Development Of A Real-Time Monitoring System For Detect The Use Of Personal Protective Equipment (PPE) From Machine Learning

Gustavo G. F. Santos[1], Flávio G. Pereira[1], Adilson R. Prado[1]

[1]*Dept. Engenharia de Controle e Automacão, Instituto Federal do Espírito Santo*
Av. dos Sabiás, 330 - Morada de Laranjeiras, 29166-630, Serra - ES, Brazil
gustavo.gregorios@gmail.com, flavio.garcia@ifes.edu.br,

Abstract. PPE is the equipment for individual use used by the worker where its main purpose is to protect against risks capable of jeopardizing their health and safety, in addition also reducing costs to the employer with personnel replacements, dismissals and indemnity processes. However, many times, either through negligence or discomfort, there is resistance to its use and/or the removal of the equipment during the performance of activities. In view of this problem, the HSE department must therefore inspect and monitor the proper use of workers' personal protective equipment mostly of the time. As an alternative to assist the security department in verifying, demand and quantifying the use of protective equipment in the workplace, this project presents a machine learning model based on the You-Only-Look-Once (YOLO) architecture to verify the workers' compliance regarding their safety behavior in real time, using images / video of a security system installed in a busy place within an industry. The algorithm uses the approach of detecting workers and basic PPE as helmet, gloves, goggles simultaneously by deep learning previously trained by a image dataset of workers in differents types of labour ambient, and next verifies that each bounding box generated is in the correct position, thus confirming if the worker is carrying PPE or not. Later, a program developed in python using the opencv library will quantify the use or not of the use of PPEs, from the bounding boxes generated by more specifically YOLOv4, providing in this way a statistical report as an output. This statistical report will be useful for the HSE (Health, Safety and Environment) department to use its statistics as indicators of reliability that will assist in decision-making and in the management of security within the company.

Keywords: PPE, Safety, Machine learning, Deep Learning, YOLO

1 Introduction

In the vast majority of work activities, from the simplest to the most complex, there will be always a risk of some type of accident occur, accidents in which they can injure the integrity of the worker, which can cause lost time accident, get some sequel or even die [1].

In order to avoid such injuries to the worker and when there are no preventive measures capable of eliminating the risks of environment where the activity takes place, in other words, when the measures of collective protection are not viable, efficient and sufficient for the mitigation of accidents at work and/or occupational diseases the use of personal protective equipment (PPE) is necessary. PPE is the first individual barrier for amortization of any injury in the event of an accident or exposure to some risk [2].

The use of protective equipment is determined by a technical standard. In Ordinance No. 3214, of June 8, 1978, which approves the Regulatory Norms (NR), Chapter V, Title II, of the Consolidation of Labor Laws, relating to Occupational Safety and Medicine, The regulatory norm number 6 (NR-6) stands out, this standard establishes several regulations evidencing responsibilities of employer, employee and manufacturers as to their supply, use and certification of the Personal Protective Equipment [3].

The rule determines that the employer's obligation, provide the PPE free of charge to the worker for the performance of their functions within the company, as well as the employee must use it only for the purpose for which it is intended and comply with the employer determinations on proper use [3].

Despite its great importance to ensure the worker integrity, there is often resistance to its use or the removal of equipment during the activities execution. For reasons such as forgetfulness, discomfort, negligence, lack of correct information about the use, mobility limitation and even lack of sensitivity lead to the removal of the equipment during the execution of activities [4].

It's obligation of supervisors and the company to demand the PPE use and ensure that professionals utilize it properly[3].

In order to guarantee the use of PPE, the company need a instructed HSE team, but to satisfy modern safety management requirements, relying only on traditional manual supervision becomes hard work. In this context, it is important to study the automatic detection and recognition of objects in the work environment [5], currently the methods of object detection via computer vision and artificial intelligence have been increasing significantly, mainly through Convolutional Neural Networks. Such a network has the function of extracting patterns and characteristics from a set of images, through previous training, and it manages to classify these characteristics according to a specific object, and later perform the detection of these same objects in images later exposed to this network [6].

Regarding of the problem of non-use of PPE linked to the difficulty of inspection, the present work aims to develop an automatic identification system for basic PPE (helmet, gloves, safety glasses) from machine learning strategies, using convolutional neural networks (CNN). Such system will help the HSE team to verify, warn and quantify the use of protective equipment in the workplace. The data generated will be used to assist the company in decision making and increase security performance.

This work is divided into five chapters, the first contains the abstract and the introduction with the context of the proposed problem. The second deals with the literature review which the project was based on. The third chapter describes the methodology used to carry out the research. In chapter four are presented the results of the project and finally in the last chapter it is defined as conclusions and future proposals.

2 Literature Review

In the fundamental guidelines of work, the worker has the right to enjoy a healthy and pleasant quality of life [7].

With the objective of reducing the risks of accidents at work, the Occupational Safety emerged, a science aimed to creating preventive methods and technical measures that aim to eradicate, or at least mitigate the occurrence of accidents at work [7].

In Brazil, were created 37 Regulatory Norms (NRs). The NRs are rules whose content covers rights, duties and obligations to be fulfilled by employers and workers, being mandatory for private and public companies that are under the Consolidation of labor laws regime. The standards bring the minimum safety requirements and impositions for work, addressed in different topics according to each NR [7].

This article is based on NR 06, which deals with personal protective equipment and its guidelines. This standard deals with the guidelines for the supply, use and certification of PPE.

2.1 Regulatory Norm 06 and PPE

The Regulatory Norm 06 defines Personal Protective Equipment - PPE, every device or product, for individual use, used by the worker, which is intended to protect against risks that have the potential to threaten the health of the worker in the work environment [3]. According to NR 06, in item 6.6.1, it is up to the employer regarding the PPE:

(a) acquire PPE suitable for the risk of each activity;
(b) demand its use;
(c) provide PPE approved by the Ministry of Manpower;
(d) guide and train the worker on the proper use, storage and conservation;
(e) replace immediately when damaged or lost;
(f) be responsible for sanitation and periodic maintenance; and,
(g) report any irregularities observed to the Ministry of Manpower.

Also according to the standard, in item 6.6.7, it is up to the employee regarding the PPE:

(a) using it only for the purpose for which it was intended;
(b) be responsible for the care and conservation;
(c) communicate to the employer any change that makes it unfit for use; and,
(d) comply with employer determinations on proper use.

Among the responsibilities mentioned, it is evident that in sub-item (a) of item 6.6.7, the employee has the obligation to use the PPE for his work activity in which he was conducted. However, according to [9], for reasons such as forgetfulness, discomfort, negligence or lack of adequate instruction, they generate resistance to its use during the execution of the work, and it is up to the employer to follow sub-item (b) of item 6.6.1, and for that, the company must have a well trained HSE department responsible for the acquisition, distribution, packaging, instruction of the PPE.

2.2 Noncompliance with Regulatory Norms

Companies must strictly follow regulatory norms to preserve the safety and health of workers; in case of non-compliance, they are subject to sanctions by the Federal Government [8].

The Ministry of Manpower is responsible for verifying that the norms are being respected; if they are not, sanctions vary from claims and public civil actions to payment of fines. The main sanctions are [8]:

- Fines imposed by the Ministry of Manpower;
- Prohibition of the establishment, machinery and equipment;
- Payment of health and hazard additional;
- Risk of death or health is characterized as a Danger Crime (Article 132 of the Brazilian Penal Code);
- Physical damage is characterized as Bodily Injury (article 129, §6, of the Brazilian Penal Code).
- Death of the worker resulting from non-compliance with safety standards, the case is treated as a homicide (Article 121 of the Brazilian Penal Code).

2.3 Convolutional Neural Networkss

A Convolutional Neural Network - CNN, is a variation of a multilayered Perceptrons network, they are a Deep Neural Network, the data processing is done from multidimensional matrices originated from the image, the network applies various visual data filters, maintaining the relationship between image pixels throughout the network processing at each layer. Figure 1 illustrates a CNN [10].

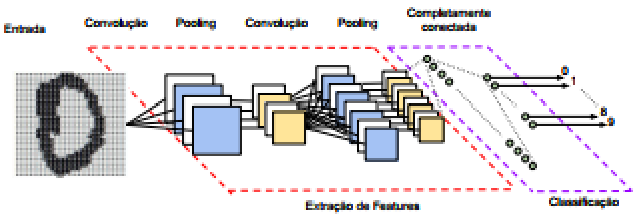

Figure 1. Architecture of a CNN model [10].

In a convolutional neural network there are multiple layers with different functions. Initially, filters called convolution are applied to the input layers. A convolution layer is composed of several neurons, each responsible for applying a filter to a specific piece of the image. Each neuron is connected to a set of pixels from the previous layer and each of these connections is assigned a weight. Using the weight previously assigned to each neuron and the combination of its inputs, an output is produced that is passed to the next layer[10].

Subsequently, an activation function is applied to the convolution filter, which performs transformations on the data received in each neuron[10].

Finally, the data is passed through a grouping layer, called pooling, which has the function of resizing the network data, increasing the network agility and creates spatial invariant [10].

For a convolutional neural network to be able to perform a classification task, at least one layer of fully connected neurons must be added after the convolution, activation and polling set, this layer traces the decision paths from the data coming from the filters of previous layers of each response class [10].

Convolutional neural networks are being widely used for image classification and object detection. At the moment, networks based on convolutional neural networks capable of real-time object detection are SSD (Single Shot Detector), YOLO (You-OnlyLook-Once), R-FCN (region-based fully convolutional network), and RetinaNet [11] .

According to Nipun D. Nath et al. [11] The challenge of achieving real-time computing is to maintain satisfactory accuracy. Although most of the time, fast algorithms compromise significantly the accuracy to get real-time computing, so far only YOLO is faster and more accurate than the alternatives described above.

YOLO predictions are performed by a network that can be easily trained end-to-end to improve performance. For industrial purposes the high efficiency combined with its higher speed make YOLO a reasonable choice for real-time image processing[11].

The resources of the application of computer vision for problem solving are becoming more and more popular, this method is being used extensively for the monitoring of unsafe behavior, particularly in the detection of PPE, given the wide range of data that can be acquired in videos and images from surveillance cameras [12].

2.4 YOLO

The YOLO network on which this project based was first proposed by Joseph Redmon et al [13]. The YOLO is a deep neural network with 24 convolutional layers followed by 2 fully connected layers. The first convolutional layers are responsible to extract features from the image while the fully connected layers predict the output probabilities and bounding box coordinates by using regression problem, direct from image pixels to bounding box coordinates and class probabilities by seeing the whole image during training testing time. The network uses full-image resources to predict each bounding box[13].

The input image is split into S × S squares. Each square within the grid predicts B number of bounding boxes and their confidence values together with the classes C. The confidence is defined as prediction of the object (Pr(Object)) x IOU, its show how confident is the model placing if in that box contain the object and how accurate the network think it is [13].

If there is no object in that cell, the confidence scores need to be zero. If the object exist the the confidence score is equal the intersection over union (IOU) [13].

To calculate the IoU its taken the intersecting area between the bounding boxes and the ground truth bounding boxes of the same area and the total area covered by the two bounding boxes—also known as the Union[14].

The intersection divided by the Union is the IOU, meaning by this way the ratio between the overlap to the total area, estimating how close prediction bounding box to the original class. As shown at figure 2[14].

$$IOU = \frac{\text{area of overlap}}{\text{area of union}} =$$

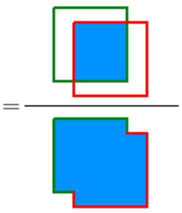

Figure 2. Intersection Over Union (IOU) [14]

At the end each bounding box is represented by five values: *bx, by, bw, bh*, and confidence, the output is a tensor of shape $S \times S\ (5 \times B + C)$. When multiple frames are predicting the same object, the network uses a non-maximum suppression technique to find the most suitable frame[15].

This article focus on YOLOv4, this was proposed by Bochkovskiy et. al. in 2020 as an improvement of the versions before invented, it follows the same theory of the firts yolo network, but include a combination of changes in architectural design and training methodologies, with a much more complex convolutional network in its architecture [15]. The high-level architecture of YOLOv4 network is composed by three main parts, the *backbone* with CSPDarknet53 CNN used as feature extractor, the *neck* a set of extra layers used to extract different features of the backbone layers and the *head* also known as dense prediction it the part to carry out the detection and the bounding boxes regression [15]. YOLOv4 outperforms the other members in the YOLO family with an Average Precision of 43.5% on COCO dataset with 65 FPS in a Tesla V100 [15].

2.5 Related works

In the literature, the use of CNN is widely used for the detection of PPE. For example, LI, Y. et al. [5] concerned with the safety of employees involved in construction civilians they have developed a method of detecting safety helmets using convolutional networks, with the aim of contributing to the monitoring of helmet use at the construction site, the authors consider manual verification as time-consuming, error-prone and not sufficient to meet standards modern building. The methodology used was to use convolutional neural networks to perform the, the authors used the YOLO and SSD network.

CHEN, S.; DEMACHI, K. [16] developed a project whose main methodology is to use convolutional neural networks to identify helmet and mask and to use a body keypoint extraction program (OpenPose) to determine if the helmet and mask are in the correct places.

NATH, N. D. et al [11], and proposes the automatic detection of work clothes and helmets as the objective. The article comments on other existing methods for analyzing the use of PPE, in which they do not use machine learning, such as PPE with RFIDs that verify the location of the worker and whether he is using the PPE or not. The project's methodology is the use of deep learning, using the YOLOv3 network to identify helmet, clothing.

In a different approach, but totally linked to safety FANG, W. et al. [12] developed a method that does not recognize PPE, but rather identifies the safe positioning of the worker at the construction site. The authors developed, using deep learning, more precisely a neural network called Mask Region Based Convolutional Neural Network (R-CNN), a method that checks if the worker is very close to the edge of some structure, whether it has a guardrail or not, thus identifying the potential hazard of fall from a height.

3 METHODOLOGY

The PPE defined for identification in this project were the helmet, goggles and gloves, the definition of identifying these protective equipment was based on the understanding during the bibliographic review that they are the most used in the work environment in various industrial segments. , and in addition to being responsible for protecting the main limbs and organs of the human body, where intact they promote the well-being of the employee during his daily and working life, the helmet protects the brain from injuries that can cause sequelae or even death, the glasses protection because it protects the eyes from exposure to particulates / radiation that can cause wounds or even loss of vision, and glove because it protects from injuries to the hands, being from cuts or even crushing.

The project was divided into six main parts: the acquisition of images, class annotation of each PPEs, quantification of the annotations in the database, the development of the neural network algorithm capable of classifying and demarcating people using or not PPE, changes in the neural network to obtain the best performance and the development of an algorithm that uses the already trained network to identify PPE frame by frame in a video.

The dataset used for the elaboration of the project contains 7555 images of people in a work environment wearing PPE or not, all images are in the public domain, with 7041 images acquired from the website "https://public.roboflow.com/" [17] and the other 514 acquired by Google image. Image resolutions range from 151x184 to 7360x4912.

For the development of the deep neural network algorithm, the Darknet framework and the YOLOv4 network were used, as mentioned before a network capable of achieving a much higher detection speed than competing techniques, without losing in accuracy.

Like any other convolutional neural network, YOLOv4 also needs to be fed with pre-classified data, by default this network uses for each image in the database a ".txt" file with the same name as the image file. This text file is responsible for informing the network about the classes of objects as well as their locations in the image. To locate the object in the image, YOLOv4 works with bounding boxes, at a specific type of coordinate, the bounding box is represented by four values [xCenter, yCenter, width, height]. All values are normalized with max width and max height.

The classes of objects in the text file are annotated by integers, in the case of the project 6 different types of classes were defined, ranging from 0 to 5, where each class is annotated respectively to "without helmet", "with helmet", "without glove", "with glove", "without glasses" and "with glasses." See Figure 3 for an example of an annotation pattern that the YOLOv4 network needs to learn correctly.

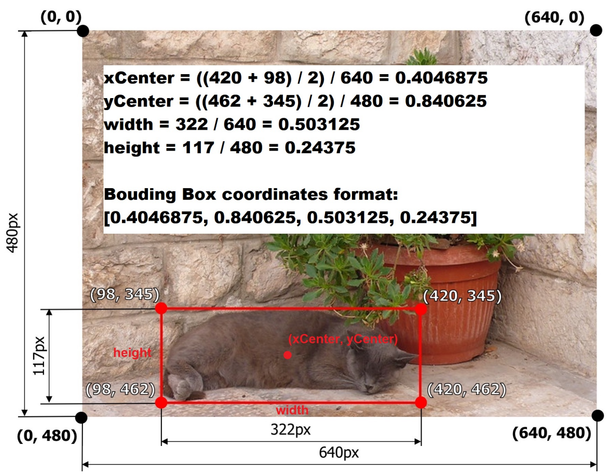

Figure 3. An example image with a bounding box from the COCO dataset [18]

To annotate each object in 7555 images referring to the YOLOv4 network standard coordinate format, it was necessary to use the program LabelImg, the program allows you to manually select the object that you want to be identified as well as its location by bounding box, after demarcation, it automatically generates the text file corresponding to the image with the correct annotation. Look at figure 4.

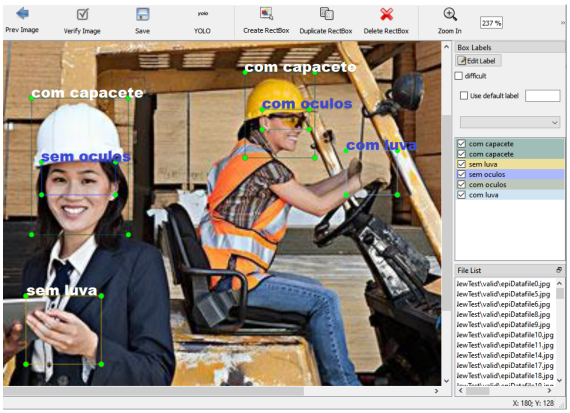

Figure 4. Image annotation using the labelImg program

After the annotation of the entire database, a python algorithm was developed to quantify the classes in each image, in order to assist in the step of segregating the database into a training set and a validation set. YOLO requires this segregation for correct operation. This set was divided into 70% for training and 30% for validation. This proportion was not applied to the number of images, due to the fact that there are several object quantities of the same class and of different classes in a single image, so probabilistic, the training and validation set would be unbalance in relation to the classes, leading to inefficient learning, to avoid this problem, the algorithm has filtered the image classes so that all classes in their training and validation sets are distributed in a 70/30 ratio. Figure 5 shows the total amount of annotations of each class distributed by the training/validation set.

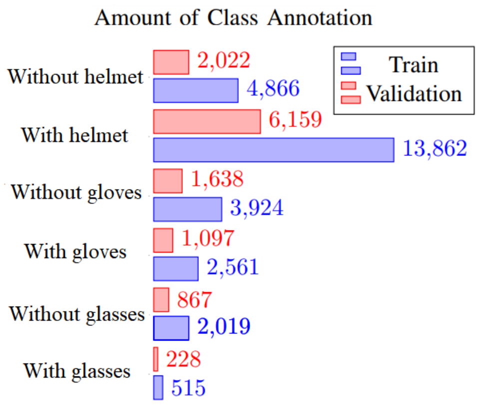

Figure 5. Amount of classes annotation

For the YOLOv4 configuration, training and validation were used the online platform Google Collaboratory. This choice was made due to the ease of installation and configuration of Darknet, an open source neural network framework written in C and CUDA, where YOLOv4 is based on, which can be quickly cloned from GitHub servers. The Google Collaboratory also provides limited access to Google servers GPUs, which during network training considerably reduces processing time. For the training, a technique called transfer of learning was used several times, where at first a public convulocional base was used already trained with a dataset called COCO, a database composed of 80 classes of everyday objects. This technique reduces the processing time of a new network since the base weights have already been pre-defined in order to identify images of people and objects in general. The following steps were taken to improve object detection tasks.

- Checking whether all the images were correctly labelled and if not label them.
- Change the input resolution of the network to verify if got a precision increase.
- Using different orientations, brightness, contrast, color changes, mirroring and mosaic of the input image as an augmentation technique.

Mean Average Precision metrics – mAP with IoU threshold = 50% and precision were used to evaluate the trained models after the various changes and modifications made.

After defining the best model, an algorithm was developed in Java Script that allowed the use of a webcam with the Google Collaboratory, and a python algorithm that processed the model of the YOLOv4 network previously trained and executed each frame captured by the webcam. Subsequently, the image containing the bounding box generated by the CNN was superimposed on the image of the next frame of the video stream. Note that in this step, the images collected by the webcam were processed in real time.

4 RESULTS

As mentioned before were made a several types of tests spent many training hours, changing different hyper-parameters and augmentation techniques as changing image input size, adding and removing HUE augmentation, with an objective to find the more accurate model.

Besides all these modifications were got much close results in each test, as can been seeing at the graphs of Mean average precision x average loss in the figure 6.

As can be seeing at figure 6, in spite of paramete's changes in the training phase the mean average precision do not vary too much and stabilized at around 70%, same happen with the average loss which decrease from the beginning to the end and stabilized on 2.8.

Even with the very close results were possible to decide the best of them by calculating the average of all mAP at each model, and the best was the model with input resolution of 416x416 without HUE augmentation.

The table 1 show how the this model classified each class in the course of interactions evaluated by the precision metric.

Table 1

Anotação	1000	4000	8000	12000
without helmet	91.65%	95.55%	95.24%	94.86%
with helmet	95.58%	97.79%	97.66%	96.91%
without gloves	54.17%	66.82%	65.05%	63.03%
with gloves	49.66%	62.76%	59.75%	57.97%
without glasses	19.60%	43.64%	60.62%	59.20%
with glasses	19.02%	50.48%	63.48%	68.52%
mAP	54.95%	69.51%	73.63%	73.41%

Figures 7 show an example of how the best neural network model managed to correctly classify and segment people who were not wearing PPE, with annotation of ”without helmet”, ”without glasses” and “without gloves” and people who were wearing, with annotation of “with helmet”, “with glasses” and ”with glove”.

Figure 8 shows the correct processing of CNN in real-time, at a rate of 25fps, YOLOv4 was able to correctly identify and segment images captured by the webcam and accurately. Also note by figure 8 that CNN correctly learned the identify the correct placement of PPE, for example the network does not segment helmets as “with helmet”, equipment that is hung or held away from the head. Also in real time could identify with high accuracy the classes related with glasses and gloves.

5 CONCLUSIONS

It was concluded that the neural network produced responses as expected, managing to correctly classify and segment the conformity or non-conformity of the use of PPE due the high mean average precision of 73.63% that the network achieved.

In this work were realised that the YOLO a its really power-full network, as can be analysed by the table 1, the precision of the classes annotation ”without helmet” and ”with helmet” was plus 90% since the beginning of training, meaning that the network can learn fast if got an huge data, in contrast the classes ”without glasses” and ”with glasses” got so much less amount of data than helmet annotation but the network could learn how to identify this PPE over the batches even with small data, bordering 60% of precision at the top mAP.

In all tests performed, it was observed that after 4000 iterations the mAP value reached 70% and stabilized, sometimes exceeding this range slightly, on the other hand the average loss value continued to decrease. Such behavior in a neural network shows a clear case of over-fitting, where the the network would no longer be learning from the data, only recording. In the case of our problem the only way to solve this, is increase data of the gloves and glasses annotation whereas all the possible changes in hyper-parameters were already made.

As shown in the previous section, due to its high processing speed, it is entirely feasible to use the YOLOv4-based neural network for real-time PPE detection, linking the network to a camera. For future steps, will be create a system that the trained YOLOv4 neural will analyse videos from a surveillance CCTV installed at workplace area that got high-traffic of people going and coming from work sites, which, when identifying the non-compliance of the employee regarding the use of PPE, the system will generate an alert for the HSE team to take the appropriate

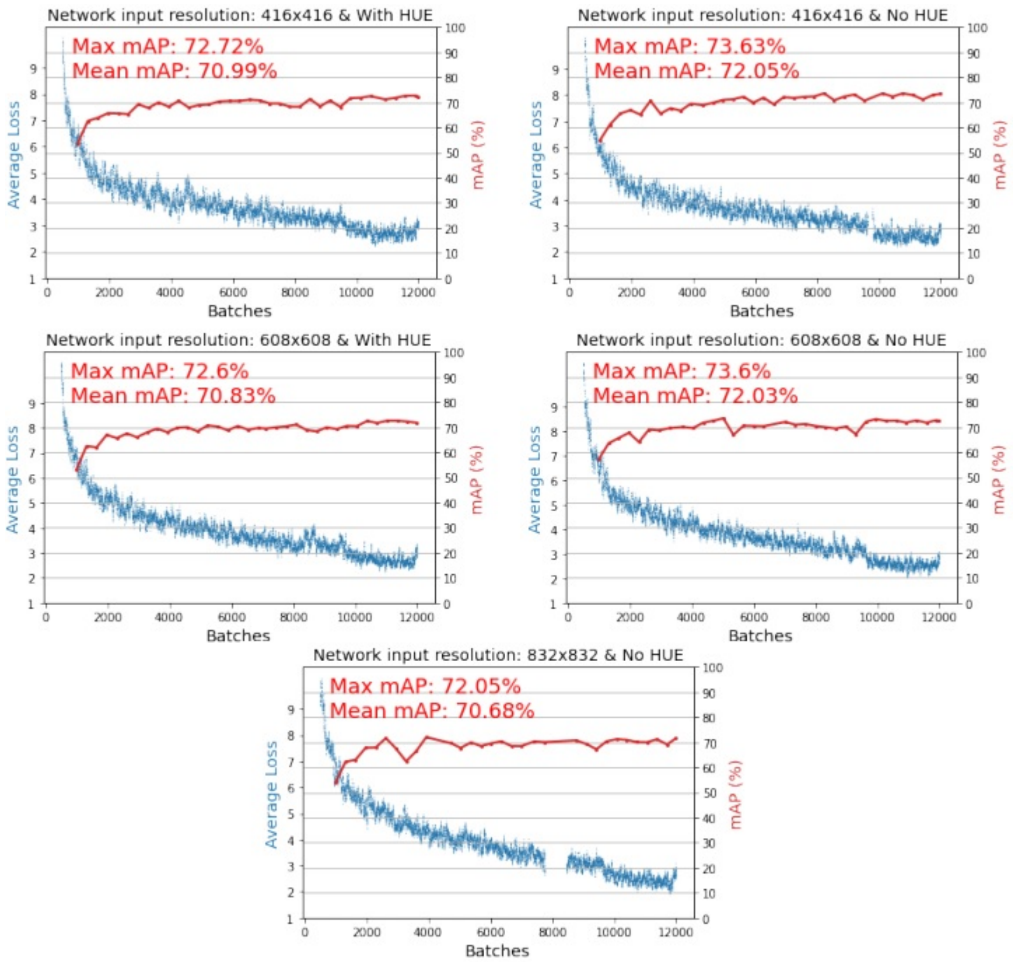

Figure 6. mAP x Average loss graph results

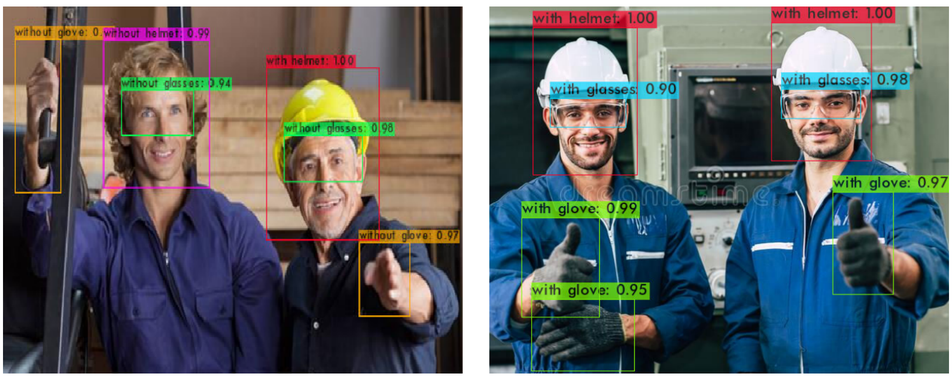

Figure 7. YOLO correct classification of Workers with and without wearing full PPE

measures, in addition will also generate logs of the number of people who are compliant or non-compliant with safety, thus generating statistics that can be used as reliability indicators that will assist in decision making and in the safety management within the company.

More images will also be annotated in addition to the collection of more annotations of workers wearing protective glasses, in order to improve the accuracy of the identification of this PPE that is so important for the integrity of the worker.

To the CAPES/FAPES - PDPG Cooperation, TIC+TAC project, for the financial support (TO 133/2021, Process Nº 2021-CFT5C).

The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

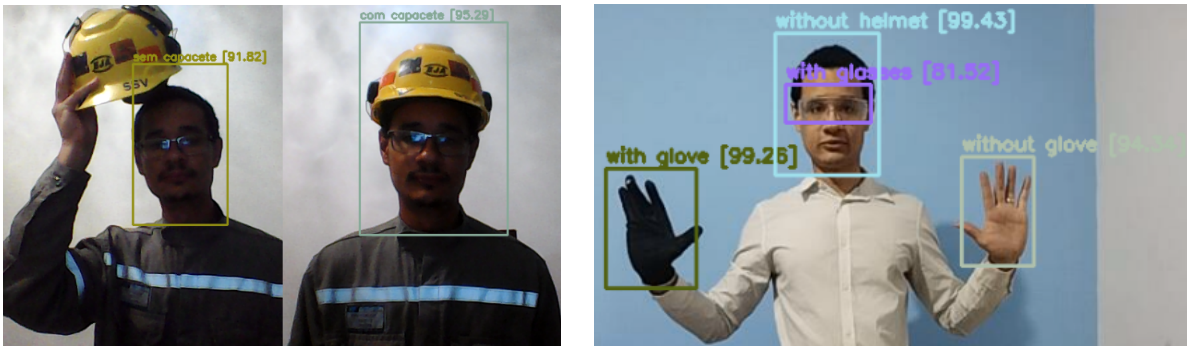

Figure 8. Identification of safe behavior by video

References

[1] VIANA, M. R. Estatística De Acidentes De Trabalho Em Uma Empresa De Coleta De Resíduos Domiciliares E Industriais – Estudo De Caso. Universidade Tecnológica Federal Do Paraná. Curitiba. p. 65, 2014.

[2] REVISTA CIPA, Cipa e Incêndio. Importância da fiscalização do uso de EPIs e EPCs, 11 Jan 2019. Disponível em: shorturl.at/akFG7. Acesso em: 17 jan. 2021

[3] BRASIL, Ministério do Trabalho. NR 06 - EQUIPAMENTO DE PROTEÇÃO INDIVIDUAL - EPI, Brasília: Ministério do Trabalho, 2018.

[4] FIORAVANTE, E.; BARONI, F. Causas da Resistência ao Uso do Equipamento de Proteção Individual (EPI). p. 9, 2012.

[5] LI, Y. et al. Deep Learning-Based Safety Helmet Detection in Engineering Management Based on Convolutional Neural Networks. Advances in Civil Engineering, v. 2020, p. 1–10, 19 set. 2020.

[6] L. M. Gonzaga e G. M. de Almeida, Artificial Intelligence usage for identifying automotive products, p. 7, 2020

[7] TRENTIN, M. L. A Importância Do EPI E Do Certificado De Aprovação (CA) Para O Servidor Público Estatutário. Universidade do Vale do Rio dos Sinos. São Leopoldo. p. 16, 2016.

[8] OCUPACIONAL, Medicina e engenharia de segurança do trabalho. Consequências do descumprimento das Normas Regulamentadoras, 19 nov. 2018. Disponível em:https://www.ocupacional.com.br/ocupacional/consequencias-do-descumprimento-dasnormas-regulamentadoras/. Acesso em: 05 out 2019

[9] NASCIMENTO, I. G. et al. Segurança No Trabalho: Motivos Que Levam O Trabalhador Da Construção Civil A Deixar De Utilizar Os EPIs. XI Congresso Nacional De Excelência Em Gestão. P. 21, 2015

[10] VARGAS, A. C. G. CARVALHO, A. M. P. VASCONCELOS, C. N. Um estudo sobre Redes Neurais Convolucionais e sua aplicação em detecção de pedestres. Universidade Federal Fluminense. Niteroi . p. 4, [s.d.].

[11] NATH, N. D.; BEHZADAN, A. H.; PAAL, S. G. Deep learning for site safety: Real-time detection of personal protective equipment. Automation in Construction, v. 112, p. 103085, abr. 2020.

[12] FANG, W. et al. A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network. Advanced Engineering Informatics, v. 39, p. 170–177, jan. 2019.

[13] REDMON, J. et al. You Only Look Once: Unified, Real-Time Object DetectionarXiv, , 9 maio 2016. Disponível em: http://arxiv.org/abs/1506.02640. Acesso em: 3 jul. 2022

[14] PADILLA, R.; NETTO, S. L.; DA SILVA, E. A. B. A Survey on Performance Metrics for Object-Detection Algorithms. 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). Anais... In: 2020 INTERNATIONAL CONFERENCE ON SYSTEMS, SIGNALS AND IMAGE PROCESSING (IWSSIP). Niterói, Brazil: IEEE, jul. 2020. Disponível em: https://ieeexplore.ieee.org/document/9145130/. Acesso em: 3 jul. 2022

[15] BOCHKOVSKIY, A.; WANG, C.-Y.; LIAO, H.-Y. M. YOLOv4: Optimal Speed and Accuracy of Object Detection,arXiv,22 abr. 2020. Disponível em: http://arxiv.org/abs/2004.10934. Acesso em: 3 jul. 2022

[16] CHEN, S.; DEMACHI, K. A Vision-Based Approach for Ensuring Proper Use of Personal Protective Equipment (PPE) in Decommissioning of Fukushima Daiichi Nuclear Power Station. Applied Sciences, v. 10, n. 15, p. 5129, 26 jul. 2020.

[17] Roboflow, Public Dataset - Hard Hat Workers Disponível em:https://public.roboflow.com/object-detection/hard-hat-workers/2/images/90d4f039570e8ddf467cd6ab24aa4883. Acesso em: 25 mai 2021

[18] Bounding boxes augmentation for object detection. Disponível em: shorturl.at/lvP26. Acesso em: 25 jul 2022