Research Assessment #8

Date: 5 November 2021

Subject: Comparison of Viola-Jones and HOG Methods for Computer Vision

MLA Citation:

Rahmad, C, et al. *Comparison of Viola-Jones Haar Cascade Classifier and Histogram of Oriented Gradients (HOG) for Face Detection*, IOP Science, 2020, https://iopscience.iop.org/article/10.1088/1757-899X/732/1/012038.

Assessment:

In my research so far, I have been focusing on deep learning models for computer vision because I had always thought that they would be more powerful and better to use. However, I learned in my last interview that traditional computer vision methods are also very useful and important, so I decided to learn more about some of them in this week's research. To do this, I learned more about two very popular methods known as the Viola-Jones Classifier (V-J) and the Histogram of Oriented Gradients (HOG).

Although these were not based on deep learning, they still seemed to be effective for detecting faces, which is a very complex task. Something that seemed particularly interesting was the cascade element of the V-J method. This could be really useful because it uses successively stricter thresholds to determine whether the object was the target object or not, which is useful because it takes less computation power.

However, the tradeoff seems to be that it takes more time, especially based on the scale factors that are used. Once again, this is a reminder that the application of the model is the most important consideration for choosing a method and what elements to include in a model.

Another interesting aspect of the study was that it was found that the V-J method wasn't as good at recognizing the object if it was partly hidden in some way, through blockages or tilt. At the same time, it did well in detecting variations of the target object, which showed me that these considerations probably extend mostly to the cascade classifier. As a result, using a cascade classifier would be most useful in applications where the target object is presented front-on to the camera with little tilting or blockages in the field of vision. For this reason, it may not be so useful in applications that detect moving objects at angles. For example, if the camera is at an angle and the object is moving past it, the cascade classifier may have trouble detecting the object. This has helped me realize that the computer can only see what is presented by the camera, so the positioning and angle of the camera can have a much greater impact on the choice of method than I had previously realized.

The sliding window method of the HOG was also very interesting to me. Although this is probably a basic method of classification and is widely known, previous articles that I have read haven't mentioned or explained it, so I never really understood how classification actually works. Now I do, and this can really help me continue to understand the process so that I will be better able to synthesize my research into a research paper. The question that I had about the sliding window is that, if there are many objects in sporadic areas of the image, wouldn't it take a long time to classify

everything because it would have many misclassified objects and would need to sort those out section by section? As a result, are there any classification methods that are faster and more efficient?

Through this article, I've gained a better understanding of how computer vision models actually work, and I've gained a lot of leads into my further research so that I can continue to make good progress on my original work. I'm excited to see what I find out next!