Katherine Stanton

ITAI 1378- Intro to Computer Vision

Professor Patricia McManus

09-26-24

Using SVM for Image Classification – Lab 05 Reflective Journal

1. **Introduction:**

* **Overview:** This week’s lab assignment introduced us to the dataset CIFAR-10, a classical machine learning dataset that contains 60, 000, 32x32 color images in 10 different classes, 3 of which we worked with in this notebook, cats, dogs, and ships. [1]
* **Purpose:** The goal of this exercise was to preprocess the data, extract features from the images, and to deploy a Support Vector Machine (SVM) model to train it to accurately guess between the three classes of provided images: a classification task.

1. **First Impressions**:

* **Initial Thoughts**: Upon first encountering the notebook, I was excited to try something new, as I have never used machine learning for image classification before. I have worked with an SVM model in a previous machine learning course, using an RBF kernel, and I recall it taking an enormous amount of computational power to complete, even after reducing the dataset to a sample. With that in mind, I was not surprised that we are now working with a subset of the original CIFAR-10 dataset.
* **Support Vector Machine**: The provided summary of what a SVM is was a helpful refresher, especially in the context of image classification, which is new territory. Coupled with an article from Medium [2] by Tanvi Penumudy titled “*A Beginner’s Introduction to SVM*”, which includes relevant visualizations, I felt adequately prepared to understand the exercise. Displayed in the snippet below is the initiation of the model using a linear kernel and training the model on the flattened *X\_train-flat* set and the *y\_train\_subset*. The *X\_test\_flat* test set is then used to make predictions on the model’s various metrics which are output below.

A screenshot of a computer program

Description automatically generated

Utilizing Gemini to translate the results, the accuracy was 54.7%, meaning the model guessed a little bit better than by chance. Precision is the ratio of true positives and false positives, with the ship class being guessed correctly by 66% versus the cat and dog classes at 48% and 49% respectively. Recall is the ratio of true positives to the sum of true positives and true negatives and had nearly identical results to precision. The F1-score is the mean or average of precision and recall which was also similar.

1. **Learning Process**:

* **Strategy**: My initial strategic approach to the exercise was to read through every step, observe the output of each cell, and then methodically go over each part of the notebook using Google Colab’s chatbot assistant Gemini to help in answering questions about parts of the code I was unfamiliar with. For example, with the code below, I was not familiar with the purpose of the numbers used to convert images to grayscale or the numbers used in normalizing images. A screenshot of a computer code

  Description automatically generated
* **Google Gemini**: A quick search with Gemini answered these questions; (regarding [0.2989, 0.5870, 0.1140]), “*these values are standard weights used for converting an RGB image to grayscale. They represent the luminance of each color channel and are based on the sensitivity of the human eye to different colors*.” While 255.0 “*is the maximum pixel value in an 8-bit grayscale image (or in each color channel of an 8-bit RGB image). Dividing by 255.0 normalizes the pixel values to be between 0 and 1, which is often helpful for machine learning algorithms.”* These quick clarifications helped immeasurably in better understanding the entire notebook as I learn better when I am clear on the finer details of things. I was familiar with the general concepts and understood the need for reducing the classes into three subsets instead of the original 10 classes. After taking an Intro to Machine Learning class this past summer, concepts like splitting the data into training and testing sets and initiating and evaluating different models have been hammered in at this point.

1. **Challenges and Triumphs**:

* **Hurdles**: The most difficult part of the exercise was truly understanding the code. I am still learning Python, and some of the called functions were new to me, particularly like np.dot() , .ravel(), and the .flatten() method. To make up for my lack of Python proficiency, I utilized the Gemini assistant to help me better understand the unfamiliar parts of code. Another challenge was understanding what the metrics used to evaluate the model mean in this context. The SVM model took about 5 minutes to complete running. No mistakes were made as this assignment didn’t require our own input.
* **Rewards**: The section of the notebook that showed the stages of the images, from original, to grayscale, to normalized was interesting and fun. This was the first time seeing visualizations outside of graphs and charts that I had created. I was surprised that, at least to the human eye, there was not much of a difference between the grayscale and the further processed normalized images.

1. **Personal Growth**:

* **Practice Makes Perfect**: The exercise has deepened my understanding of image classification and machine learning in general. I was introduced to a greater breadth of the NumPy library than I was originally familiar with. I would be somewhat confident in my ability to apply SVM to other image classification tasks, though I am still actively learning Python and would seek help from outside sources to ensure my code was legible.

1. **Looking Ahead**:

* **Remaining Questions**: One question I do have is wondering why the ship class was more accurately guessed than the other two, is it that it’s harder for a computer to differentiate between two animals than a vessel that is not similar in appearance?
* **Future Endeavors**: I would love to continue to explore the CIFAR-10 dataset on my own and to attempt to tweak certain parameters to see if I could return a better result. This exercise has been another great addition to the foundation I am trying to build to prepare for a future in the tech workforce. I know at the very least, I will be working in Python, so these labs are invaluable with exposure to this particular programming language.

1. **References**:

* 1. <https://en.wikipedia.org/wiki/CIFAR-10>
* 2. <https://medium.com/analytics-vidhya/a-beginners-introduction-to-svm-c641c3ff2769>