**PERSONAL PROTECTIVE EQUIPMENT DETECTION IN INDUSTRIAL FACILITIES USING CAMERA VIDEO STREAMING**

Lucas Nascimento1; July Macêdo1; Caio Souto Maior1; João Santana1; Márcio Moura2; Ísis Lins2

Center for Risk Analysis, Reliability and Environmental Modeling; Department of Production Engineering; Federal University of Pernambuco

Enrique Droguett3

Mechanical Engineering Department; University of Chile

**ABSTRACT**

Organizations and industries must ensure safe operation of their facilities, employing rigorous risk management techniques for planning and executing their activities. The use of personal protective equipment (PPE) represents the closest layer of protection to workers and can considerably reduce the risk of exposure to hazards, being critical for safety in industrial environments. The hazards addressed by protective equipment include physical, acoustic, electrical, heat and chemicals. Despite the substantial efforts in increasing awareness about the benefits of PPE to strive towards zero accident philosophy, operators often neglect its use when not being supervised. However, organizations commonly have surveillance cameras installed which might provide useful visual information on correct usage of PPE. In this context, computer vision is an interdisciplinary field that seeks to automate tasks that the human visual system can do and includes domains of signal and image processing, pattern recognition and artificial intelligence. Moreover, object recognition is a prominent technology from computer vision for finding and identifying objects in an image or video sequence. Then, this work aims to create an automatic PPE detection from surveillance cameras and other video streams using computer vision and machine learning. Equipment such as helmets, safety glasses, earplugs and other garments are checked for whether they are being used by operators in a real-time monitoring, alerting supervisors to prevent accidents and ensure a safer environment

1. **INTRODUCTION**

Even with the scientific and technological progress, statistics provided by the International Labour Organization (ILO) demonstrate that working conditions in many countries (e.g. European Union) have not changed to such a degree as to significantly reduce the problem of occupational injuries [1]. Therefore, every effort to decrease the number of accidents or, at least, maintain its rate at an acceptable range is highly important, and can be employed either by organizational actions, collective training or individual safeguard.

The traditional approach to avoid loss is the implementation of barriers, which plays a central role in the prevention of accidents. [2] defines safety barriers as ‘physical and/or non-physical means planned to prevent, control, or mitigate undesired events or accidents’.

Indeed, there are many opportunities to interrupt or change an accident sequence of events before it evolves into a loss. First, an answer is to change the preconditions for an accident to occur by eliminating the energy source or modifying the energy characteristics from the hazard. Second, barriers may interrupt, dilute, or redirect the energy flow during the latter part of the accident process (e.g. separating the victim from the energy flow). As last barrier, it is possible to improve the victim’s ability to endure the energy flow (e.g. wearing some protective equipment), which is the ultimate protection to avoid damage [3].

In this context, Personal Protective Equipment (PPE) is usually adopted to protect the individual against health or safety risks at work. It includes items related with protection of head, face, eye, hand, arms, and legs [4]. There are consolidated regulations for the usage of PPE in industries that aim to decrease the frequency of misuse or absence of PPE[5][6]. Also, PPE’s positive impacts are very significative (e.g. rate of eye injury and lost work time can be reduced by 50% or more when PPE is worn [7].

Head, as a vital body part containing possibly the most important human organ, needs appropriate attention. Every year, approximately 1.7 million people are hospitalized or die as a result of a traumatic brain injury (TBI) only in the United States [8]. Protective headgear and helmets decrease the potential for severe TBI following a collision by reducing the acceleration of the head upon impact, thereby decreasing both the brain-skull collision, as well as the sudden deceleration induced axonal injury [9].

There are several types for head protection such as industrial safety helmets, bump caps and firefighters’ helmets. The use of those equipment is necessary in activities like low-level fixed objects with risk of collision (e.g. pipework, machines, scaffolding) and transport activities involving the risk of falling material (e.g. hoists, lifting plant, conveyors) [10]. The energy absorbing material of a helmet compresses itself to absorb force during the collision and slowly restores itself to its original shape. This compression and restoration has the effect of prolonging the duration of the collision, while reducing the total momentum transferred to the head [11].

The problem relies on the fact that, even with understanding about the safety improvement that the usage of PPE leads, its usage is often neglected in industry. The report of the ILO estimates that 2.34 million people die every year in the world due to occupational accidents, some of these deaths caused by non-use of PPE [12]. A common approach is to impose fines and penalization to workers, who do not wear the required PPE when performing specific activities. However, supervision to guarantee its use is normally performed in person by a higher-level employee, which makes almost impossible to control all operators during the whole labor time.

Actually, there is an extensive discussion concerning ethical issues in workplace surveillance, referring to management’s ability to monitor, record and track employee performance, behaviors and personal characteristics in real time [13]. Most of the discussion involves the so-called Electronic Performance Monitoring (EPM) about employee’s control in social and technological forms (e.g. Internet and email monitoring, location tracking, biometrics) and the understanding of privacy boundaries surrounding employee information[14][15] . However, our discussion is to assure that proper safety protocol is followed, preventing injury to employees, as well as avoiding damage to the assets through a consistent and trustworthy model.

Therefore, an automatic method for monitoring PPE usage presumably is significant worthy for industrial safety, representing an impactful opportunity for the use of Computer Vision (CV). CV is an interdisciplinary field aiming to investigate and develop computers with high-level understanding from digital images or videos, describing the world that we see and to reconstruct its properties [16]. From the perspective of engineering, it seeks to automate tasks that the human visual system can do [17]. The development of high-powered computers, the availability of high quality and inexpensive video cameras, and the increasing need for automated video analysis has generated a great deal of interest in object tracking algorithms in CV field [18]. Therefore, this paper aims to develop a model for automatic PPE detection from industrial video streams using computer vision and machine learning, employing modern technologies to create tools capable of modifying and innovating methods currently used in industries.

The rest of this paper is organized as follows: Section II introduces some ideas and concepts on CV, while Section III presents some works of object/person detection, describing the methodology applied in PPE detection. Section IV demonstrates the developed model, Section V provides possible usages as decision supporter and Section VI concludes remarks

1. **COMPUTER VISION**

Computer vision (CV) studies the automated extraction of information from images and videos. Information can mean anything from 3D models, camera position, object detection and recognition to grouping and searching image content [19]. CV gathers knowledge from many fields, such as image processing, pattern recognition, mathematics and artificial intelligence. One of its main goal is to enable computers to reproduce core functions of human vision, such as motion perception and scene understanding.

Hence, visual object tracking have been constantly studied and presents three key steps for detection in video analysis: detection of movement of objects, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior [18]. Essentially, the basis of visual object tracking is to robustly estimate the motion state (i.e., location, orientation, size, etc.) of a target object in each frame of an input image sequence [20].

Specifically, intelligent visual surveillance systems deal with the real-time monitoring of persistent and transient objects within a specific environment [21]. The goal of these systems is not only to put cameras in the place of human eyes, but create an entire surveillance system as automatically as possible [22].

There exist some well-known visual surveillance systems such as W4 [23]; Haar-wavelet Adaboost[24] and ViBe[25], mainly developed to detect different vehicles types, groups of people, pedestrians, people access control. Every system is developed seeking to compensate the capability limitation of human operators in monitoring enormous number of cameras at the same time. Thus, exploring similar tools and challenges to detect usage of PPE in order to avoid accidents in industries represents an interesting case.

1. **REAL TIME OBJECT DETECTION**

Techniques from statistical pattern recognition have, since the revival of neural networks, obtained a widespread use in digital image processing [26]. Due to the outstanding work of [27] for image classification, Deep Neural Networks (DNN) have been successfully studied in different fields of application such as speech recognition [28], vibration analysis [29], electronic nose data [30] and physiological data [31]. However, surely, the most promising results are found in the field of computer vision, bringing impressive developing in tasks like automatic object and face recognition.

One promising, open and free project that uses DNN for object detection is You only look once (YOLO). YOLO is a system for detecting objects and was first created on the Pascal VOC 2012 dataset, detecting the 20 Pascal object classes, such as person, birds, dogs, car, bicycle, bottle, table and chair, as could be seen in Fig. 1 [32]

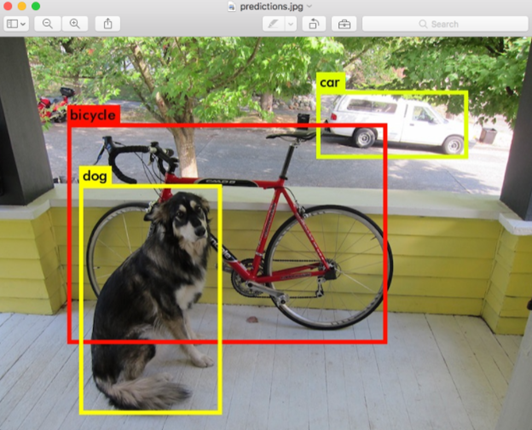


Figure 1 - Example of object detection using YOLO. Adapted from Redmon et al. (2015).

The developers adopt a different approach than the standard object detection models that uses classifier based-systems applied at multiple locations and scales in an image, which typically considered as detections high scored regions of the image. In YOLO, a single DNN is executed to the full image. This network divides the image into regions and predicts bounding boxes and probabilities for each region, with these bounding boxes being weighted by the predicted probabilities. It has considerable advantages to other object detection models, once it looks at the whole image, and then its predictions are informed by global context in the image [32]. It also makes predictions with a single network evaluation, making YOLO extremely fast, allowing usage even for computers without a Graphics Processing Unit (GPU).

Still, an improved model, YOLOv2, has already been developed. More robust, detecting more than 9000 objects without losing real-time performance, YOLOv2 is a state-of-art object detection system with results comparable or with even better than many other systems [33]. Moreover, the YOLO project is open, well described, easily explained and user-friendly to anyone, who has some basis in computer programming. It even demonstrates how to include objects that were not on its detection basis, how to process and train a new model, allowing adaptation for different purposes.

Hence, using YOLOv2 as a key tool, we trained a new model to automatically detect PPE usage. Specifically, we were interested in identifying whether workers were wearing or not a safety helmet when performing some activities in which the protection was required

1. **HELMET DETECTION USING YOLO**

YOLO project easily provides a pre-trained model, which could be used as a basis for detecting new types of objects. As any machine learning algorithm, YOLO requires a training dataset that will ‘teach’ the machine how an unknow object looks like. For our specific goal, 731 images containing helmets were used to give sufficient information about its appearance. All images were obtained from ImageNet [34], and location of helmets in images were annotated manually. ImageNet is an image database organized according to the WordNet [35] hierarchy, in which each node of the hierarchy is depicted by hundreds and thousands of images. It presents useful resource for researchers that needs image data, containing innumerous classes of items.

The network was trained for about 8 hours, running in a Nvidia GeForce GTX 960m GPU, with 4GB of video random access memory (VRAM). Once the algorithm finished its training, our helmet detection model could be applied to a specific imageor to a video stream, such as a camera feed, processing every frame. Our model runs in real-time, maintaining the frame rate of the camera (30 frames per second – FPS). Fig. 2 depicts the model applied to a standard web camera video streaming.

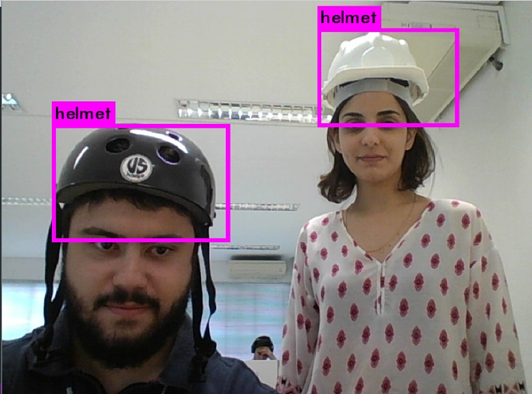


Figure 2 - Helmet detection model applied for video streaming

Then, a script was created to alert surveillance operators whether an abnormal situation appears (i.e. helmets were not detected). Due to simplifications proposed, the model aims to be applied in a room containing specific number of employees that should be using helmets. For each frame, our algorithm detects the use of helmet and counts how many are present in the scene. If the number of detected helmets is different from number of previously defiRned people in the room for more than a brief period (e.g. 10 seconds, or 30 seconds), then an alert was emitted. Fig. 3 shows a computer screen when an alert was presented (i.e. one person is not using helmet in the image).

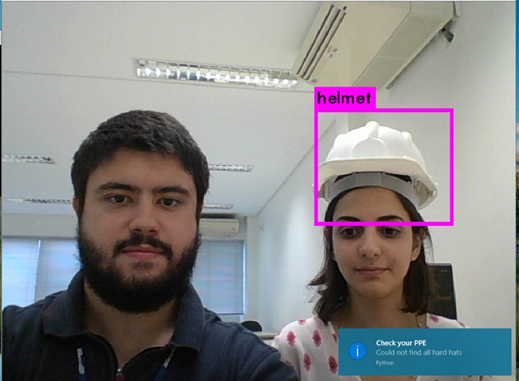


Figure 3 - Alert emitted when model detect anomaly situation

Moreover, if the number of detected helmets does not return to its normal value, alerts may continue to be emitted, but the period between alerts may be defined as desired (e. g. 30 seconds, or two minutes). Properly adjusting this period avoids unnecessary alerts that will continually distract the surveillance operator even after recognition of first anomaly situation, while still reminding that an abnormal situation is ongoing.

1. **SUPPORTING DECISIONS**

In practice, the presented model could be explored as a tool in different contexts, supporting decisions for the safety manager. The idea was to develop the model to be highly adaptable and manageable for various situations, providing specific information accordingly.

For example, as aforementioned, the alert period is easily adjustable to avoid unnecessary warnings. Still, other types of warnings (e.g. depending on how long operators remains without PPE; how many operators are not wearing the PPE) could be easily implemented and customable, providing information for the decision-maker to determine whether or not someone must be notified. Moreover, it is also possible to use information and statistics provided by the model (e.g. how many times alerts were displayed per day; who long operators had remained without PPE) as a safety indicator.

Still, implementation of a real time alert for the operators (e.g. a particular warning light is lit somewhere in the room) connected with the model would emphasize (or create) the sense of autoregulation among them, reducing (or sharing) the surveilling word load expected for the supervisor.

With further and wider adaptations, this surveillance technology could be also implemented to monitor other barriers than PPE. Related with the initial barriers, for example, CV could be used to detect hazards (e.g. fire, toxic gases) and serve as redundancy to sensors responsible for interrupting the energy flow in case of accidents. For inner barriers, it is possible to create alerts and warnings for whether an operator approaches a danger zone based in images of the area. In both previously mentioned barrier, CV would help to eliminate or reduce the consequences of unwanted energy flow, rather than dealing with the last safety impediment represented by the PPE.

1. **CONCLUSIONS**

This paper presented an approach for automatically detecting PPE usage in a controlled environment, using object detection with YOLO. By using YOLO, this method achieves a reasonable balance between speed and confidence, running in real-time, which results in relative low computational resource usage. Moreover, it is possible to adjust the model to different scenarios according to specific requirements. This could lead to beneficial results to safety engineering since the detection is performed automatically and does not require constant human attention.

As a matter of our current research, we aim at extending this model for application in a wider range of situations. For instance, YOLO can be trained to identify other types of PPEs, so that it could be used for simultaneously monitoring usage of different PPEs. Also, the script used for alerts could be improved, allowing this method to cover a wider range of scenarios. A further step is to use real surveillance videos as input, detecting the usage of PPE in realistic environment, preventing accidents and providing an improvement on the safety monitoring system of industries.

1. **REFERENCES:**

[1] Cavazza N, Serpe A. Effects of safety climate on safety norm violations: exploring the mediating role of attitudinal ambivalence toward personal protective equipment. *Journal of Safety Research*. 2009. 40(4), 277–283.

[2] Sklet S. Safety barriers: Definition , classification , and performance. *Journal of Loss Prevention in the Process Industries*. 2006. 19, 494–506.

[3] Kjellén U, Albrechtsen E. *Prevention of Accidents and Unwanted Occurrences*. 2017. CRC Press.

[4] Health and Safety Executive. *Personal protective equipment (PPE) at work: A brief guide*. HSE Books. 2013.

[5] Occupational Safety and Health Administration (OSHA). Personal Protective Equipment. (2004).

[6] U.S. Department of Homeland Security. Guide for the Selection of Personal Protective Equipment for Emergency First Responders. 2002. I(November), 1–118.

[7] Lipscomb H J. Effectiveness of interventions to prevent work-related eye injuries. *American Journal of Preventive Medicine*. 2000. 18(4 Suppl), 27–32.

[8] McCrory P, Meeuwisse W, Johnston K, Dvorak J, Aubry M, Molloy M, Cantu R. Consensus statement on concussion in sport: The 3rd International Conference on Concussion in Sport held in Zurich, November 2008. In Journal of Athletic Training. 2009. (Vol. 44, pp. 434–448).

[9] Newman J A, Beusenberg M C, Shewchenko N, Withnall C, Fournier E. Verification of biomechanical methods employed in a comprehensive study of mild traumatic brain injury and the effectiveness of American football helmets. *Journal of Biomechanics*. 2005. 38(7), 1469–1481.

[10] Health and Safety Executive. *Personal protective equipment at work*. 2015. (Vol. 25).

[11] Pellman E J, Viano D C, Withnall C, Shewchenko N, Bir C A, Halstead P D. Concussion in professional football: Helmet testing to assess impact performance - Part 11. *Neurosurgery*. 2006. 58(1), 78–95.

[12] International Labour Office. *ILO introductory report: global trends and challenges on occupational safety and health: XIX World Congress on Safety and Health at Work*. 2011.

[13] Ball K. Workplace surveillance: an overview. *Labor History*. 2010. 51(1).

[14] Alder G S. Ethical Issues in Electronic Monitoring: A Consideration of Deontological Perspectives. *Journal of Business Ethics*. 1998. 17(7), 729–743.

[15] Allen M W, Coopman S J, Hart J L, Walker K L. Workplace Surveillance and Managing Privacy Boundaries. *Labor History*. 2015. 51(1), 172–200.

[16] Szeliski R. *Computer Vision : Algorithms and Applications*. 2010. Springer.

[17] Sonka M, Hlavac V, Boyle R. *Image Processing, Analysis, and Machine Vision*. Thomson Learning. 2008.

[18] Yilmaz A, Javed O, Shah M. Object tracking: A survey. *ACM Comput. Surv.* 2006. 38(4), 13.

[19] Jan Erik Solem. Programming Computer Vision with Python. *Programming Computer Vision with Python*. 2012. 264.

[20] Li X, Hu W, Shen C, Zhang Z, Dick A, Hengel A Van Den. A survey of appearance models in visual object tracking. *ACM Transactions on Intelligent Systems and Technology*. 2013. 4(4), 1–48.

[21] Valera M, Velastin S A. Intelligent distributed surveillance systems: a review. *IEE Proceedings - Vision, Image, and Signal Processing*. 2005. 152(2), 192.

[22] Hu W, Tan T, Wang L, Maybank S. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*. 2004. 34(3), 334–352.

[23] Haritaoglu I, Harwood D, Davis L S. W4: Real-time surveillance of people and their activities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000. 22(8), 809–830.

[24] Enzweiler M, Gavrila D M. Monocular pedestrian detection: Survey and experiments. In IEEE Transactions on Pattern Analysis and Machine Intelligence. 2009. (Vol. 31, pp. 2179–2195).

[25] Barnich O, Van Droogenbroeck M. ViBe: A universal background subtraction algorithm for video sequences. *IEEE Transactions on Image Processing*. 2011. 20(6), 1709–1724.

[26] Egmont-Petersen M, de Ridder D, Handels H. Image processing with neural networks—a review. *Pattern Recognition*. 2002. 35(10), 2279–2301.

[27] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks. *Advances In Neural Information Processing Systems*. 2012. 1–9.

[28] Hinton G, Deng L, Yu D, Dahl G, Mohamed A R, Jaitly N, Senior A, Vanhoucke V, Nguyen P, Sainath T, Kingsbury B. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*. 2012. 29(6), 82–97.

[29] Guo X, Chen L, Shen C. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Measurement: Journal of the International Measurement Confederation*. 2016. 93, 490–502.

[30] Längkvist M, Coradeschi S, Loutfi A, Balaguru Rayappan J B. Fast classification of meat spoilage markers using nanostructured ZnO thin films and unsupervised feature learning. *Sensors (Switzerland)*. 2013. 13(2), 1578–1592.

[31] Mirowski P W, LeCun Y, Madhavan D, Kuzniecky R. Comparing SVM and convolutional networks for epileptic seizure prediction from intracranial EEG. In Proceedings of the 2008 IEEE Workshop on Machine Learning for Signal Processing, MLSP 2008. 2008. (pp. 244–249).

[32] Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. 2015.

[33] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger. *Conference on Computer Vision and Pattern Recognition*. 2017. 7(3).

[34] Jia Deng, Wei Dong, Socher R, Li-Jia Li, Kai Li, Li Fei-Fei. ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 2009. 248–255.

[35] Fellbaum C. WordNet: An Electronic Lexical Database. *MIT Press, Cambridge, London, England*. 1998.