Motorcycle Helmet Detection Benchmarking

Kunal Agrawal[1], Vatsa S. Patel[1], Ian Cannon[1,2], Minh-Triet Tran[3], and Tam V. Nguyen[1](✉)

[1] Department of Computer Science, University of Dayton, Dayton, USA
tamnguyen@udayton.edu
[2] Applied Sensing Lab, University of Dayton Research Institute, Dayton, USA
[3] Faculty of Information Technology, University of Science, VNUHCM, Ho Chi Minh City, Vietnam

Abstract. In this paper, we focus on evaluating the robustness of helmet detection in the context of traffic surveillance, achieved through state-of-the-art deep learning models. This aims to contribute significantly to motorcycle safety by implementing intelligent systems adept at accurately identifying helmets. An integral component of this inquiry entails a meticulous benchmark of cutting-edge object detection models and the integration of advanced techniques, aiming not only to bolster accuracy but also to improve the overall practicality and effectiveness of helmet detection systems. The experimental results highlight the effectiveness of the state-of-the-art object detection methods in detecting helmets and the potential of transferring from the traffic domain to the construction site domain.

Keywords: Helmet · Object Detection · Benchmarking · Robustness

1 Introduction

In the field of computer vision and intelligent transportation systems, the precise identification of safety equipment, particularly helmets, is pivotal for advancing road safety. This research embarks on a transformative journey to push the boundaries of helmet detection, harnessing the power of sophisticated deep-learning methodologies. The imperative for robust and efficient helmet detection becomes particularly pronounced in the domain of traffic surveillance, where traditional methods often prove inadequate in addressing the multifaceted challenges posed by real-world scenarios.

As urban landscapes undergo a notable surge in the prevalence of motorcycles and electric bikes [1], the imperative to ensure the safety of riders has become an increasingly critical concern in contemporary society. Helmets, recognized as fundamental safety accessories, play a crucial role in mitigating the risk of head injuries during accidents. However, the effectiveness of helmets is intricately linked to their proper usage, emphasizing the urgent need to develop advanced systems capable of precisely and reliably identifying the presence of helmets in various scenarios.

This paper introduces a diverse array of innovative approaches to helmet detection, as depicted in Fig. 1, with a deliberate focus on creating a new dataset and harnessing

W. Buntine et al. (Eds.): SOICT 2024, CCIS 2350, pp. 203–215, 2025.
https://doi.org/10.1007/978-981-96-4282-3_17

the capabilities of different object detection models such as YOLO [2] (You Only Look Once), Faster RCNN [3], RT-DETR (Real Time Detection Trans-former) [4] and Detectron2 [5]. These real-time object detection algorithms are strategically selected for their ability to swiftly and efficiently identify objects in dynamic scenarios, rendering them especially well-suited for applications such as traffic surveillance. The inherent robustness of these models is further emphasized through the incorporation of advanced techniques like Spatial Pyramid Pooling, thereby augmenting their effectiveness in intricate and varied environments.

Fig. 1. Different examples of helmets under different conditions/viewpoints/clarity: a) 1-rider bike, b) 2-rider bike, c) rear view, d) side view, e) night-time view and, f) blurry view.

Moreover, the research extends its exploration into the domain of ensemble methods, aiming to fortify the overall robustness and reliability of the helmet detection system. This involves the integration of multiple models within an ensemble framework, with the strategic objective of synergizing their individual strengths. By doing so, the system's performance is enhanced across a broad spectrum of conditions, solidifying its position as a comprehensive solution for accurate helmet detection in settings that continuously evolve and present dynamic challenges.

In response to the escalating prevalence of motorcycles and electric bikes, the innovative approaches presented in this paper not only address the immediate concerns surrounding helmet detection but also contribute to the broader narrative of rider safety in urban environments of developing countries. Beyond the field of technical intricacies, the research anticipates and responds to the evolving landscape of transportation, where intelligent systems are essential components in the quest for enhanced safety and efficiency.

The significance of this research transcends its technical intricacies and resonates within the broader domain of intelligent transportation systems. By elevating the precision of helmet detection in challenging conditions, the outcomes directly align with the overarching goal of such systems – the reduction of accidents and the improvement of road safety. Furthermore, envisioning seamless integration, the proposed methods could easily be incorporated into existing traffic surveillance and planning systems [6, 7], thereby fostering a safer environment for riders, and enhancing the overall efficacy of traffic management.

This paper, as a comprehensive exploration of helmet detection, meticulously navigates through the intricacies of deep learning models, the integration of novel techniques, and their practical application in real-world scenarios. Subsequent sections will systematically unveil the methodology, experiments, and results, offering a holistic understanding of the pioneering advancements achieved in the field of helmet detection and their far-reaching implications for the evolution of intelligent transportation systems. The seamless integration of these advancements into existing frameworks is poised to revolutionize road safety practices and contribute significantly to the broader landscape of intelligent transportation systems.

2 Related Work

The field of helmet detection has undergone profound transformations driven by the continuous evolution of computer vision and deep learning techniques. This section strives to offer a comprehensive review of relevant literature, shedding light on key contributions in the domain. This exploration not only synthesizes existing knowledge but also establishes a contextual foundation for the proposed architecture.

A noteworthy aspect of recent research involves the exploration of YOLOv5s in the realm of object detection tasks. Huang et al. [8] pioneered an advanced YOLOv5s-based method specifically tailored for electric bike helmet recognition. Their innovative approach yielded enhanced detection efficiency and practicality, acting as a catalyst for further investigations in specialized domains. This underscores the adaptability of YOLOv5s in addressing nuanced challenges within the realm of helmet detection.

Chen et al. [9] embarked on the development of lightweight helmet detection algorithms, a crucial pursuit for ensuring real-time processing in safety applications. Their work placed significant emphasis on safety helmet-wearing detection in industrial settings, advocating for algorithms that offer swift and accurate recognition. This research substantially contributes to the intersection of real-time safety applications and computer vision, recognizing the importance of expeditious and precise helmet detection in critical environments.

In a parallel vein, Fan et al. [10] delved into the application of ensemble methods in helmet detection. Their deep learning-based ensemble method showcased advancements in minimizing false positives, ensuring a more reliable helmet detection system. This work not only addresses the challenges of false positives but also makes valuable strides in enhancing the overall robustness of object detection models. The integration of ensemble methods adds a layer of complexity and efficacy to helmet detection systems.

The YOLO series has gained popularity for real-time object detection due to its effective balance between speed and accuracy, but its performance is hindered by the Non-Maximum Suppression (NMS) step. Transformer-based detectors like DETR offer an alternative by eliminating NMS but suffer from high computational costs that limit their practicality. To address these issues, Lv et al. [4] proposed the Real-Time DEtection

TRansformer (RT-DETR), an end-to-end object detector that maintains high speed and accuracy by employing an efficient hybrid encoder for rapid multi-scale feature processing and an uncertainty-minimal query selection to enhance initial query quality, while also allowing flexible speed tuning through adjustable decoder layers.

Recent advancements in deep learning have significantly improved image classification, segmentation, and object detection, including detecting helmets on bike riders to enhance road safety. Singh et al. [5] analyze various approaches and experiments with state-of-the-art models like Detectron2 and EfficientDet, demonstrating their effectiveness in helmet detection.

The synthesis of the reviewed literature underscores the diverse approaches employed in helmet detection, with a particular emphasis on the YOLOv5s architecture, lightweight algorithms, integration of SPP, utilization of ensemble methods, and the significance of curated datasets. Building upon these insights, the proposed architecture aspires to contribute to ongoing advancements in intelligent transportation systems and road safety. By amalgamating strengths and addressing the limitations highlighted in the literature, the proposed architecture seeks to elevate the precision, efficiency, and adaptability of helmets detection in dynamic real-world scenarios. This endeavor aligns with the broader trajectory of advancements in computer vision and deep learning, fostering a safer and more intelligent future for transportation systems.

Fig. 2. The user interface of the annotation tool.

3 Proposed Work

In this section, we outline a thorough research methodology for benchmarking a resilient helmet detection system in traffic videos related to helmets. The proposed work includes data collection, data augmentation and preprocessing, model development, training, and evaluation, and cross-domain adaptation using the Hardhat Construction Dataset [11], iterative refinement, ethical considerations, and documentation. Leveraging pertinent literature and best practices guides our research at every step.

3.1 Motorcycle Helmet Detection Dataset (MHDD)

Our methodology started with the collection of the Motorcycle Helmet Detection Dataset (MHDD), a foundational element crucial for developing a robust helmet detection system capable of adapting to a myriad of diverse environmental conditions and traffic scenarios. There are a few similar datasets available such as the Caltech Pedestrian dataset [12] (taken from a vehicle driving) and multiple datasets focusing on biker's helmets. No such public dataset is readily available that focuses on motorcyclist helmets from a traffic camera view. This made us create a new dataset that overcomes these issues and provides a readily available dataset for future use.

Compiling our dataset involves a comprehensive sourcing strategy, tapping into various channels to ensure a rich and diverse representation. We leverage public video feeds from traffic cameras and surveillance systems in Vietnam from different regions. This exhaustive selection ensures the inclusion of a broad spectrum of traffic scenarios and environmental conditions in developing countries, significantly contributing to the robustness and adaptability of our helmet detection system.

The cornerstone of our methodology lies in rigorous data annotation performed by trained annotators. This involves the meticulous marking of regions of interest (ROIs) containing motorcyclists and the precise indication of helmet presence [13]. In our work, we use Roboflow [14] for annotation. In particular, this tool empowers annotators to create high-quality annotations efficiently, thereby contributing to the depth and accuracy of our dataset. The graphical user interface of this tool is illustrated in Fig. 2.

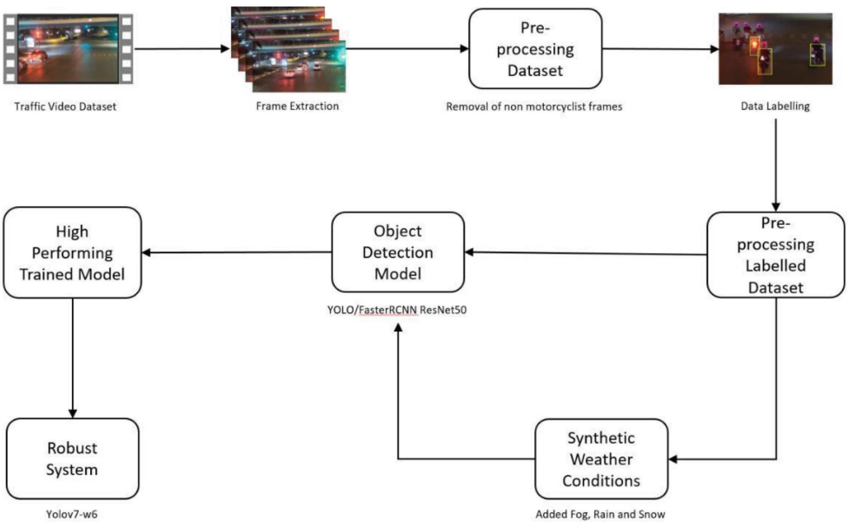

Fig. 3. The flowchart of the computational framework.

3.2 Data Processing and Augmentation

To ensure uniform data representation, videos undergo segmentation into individual frames or clips, effectively addressing variations in frame rates and mitigating compression artifacts [15]. In our pursuit of an enriched and diverse dataset, we employ a spectrum of data augmentation techniques, encompassing random rotations, flips, brightness adjustments, and cropping [16].

In the initial stages of data preprocessing, we prioritize image enhancement techniques such as histogram equalization and adaptive contrast enhancement. These methods are pivotal in refining image quality, particularly for videos captured in challenging lighting conditions [17]. Subsequently, we implement normalization, wherein pixel values in the images are adjusted to attain a zero mean and unit variance, thereby instrumental in facilitating model convergence and enhancing the overall stability of the system [18].

As we immerse ourselves in simulating diverse weather conditions during the data augmentation process, our primary objective is to fortify the robustness of the helmet detection system. Synthetic scenarios involving rain, fog, and snow are introduced, effectively replicating adverse weather conditions. By leveraging Pillow and NumPy libraries we added synthetic weathers such as creating an overlay of small white circles (snowflakes) with random positions and some transparency to simulate snowfall, generating semi-transparent white ellipses (fog clouds) randomly placed over the image and then applying Gaussian blur to soften and blend them to create a foggy appearance, and adding vertical white lines (raindrop) of varying lengths and positions onto a transparent overlay to mimic falling rain.

Our dataset becomes more resilient by seamlessly integrating these simulated weather conditions into the training and evaluation phases. This ensures that the helmet detection system is not only adept at handling real-world challenges posed by dynamic weather scenarios but also excels in accurately identifying helmets in adverse conditions. This comprehensive approach to data augmentation not only augments the model's adaptability but also significantly contributes to its precision, reliability, and overall effectiveness in varied environmental conditions.

The overall size of the dataset consists of 8,000 images including different weather scenarios, i.e., normal conditions, rain, fog, and snow. The dataset is broadly divided into two sections Colored and Grayscale each comprising of 4000 images. Furthermore, both these sections have 4 subsections based on conditions such as normal conditions, synthetic snow added, synthetic fog added, and synthetic rain added. Each of these subsections has 1000 images each respectively in both colored and grayscale format making the total size of the dataset 8000 images. (2 broad sections x 4 subsections x 1000 images each).

3.3 Computational Framework

Our envisioned framework for robust helmet detection in traffic videos is strategically crafted to surmount challenges posed by fluctuating weather conditions, ultimately ensuring the safety of motorcyclists. This comprehensive framework consists of key components, each playing a pivotal role in the overall system, as illustrated in Fig. 3.

To initiate the framework, traffic videos captured by surveillance cameras serve as the primary input source. These videos constitute the foundational data for the helmet detection system. Frame extraction is the next crucial step, where frames are extracted from the input traffic videos. Each frame represents a snapshot of the traffic scene, forming the basis for subsequent analysis. Extracted frames undergo essential preprocessing tasks, including noise reduction, contrast enhancement, and resizing as part of the preprocessing dataset phase. These tasks aim to ensure the consistency and high quality of the input data, laying the groundwork for accurate analysis.

Moving forward, the data labeling stage is imperative. For supervised training, the dataset is meticulously annotated, with each helmet within the frames being labeled. This annotated dataset becomes the bedrock for training the helmet detection model. Subsequently, the labeled dataset undergoes further refinement to align it with real-world scenarios. Synthetic weather scenarios, such as rain, fog, and darkness, are introduced to simulate various environmental conditions, making the dataset more robust and reflective of diverse challenges.

The object detection model, the final and pivotal element in the framework, is a deep learning-based model specifically designed to identify helmets within the frames. Crafted for high accuracy and capable of handling challenging conditions, this model is finetuned to excel under various weather conditions. The finetuned object detection model stands as the core component for robust helmet detection, ensuring the system's adaptability to dynamic scenarios. The object detector locates and highlights helmets within the frames, contributing significantly to motorcyclist safety in diverse traffic scenarios.

Fig. 4. Object Detection results of all models.

3.4 Implementation

We adopt pre-trained deep convolutional neural networks (CNN) that serve as the base model for feature extraction. Consideration will be given to well-established architectures such as ResNet [19], VGG [20], or Inception [21]. Then, the selected base model will undergo fine-tuning using our annotated traffic video dataset. Transfer learning techniques will be applied, leveraging knowledge from large-scale datasets like ImageNet to adapt the model for helmet detection. An object detection head, such as a Region Proposal Network (R-CNN), will be added to the base model to identify ROIs containing motorcyclists' heads. For the helmet detection head, a subnetwork for helmet detection will be integrated within the ROIs identified in the previous step. Architecture options include Faster R-CNN [3] or YOLO [2] for object detection.

Regarding the loss function, an appropriate loss function, combining classification loss and localization loss (e.g., Smooth $\mathscr{L}1$ loss), will be defined for helmet detection. The model will be trained using annotated data, employing an optimizer such as Adam [22]. Training progress will be monitored using validation data, and techniques like learning rate schedules and early stopping will be used for optimization.

4 Benchmarking

In this section, we conduct a comprehensive evaluation of robust helmet detection benchmarking, aiming to scrutinize the performance of multiple models across diverse scenarios. Our objective is to benchmark and compare these models using key metrics: precision, recall, F1-Score, and mean average precision (mAP) [18] to gauge their effectiveness in helmet detection.

Fig. 5. The illustration of different weather conditions.

4.1 Metrics Performance

We employed the following evaluation metrics. Precision indicates the percentage of correctly detected helmets out of all identified by the model. The recall represents the percentage of actual helmets correctly detected by the model. F1-Score is a balanced measure between precision and recall. Finally, Mean Average Precision (mAP) reflects the average precision scores for different classes, in this case, helmets.

4.2 Modal Comparison

Several models designed for helmet detection underwent training and testing. Table 1 & Table 2 summarize their performance. Table 1 streamlines the decision-making process for selecting the most appropriate model for real-world applications as discussed in further sections in more detail.

4.3 Visual Assessment

Visual assessment, coupled with quantitative metrics, plays a pivotal role in conducting a thorough evaluation of the model's performance. The inclusion of visual examples, as illustrated in Fig. 4, enriches our understanding of how each model operates in diverse scenarios, introducing a qualitative dimension to the evaluation process.

Within these visual representations, the models' effectiveness in identifying helmets under various conditions is vividly highlighted. Each image encapsulates the model's responses to real-world challenges, encompassing factors such as varying lighting conditions, adverse weather scenarios, and complex traffic scenes. This visual evidence not only validates the quantitative data derived from metrics but also provides a comprehensive and nuanced perspective on the models' resilience and adaptability.

Table 1. The experimental results of different models trained on traffic data and tested on traffic data. Both training and testing data are included in MHDD dataset. The best performance is marked as **boldface**.

Sr. No	Model	mAP
1	Yolov5s (Colored)	0.857
2	Yolov5s (Grayscale)	0.863
3	Yolov7-w6 (Colored)	0.881
4	Yolov7-w6 (Grayscale)	**0.899**
5	Yolov6s (Colored)	0.809
6	Yolov6s (Grayscale)	0.795
7	Yolov8s (Colored)	0.863
8	Yolov8s (Grayscale)	0.847
9	FasterRCNN-ResNet50 (Colored)	0.870
10	FasterRCNN-ResNet50 (Grayscale)	0.871
11	RT-DETR (Colored)	0.790
12	RT-DETR (Grayscale)	0.770
13	Detectron2 (Colored)	0.705
14	Detectron2 (Grayscale)	0.757

These visual examples act as a qualitative supplement to the quantitative assessment, fostering a deeper comprehension of the models' strengths and limitations. Beyond numerical scores and metrics, these images furnish the research community with tangible proof of the model's performance in practical, real-world situations. Such visual assessments contribute significantly to a holistic and insightful interpretation of the models' overall effectiveness and suitability for deployment in dynamic environments.

4.4 Discussions

The comprehensive evaluation has unveiled valuable insights into the models' performance as shown in Table 1. The critical metric of precision highlights the exemplary capabilities of Models 3 and 10 (different families). Model 4 (Yolov7-w6 Grayscale) distinguishes itself with an impressive mean Average Precision (mAP) of 0.899, closely trailed by Model 10 (FasterRCNN-ResNet50 Grayscale) with a commendable mAP of 0.871.

In the realm of Recall, Model 3 demonstrates outstanding performance, surpassing its counterparts with a mAP of 0.881. Model 10 exhibits competitive recall capabilities, boasting a mAP of 0.871. Meanwhile, Model 1 (Yolov5s Colored) achieves the highest F1-Score, with a mAP of 0.857. This metric signifies a harmonious blend of precision and recall, positioning it as a noteworthy contender adept at balancing these two crucial aspects of helmet detection.

A pivotal consideration lies in the mAP, where Model 10 (FasterRCNN-ResNet50 Grayscale) outperforms others with a score of 0.871. This underscores the model's consistency and effectiveness across a spectrum of conditions, reinforcing its reliability in diverse helmet detection scenarios. Considering all the results we can broadly say the model works better in grayscale over colored ones.

These insights empower users to make informed decisions by comprehending the trade-offs between precision, recall, and adaptability. Although RT-DETR and Detectron2 performed poorly compared to all other models in the range of mAP 0.7–0.8 but whether prioritizing precise helmet identification or comprehensive coverage, the varied strengths of each model facilitate customized selections aligned with specific application needs.

Table 2. The experimental results on the test traffic data with different models and training datasets.

Model	Train Dataset	mAP
Yolov7-w6 (Colored)	HardHat + Traffic	0.904
Yolov7-w6 (Grayscale)	HardHat + Traffic	0.892
Yolov7-w6 (Colored)	HardHat	0.904
Yolov7-w6 (Grayscale)	HardHat	0.865
FasterRCNN-ResNet50 (Colored)	HardHat	0.878
FasterRCNN-ResNet50 (Grayscale)	HardHat	0.904

4.5 Robustness Evaluation

This section emphasizes the robustness of hardhat helmet detection, achieved by integrating a hardhat dataset with the traffic dataset and training it using the best-performing models (Model 4 - Yolov7-w6 Grayscale) and (Model 10 – FasterRCNN-ResNet50) as discussed in the previous section. Evaluations showcase the model's adaptability and performance in real-world scenarios as mentioned in Table 2 with an impressive mAP of 0.904. Though both the models were able to score the mAP of 0.904 the preferred model will be Model 4 (YOLOv7-w6) as it's a faster model compared to Model 10 (FasterRCNN-ResNet50).

Weather Robustness Assessment. Model 4's adaptability under different weather conditions is tested, showing consistent precision and recall across adverse scenarios as shown in Fig. 5.

Real-World Implications. The benchmarking results have significant implications for helmet detection system deployment. Depending on application demands, users can select a suitable model, considering critical factors such as precision and recall. The adaptability of models across different weather conditions underscores the necessity for versatile, robust systems ensuring safety in adverse environments The benchmark has significantly advanced our understanding of helmet detection models in diverse conditions, emphasizing the critical need for adaptability to ensure motorcyclist safety. Beyond traffic scenarios, the application scope extends to industrial safety and various domains, showcasing the broader societal impact of intelligent transportation systems. Our contributions to the fields of computer vision, deep learning, and transportation safety underscore the importance of tailored model selection to meet specific application needs, recognizing the nuanced challenges presented in real-world scenarios. Despite inherent limitations, this paper serves as a foundational step toward future advancements in helmet detection research, shedding light on the necessity for diverse datasets and effective domain adaptation strategies.

a b c d

Fig. 6. Visualization of failure cases: a) wrongly detected helmet b) wrongly detected helmet and motorcycle c) dark hair detected as helmet and d) cap detected as helmet.

5 Conclusion and Future Work

This paper advances helmet detection systems strategically focusing on optimizing real-time processing capabilities for practical deployment in dynamic traffic scenarios. We address the ongoing challenge of enhancing model performance in adverse weather conditions which demands dedicated research efforts and exploring innovative techniques and technologies. We plan to explore multimodal data fusion, where integrating information from various sensors could significantly elevate detection accuracy, particularly in scenarios with challenging visibility conditions as shown in Fig. 6. The expansion of benchmark datasets and the exploration of advanced domain adaptation techniques are pivotal steps toward creating more robust models capable of handling diverse and complex environments.

For future work, we focus on optimizing helmet detection systems for real-time processing, facilitating practical deployment in traffic scenarios. Enhancing model performance in adverse weather conditions remains a critical challenge, warranting further investigation. Multimodal data fusion, encompassing data from various sensors, could enhance detection accuracy, especially in challenging visibility conditions. Expanding benchmark datasets, exploring domain adaptation techniques, and broadening the scope to include anomaly detection for comprehensive traffic safety are avenues for future exploration.

Acknowledgment. This research was supported by the National Science Foundation (NSF) under Grant 2025234.

References

1. Huang Ma, C., Yang, D., Zhou, J., Feng, Z., Yuan, Q.: Risk riding behaviors of urban e-bikes: a literature review. Int. J. Environ. Res. Public Health **16**(13), 2308 (2019)
2. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection. arXiv preprint arXiv:1506.02640 (2016)
3. Redmon, J., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788 (2016)
4. Lv, W., et al.: DETRs Beat YOLOs on Real-time Object Detection (2023)
5. Singh, R., Shetty, S., Patil, G., Bide, P.J.: Helmet detection using detectron2 and efficientdet. In: 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, pp. 1–5 (2021)
6. Nguyen, X.-D., et al.: Adaptive multi-vehicle motion counting. J. Signal Image Video Processing **16**(8), 2193–2201 (2022)
7. Nguyen, T.V., et al.: Data-driven city traffic planning simulation. ISMAR Adjunct, pp. 859–864 (2022)
8. Huang, B., et al.: An improved YOLOv5s-based helmet recognition method for electric bikes. Appl. Sci. **13**(15), 8759 (2023)
9. Chen, J., Deng, S., Wang, P., Huang, X., Liu, Y.: Lightweight helmet detection algorithm using an improved YOLOv4. Sensors **23**(3), 1256 (2023)
10. Fan, Z., Peng, C., Dai, L., Cao, F., Qi, J., Hua, W.: A deep learning-based ensemble method for helmet-wearing detection. PeerJ. Computer Sci. **6**, e311 (2020)
11. Shen, J., Xiong, X., Li, Y., He, W., Li, P., Zheng, X.: Detecting safety helmet wearing on construction sites with bounding-box regression and deep transfer learning. Computer-Aided Civil and Infrastructure Eng. **36**(2), 180–196 (2021)
12. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (VOC) challenge. Int. J. Comput. Vision **88**(2), 303–338 (2010)

13. Zhou, Y., Liu, L., Shao, L., Mellor, M.: Fast automatic vehicle annotation for urban traffic surveillance. IEEE Trans. Intell. Transp. Syst. **19**(6), 1973–1984 (2017)
14. Dwyer, B., Nelson, J., Solawetz, J., et al.: Roboflow (Version 1.0) (2022). https://roboflow.com.
15. Wang, W., Zhou, T., Porikli, F., Crandall, D., Van Gool, L.: A survey on Deep Learning Technique for Video Segmentation. arXiv e-prints, arXiv-2107 (2021)
16. Li, J., Wang, D., Li, S., Zhang, M., Song, C., Chen, X.: Deep learning based adaptive sequential data augmentation technique for the optical network traffic synthesis. Opt. Express **27**(13), 18831–18847 (2019)
17. Ren, D., Sun, T., Yu, C., Zhou, C.: Research on safety helmet detection for construction site. In: 2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI), pp. 186–189. IEEE (2021)
18. Flach, P., Kull, M.: Precision-recall-gain curves: PR analysis done right. Advances in Neural Information Processing Systems **28** (2015)
19. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016)
20. Simonyan, K., Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556 (2014)
21. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818–2826 (2016)
22. Kingma, D.P., Ba, J.: Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980 (2014)