SSRG International Journal of Electronics and Communication Engineering Volume 12 Issue 4, 159-166, April 2025
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V12I4P115

Original Article

Design of a Personal Protective Equipment detection system using Computer Vision and Convolutional Neural Networks

Brayan Kevin Medina Villagomez[1], Paul Andy Quispe Huachaca[2], Jesús Talavera Suarez[3]

[1,2,3]*Universidad Nacional de San Agustín de Arequipa, Arequipa, Perú.*

[1]*Corresponding Author : bmedinav@unsa.edu.pe*

Received: 11 February 2025 Revised: 13 March 2025 Accepted: 15 April 2025 Published: 29 April 2025

Abstract - *In work environments, the identification and correct use of Personal Protective Equipment (PPE) is essential to avoid incidents and safeguard employees' health. In this paper, an innovative strategy for the automatic identification of PPE using computer vision techniques and Convolutional Neural Networks (CNN) is presented. The method uses a specially trained CNN to interpret images of the equipment and a labeled dataset that was developed especially for PPE detection. Standard item detection criteria were used to evaluate the system's performance, and they were shown to be effective in correctly identifying Personal Protective Equipment (PPE) in photographs of industrial environments. The results of this study show a great degree of sensitivity and accuracy in the identification of several kinds of Personal Protective Equipment (PPE), indicating that the development of this technology can improve automated inspection tasks and safety in the industrial workplace, avoiding dangerous circumstances by providing better control.*

Keywords - *Computer vision, Convolutional Neural Networks, Personal protective equipment, Detection System.*

1. Introduction

In industrial and building contexts, safety at work is of utmost importance [1]. It is essential to prevent accidents and safeguard the health of employees to ensure a safe and productive working environment. In this situation, Personal Protective Equipment (PPE) plays a vital role in establishing a barrier between employees and the possible risks existing in their work environment, such as mechanical impacts, exposure to harmful chemicals or electrical hazards [2]. PPE can be difficult to use properly and monitor worker compliance in complex and changing industrial environments despite their importance. It can be challenging for safety teams and managers to ensure that all workers consistently use the right equipment [3]. Because traditional visual detection procedures are subjective, extensive and prone to human error, their effectiveness can be compromised, which is why workplace accidents are often seen. The use of state-of-the-art machine vision and deep learning technology in this framework creates new possibilities to maximize autonomous and accurate tracking and identification of PPE [4, 5] and eliminate the human factor altogether. Computer vision is a branch that belongs to artificial intelligence and focuses on creating systems capable of interpreting and understanding digital images and videos thanks to its algorithms. Convolutional neural networks (CNNs), on the other hand, have shown remarkable success in tasks involving object identification and visual recognition, important issues when wanting to develop this system, and due to its ability to immediately learn hierarchical representations of visual information from the input, it has allowed to increase the accuracy of detection [6].

This paper provides a new approach using CNN and machine vision to identify personnel not wearing Personal Protective Equipment (PPE) in industrial environments. This study focuses on creating and implementing an automated system that can recognize and confirm that PPE is not present in real-time video frames. This technology has great potential to improve the safety of personnel and those around them at work by providing an objective and accurate tool for monitoring and enforcing safety regulations related to the mandatory use of PPE in industrial environments and high-risk situations, such as working at heights or with hot material.

In this work, Convolutional Neural Networks (CNNs) are combined with contemporary computer vision and deep learning technologies to promote both technical innovation and worker safety in an industry that performs PPE inspection manually and visually. This allows automation of inspection tasks and continuous improvement of safety procedures in industrial environments, thus avoiding any type of work-related mishap related to the non-use of Personal Protective

Equipment (PPE) when required. The distribution of the information in this document is divided as follows: The work related to this research is presented in Section 2. Subsequently, Section 3 develops and explains the complete methodology of this proposed system. Section 4 details the design and construction of the hardware and software system and its implementation in a real agricultural environment. The analysis of the results and discussions obtained are presented in Section 5. Finally, the conclusions of this research are found in Section 6.

2. Related Works

Automated detection of Personal Protective Equipment (PPE) in industrial environments has been the subject of several investigations in recent years due to the growing need to improve occupational safety through advanced computer vision and deep learning technologies, highlighting their most significant approaches, methodologies and findings. In this part, the use of computer vision and Convolutional Neural Networks (CNN) is emphasized, so some of the most relevant research in this area is reviewed. Authors Hayat et. Al. presented a real-time, computer vision-based automatic safety helmet detection system, "You Only Look Once" (YOLO), at a construction site [7]. They showed that their experimental results with the YOLO architecture achieved the best mean average accuracy of 92.44%, thus showing excellent results in safety helmet detection.

However, the technology presented shortcomings in partial occlusions in different illumination situations. Lema [8] developed a cost-effective solution for real-time monitoring of PPE use. Their work seeks to improve worker safety by verifying its use. Their proposed methodology achieves a 6% improvement in the average accuracy of PPE detection. Their system facilitates continuous monitoring of safety parameters in industrial facilities, immediately notifying management when these parameters exceed predetermined thresholds. Moreover, their system was designed to work on really inexpensive devices, which facilitates its application in a variety of companies and industries. To detect and categorize EPPs in high-resolution photographs, Wang et al. [9] proposed a new novel method based on semantic segmentation approaches. They identified different types of EPPs with an accuracy of 90,3 % using a modified U-Net network.

This approach worked quite well, but due to the complexity of the required processing, real-time implementation is still difficult. This work demonstrates the potential of computer vision and deep learning technologies for automated EPP detection. However, there are still issues with system robustness in hostile environments common in industry sectors and optimization for low-cost, real-time devices. By developing a CNN-based PPE detection system designed for use in a variety of dynamic industrial environments, this study aims to overcome these limitations in the state of the art and achieve a system that ensures the safety of personnel in an industry or company where Personal Protective Equipment (PPE) is used.

The application of artificial intelligence has been fundamental for developing systems to monitor the use of personal protective equipment (PPE) in real time and with high accuracy. For example, during the time of the COVID-19 pandemic (the year 2020), a web application was created to monitor the use of protective masks in work environments, employing artificial vision and deep learning techniques to automatically detect the presence of masks on workers [10]; this action significantly helped healthcare workers to detect people without face masks automatically. Likewise, computer vision algorithms for PPE detection have been developed, such as the one presented by a group of researchers in 2024, which employs image processing and deep learning techniques to identify the correct use of masks and gloves in industrial environments [11]. The availability of specific datasets is crucial and fundamental to training models and systems for automatic detection of EPP equipment. Ahmad and Rahimi [12] presented the SH17 dataset, which contains more than 8000 annotated images with 17 classes of PPE collected in various industrial environments, which is important data for starting the training of a system. They trained object detection models and achieved an accuracy of 70.9 % in PPE detection [12], a high and relevant value in research. On the other hand, Sandru et al. [13] proposed the SuPEr-SAM model, which uses a supervisory signal from a pose estimator to train a spatial attention module, which improves the accuracy of EPP recognition in images. This innovative approach combines pose estimation with object detection to improve the accuracy of EPP identification [13**].**

In another investigation, authors Ahmed et al. [14] used a deep convolutional neural network to recognize various PPE, such as gloves, vests and helmets. On a dataset labeled especially for this use, their model, which is based on the YOLO architecture, obtained an acceptable mean average precision of 96%, improving the accuracies of other related studies. The system's high computational resource requirements prevented it from being implemented on devices with little computing power, even if the findings were encouraging. The application of CNN for real-time EPP identification in videos was investigated by Lo et al. [15]. The model the research's authors developed was based on the Faster CNN architecture. The company provided the dataset of industrial surveillance videos used to train and evaluate the algorithm. At a processing speed of 25 frames per second and 97% accuracy, the system successfully detected EPP very instantly. However, the system faced challenges detecting PPE in highly dynamic environments with high worker density. Massiris et al. [16] proposed using the YOLO neural network to monitor the use of PPE, such as gloves, helmets and high-visibility clothing, in industrial environments. They used a collection of movies taken using sports cameras to train the algorithm, and it has shown encouraging results in identifying these objects. Honggang presented a computer vision-based method in [2] for identifying safety helmets on building sites. The researchers in that paper identified helmets in static photographs with 95% accuracy by combining image processing methods with support vector machine (SVM) classifiers, achieving a high accuracy value. However, their proposed system was challenging for partial occlusions and fluctuating illumination conditions. A related study created an automated system using machine vision techniques to produce indicators of the correct use of safety equipment to monitor PPE use in the construction industry. This approach allowed for more efficient and accurate monitoring in work environments [2], thereby reducing workplace accidents. More recent state-of-the-art studies have investigated the use of preferred models for PPE detection. Using the YOLOv4 object detector, Karlsson et al. [17] built a system that can recognize various types of PPE in photographs of industrial workers, including hard hats, safety vests, gloves, and safety glasses, among other PPE worn by industrial personnel. By recognizing these items with high accuracy, the system helped increase worker safety [17]. Islam et al. [18] introduced a YOLOv7-based method to identify construction workers wearing all necessary safety equipment, such as hard hats, safety glasses, jackets, gloves, and shoes, among other mandatory PPE. With an average mean accuracy (mAP) of 94%, the model proved to be successful at recognizing personal protective equipment (PPE) in construction settings [18].

Despite significant advances in automatic PPE detection, some challenges need to be addressed. Variability in lighting conditions, partial occlusions, and the diversity of industrial environments can affect the accuracy of the models. To further improve the generalizability of the models, it is clear that larger and more diverse datasets are needed as training databases. To improve detection in hard-to-reach locations and augment optical data, future research can use other sensors, such as LiDAR systems or infrared cameras, to support conventional cameras in detecting workers without PPE. Similarly, implementing more sophisticated deep learning methods, such as attention models or Generative Adversarial Neural Networks (GANs), could improve automated detection systems and increase the accuracy of current PPE detection systems.

3. Methodology

In this system, a specially labeled dataset was used for this work to build the Personal Protective Equipment (PPE) detection system. The video frames, which were taken in current industrial environments, show a variety of lighting configurations, perspectives, and PPE usage, for example, helmet, gloves, safety glasses, reflective vests, and face masks, which were the five main categories into which the protective components were divided. The block diagram of the created system is shown in Figure 1.

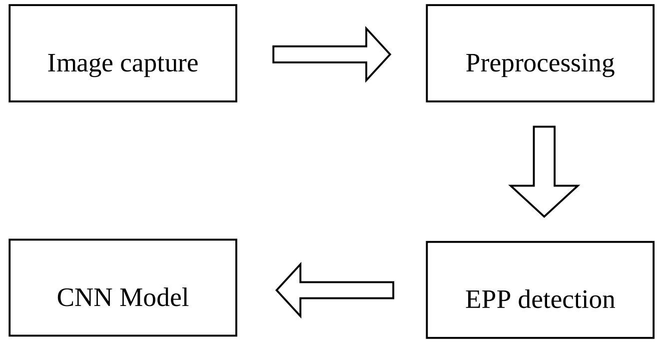

Fig. 1 Block diagram of the developed system

3.1. Image Capture

In this section, high-resolution cameras are used to capture images of industrial environments and personnel. To ensure their quality and speed up the detection procedure, the acquired photographs must have the appropriate lighting and location; in this way, the system avoids false positives due to lighting changes. Various viewpoints and situations are chosen to illustrate the unpredictability of the industrial environment. An example of a video frame of the workplace prior to the system is displayed in Figure 2.

Fig. 2 Video frame before processing

3.2. Preprocessing

In this second stage, the collected images undergo a series of preprocessing steps necessary in image processing to improve the quality and adjust them to the detection model. These methods include normalizing pixel values to a range of 0 to 1, resizing images to 224x224 pixels, the standard resolution most frequently used in scientific papers, and employing data augmentation methods like lighting adjustments, rotations, and reflections. When these methods are combined, improved post-processing is ensured, and increased model resilience to environmental changes is observed.

3.3. CNN Model

In the built CNN model, several convolutional layers are used to extract features, and then pooling layers are used to reduce dimensionality; this was done after proving that it performed better when training the model. The output layer groups the photos into the specified PPE categories, while the dense layers allow the analysis of the extracted features. For multiclass classification, activation functions such as ReLU are used in the hidden layers, and softmax is used in the output layer. Accuracy, sensitivity, F1-score and mAP are among the quantitative measures used to evaluate the performance of the trained model. Sensitivity evaluates how well the model can identify PPE items in the photos, while accuracy calculates the percentage of accurate predictions. While mAP provides a summary of the model's performance across multiple classes, the F1-score integrates both metrics, and these measures will be evaluated and discussed in section 5 of this paper.

3.4. Detection of Lack of PPE

Following training and evaluation, the methodology suggested in this study is put into practice for automatically identifying Personal Protective Equipment (PPE) on workers in industrial settings. After analyzing the input photos, the system produces a result that shows whether PPE was found. Supervisors can be notified using this data, and automated reports on safety rule compliance can be produced.

4. Experimental Development

The experimental development of the system was carried out in a controlled environment that simulated real industrial conditions, and some other data sources were taken from a real industrial environment. Exhaustive tests were carried out in various scenarios to evaluate the accuracy and robustness of the model. The results obtained in each phase of the process were documented, from image capture to model evaluation, ensuring data traceability and reproducibility of the experiments. In addition, processing times and the capacity of the system to operate in environments with limited computing resources were analyzed; for this reason, the video images were reduced to a small resolution of 224x224 pixels.

4.1. Preparing the Work Environment

Implementing a Personal Protective Equipment (PPE) identification system based on artificial vision and Convolutional Neural Networks (CNN) was the main objective of the experimental development of this project, especially this research. A robust computing environment suitable for processing video images was initially established to provide efficient training and reproducible results. The advanced hardware elements of this environment included an NVIDIA RTX 3090 GPU, along with an Ubuntu 20.04 operating system configured to maximize compatibility with necessary libraries and frameworks. Software-wise, CUDA 11.4 was utilized to speed up processing using the GPU, while PyTorch served as the primary foundation for model construction. In addition, tools such as OpenCV for image preprocessing and Matplotlib for results visualization and analysis were integrated, ensuring a well-structured workflow for the training and validation stages of the system.

4.2. Developed Algorithm

Convolutional Neural Networks (CNNs), in particular the YOLOv5 model, are the basis of the algorithm designed for the identification of personal protective equipment (PPE). This model has been modified to recognize various types of PPE in real-time or previously recorded photographs, both of which were evaluated in this paper. The flow diagram of the process is shown in Figure 3.

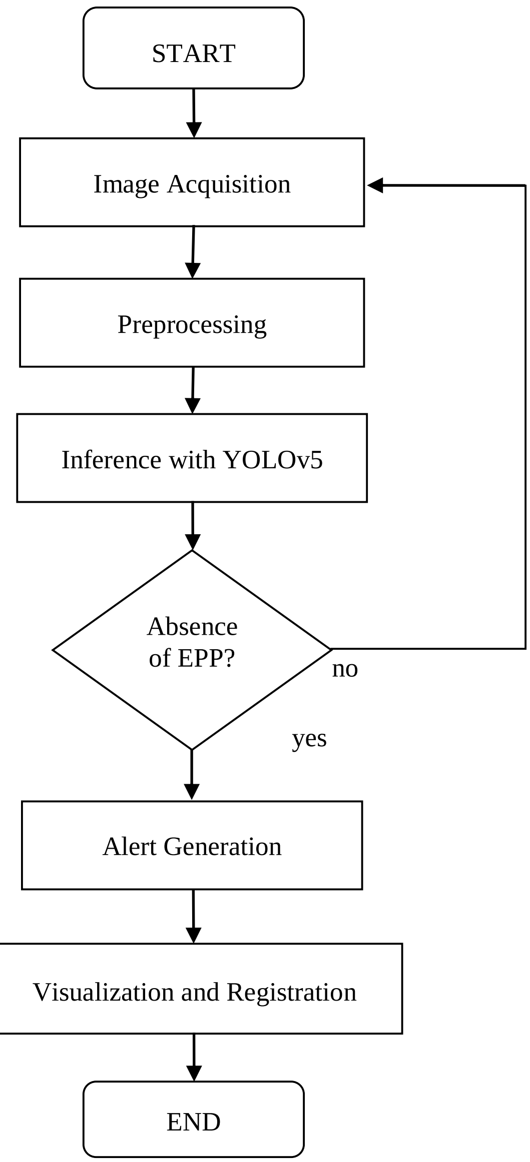

Fig. 3 Algorithm flowchart

The logical flow of the algorithm is described below:

- Image Acquisition: The system captures images from video cameras in real time or processes previously stored images.
- Preprocessing: Images are resized to 640x640 pixels and normalized to meet the requirements of the YOLOv5 model. Quality improvement techniques are also applied in case the images present lighting deficiencies.
- Inference with YOLOv5: The model analyzes the image to identify and classify objects of interest (helmets, glasses, vests, etc.) within the predefined class framework.
- Post-processing: The model outputs are evaluated to filter predictions with low confidence and to identify possible false positives.
- Alert Generation: In case the absence of PPE is detected, the system issues visual and auditory alerts and records the events in a database for later analysis.
- Visualization and Registration: The results are presented to the system operator in a graphical interface and are stored for future audits.

4.3. Building the Dataset

One of the most crucial stages in the development of this proposed system was the creation and use of the dataset for training the learning model, which was specifically created for the PPE identification challenge and included nearly 5,000 images of industrial environments, such as factories, construction sites, assembly plants, among other sources collected by the research team. These images, which were edited to meet the project's specific needs, were collected from public databases and real-time security camera records. The cameras were placed in various industries. The images were painstakingly tagged with YOLO-friendly remarks. The labels included categories like helmets, vests, glasses, and an extra class to indicate a lack of PPE to aid the model in learning to differentiate between various items. To improve the robustness of the model to variations in the environment, data augmentation techniques such as rotation, brightness and contrast changes, and random scaling were applied, thus increasing the diversity of the training images and making them versatile against sudden changes in lighting.

4.4. Model Architecture and Training

Since speed and accuracy are essential factors for real-time applications of this nature, YOLOv5 was selected as the model architecture for this system. The final layers of the network were tuned to match the number of EPP classes in the dataset, thus adapting it to the specific purpose required for this research. In addition, a transfer learning strategy was applied, speeding up the process and improving convergence by starting the training with a model previously trained on the COCO dataset. The hyperparameters were carefully tuned to optimise the performance: a learning rate of 0.001, a batch size of 16, and a total of 50 epochs were achieved. The Adam optimizer and a weight regularization 0.0005 were used to ensure stable training.

The training process was divided into two phases. In the first, the upper layers of the network were fine-tuned with a portion of the dataset, which took approximately six hours of intensive computation. Utilizing the data set that had previously been collected by the cameras and the movies, the entire model was enhanced in the second phase, which required an additional training day. Cross-validation was used in both phases to avoid overfitting. The accuracy of the technique developed, and the loss at each epoch was used to monitor the performance of the model developed by the research team. This method allowed training problems to be identified and real-time adjustments to be made to improve the performance of the proposed model, thus eliminating the need to stop the system for optimization whenever improvements were made and with increasing data.

4.5. Real-Time Testing

Extensive testing was conducted in controlled, real-world industrial environments to evaluate the system's effectiveness developed by the research team in this paper. Test photographs taken under optimal lighting and camera angles were used to evaluate the model in controlled testing. On the other hand, in real scenarios, such as active factories and construction sites, the system faced additional challenges, such as variable lighting, partial occlusion of workers, and rapid movement.

4.6. Integration with the Security System

Finally, an additional module was developed that integrates a real-time alarm system with the detection system. When PPE use is not detected, this module records the incidents in a database for later analysis and generates visual and audible alerts. This additional feature worked well in an industrial setting by automating the monitoring of workers to automatically detect PPE, greatly reducing the need for human intervention and thus decreasing human error. While the model had several drawbacks, such as reduced accuracy in extremely low lighting conditions, the overall results show how this technology has the potential to revolutionize workplace safety by automating crucial activities and reducing losses due to workplace accidents.

5. Results and Discussion

5.1. Model Performance Evaluation

Standard metrics for object detection systems, such as false positive rate (FPR), inference time per image, and mean average precision (mAP), were used to evaluate the performance of the proposed model. The results demonstrated the system's excellent accuracy in recognizing personal protective equipment (PPE), such as helmets, vests, and goggles, achieving an average accuracy of 89.3% on the test set. The model demonstrated a remarkable ability to reliably distinguish between different classes, even in complex cases when multiple workers were present in the same image. Its effectiveness in reducing false positives, which is crucial in systems where errors could generate unnecessary alarms or interfere with real-time operation, was further demonstrated by the false positive rate, which was less than 4%. In terms of processing speed, the model-averaged 45 Frames Per Second (FPS) on high-performance hardware, making it ideal for deployments in dynamic industrial environments requiring real-time analysis.

Table 1. Summary of the model's performance

Metric	Value	Description
Mean Average Precision (mAP)	89.3%	Measures the model's accuracy in detecting PPE, calculated as the area under the precision-recall curve.
False Positive Rate (FPR)	< 4%	Percentage of incorrect predictions where non-existent PPE was detected.
Inference Speed	45 FPS	Number of frames processed per second on high-performance hardware.

Further tests were also carried out to assess the model's performance under more difficult circumstances, including fast motions, partial worker occlusion, and changes in illumination. The accuracy of the suggested model remained above 90% under optimal illumination. However, accuracy dropped a little in poor light, only reaching about 80%. Poor visual quality affected PPE identification by 10%, highlighting system deficiencies and paving the way for improving the system for unfavorable lighting conditions.

The model demonstrated an acceptable level of accuracy in identifying partial occlusions, such as when workers were partially obscured by equipment or other objects, depending on the visibility of the PPE. These results suggest that an approach based solely on computer vision could benefit from incorporating other technologies, such as LiDAR or heat sensors, to improve identification in challenging circumstances. Figure 4 includes some images taken from the PID detection system, made in a real environment of an industry dedicated to manufacturing mechanical parts and structures.

(a)

(b)

(c)

Fig. 4 Video frame of PPE detection system

5.2. Comparative Analysis with Traditional Methods

The Convolutional Neural Network (CNN)-based system was compared with traditional computer vision methods, such as Support Vector Machine (SVM)-based classifiers and Histograms of Oriented Gradients (HOG), to understand the proposed system better. The comparative findings illustrated the outstanding advantages of the YOLOv5-based model in this study developed by the research team. In detecting PPE, the CNN-based method performed 19.3% better than traditional methods, which had an average accuracy of 70%.

This suggests that when implementing this system in a current industrial setting, the CNN-based method can better handle complex situations, such as the presence of multiple people and changes in environmental conditions. In addition, traditional methods showed a higher susceptibility to false positives and false negatives, making them less reliable for real-time applications, again demonstrating that the developed system is better suited for this type of testing. For example, in environments with changing lighting, the HOG+SVM methods recorded a false positive rate close to 15%, in contrast to the 4% achieved by YOLOv5. These findings highlight the impact of modern neural network architectures in improving critical tasks such as safety monitoring in the workplace. Table 2 presents the comparison between the CNN model and the traditional method.

Table 2. Significant advantages of the CNN-based model over traditional methods

Method	Mean Average Precision (mAP)	False Positive Rate (FPR)
HOG + SVM	70.0%	~15%
YOLOv5 (Proposed CNN-based Model)	89.3%	<4%

5.3. Discusión

The results of this study show how convolutional neural networks can be used to recognize personal protective equipment (PPE) in commercial and industrial environments with several people present. Compared to manual or semi-automated approaches used in many industries, the high accuracy and speed of the model allow its use in real-time, improving plant safety. However, to maximize the system's utility, certain obstacles were observed that need to be addressed in future research. For example, limitations in partial occlusion scenarios and dependence on image quality indicate the need to investigate hybrid approaches that integrate computer vision with other sensor technologies. Implementing these future improvements will be essential to increasing the system's effectiveness in various industrial contexts dedicated to fabricating steel structures. Another crucial research component is the system's ability to adapt to diverse industrial conditions. While the model performed well in factories and construction sites, it may not perform well in sectors such as mining or chemical facilities, which have distinctive visual characteristics. Testing in other environments is essential to validate the proposed system further. This emphasizes the importance of using data sets representing various situations during model training. Last but not least, although the system efficiently automates PPE detection, its use in real-world environments must also consider elements such as worker privacy and compliance with regional laws, which may require additional system architecture and functionality modifications.

6. Conclusion

Based on computer vision methods and Convolutional Neural Networks (CNN) using the YOLOv5 model, this study presents a revolutionary system for identifying Personal Protective Equipment (PPE). The suggested approach demonstrated remarkable effectiveness in the automated identification of various types of PPE in industrial environments during tests conducted under controlled conditions simulating real scenarios. The results showed that the system is robust and reliable for testing in real situations, with an average accuracy (mAP) of 89.3% and a false positive rate of less than 4%. One of the main achievements of this research is the incorporation of an optimized workflow covering image preparation and real-time alert generation. By automating the control of proper PPE use, this technology increases workplace safety and reduces the need for human intervention in repetitive and error-prone operations. Furthermore, by processing up to 45 frames per second (enough for most dynamic industrial environments), the model's implementation on high-performance hardware provided its viability for real-time applications. However, the survey also highlighted crucial areas for improvement.

The need to investigate complementary alternatives, such as incorporating more sensors or improving learning strategies for adverse situations, is highlighted by the model's dependence on illumination and its poor performance in partial occlusion environments. To improve the generalization and flexibility of the system, it is also recommended that the training dataset be extended to cover a wider range of industrial environments and types of PPE. In terms of application, the technology has the potential to revolutionize safety monitoring in sectors such as mining, manufacturing and construction.

Acknowledgments

The authors would like to thank Universidad Nacional de San Agustin de Arequipa.

References

[1] Deming Yang, and Ling Yang, "A Deep Learning-Based Framework for Vehicle License Plate Detection," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 1, pp. 1009-1018, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Honggang Wang, "Detection of Personal Protective Equipment (PPE) Using an Anchor Free-Convolutional Neural Network," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 2, pp. 366-374, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Yan Wang, "A Deep Learning-Based Approach for Vision-Based Weeds Detection," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 12, pp. 75-82, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Deepa, K. Jyothi, and Abhishek A. Udupa, "An Improved Depth Estimation Using Stereo Matching and Disparity Refinement Based on Deep Learning," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 11, pp. 552-559, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] William Percy Cuno Zuniga et al., "Design of a Terrain Mapping System for Low-Cost Exploration Robots Based on Stereo Vision," *Przegląd Elektrotechniczny*, vol. 99, no. 5, pp. 270-275, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Abdulaziz Alorf, “A Robust Deep Learning Model for Terrain Slope Estimation,” *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 1, pp. 1231-1245, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Ahatsham Hayat, and Fernando Morgado-Dias, "Deep Learning-Based Automatic Safety Helmet Detection System for Construction Safety," *Applied Sciences*, vol. 12, no. 16, pp. 1-14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Darío G. Lema, Rubén Usamentiaga, and Daniel F. García, “Low-Cost System for Real-Time Verification of Personal Protective Equipment in Industrial Facilities Using Edge Computing Devices,” *Journal of Real-Time Image Processing*, vol. 20, pp. 1-18, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Hongzhen Wang et al., "Gated Convolutional Neural Network for Semantic Segmentation in High-Resolution Images," *Remote Sensing*, vol. 9, no. 5, pp. 1-15, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[10] Fabio Oliveira-Teixeira, Thiago Pedro Donadon-Homem, and Alexandre Pereira-Junior, “Artificial Intelligence Application to Monitor the Use of Protective Masks,” *Emerging Risks and Opportunities in the Latin American and Caribbean Region*, vol. 19, no. 33, pp. 205-222, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Ryan Abraham León Montero, and Jessica Ayme León Montero, “Development of an Artificial Vision Algorithm for the Recognition of the Correct Use of Epp,” *4th LACCEI International Multiconference on Entrepreneurship, Innovation and Regional Development - LEIRD 2024*, pp. 1-11, 2024. [CrossRef] [Publisher Link]

[12] Hafiz Mughees Ahmad, and Afshin Rahimi, “SH17: A Dataset for Human Safety and Personal Protective Equipment Detection in Manufacturing Industry,” *Journal of Safety Science and Resilience*, pp. 1-32, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Adrian Șandru et al., “SuPEr-SAM: Using the Supervision Signal from a Pose Estimator to Train a Spatial Attention Module for Personal Protective Equipment Recognition,” *Arxiv*, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Mohammed Imran Basheer Ahmed et al., “Personal Protective Equipment Detection: A Deep-Learning-Based Sustainable Approach,” *Sustainability*, vol. 15, no. 18, pp. 1-18, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Jye-Hwang Lo, Lee-Kuo Lin, and Chu-Chun Hung, “Real-Time Personal Protective Equipment Compliance Detection Based on Deep Learning Algorithm,” *Sustainability*, vol. 15, no. 1, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] M. Massiris et al., “Automated System for Monitoring the Use of Personal Protective Equipment in the Construction Industry,” *Ibero-American Journal of Automation and Industrial Computing*, vol. 18, no. 1, pp. 68-74, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Jonathan Karlsson et al., “Visual Detection of Personal Protective Equipment and Safety Gear on Industry Workers,” *Arxiv*, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Md. Shariful Islam et al., “A Deep Learning Approach to Detect Complete Safety Equipment for Construction Workers Based on YOLOv7,” *Arxiv*, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]