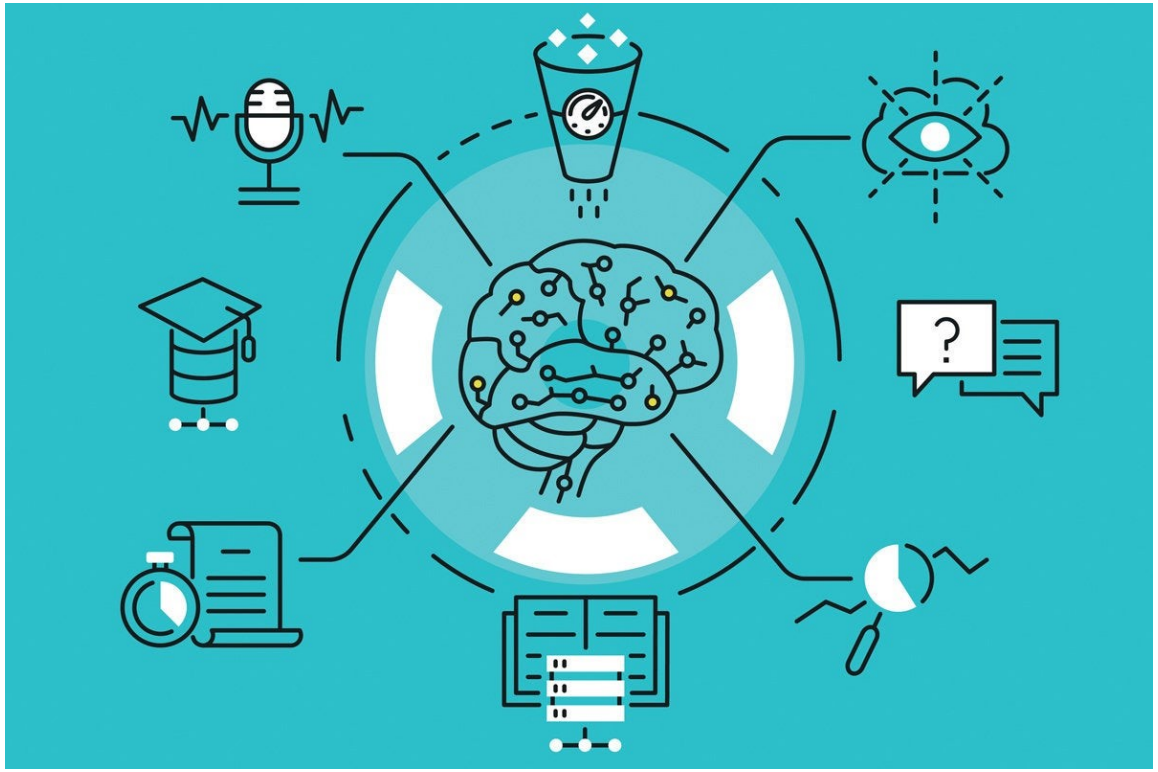
Selecting Image Features for Machine Learning Tasks

Michael P Gonzalez

Table of Contents



Summary

This project is based on creating features from images to be used in machine learning tasks. This is one of many use cases for displaying machine learning tasks. This project will have three parts to it. These three parts are needed, when you want to create features from images. I have an example image of an airplane that is 256 pixels by 256 pixels.

I will use this image to demonstrate the first two parts of this process. I will demonstrate each part and give a brief description on why it is useful to create features from images. The first part is called creating features for machine learning. The second part is called encoding mean color as a feature.

The third part is called encoding color histograms as features. I used the python environment within Jupyter Notebook to perform the machine learning tasks. I also used the OpenCV-Python, Numpy, and Matplotlib packages for this project.

Findings

I was tasked to create features from images that will be used in machine learning tasks. I broke down this project to three sections.

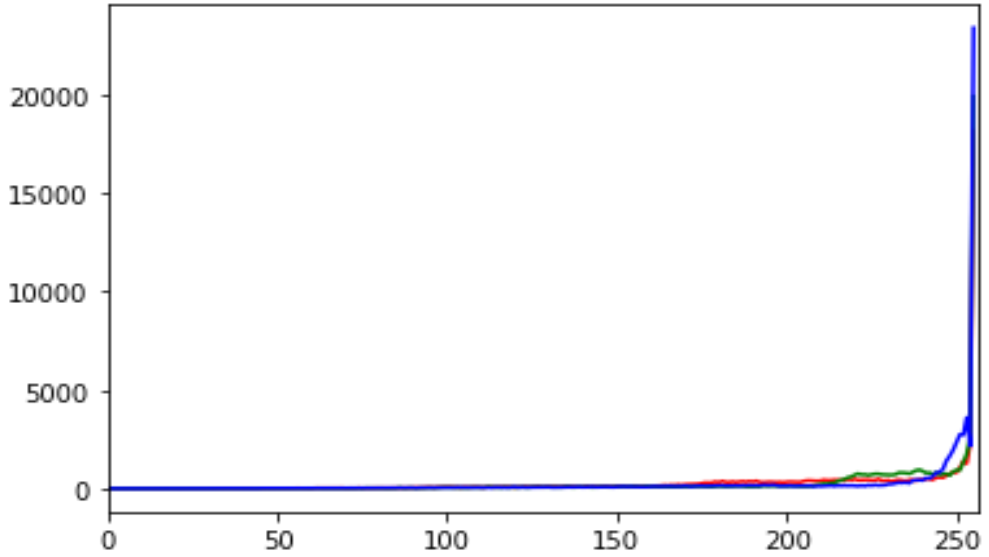
For the first section, I had to convert an image into an observation for machine learning. The second section, I needed to find the mean color values that was related to the newly created features.

For the third section, I created a list of features that were related to the color values in an image. I used those values to create a color histogram.

Color image features snapshot

Challenges

I had one challenge that I did not foresee when I was working on this project. The challenge for me was to find the correct version of OpenCV-python package. I had a difficult time finding the correct version that would work with my python environment and Anaconda. After researching, I found that others had similar issues and I was able to find a solution.



Color Value Histogram

Conclusion

I did not know that there was a way to create and use features from an image. This is a useful method when it comes to machine learning tasks that includes features from images. I could see more complex problems utilizing this with other methods. It is an interesting start to produce machine learning solutions.