Review of: "An Intelligent Analytics for People Detection Using Deep Learning"

Brandon Victor[1]

1 La Trobe University

Potential competing interests: No potential competing interests to declare.

This work is not a useful contribution to the field. The dataset and results are nonsense, and it spends an inordinate amount of space repeating that deep learning and autonomous vehicles are “revolutionising” this or that.

First, the good: it proposes comparing some classic detection algorithms on what appears to be a novel dataset. As a goal, that's fine. Everything else, however, falls apart.

Let's start with the results. The results are nonsense. The graphs show performance degrading from 100% to 0% as more instances were added. Every part of that is wrong. Performance will never be 100%. It should improve as instances are included. It should never reach 0%.

The dataset is never properly described, and the figures don't show a detection dataset at all. Honestly, I don't know what that's supposed to be. One image of grey blobs and a closeup of someone's face.

They compare YOLO and Faster RCNN to “CNN”. Theoretically, they could create their own architecture, but they never described doing so. So they end up making ridiculous claims that seem to refer to ALL CNNs.

The writing does not have many grammatical errors, but meanders and loses the point frequently. It talks a big game about deep learning revolutionising this or that, but doesn't provide references. Such language is not appropriate for a scientific paper because it is so non-specific as to mean nothing at all. And it was repeated, and repeated, for no reason.

Many statements did not logically follow from one another. For example, section 1.3 starts: “Despite the advantages of people detection, there are its drawbacks such as High Accuracy…”. This sentence typifies the sentences throughout. It is a mostly technically correctly formed English sentence, but the content is barely coherent and clearly false.

Even the structure of the points was poor. Early on, the focus was on autonomous vehicles. But about halfway through, it started referencing other use cases for person detection. And this eventually devolved into randomly mentioning use cases for CNNs in general.

The author should:

1. Consider what point they are trying to make with each statement, consider what purpose it serves, and whether it is justified and makes sense.
2. Figure out your focus and stick to it. Don't rely on praising deep learning for the majority of your content.

3. Don't hedge where you can avoid it. Instead of saying “we did some data preprocessing like X, Y, Z”, give a complete, exhaustive, definitive list of preprocessing steps.
4. Describe the exact steps taken to create the dataset. Where did the images come from? How many are there? What was the process of labeling?
5. Describe the exact architecture used for the “CNN” option.
6. Show better examples of the dataset. Show examples of the models working and not working.
7. Set up a legitimate test. Usually, showing mAP on the full test as a table is a good start.