Tales Araujo Leonidas
Professor Patricia McManus
ITAI 1378: Computer Vision
March 20, 2024

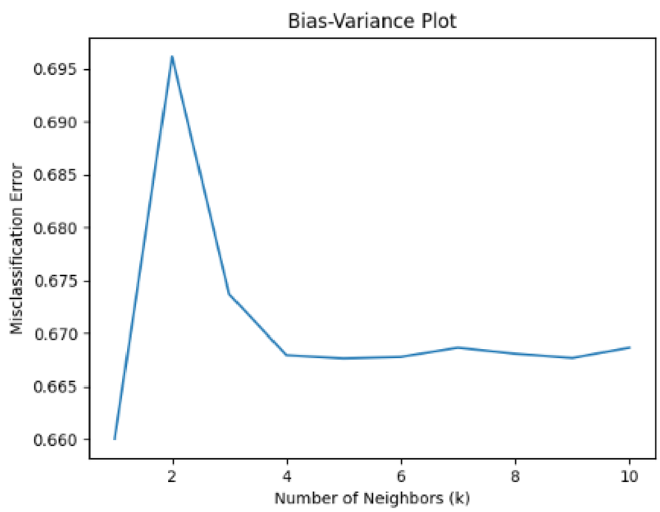
Object Classification and Detection in a Real-World Scenario: Reflective Journal

Introduction

This journal narrates my journey through the Lab “Code-Driven Object Classification Showdown” as a Computer Vision class assignment. This Lab equipped me with an invaluable understanding of the process of training, addressing bias, and optimizing a k-nearest neighbors (KNN) classifier and a convolutional neural network (CNN) model. This overview aims to summarize the main topics covered and reflect on the workflow and challenges of applying object detection and classification models in real-world scenarios using basic coding skills.

Experience Description

Although I often have seamless experiences on Google Colab, I preferred to run this Lab on Sagemaker Studio Lab due to GPU demand. For this reason, I had to install the TensorFlow framework before importing the libraries and using TensorFlow's Keras API to load the CIFAR-10 dataset, which is an extensive dataset of 60.000 images divided into 10 classes. By splitting it into train and test datasets, and flattening the labels, I ensured they match the output layer of the CNN. In the preprocessing phase, I reshaped the data to have a mean of 0 and a variance of 1, a good practice when working with a KNN classifier and reshaped the images from 3D arrays to 2D arrays to be processed by scikit-learn.



Training the model took a considerable time even using GPU, but plotting the bias-variance enabled the visualization of the model's behavior according to the "number of neighbors" and the "misclassification error" metrics. The same challenge of training time applies to fine-tuning the model using GridSearchCV to find the optimal k value that minimizes the misclassification error, which happened to be the "1", with an accuracy of 35.67%.

```
k=1, Accuracy: 35.67%
k=2, Accuracy: 31.45%
```

The last part illustrates building a CNN model with efficient hyperparameters, optimizing, training, and evaluating it to compare it with the KNN model.

Personal Reflection

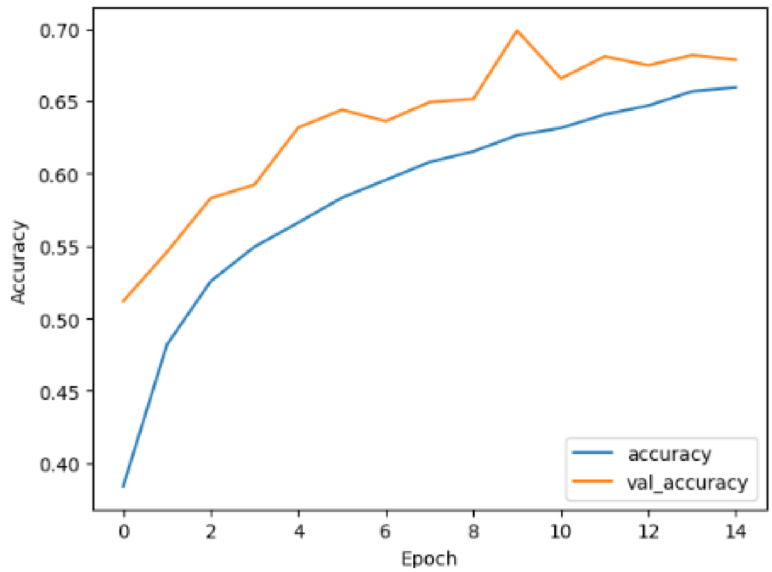
The experience was enlightening and provided a great knowledge and overview of techniques used to efficiently build, train, optimize, and evaluate both KNN and CNN models. The Lab solidified my understanding of normalization and regularization and rehearsed key concepts when dealing with image data, such as flattening and data augmentation. When training the CNN model, it was clear that it outperformed the KNN approach, showcasing an accuracy of 46.7% on the second epoch,

```
782/782 ━━━━━━━━━━━━━━━━━━━━ 44s 53ms/step - accuracy: 0.3281 - loss: 2.0721 - val_accuracy: 0.5119 - val_loss: 1.3361
Epoch 2/50
782/782 ━━━━━━━━━━━━━━━━━━━━ 42s 53ms/step - accuracy: 0.4671 - loss: 1.4858 - val_accuracy: 0.5459 - val_loss: 1.2965
```

ending up with 62.2% before stopping due to the "early stopping", which ends the model training to prevent overfitting and save training time when the validation loss does not improve for 5 consecutive epochs.

Improvements and Learning

The aforementioned technique "early stopping" is one of my greatest learnings from this Lab. Additionally, I found great value in understanding how bias and variance can be identified and addressed, as mentioned in the Notebook, "If a model has high error on the training set, it's likely due to high bias. If it performs well on the training set but poorly on the validation set, it's likely due to high variance." In the evaluation image to the side, it is possible to see how accurately the CNN model performed with the validation dataset (yellow) and with the training dataset (blue) after each epoch, still indicating a possible high bias despite the higher accuracy compared to the KNN model. With this practical experience, I am more confident about my mid-term project of image classification to distinguish between real and GenAI profile pictures, identifying fake profiles in digital media while minimizing bias.



Conclusion

In conclusion, my journey through the "Code-Driven Object Classification Showdown" Lab has been a profoundly enriching experience. The hands-on exploration of KNN and CNN models deepened my understanding of object classification and detection and highlighted the critical roles of bias identification, model optimization, and the application of techniques like early stopping in enhancing model performance and saving training time. As I reflect on this journey, I am reminded of the limitless potential of artificial intelligence to transform our interaction with digital media, making it safer and more authentic.