Response to the Reviewer Comments (Author Responses in RED)

Dear reviewers, thank you very much for your time and effort on reviewing our paper. Based on your valuable comments, we carefully modified our paper and addressed all the concerns mentioned in your comments. These comments largely improved the quality of our paper. The following are the detailed replies to all the comments.

Review 1

========

Contribution of the Submission

The paper describes an interesting concept in detecting protective glasses in shop floor risk areas using deep learning. The use of the concept has been illustrated using an experimental setup. The architecture of the setup includes cloud based solutions, IoT, vision systems, interfaces with devices. The architecture and the application utilizes the state of the art in the technology and a relevant topic in Learning Factories context.

Comments for the Authors

The paper is well written with a good structure. The concepts are described appropriately. The improvement gained in the performances indicated need to be elaborated.

> Reply: We modified the text in section 4.2 to further elaborate on the improvement gained – "Results show a 155% improvement in overall model precision, 79.1% improvement in recall, and 233% improvement in the mean average precision compared to the base model. Although there is improvement in the overall performance, for applications involving PPE compliance detection, these numbers are still low."

The statements saying 'poor performance of …..the training data' is vague to understand. Either its due to lack of sufficient data to train or the quality of training data is not good enough. Explain more on this.

> Reply: We modified the text to explain it more clearly in section 4.2–"The poor performance of the base model was likely due to considerable variations in the images obtained from Google search in terms of the image background, prominent colors, and relative size of the person in the images. Due to the poor prediction performance of the base model, the trained model deployed on AI cameras in the lab environment was not able to detect safety glasses from its live feed. In order to address these shortcomings due to inconsistency between training and test images, an improved model was developed by training only on the images obtained from the lab environment."

Review 2

========

Contribution of the Submission

Assisted Detection and Check if Safety Glasses are worn or not. This by use of AI, machine vision and IoT. Executed in a Learning Factory.

Comments for the Authors

Highly interesting and important topic!

Clear structure, string and readability.

SUGGESTED IMPROVEMENTS

Fig. 1: use bigger characters in the headline of the boxes for better reading.

Reply: Thank you for the suggestion. We modified Fig. 1 accordingly.

First 2 paragraphs of point 4 are still part of the Experimental Setup and the Framework than Results. Think about it.

Reply: Agreed. These paragraphs have been moved as suggested.

It would be good to give answers also to these questions:

If there is an mAP, why is there no mean Average Recall?

Reply: The Custom Vision toolbox does not natively report this data. This suggestion has been passed to Microsoft for consideration. This is outside of the ability of the authors to control.

Can the application differ between safety glasses and normal spectacles?

Reply: Our preliminary analysis suggests that the application can distinguish between safety glasses and normal spectacles to some extent as one of the co-authors involved in physically training and testing the AI model was wearing spectacles. However, the capability of distinguishing between the safety and normal glasses has not been thoroughly evaluated in this study and is a topic for future study. It seems intuitive that detecting safety glasses for a person already wearing normal glasses would be a harder problem to tackle (particularly when the safety glasses have transparent or light color borders) and would require extensive amount of training data for the AI model. One of the possibilities is to use the safety glasses with a unique

color border, so that the distinguishing boundaries are well detected and the recall and precision of the AI model is good.

Comparison of Fig.3 and 4: Please use the same order naming the tags!

Reply: Thanks for the comment. We modified the naming order.

Use the same scale!! Image count: A bar with length 525 (Fig. 4) must be about 4 times longer than the bar in Fig. 3 ! The size of characters should be enlarged for better reading.

Reply: Thanks for the suggestion. We removed the bar to enlarge the text for better reading.

You could have done this job in any laboratory. Where ist the connex to a Learning Factory? Where is the Teaching and Learning Concept. What is prepared by experts and what are the students able to work on? Find some lines also for this.

Reply: This is prototype of a system which will be implemented and actively improved in the learning factory by students and faculty. The initial architecture utilizes previously untested technologies, hardware and a hybrid computing architecture which required significant interfacing with faculty, Microsoft and University Enterprise IT. The combination of factors lead faculty to develop and implement the initial proof of concept model in a controlled location (supporting laboratory) with minimal activity and isolated IT systems.

The list of references is rather poor. AI and Vision detection is already a rich area in research. Add at least 10 references.

Reply: The list has been expanded to focus more broadly on AI Vision and AI enabled safety systems while also providing additional of the specific technology deployed in the system.