2023-06

Medical Personal Protective Equipment detection using YOLOv7

Ali, Md. Anas

Independent University, Bangladesh

Medical Personal Protective Equipment detection using YOLOv7

Md. Anas Ali[1], Mahmudul Haque[2], Saadia Binte Alam[3], Rashedur Rahman[4], M Ashraful Amin[5], Syoji Kobashi[6]

[1,4,6]Graduate School of Engineering University of Hyogo, 2167 Shosha, Himeji, Hyogo, Japan.
[2]Department of Computer Science and Engineering, Independent University, Bangladesh, Dhaka 1299, Bangladesh.
[3,5]Center for Computational & Data Sciences, Department of Computer Science and Engineering, Independent University, Bangladesh, Dhaka 1299, Bangladesh.

E-MAIL: [1]anasalilimon@gmail.com, [2]mahmud.eece@gmail.com, [3]saadiabinte@iub.edu.bd
[4]rashed.riyadh14@gmail.com, [5]aminmdashraful@iub.edu.bd, [6]kobashi@eng.u-hyogo.ac.jp

Abstract:

In recent years, healthcare and safety have been a major focus of deep learning research. This paper focuses on the detection of Medical Personal Protective Equipment (MPPE) in the healthcare sector using YOLOv7. Improper use of personal protective equipment (PPE) can result in the contamination and cross-contamination of infectious diseases, so it is crucial for healthcare professionals to use it correctly. The CPPE-5 dataset was used to train the model, which contains 1029 high-quality images divided into five categories: *coveralls*, *face shield*, *gloves*, *masks*, and *goggles*. The objective of this research is to create an accurate model for future applications and development using a suitable medical PPE dataset. The proposed model outperforms previous studies, with an optimal mAP of 90.93%, indicating that it is a promising method for detecting MPPE in the healthcare sector.

Keywords:

MPPE; YOLOv7; object detection; deep-learning; CPPE-5

1 Introduction

The healthcare sector is a critical component of public welfare, with healthcare professionals working tirelessly to ensure patient safety. Despite their efforts, infectious diseases can still spread from person to person via a range of means, including physical contact, air, body fluids, and other carriers, putting healthcare professionals at risk of contamination. Hospital-acquired infections affect one out of every ten patients, contributing to longer hospital stays, morbidity, and antibiotic resistance [1, 2, 3, 4, 5]. The improper use of Medical Personal Protective Equipment (MPPE) is one of the primary sources of disease transmission in healthcare settings.

MPPEs are crucial for protecting healthcare professionals from droplets from infected patient discharges, as well as contaminated surfaces [6]. Incorrect use of such equipment can result in cross-contamination and infection of healthcare professionals or patients. As a result, it is critical to constantly monitor whether these healthcare professionals are using MPPE correctly. On the other hand, manual supervision of their correct use is a time-consuming and exhausting process, and not every healthcare professional can be constantly monitored. Hence, to address the challenges associated with PPE detection and enforcement, artificial intelligence (AI) based solutions need to be considered.

In a work on real-time detection, Khosravipour et al. [7] implemented two popular algorithms, YOLOv3 and SSD MobileNet, to the multi-class detection problem of whether people were wearing PPEs like masks and gloves in public areas during COVID-19, and evaluated their performance using an 8250-image dataset. In terms of mAP, the results show that YOLO outperforms SSD MobileNet. Similarly, Nath et al. [8] presented three deep-learning models based on the YOLO architecture for determining worker PPE compliance in a separate work on construction site safety. They used the Pictor-v3 dataset, which included 1,500 annotated images and 4,700 instances of workers dressed in various PPE components. The first approach was the fastest, with 13 FPS and 63.1% mAP.

On a different domain, Kumar et al. [9] also proposed a YOLO-based approach for fire and PPE detection using the YOLOv4 and YOLOv4-tiny algorithms. A dataset of 14,500 samples was used to train the model, and the results showed a maximum mAP of 76.86% for real-time detection and surveillance.

For the medical domain, Dagli et al. [10] discovered no publicly available datasets for MPPE detection and proposed the CPPE-5 dataset, which they compared to other popular datasets with a wide range of PPE-related categories in their research. Furthermore, Wu et al. [11] proposed an improved one-stage detector-based model called ME-YOLO which met real-time detection requirements by ensuring a balance between performance (97.2% mAP) and efficiency (53 FPS). The authors created a detector that detects whether or not people are wearing PPE by using the YOLOv4. Again, Protik et al. [12] developed an MPPE detector using a combined dataset, annotated images, and image augmentation techniques, achieving an mAP of 79%. The authors suggest that their method, which uses the YOLOv4 model in conjunction with a merged dataset and other techniques, can be employed for PPE detection to prevent the spread of COVID-19 and other similar diseases.

Furthermore, Kwak et al. [13] a YOLOv5 based method for determining whether workers were wearing safety helmets, with an mAP of 0.959. Besides, Bhing et al. [14] used the YOLO object detection algorithm for real-time PPE detection, achieving 84.5% accuracy on their PPE dataset with seven classes in over 1,300 images.

Recently, Wang et al. [15] presented the YOLOv7 object detector that outperformed all known state-of-the-art object detectors on the MS COCO dataset in terms of speed and accuracy, with 56.8% AP ranging between 5 FPS to 160 FPS. They proposed a new approach for real-time object detection called "trainable bag-of-freebies" which improves object detection accuracy. Using this method, YOLOv7 outperforms other well-known object detectors like YOLOR, YOLOX, YOLOv5, and many others. With the development of the YOLOv7 series of object detection systems, they have made a significant impact in the field and achieved state-of-the-art results in real-time object detection.

The objective of this research is to determine the suitability of using AI-based solutions for MPPE detection in real-world healthcare settings. The study will concentrate on the difficulties encountered in MPPE detection, as well as how YOLOv7 can be used to overcome these difficulties. Hence, the significance of this research cannot be overstated, as the healthcare sector faces significant challenges in the detection and enforcement of MPPE. The use of YOLOv7 for MPPE detection will ensure the safety of healthcare professionals and their patients.

2 Dataset

The CPPE-5 dataset was used in this experiment. Unlike most other popular datasets, which focus on broad-level categories, this dataset was constructed with the goal of facilitating MPPE detection and classification.

2.1 Description

The CPPE-5 dataset comprises of 1029 images, of which, 1000 were used for training and the remaining 29 were used for testing. The dataset was categorised into 5 classes: *coveralls*, *face shields*, *gloves*, *masks*, and *goggles*. The images were annotated using a set of bounding boxes and labels. These images were acquired from a range of sources to ensure high quality and diversity. The sources of the images include open-source datasets, publicly available images, and images captured from real-life scenarios.

2.2 Pre-processing of the Images

Before training the model, each image of the dataset was resized to 640×640 pixels for efficient processing and consistent feature learning. Following that, augmentation parameters were set to add the required diversity to the dataset.

The hue, saturation, and image values were all shifted by 1.5%, 70%, and 40%, respectively. Furthermore, the images in the dataset were rotated by up to 20% and scaled by up to 50%. There was also a 50% chance that the image would be flipped left to right for further augmentation. Finally, mosaic and loss OTA were used on the dataset during training to increase the effectiveness of the model.

3 MPPE Detector

In this study, we use the CPPE-5 dataset to detect and classify different MPPEs. The methodology outlined below outlines the steps involved in conducting object detection experiments on the CPPE-5 dataset using the YOLOv7 architecture.

3.1 YOLOv7

The YOLO model uses a three-step process in which image frames are featured through a backbone, combined and mixed in the neck, and finally passed into the network's head to predict the positions and classes of the objects around which bounding boxes should be shown. YOLOv7 seeks to raise the standard in object detection by predicting bounding boxes more effectively than its contemporaries at comparable inference speeds.

Efficient inference speed is critical to the YOLO network's performance. Wang et al. [15] sought to improve the efficiency of the backbone's convolutional layers by building on the Efficient Layer Aggregation Networks (ELAN). They chose an extended ELAN computational block that reduces gradient distance and improves learning.

In YOLOv7, Wang et al. [15] tackled the challenge of scaling the network depth and width while maintaining an optimal model architecture for different sizes. They accomplished this by concatenating layers and scaling the model architecture simultaneously, resulting in a versatile model that can be adjusted to meet the needs of various applications. To accommodate varying accuracy and inference speed requirements, object detection models are typically released in a series of scaled models. Gradient flow propagation channels were later employed to decide which network modules should and should not employ re-parameterization techniques.

3.2 Model Training

The YOLOv7 model was trained with a batch size of 12, an early stopping criterion, and a learning rate of 0.001. If the training loss did not improve after 30 epochs, the training was terminated. The model was pre-trained with the COCO dataset [16] to improve performance. Table. 1 shows all of the hyperparameters used to train the model, and was then trained in a system with the configuration shown in Table. 2.

3.3 Model Evaluation

The trained YOLOv7 model was tested on the 29 images in the test dataset to validate its ability to detect the five object categories in the images in terms of different evaluation metrics. The evaluation of the object detection model's performance is an important aspect of the experiment. In this study, numerous metrics, including mAP, recall, and F1-score, were employed to assess the performance of the model.

The Precision x Recall curve is a widely used metric to evaluate object detectors. It shows how precision and recall vary as the confidence threshold changes. Precision represents the percentage of detected objects that are actually positive, while recall represents the percentage of positive objects that are detected. A good object detector for a specific class should have high precision and recall for various confidence thresholds.

TABLE 1. Hyperparameter for training the model.

Hyper-parameter	**Value**	**Hyper-parameter**	**Value**
lr0	0.001	cls	0.3
lrf	0.01	cls_pw	1.0
momentum	0.937	obj	0.75
weight decay	0.0005	obj_pw	1.0
warmup epochs	3.0	iou_t	0.2
warmup momentum	0.8	anchor_t	4.0
warmup bias lr	0.01	paste_in	0.0
box	0.05	loss_ota	1

TABLE 2. YOLOv7 training system configuration.

Device Name	**Configuration**
Operating System	Windows 10 (64-bit)
RAM	32 GB
CPU	AMD Ryzen 7 2700 8 Core 3.20 GHz
GPU	NVIDIA GeForce RTX 2080 Ti, 11 GB Video Memory
Deep Learning Framework	Python = 3.9.13
	CUDA = 11.3.1
	PyTorch = 1.12.1

F1 score is another widely used metric that combines precision and recall. It is the harmonic mean of precision and recall. It is calculated using equation 1.

$$F_1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (1)$$

F1 score ranges from 0 to 1, where 1 represents perfect precision and recall, and 0 represents poor performance. A high F1 score indicates a good balance between precision and recall. Moreover, mAP is a commonly used metric for evaluating object detectors that consider both precision and recall. It is calculated as the average of AP across all object classes. AP is calculated as the area under the Precision x Recall curve.

The evaluation metrics were used to objectively evaluate the performance of the YOLOv7 model on the CPPE-5 dataset. These metrics provide a comprehensive evaluation of the model's ability to accurately detect the different categories of MPPE in the images.

4 Experimental Result

The best weights were obtained at the 34th epoch. The performance of the object detection model was evaluated on the test dataset based on a Precision x Recall curve and mAP.

The Precision x Recall curve for each class of object is plotted in Figure. 1. We have used the mAP as a summary of the AP of each class to evaluate the overall performance of the object detection model. The model has an mAP of 90.93% on the test dataset. We also calculate the AP for each class of object detected by the model. The numerical results for the F1 score,

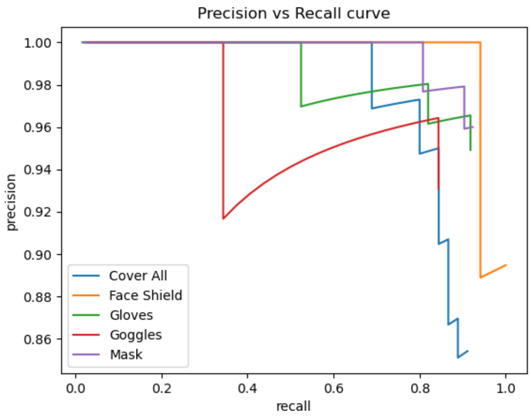

FIGURE 1. Precision x Recall curves for each of the five MPPE classes.

TABLE 3. Performance of the trained YOLOv7 model on the test dataset.

Metrics	**Mask**	**Face shield**	**Gloves**	**Goggles**	**Cover-alls**
Instances	52	17	61	32	45
T	48	17	56	27	41
FP	2	2	3	2	7
Precision	0.96	0.89	0.95	0.93	0.85
Recall	0.92	1	0.92	0.84	0.91
F1	0.94	0.94	0.93	0.88	0.88
Average precision	0.92	0.99	0.91	0.83	0.9
Mean AP			0.9093		

Precision and AP for each class are presented in Table. 3 and the performance metrics have been shown in Figure. 2.

This curve is constructed by taking into account all kinds of objects detected by the model. We can see that the model has high precision at low recall values, which indicate that the model is capable of detecting objects accurately even at low confidence thresholds. The precision decreases as recall increases, which is expected, as it becomes harder to detect objects accurately as the number of objects in the image increases.

Overall, the experimental results indicate that the object detection model is effective in detecting various objects as shown in Figure. 3, with high accuracy and efficiency. However, it's worth noting that object detection models are not perfect and may sometimes make mistakes. For example, in cases where objects are partially occluded or have similar appearances, the model may have difficulty distinguishing between them. Additionally, the performance of object detection models can be affected by factors such as lighting conditions, image quality, and object scale.

TABLE 4. Comparison of the model with other contemporary approaches.

Study	**Algorithm**	**Dataset**	**Precision**	**Recall**	**mAP**
Kumar et al. [9]	YOLOv4	Self-made	-	-	76.86%
Nath et al. [8]	YOLO	Pictor-v3	-	-	63.1%
Protik et al. [12]	YOLOv4	Hybrid	-	-	79%
Bhing et al. [14]	YOLO	Medical PPE	84.5%	-	-
Wang et al. [15]	YOLOv7	MS COCO	57.2%	58.5%	55.9%
This work	**YOLOv7**	**CPPE-5**	**92.2%**	**91.3%**	**90.93%**

5 Result and Discussion

PPEs are a crucial part of ensuring the safety of healthcare professionals. Several previous studies have utilised deep learning algorithms to detect PPE compliance. As shown in Table. 4, Kumar et al. [9] utilised the YOLOv4 algorithm to detect PPE in real-time and achieved an mAP of 76.86% at construction sites. Nath et al. [8] developed three deep learning models for determining worker PPE compliance using the YOLO architecture, with their fastest approach having 63.1% mAP. Bhing et al. [14] utilised the YOLO object detection algorithm to detect the presence of MPPE. The model achieved a precision of 84.5% on their PPE dataset. Finally, Protik et al. [12] developed the YOLOv4 object detector, which outperforms all other known object detectors in terms of performance.

In this study, we proposed a novel approach towards MPPE detection utilising the YOLOv7 algorithm, which has been demonstrated to be the fastest and most accurate object detector currently available. Our proposed model achieved a maximum mAP of 90.93%, significantly higher than the results reported in previous studies. Our approach outperforms the approach by Protik et al. [12] with a 11.93%, Nath et al. [8] with a 13.17% and Kumar et al. [9] with a 27.83% improvement in mAP respectively, making it suitable for real-time detection in healthcare settings.

The results of this study demonstrate the potential of the YOLOv7 algorithm in automating MPPE detection. The proposed model achieved state-of-the-art performance in terms of

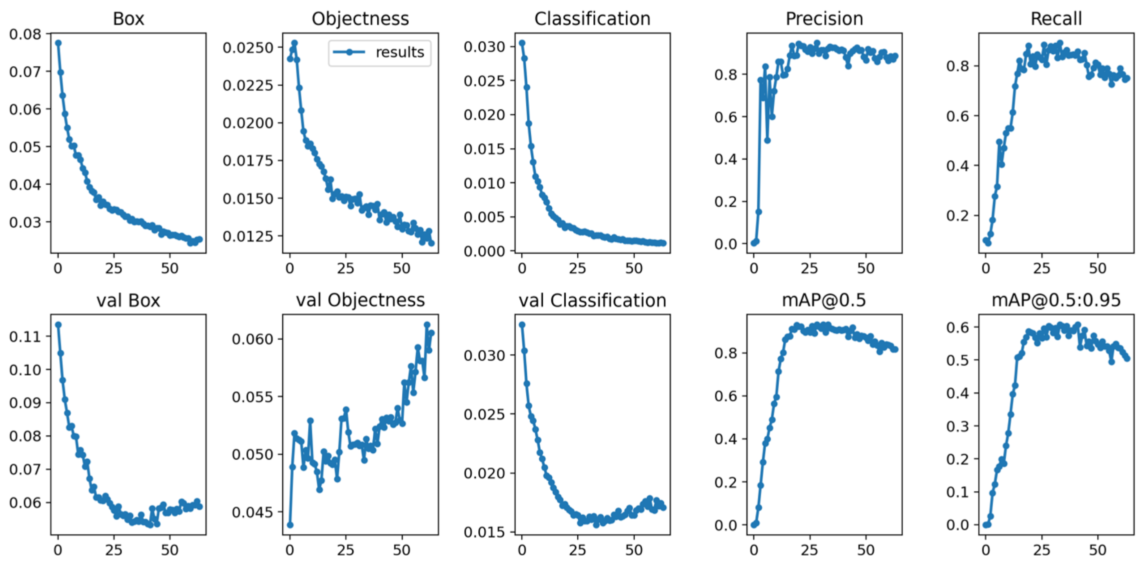

FIGURE 2. Performance metrics for the YOLOv7 model.

accuracy. The improved performance can be attributed to the fact that the YOLOv7 algorithm has been specifically designed for accurate object detection, with better feature extraction capabilities and a more efficient architecture compared to previous versions of the YOLO algorithm.

FIGURE 3. Some instances of MPPE detection (Green box is the ground truth and red box for detection).

Our approach achieves the most optimum mAP of 90.93%, outperforming all the previous works on MPPE detection discussed above. The higher mAP values of our approach indicate that our proposed approach is more effective in identifying and localizing MPPE objects in images.

6 Conclusion

In conclusion, our study evaluated various models and introduced a novel approach to MPPE detection based on the YOLOv7 algorithm, which currently stands as the fastest and most accurate object detector. Our proposed model achieved an optimal mAP of 90.93% on the CPPE-5 dataset with five classes and over 1029 images, surpassing the results of previous studies. The significance of our findings lies in the effective and accurate detection of MPPE, which is critical for the health and safety of healthcare professionals and the general public. Our study has significant implications for future research, as our proposed model can be extended to other medical image analysis tasks requiring speed and precision, such as patient diagnosis and medical equipment detection. Additionally, future research can focus on developing larger datasets with more classes to expand the scope of our model to detect other medical objects, such as instruments and devices. We recommend further research into integrating our proposed model into existing healthcare systems to enhance the safety of healthcare professionals and patients. Improving the model's interpretability and addressing complex issues will aid in building credibility and public confidence in the model. Our research contributes to ongoing efforts to improve the safety of healthcare professionals and the general public by providing an effective and accurate model for MPPE detection.

References

[1] P. L. Russo, A. J. Stewardson, A. C. Cheng, T. Bucknall, and B. G. Mitchell, "The prevalence of healthcare associated infections among adult inpatients at nineteen large australian acute-care public hospitals: a point prevalence survey," *Antimicrobial Resistance & Infection Control*, vol. 8, pp. 1–8, 2019.

[2] A. van Buijtene and D. Foster, "Does a hospital culture influence adherence to infection prevention and control and rates of healthcare associated infection? a literature review," *Journal of infection prevention*, vol. 20, no. 1, pp. 5–17, 2019.

[3] B. Allegranzi, C. Kilpatrick, J. Storr, E. Kelley, B. J. Park, and L. Donaldson, "Global infection prevention and control priorities 2018–22: a call for action," *The Lancet Global Health*, vol. 5, no. 12, pp. e1178–e1180, 2017.

[4] T. Karopka, G. Mernitz, M. L. Mogensen, S. Reich, R. Lövström, H. Dalianis, H. Leslie, and M. Behnke, "Minimum data set and standards in the context of nosocomial infection surveillance and antimicrobial stewardship," *InHEC (Health-Exploring Complexity)*, 2016.

[5] R. Barratt, M. Wyer, S.-y. Hor, and G. L. Gilbert, "Medical interns' reflections on their training in use of personal protective equipment," *BMC medical education*, vol. 20, pp. 1–9, 2020.

[6] J. H. Verbeek, B. Rajamaki, S. Ijaz, R. Sauni, E. Toomey, B. Blackwood, C. Tikka, J. H. Ruotsalainen, and F. S. K. Balci, "Personal protective equipment for preventing highly infectious diseases due to exposure to contaminated body fluids in healthcare staff," *Cochrane database of systematic reviews*, no. 4, 2020.

[7] S. Khosravipour, E. Taghvaei, and N. M. Charkari, "Covid-19 personal protective equipment detection using real-time deep learning methods," *arXiv preprint arXiv:2103.14878*, 2021.

[8] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Automation in Construction*, vol. 112, p. 103085, 2020.

[9] S. Kumar, H. Gupta, D. Yadav, I. A. Ansari, and O. P. Verma, "Yolov4 algorithm for the real-time detection of fire and personal protective equipments at construction sites," *Multimedia Tools and Applications*, vol. 81, no. 16, pp. 22 163–22 183, 2022.

[10] R. Dagli and A. M. Shaikh, "Cppe-5: Medical personal protective equipment dataset," *arXiv preprint arXiv:2112.09569*, 2021.

[11] B. Wu, C. Pang, X. Zeng, and X. Hu, "Me-yolo: Improved yolov5 for detecting medical personal protective equipment," *Applied Sciences*, vol. 12, no. 23, p. 11978, 2022.

[12] A. A. Protik, A. H. Rafi, and S. Siddique, "Real-time personal protective equipment (ppe) detection using yolov4 and tensorflow," in *2021 IEEE Region 10 Symposium (TENSYMP)*. IEEE, 2021, pp. 1–6.

[13] N. Kwak and D. Kim, "A study on detecting the safety helmet wearing using yolov5-s model and transfer learning," *International Journal of Advanced Culture Technology*, vol. 10, no. 1, pp. 302–309, 2022.

[14] N. W. Bhing and P. Sebastian, "Personal protective equipment detection with live camera," in *2021 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. IEEE, 2021, pp. 221–226.

[15] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv preprint arXiv:2207.02696*, 2022.

[16] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 2014, pp. 740–755.