Fraud Face Detection at ATM using YOLOv5

Modalavalasa Divya[1*],K.B.Anusha[2],Ch.KrishnaVeni[3], T.Sai Sriya[4],K.Jagadeesh Kumar[5], S.Siresha[6], R.Bala Vinoth[7]

[1,2,3,4,5,6,7]Department of Computer Science & Engineering, Aditya Institute of Technology and Management, Tekkali-532201, India.

*divya.modalavalasa@gmail.com

Abstract. The increasing of fraudulent activities including wearing helmets and masks in the ATM premises. This System is designed to enhance security and efficiency in automated teller machine operations. This project focuses on real-time surveillance, wearing masks, helmets, anomaly detection, multiple face detection mitigate risks associated with fraudulent activities.It employs advanced technologies like video analytics and machine learning to identify suspicious activities, ensuring a secure and seamless banking experience for users. This System includes features such as remote monitoring through a centralized dashboard, alert notifications for unusual transactions or security breaches and a comprehensive reporting module for analyzing ATM performance and user behavior. By integrating cutting-edge technologies, this project aims to provide a robust solution for ATM management and security in the evolving landscape of banking services.

Keywords: Artificial Intelligence, Prompt Engineering, Natural Language Processing, MERN Stack.

1.Introduction

The ATM examining System is a inclusive solution originate to build up the coherence and compact of automated teller machines(ATMs) .This system enlists advanced technologies to monitor and manage ATMs in real time, ensuring optimal functionality, timely maintenance, and robust security measures. By integrating devices, connectivity, and data analytics, the ATM Monitoring System aims to minimize downtime, prevent fraudulent activities, and provide a seamless banking experience for both financial institution and their customers. One of the emerging challenges is the occurrence of fraud by individuals wearing masks and helmets, making it difficult to identify and track suspicious activities.

In the dynamic landscape of banking and finance, the Automated Teller Machine (ATM) stands as a crucial channel for delivering seamless and convenient services to customers. However, the increasing sophistication of security threats, coupled with the imperative for operational efficiency, calls for a robust solution – the ATM Monitoring System. The System is designed to address two main security scrutinize first one is presence of individuals wearing masks and helmets, and the event of more than one person inside the ATM. though real time video analysis from surveillance cameras, our solution engage computer vision to detect helmet and mask-wearing individuals, on the nail of generating alerts to notify pertinent authorities.

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_44

1.1 Problem Statement

Our ATM Monitoring System's main goal is to spot possible security risks within ATMs by looking for people wearing masks and helmets. This is important because fraudsters frequently use these kinds of disguises to hide their identity when engaging in illegal activity. Furthermore, the technology is designed to sound a warning in the event that multiple people are within the ATM, as this could be a sign of an attempted breach of security or unapproved entry.

1.2 Existing System

In the realm of Artificial Intelligence (AI), the ability to understand and respond accurately to prompts has emerged as a pivotal challenge in recent times. This difficulty often leads to inaccurate results, hindering AI's overall performance and applicability across various domains. Addressing this issue requires innovative solutions that empower AI systems to comprehend prompts effectively, thereby enhancing their utility in real-world applications

1.3 Proposed System

The contemplate of the proffer ATM surveillance network is to use cutting-edge computer vision techniques more particularly, the YOLOv5 object identification algorithm to solve the shortcomings of the current systems. The following are the suggested system's salient features:

Helmet and Mask Detection: In real-time, the system employs YOLOv5 to reliably identify people donning masks and helmets. The model's capacity to recognize minute differences in clothing is guaranteed by specialized training on a wide range of datasets.

Multi-Person Detection: If more than one person is inside the ATM, the system uses sophisticated computer vision to identify and sound an alert. By detecting possible unauthorized access or questionable activity involving numerous people, this feature improves security.

Real-time Monitoring: The YOLOv5 model integration enables the system to offer real-time monitoring features.

Alert Generation: The system sounds an alert if it sees people wearing masks and helmets or if it recognizes more than one person within the ATM. Alerts are intended to quickly inform pertinent authorities and interested parties.

The suggested technology offers a comprehensive approach to prevent fraudulent activity, which should greatly improve ATM security. Proactive security is aided by real-time monitoring and alarm production, and excellent object detection accuracy is ensured by the integration of YOLOv5.

2 Literature Survey

You Only Look Once: Unified, Real-Time Object Detection:
[1]offers an overview of face mask detection methods and approaches, especially pertinent in light of the COVID-19 epidemic. Accurate detection methods are becoming more and more necessary as face masks are being worn widely as a prophylactic strategy. In the end, this review is a useful tool for scholars, professionals, and decision-makers who are trying to stop COVID-19 from spreading by using new technologies.

Face Mask Detection in the Era of COVID-19:

[3] presents the YOLO framework, actual device recognition system with excellent accuracy and efficiency that scans photos in a single pass. It offers a thorough approach to challenges involving object detection and location. YOLO detects things with amazing speed and efficacy by processing the entire image at once.

Table 1.Researchers Expansion On Different Representations

S.No	**Source**	**Technologies used**	**Draw Backs**
1	[6]	Artificial Intelligence are used for this system.	It cannot detect the multiple person in the ATM area premises.
2	[1]	YOLOv4 and YOLOv5 is used.	Other than public area it can't detects the frauds.
3	[3]	Viola Jones Algorithm, Image Processing.	It detects the fraud but cannot send alert to the nearest Security or organization.
4	[5]	GSM Model, Atmelm ATS952 micro controller.	This System requires heavy equipment including sensors GSM device and cost is also high.
5	[7]	CNN, Algorithm, C-LSTM.	The System can generate alert only in abnormal behaviour of the person.
6	[2]	CNN,Deep Learning(YOLOv5).	It cannot detects the multiple person in the background.

3 Methodology

Dataset: It can be trained on various datasets depending on the specific job or application. Common datasets used for gadget identification tasks with YOLOv5 are the Pascal VOC, COCO, and custom datasets designed for specific use cases. You must first gather or obtain an appropriate dataset for your application, then get ready for training before you can begin using it.

Table 2. Comparison Between Existing And Proposed Result Analysis

S.No	Source	Existed System	Proposed System
1	[6]	It is unable to identify more than one person within the ATM area.	It uses sophisticated multi-person detection systems for on-site security. It is possible to detect and follow individuals within the ATM area premises by deploying cameras that are equipped with computer vision algorithms.
2	[1]	It cannot identify scams anywhere other than public areas.	Unlike existing systems, it can detect frauds more accurately and effectively in a variety of domains and not only public areas.
3	[3]	Although it recognizes the fraud, it is unable to notify the closest security or organization.	The system is scalable to accommodate changing security requirements since it is built to manage warnings effectively even during busy times or spikes in activity. .
4	[7]	Only when a person exhibits odd behaviour may the system trigger an alert.	By continuously tracking and monitoring people's movements in the background, the system enables security professionals to always be aware of the situation.
6	[2]	It is unable to recognize the numerous people in the backdrop.	It continuously keeps an eye on transactional activity and instantly adjusts fraud detection algorithms, identifying and averting possible security risks.

3.1 Proposed System:

Data Gathering: Collected a wide range of annotated photos showing people in different situations wearing masks and helmets. In order to improve the model's resilience, we made sure the dataset accurately depicts real-world situations.

Data Preprocessing: Used techniques for picture augmentation to increase the training dataset's diversity, such as scaling, rollover, and rotation. To improve model convergence during training, data normalization was carried out to standardize pixel values. Split the dataset into grounding and validation adjust in priority to precisely evaluate the model's representation

Training: To speed up convergence, the YOLOv5 model was initialized using pre-trained weights on the dataset. Adjusted the model to meet particular detection needs using the customized ATM dataset.

Model Assessment: Assessed the generalization properties of the trained YOLOv5 model using a different validation dataset.

Integration of Real-Time Monitoring: Added the trained YOLOv5 model to the live video analysis real-time monitoring system. Made use of a video streaming pipeline to process frames instantly and sound an alarm when people wearing masks and helmets or several people were identified.

3.2 YOLOv5:

A technique for detecting objects in real time is called YOLO (You Only Look Once). YOLOv5 is a particular version that builds on earlier iterations and is intended to be quicker and more pre-cise. It predicts several bounding boxes and their class probabilities for an image simultaneously using a single neural network.

3.3 Labelimg Tool:

Training a YOLO (You Only Look Once) model involves several steps, including data preparation, model architecture selection, training, and evaluation. Labelimg images is an essential part of the data preparation process for training a YOLO model. Here's a general overview of the labelimg process:

Loading Images: After launching LabelImg, load the images you want to annotate into the tool. These images should represent the data you'll use to train your YOLO model.

Annotating Objects: In LabelImg, you'll see your loaded image displayed on the screen. To annotate an object in the image, follow these steps:

Repeat for All Images: Continue loading images one by one and annotating objects until you've labeled all the relevant data in your dataset.

Dataset Splitting: After labelimg, split dataset into training, validation, and test sets. This is crucial for evaluating model's performance accurately.

Training: Once dataset is labeled and prepared, we can proceed with training your YOLO model using the labeled data.3.4 Architecture:

In a single appraise, it uses a single simulated complex to predict leaping boxes and class stochastric straight from entire images. This architecture consists of an input shot feature extraction corner stone open work, usually an EfficientNet variation. It mainly consists of three parts: the Back Bone (CSP Dark-net), the Neck (PANet), and the Head (Layer of the structure). Each layer has a unique independent task to execute the all functions.

.

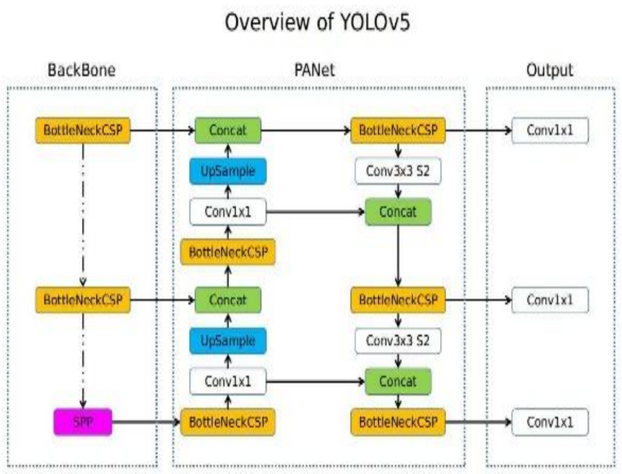


Fig 1 :Building Design of the three parts combining

In this System,Firstly we give the input to our YOLOv5 model. It processes the input through the camera and process the image. After processing it analyzes the input and detects.

Count of the person:If person are more than one it doesn't allow the person to utilize ATM services.

Detects Helmet:If the person wears helmet it gives alert, otherwise it allows us to utilize ATM.Detects Mask: If the person wears mask it gives alert, otherwise it allows us to utilize ATM.

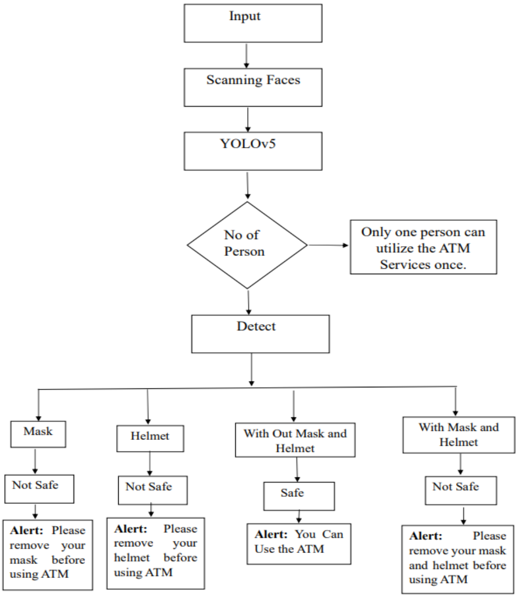
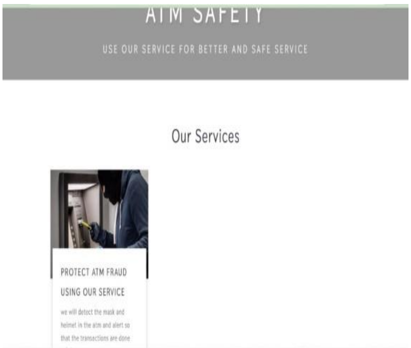


Fig 2: Proposed methodology Architecture

4 Results and Discussions

4.1 Web Application: We created one user friendly web application. In this everyone can utilize our services. Easily accessible by simply clicking on Monitor Page to start monitoring.



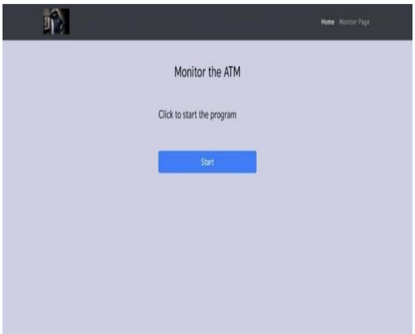


Fig 3:ATM Safety webpage frontend

Fig 4:Monitor Page of ATM safety

4.2 Predictions: Our YOLOv5 model begins identifying different faces, both with and without masks, after it has been trained.
Some of the visuals that the model predicts are seen above.

Frame (Bounding Box): The frame represents the bounding box that surrounds the object that has been identified. It determines the object's actual spatial extent within the image. This bounding box is useful to visually interpret where the object has been discoveredby the model.

Confidence Score Value: A numerical value (such as 0.88) indicates the degree of confidence or confidence score associated with the object that has been discovered. It demonstrates the degree of confidence the model has in the object it has identified as belonging to the labeled class. Increased confidence in the prediction is suggested.

4.3 Outcomes Of YOLOv5 Model:

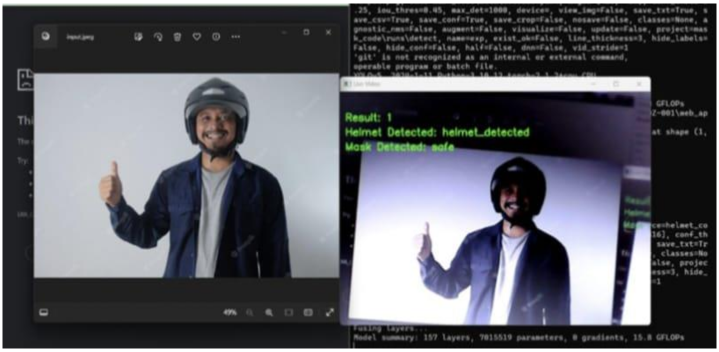


Fig 5: Application Detects Helmet

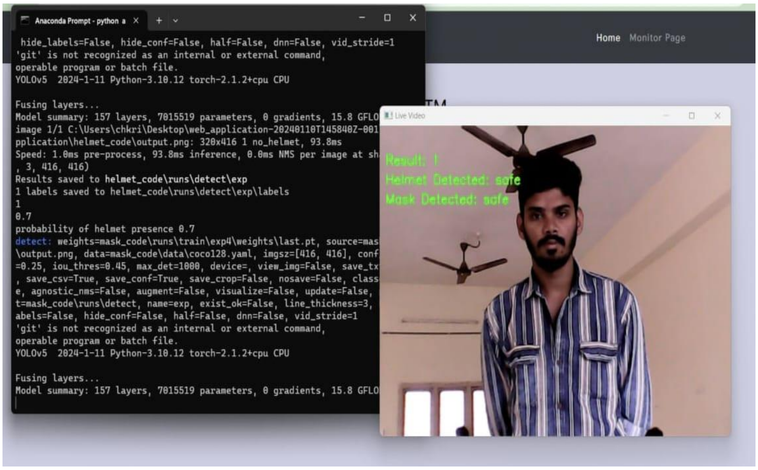


Fig 6: Picture With Without Mask And Helmet

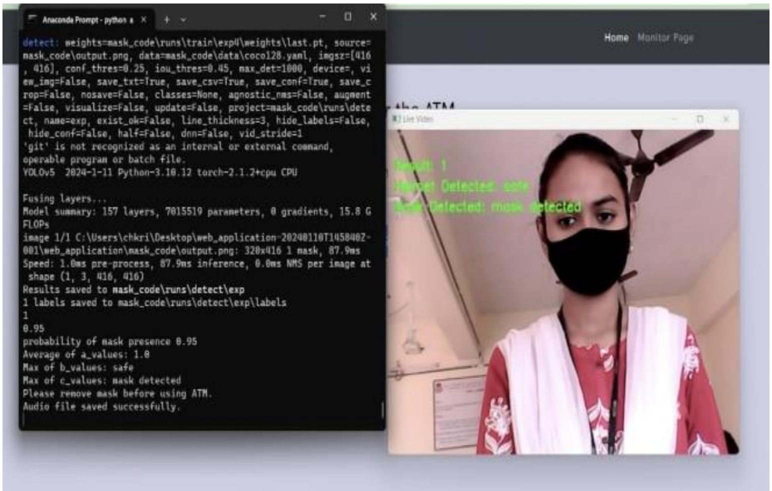


Fig 7: Application Detects Mask

Developed model can predict whether the person is wearing mask or helmet, if anyone wears mask or helmet it will gives the respective alert messages.Like:

```
     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     89/99      1.68G    0.02462     0.0246  0.0009762        105        416: 100% 45/45 [00:08<00:00,  5.29it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:01<00:00,  3.75it/s]
                   all        100        327      0.853      0.827      0.849       0.49

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     90/99      1.68G    0.02477     0.0243   0.001028         60        416: 100% 45/45 [00:07<00:00,  6.20it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.20it/s]
                   all        100        327      0.887      0.799      0.845      0.471

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     91/99      1.68G    0.02379     0.0237  0.0009083         74        416: 100% 45/45 [00:09<00:00,  4.98it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.19it/s]
                   all        100        327      0.848      0.843      0.846      0.488

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     92/99      1.68G    0.02375    0.02225  0.0008422         72        416: 100% 45/45 [00:06<00:00,  6.66it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.36it/s]
                   all        100        327      0.835      0.846      0.848      0.487

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     93/99      1.68G    0.02328    0.02262  0.0007815         92        416: 100% 45/45 [00:09<00:00,  4.96it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.25it/s]
                   all        100        327      0.827      0.834      0.838      0.487

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     94/99      1.68G    0.02395    0.02384   0.001046         66        416: 100% 45/45 [00:06<00:00,  6.50it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.30it/s]
                   all        100        327      0.913      0.807      0.875      0.495

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     95/99      1.68G    0.02361    0.02255  0.0009459         60        416: 100% 45/45 [00:09<00:00,  4.90it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.20it/s]
                   all        100        327      0.845      0.843      0.856      0.496

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     96/99      1.68G    0.02358    0.02323  0.0009543         82        416: 100% 45/45 [00:06<00:00,  6.60it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.17it/s]
                   all        100        327      0.853      0.847      0.852      0.494

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     97/99      1.68G    0.02401    0.02397   0.000893         75        416: 100% 45/45 [00:09<00:00,  4.86it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.23it/s]
                   all        100        327      0.859      0.832      0.853      0.494

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     98/99      1.68G    0.02357    0.02373   0.001002         57        416: 100% 45/45 [00:07<00:00,  6.24it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  4.20it/s]
                   all        100        327        0.9      0.807      0.855      0.494

     Epoch    GPU_mem   box_loss   obj_loss   cls_loss  Instances       Size
     99/99      1.68G    0.02356    0.02241   0.000876         70        416: 100% 45/45 [00:08<00:00,  5.37it/s]
                 Class     Images  Instances          P          R      mAP50   mAP50-95: 100% 4/4 [00:00<00:00,  5.16it/s]
                   all        100        327      0.871      0.814      0.854      0.496

100 epochs completed in 0.251 hours.
Optimizer stripped from runs/train/exp2/weights/last.pt, 14.3MB
Results saved to runs/train/exp2
```

Fig 8: Epochs of Fraud Face detection mode

5 Conclusion

In conclusion, our innovative ATM Monitoring System, utilizing the You Only Look Once (YOLO) algorithm, addresses the escalating threat of fraudulent activities involving masked individuals and multiple persons within ATM premises. By employing cutting-edge object detection techniques, the system offers real-time surveillance, promptly alerting authorities to potential security risks. The project report provides a comprehensive overview of the system's scope, objectives, and technical aspects, emphasizing its contribution to enhancing ATM security. Despite development challenges, strategic solutions were implemented, highlighting the system's evolution. With a user-friendly web interface, stakeholders can seamlessly access and manage real-time monitoring, fostering a more secure and resilient ATM environment. Ultimately, this project envisions a safer financial landscape, ensuring the integrity of transactions and prioritizing the well-being of ATM users. The project encompasses the development of a real-time monitoring system that integrates YOLO-based object detection algorithms to identify and analyse helmet and mask-wearing individuals within the ATM premises.

References

1. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). "You Only Look Once: Unified, Real-Time Object Detection." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

2. Redmon, J., & Farhadi, A. (2017). "YOLO9000: Better, Faster, Stronger." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

3. Pham, T. D., Dao, T. V., & Dang, T. H. (2020). "Face Mask Detection in the Era of COVID-19: Review." In Proceedings of the 10th International Conference on Knowledge and Systems Engineering (KSE).

4. https://iq.opengenus.org/yolov5.

5. Bernardin, K., & Stiefelhagen, R. (2008). "Multiple Object Tracking: A Literature Review." In Computer Vision and Image Understanding.

6. Di Mauro, C., Melodia, T., Palazzo, S., & Trunfio, P. (2013). "ATM Security: A Survey and Open Problems." In Journal of Computer Science and Technology.

7. Pouyan, S., Charmi, M., Azarpeyvand, A. and Hassanpoor, (2023).: Propounding first artificial intelligence approach for predicting robbery behaviour potential in an indoor security camera. IEEE Access.

8. Reddy Madhavi, K., A. Vinaya Babu, G. Sunitha, and J. Avanija. "Detection of concept-drift for clustering time-changing categorical data: An optimal method for large datasets." In Data Engineering and Communication Technology: Proceedings of 3rd ICDECT-2K19, pp. 861-871. Springer Singapore, 2020.

9. Dey, S., Chakraborty, S., Saha, S., & Dey, C. (2019). "Surveillance Video Analytics: Recent Advances, Challenges, and Future Directions." In Journal of Visual Communication and Image Representation.

10. Ghosal, S., Das, D., Nasipuri, M., & Basu, D. K. (2018). "A Survey on Real-Time Video Analytics." In ACM Computing Surveys.

11. Eltabakh, M. Y., Ben Hamza, A., & Karray, F. (2018). "Helmet Detection for Construction Safety Using Convolutional Neural Networks." In Proceedings of the International Conference on Image Analysis and Recognition (ICIAR).