Investigating Overfitting in Convolutional Neural Networks

May 19, 2025

1. Introduction

Neural networks are a form of machine learning models, roughly inspired from how the human brain processes information. They have become increasingly prevalent in our developing world, providing powerful and unique solutions to a wide variety of problems. However, with power also comes risk, as training a neural network too much can cause overfitting. Overfitting occurs when a model learns the training data too well and performs poorly on unseen data. In this study, we investigate whether increasing the number of parameters in a convolutional neural network leads to greater overfitting, as measured by the difference between training and test accuracy.

2. Statistical Question

Does increasing the number of parameters in a fixed CNN architecture cause a greater difference between training and testing accuracy?

Hypotheses

$H_0 : \beta = 0$
$H_a : \beta \geq 0$

Where β is the true slope of the population least-squares regression line that relates number of parameters of the model to the difference in accuracy of the model on the train dataset and the test dataset.

3. Data Collection

We trained x convolutional models on the CIFAR-10 dataset, each with the same architecture (pictured below) but different numbers of parameters ranging from approximately 1 million to 50 million.

For each model, we calculated:

$$\Delta\text{Accuracy} = \text{Train Accuracy} - \text{Test Accuracy}$$

We then analyzed the relationship between ΔAccuracy and the model size.

Data Display

Data Analysis

Conclusion

Reflection