Lab: CIFAR-10 and SVM

Devin Prikryl-Martin

Houston Community College

 ITAI 1378 Computer Vision

Professor Patricia McManus

September 26, 2024

CIFAR-10 and SVM Lab

This lab was a bit intimidating to start as someone who is new to programming. My only experience with training models on data comes from an assignment in the History of AI class I took over the summer where we used tensorflow playground to visualize the process of training a model. This lab is my first attempt at using a data set to train a model and achieve an output. I began by seeking to understand what an SVM algorithm is and why this particular algorithm is important to image classification. I discovered that SVM, which stands for Support Vector Machine, is a type of machine learning algorithm that is used to separate different classes of data points. To visualize this, one could imagine spilling a bag containing both red and blue marbles. The SVM would draw the best possible line it could in order to separate the two colors of marbles, with the greatest distance between the line and the nearest red and blue marble. The distance between the two marbles is called the margin, and the marbles on the edge of the margin would be considered “support vectors”. How accurately it is able to draw a precise line between the colors would indicate how accurately it is performing the classification task. (Visually Explained, 2021)

With a better understanding of the algorithm, I proceeded to download the Jupyter notebook provided and open Google colab to begin the process of using the CIFAR-10 data to train the SVM algorithm to properly identify the different types of images contained within the data. I followed the steps and did my best to understand the purpose of each step. I began by installing and importing the necessary libraries. These “libraries” as I understand it, are essentially repositories of prewritten code and other information that are available to be used by the general public in order to accomplish tasks without having to write code from scratch. I loaded the four libraries necessary for the project, noting that each has a specific purpose and must be used together to accomplish our task. The next step in the exercise was to load and preprocess the CIFAR-10 dataset. Not knowing what it means to preprocess data, I decided to ask my current preferred resource for understanding new concepts, Claude.ai, to explain it to me. I learned that there are multiple benefits to and reasons for preprocessing data, but that the ultimate goal is to alter the raw data so that it is more easily digestible by the algorithm. Preprocessing improves the quality and consistency of the data, improves the performance and efficiency of the model, reduces training time and computational power required. Each type of data requires different methods of preprocessing; one of the things that helped me to understand this concept was learning that tokenization is a type of preprocessing for natural language processing, because tokenization is a concept I am already familiar with. (Anthropic, 2024) In the context of this lab I understand preprocessing to be converting the images into grey-scale, normalizing them and flattening them. I can see how this would simplify the images in a way that could be useful to enhancing the features the algorithm is meant to detect. After displaying the sample image, which gave me a better understanding of what the images I preprocessed look like, I moved on to the training step. Honestly, I was a little surprised the training step was as simple as it was. My greatest take-away from this step was what I learned about “the kernel trick” which I came across in my journey to understand SVMs. Without necessarily having a full understanding of what changing the kernel type does aside from the high-level understanding that it draws the line, or plane in different dimensions to more easily distinguish the classes, I experimented with changing the kernel type to see if it would affect the accuracy, and other data in the classification report. (Visually Explained, 2022) There were changes to the accuracy, but they seem marginal to me based on my limited understanding. I am curious to learn more about the different kernel types and under what circumstances one is preferred over another. The final step, which displayed images from the dataset in different stages I found to be very valuable in that it is much easier for me to understand what the algorithm might be trying to glean from the images when I can see them myself. It also becomes instantly clear why grey scale would be an important preprocessing step because the pixelated images are much noisier with color than without, and I can see how that would increase the difficulty of the task.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

The most difficult part of this lab was facing the trepidation that comes with attempting to do something outside of my current skillset. Because essentially all of these concepts are new to me it can be a bit overwhelming to look at blocks of code without breaking into a nervous sweat. But when I sit down to actually work through the coursework I am always pleasantly surprised at how well I am able to understand the new concepts. I think the most valuable thing about this lab for me in particular is that it helped to demystify the process of training a model. Although I still don’t understand the process at the depth I am aiming to, like you said this was the process with training wheels on, I do feel more confident that I am capable of learning how to train, run and manipulate models. I look forward to understanding more about how image processing works, especially when it comes to understanding video and real-time image input. I recently saw an interesting computer vision project someone had done to bring attention to the current state of publicly accessible video surveillance. The software, created by Dries Depoorter, records selections of open cameras for a certain length of time, then scrapes Instagram photos tagged in the locations with the open cameras and then matches the photos with the moment they were taken on video. The result is a surreal glance into just how surrounded by surveillance we are. (Dries, 2022) A person standing in front of a red building

Description automatically generated Creative applications of computer vision and machine learning like the above example are incredibly fascinating to me. There is something about the intersection of computer science, psychology, sociology, philosophy and art that really excites me. I’m not sure where this academic journey is leading me since my goals are not very defined at this early stage, however I am confident that what I learned during this lab and will continue to learn in this course will be foundational to my career in AI.

References

* Visually Explained. (2021, September 9). Support Vector Machine (SVM) in 2 minutes [Video]. YouTube. <https://www.youtube.com/watch?v=_YPScrckx28>
* Visually Explained. (2022, May 9). The Kernel Trick in Support Vector Machine (SVM) [Video]. YouTube. <https://www.youtube.com/watch?v=Q7vT0-SVII>
* Anthropic. (2024, September 26). Claude “What does it mean to preprocess data in the context of machine learning and why is it an important step?” <https://claude.ai/chat/e0265171-4ccd-492a-ae2f-b031336fe757>
* Dries, D. (2022, September 12). *The follower, 2023-2024*. Dries Depoorter. https://driesdepoorter.be/thefollower/