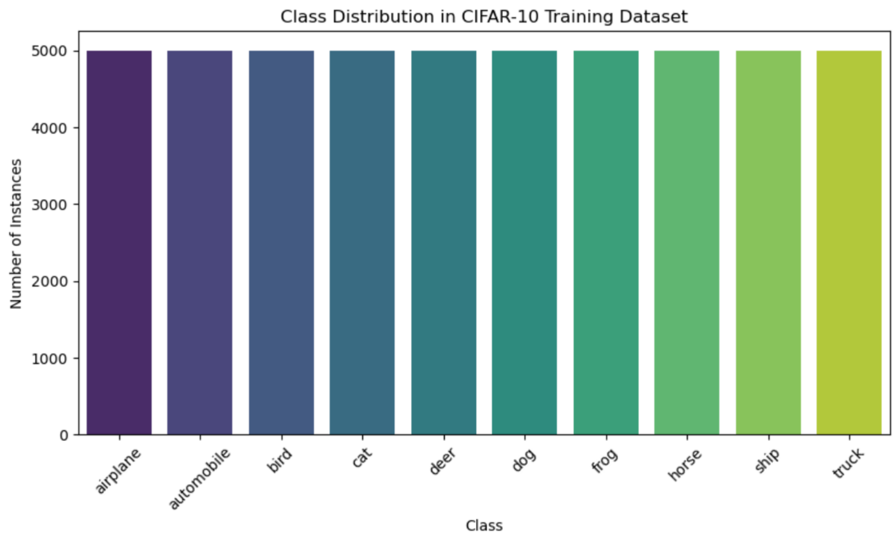

Figure 2: Class Distribution

2 Method

Specifically designed to counter occurring problems with vanishing gradients and degradation through residual blocks (Jia 2018), the ResNet-18 model is a collection of 18 layers (17 convolutional layer, 1 fully connected layer) split into 5 blocks. The residual blocks implements functionality to skip certain blocks and therefore provide different weighting across the whole training process (Al-Haija, Smadi, and Zein-Sabatto 2020).

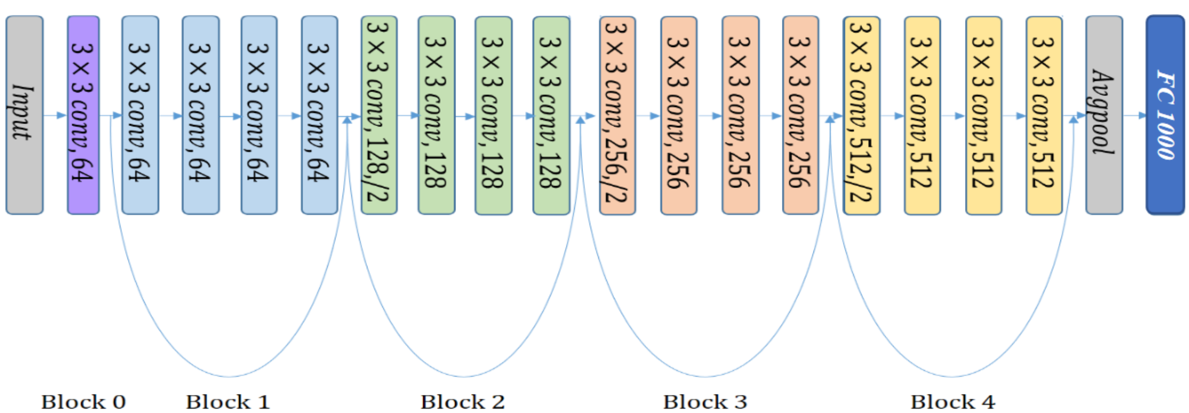

Figure 3: ResNet-18 Compilation (Al-Haija, Smadi, and Zein-Sabatto 2020)

Starting with a baseline implementation serving as fundamental comparison, values are chosen based on experience of previous projects. Starting of with Adam optimizer, learning rate 0.001, cross entropy loss function and batch size 64, the initial model is clearly overfitting.

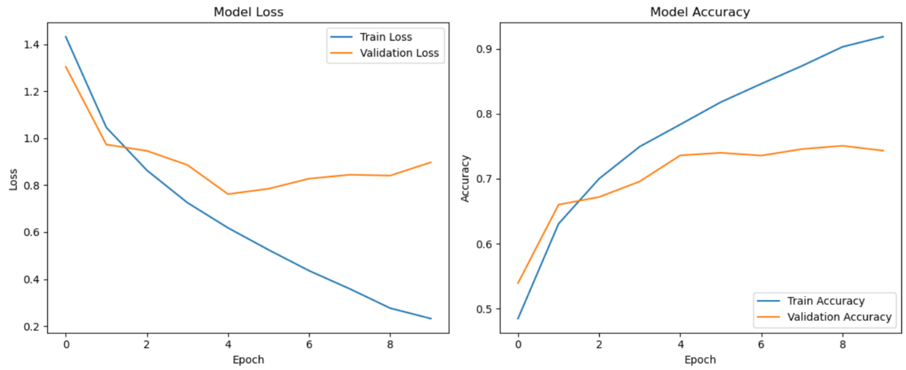

Figure 4: Baseline ResNet-18 Implementation

In a first optimization iteration, different learning rates for the Adam optimizer are tested. While a learning rate of 0.1 is struggling to find the maxima, a learning rate of 0.01 appears to result in the quickest convergence and can therefore be named the best choice.

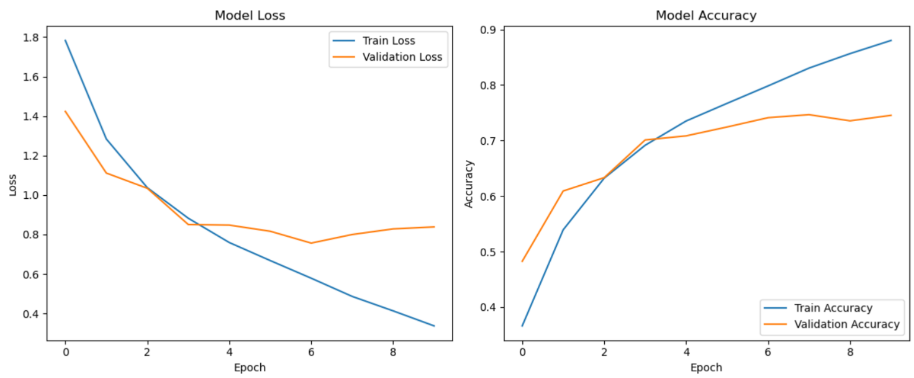

Figure 5: Adam Learning Rate = 0.01

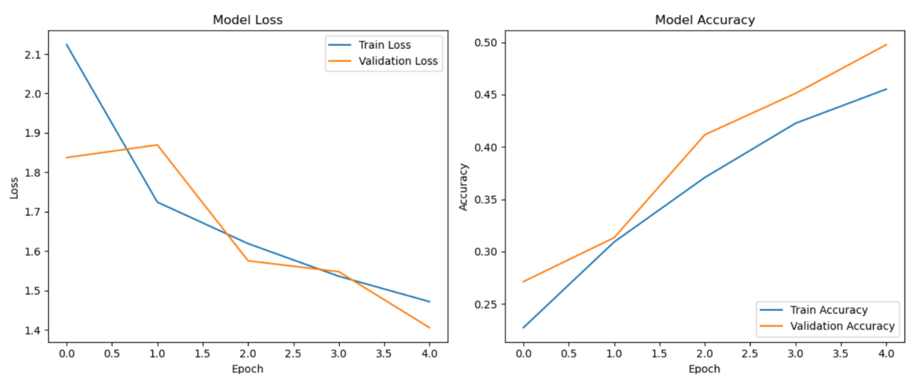

Figure 6: Adam Learning Rate = 0.1

The same procedure is repeated for SGD optimizer, testing different values in search of the best optimizer/learning rate combination.

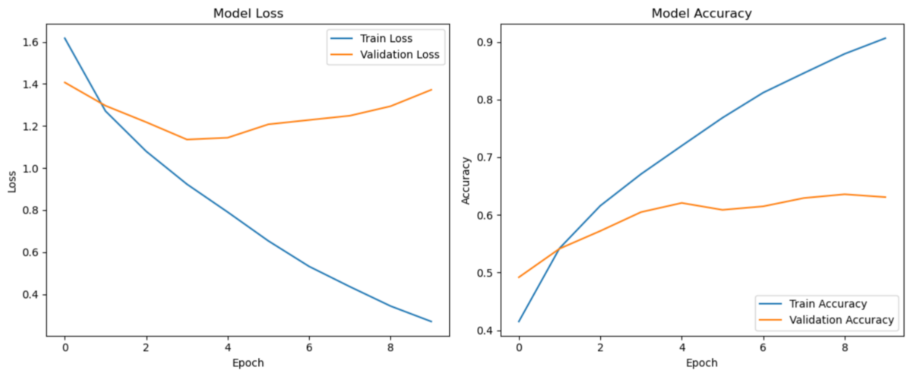

Figure 7: SGD Learning Rate = 0.01

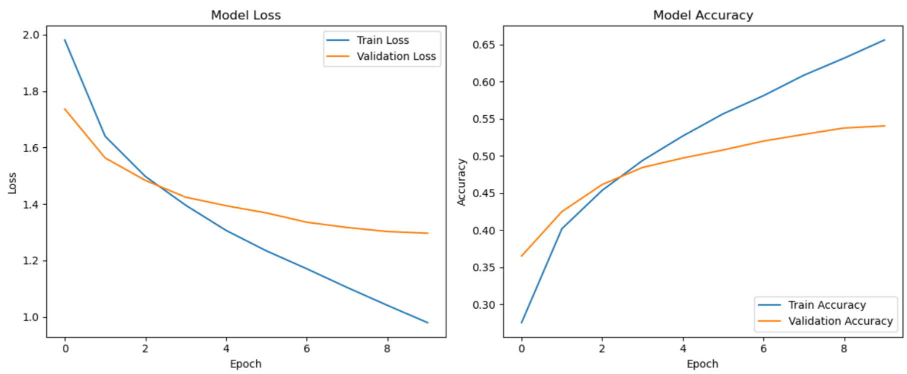

Figure 8: SGD Learning Rate = 0.001

Comparing all figures, Adam optimizer in combination with a 0.01 learning rate shows the most promising result. Compared to the other options, the training and validation graphs are diverging at the latest epoch, while simultaneously achieving the highest accuracy and lowest loss.

Nonetheless, the training data is overfitting, which is why the next task is focusing on further improvement of the still existing problem. In an extended step, different data augmentation methods, as a widely used method to increase the diversity of the training data and improve generalization. Implementing a combination of random flipping, cropping, color scheme modification (Salehin and Kang 2023), as well as a dropout imitation at the loading of the data, as the model itself is pre-defined, promises the desired effect.

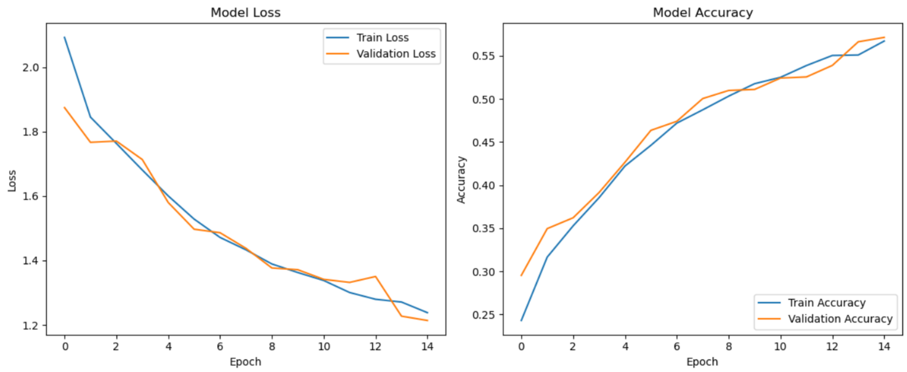

Figure 9: Adam, Learning Rate 0.01, Data Augmentation, 15 Epochs

3 Reflection

Starting from the baseline fit, the final testset result finishes with an accuracy of 70% after 15 epochs. While this result appears to be low, it is due to the limitation of resources preventing a more extended fit. Observing the graph in Fig. 9 shows a near perfect fit of training and validation data. Proven by the comparable high score of the test data, the model is performing exceedingly well with unseen data, and, given higher resources, a much better result can be achieved by simply running the training process over a greater period of time.
Overall the project for optimizing the ResNet-18 model for the CIFAR-10 dataset was successfully implemented.

References

Al-Haija, Qasem Abu, Mahmoud A. Smadi, and Saleh Zein-Sabatto (2020). "Multi-Class Weather Classification Using ResNet-18 CNN for Autonomous IoT and CPS Applications". In: *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 1586–1591. DOI: 10.1109/CSCI51800.2020.00293.

Hussain, Mahbub, Jordan J. Bird, and Diego R. Faria (2019). "A Study on CNN Transfer Learning for Image Classification". In: *Advances in Computational Intelligence Systems*. Ed. by Ahmad Lotfi et al. Cham: Springer International Publishing, pp. 191–202. ISBN: 978-3-319-97982-3.

Jia, S (2018). "Vanishing Gradient vs Degradation". In: *Towards Data Science, A Medium publication sharing concepts, ideas, and codes*.

Krizhevsky, A. (2009). *CIFAR-10 and CIFAR-100 datasets*. https://www.cs.toronto.edu/~kriz/cifar.html(Accessed:10.11.2024).

Salehin, Imrus and Dae-Ki Kang (2023). "A review on dropout regularization approaches for deep neural networks within the scholarly domain". In: *Electronics* 12.14, p. 3106.