G5 PHAM TRUNG KIEN, DO VAN QUYET

Topic: Workers Supervision for Construction Safety

Report: 4.5

Report outline is clear. Methodology, data preprocessing and training process is clearly and succinctly described with relevant quantitative information (eg number of epochs, learning rate, etc.) There is a clear analysis of challenges faced and how the model could be improved (workers in low-light conditions not being detected.) There are some minor grammatical errors (mainly subject-verb agreement) are present in the text but they do not significantly impede readability, ex. in section 3.2, writing "EfficientNetV2 also surpass" instead of "EfficientNetV2 also surpasses." In section 2, "In fact, HABBOF has number of frames and people within a frame limited, also lack of challenging cases" is a bit unclear and could be rewritten as "Images in the HABBOF dataset have a small number of frames and a limited number of people within those frames, as well as a lack of challenging cases."

Presentation: 5

Good pacing and pronunciation from both speakers and content was well organized. Intuitive explanation of ad-hoc fisheye augmentation and how to construct large dataset (CEPDOF/HABBOF always no helmet/jacket, videos from construction sites always have helmet/jacket etc), as well as differences in what they changed between their midterm progress report and now.

Creativity: 5

End-to-end model results have good F-scores for both fisheye and non-fisheye case that matches the original hypothesis in the introduction section. The work proposes a high-performance solution to an intriguing and novel issue. I noticed that the loss in Figure 7 seems to converge in as little as around 40 epochs. Perhaps the team could use early stopping to prevent overfitting in the future?

Overall: 5

Confidence: 3