IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ≍ Peer-reviewed & Refereed journal ≍ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13830

Real-Time Detection of Helmet Non-Compliance Using YOLOv3 and EasyOCR

Dhivya P[1], Ramya M[2]

Assistant Professor, Department of Artificial Intelligence and Data Science, Karpagam College of Engineering, Coimbatore, India[1]

Assistant Professor, Department of Information Technology, Velalar College of Engineering and Technology, Erode, India[2]

Abstract: Nowadays, wearing a helmet is crucial for safety, providing essential head protection and reducing the risk of severe injuries, potentially saving lives for motorbike and bicycle riders. Conversely, not wearing a helmet poses serious risks, including increased vulnerability to head injuries, fatalities, and elevated accident susceptibility due to reduced visibility and non-compliance with traffic regulations. A real-time system can leverage deep learning to detect instances of helmet non-compliance. By employing the YOLOv3 algorithm, helmet non-compliance can be detected with an accuracy ranging between 80% and 95%. Furthermore, the XML framework is utilized for precise number plate extraction from vehicles of violators, ensuring similar accuracy levels. The EasyOCR algorithm converts these number plate images into text, facilitating their storage in a database.The system ensures comprehensive documentation and monitoring by securely storing the data of helmet non-compliance, including vehicle numbers, in a database. This supports detailed record-keeping, reporting, and further analysis, ultimately contributing to improved road safety and enforcement of helmet regulations.

Keywords: Deep learning, YOLOv3, XML framework, Easy OCR, Helmet non-compliance , Number plate extraction.

I. INTRODUCTION

Safety on the roadways is more important than ever in the rapidly changing environment of today, especially for riders of motorbikes and bicycles. The use of helmets plays a pivotal role in safeguarding individuals by providing essential head protection and significantly reducing the risk of severe injuries. This project focuses on addressing the vital issue of helmet compliance among riders, recognizing its direct impact on saving lives. Conversely, the absence of helmet usage exposes riders to serious risks, such as heightened vulnerability to head injuries, increased fatalities, and a greater likelihood of accidents due to reduced visibility and non-compliance with traffic regulations. To tackle this challenge, the project harnesses the power of deep learning and cutting-edge technology. Through the implementation of deep learning algorithms, the system developed for this project can swiftly detect instances of helmet non- compliance. It employs advanced image recognition techniques to identify and extract number plates from vehicles where helmets are not being worn. The collected information, including the identified number plates, is securely stored in a database. This innovative approach goes beyond mere identification; it contributes to the enhancement of safety enforcement by automating the monitoring of helmet regulations. Additionally, the system facilitates the tracking of non compliant vehicles over time. Ultimately, this project aims to create a more efficient and proactive road safety environment by utilizing technology to address and rectify instances of helmet non-compliance.

II. LITERATURE SURVEY

Numerous research works have demonstrated the effectiveness of technology-based strategies in enhancing traffic safety. In [1] construction, worker safety, especially concerning head injuries, is crucial. While RFID tags with sensors fail to confirm proper helmet usage, machine learning methods, despite their strengths, struggle with similar-looking objects. Deep learning, particularly using convolutional neural networks (CNNs) like the SSD-MobileNet algorithm, shows promise in handling complex situations more effectively. Despite challenges such as detecting helmets in unclear images or amidst complex backgrounds, deep learning-based approaches seem promising for enhancing safety by monitoring helmet usage on construction sites. The evolution of object recognition [2] in images has seen significant transformations. Initially, techniques like SIFT and HOG were promising but struggled with complex images. The "neocognitron," introduced by Fukushima, was an early concept that did not fully meet expectations. However, Convolutional Neural Networks (CNNs), pioneered by LeCun and later advanced by Krizhevsky, achieved exceptional performance,

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13830

particularly with ImageNet. R-CNN further advanced object detection by integrating region-based proposals and CNN features, leading to better image comprehension and object delineation. CNNs have since become pivotal in modern image analysis and object recognition. The DH-SiamRPN algorithm [3] addresses challenges in object tracking within obscured visual environments. It combines techniques for identifying the darkest areas in images and refining image clarity, which enhances tracking accuracy. The algorithm's adaptability allows it to dynamically adjust parameters based on visual data quality, optimizing image conditions for precise tracking. Its integration of the Distance-IoU concept significantly refines object selection in challenging conditions, making it effective in environments with visual obstructions like fog or haze. Enhanced deep-learning approaches [4] have also been developed for detecting safety helmet usage at construction sites. By incorporating multi-scale training, additional anchors, and Online Hard Example Mining (OHEM), these approaches address issues like partial occlusion and varying object sizes, outperforming the Faster RCNN algorithm by 7% in detection accuracy. These methods demonstrate robustness in poor lighting and occluded images, showing potential for practical safety applications. Dense Convolutional Networks (DenseNet) [5] represent another advancement in deep learning. DenseNet improves deep neural networks by establishing forward connections between each layer and all subsequent layers, enhancing information flow, feature propagation, and parameter efficiency. This design reduces parameters, encourages feature reuse, and addresses the vanishing-gradient issue. DenseNet consistently outperforms previous models on object recognition benchmarks with fewer parameters, providing a promising approach to efficient deep learning. In [6] human detection, linear SVM-based methods, particularly HOG descriptors, have shown effectiveness. HOG descriptors on a dense grid outperform traditional methods, offering better human isolation in complex datasets. The suggested method achieves near-perfect results on the MIT walking people database, significantly reducing false positives. Future work aims to optimize efficiency further and incorporate motion information into HOG-based detectors. Lastly, [7] a study comparing CNN-based models (Faster R-CNN, YOLO v3, and SSD) for detecting agricultural greenhouses using high-resolution satellite images found YOLO v3 to be the best performer in terms of mAP, FPS, and visual inspection. Transfer learning proved effective in enhancing detection accuracy. YOLO v3 is recommended for operational AG detection due to its balance of speed and accuracy. Future research may explore multispectral and hyperspectral data for detailed AG information and create high-quality datasets for agricultural and land management applications.

III. METHODOLOGY

A. *Data Collection*

Creating a diverse dataset for people on vehicles involves collecting images showcasing various scenarios, encompassing different demographics, environments, and vehicle types. The dataset should capture individuals both wearing and not wearing helmets, ensuring a broad representation of real-world situations. Annotations will be added to each image, clearly indicating whether a helmet is present or absent. Additionally, the dataset will include the extraction of vehicle number plates, enhancing its utility for comprehensive analysis. This meticulously curated dataset is used to develop and training the model for our research which detect helmet and license plate contributing to enhanced safety and surveillance applications.

Fig. 1 Data Collection for detection of the number plate of non-wearing helmet

B. *Preprocessing*

In the preprocessing phase, standardizing the dataset is crucial for effective Convolutional Neural Network (CNN) training. Resizing all images to a uniform size ensures consistent input dimensions, facilitating model convergence. Normalizing pixel values within a specific range, such as 0 to 1, aids in stable training by mitigating the impact of varying intensity levels. Augmentation techniques, including rotation, scaling, and flipping, are applied to diversify the dataset, enhancing the model's generalization ability. This comprehensive preprocessing strategy optimizes the CNN's performance, fostering resilience against variations in input data and promoting robust feature extraction for improved accuracy in tasks like detecting helmet and license plate.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13830

C. *Model Architecture*

Multiple convolutional layers are used in the CNN architecture for helmet detection to extract features, and then pooling layers are used for spatial downsampling. Striding towards abstraction, these layers are complemented by fully connected layers, enabling comprehensive pattern recognition. The model is trained using the meticulously annotated dataset, where helmets' presence or absence is labeled. Leveraging optimization algorithms like stochastic gradient descent, the network refines its parameters through iterative training epochs, gradually enhancing its ability to discern helmet- related patterns. The aim is to achieve a robust, accurate model capable of real-time helmet detection, contributing to improved safety measures in various settings, such as traffic surveillance and personal protection scenarios.

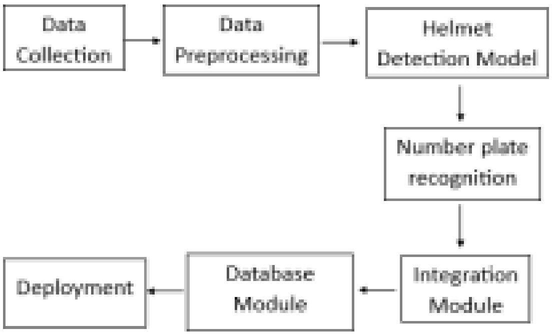

Fig. 2 Model Architecture

D. *Helmet Detection*

Deploying the trained helmet detection model involves integrating it into a system for either real-time or batch image processing. In real-time scenarios, the model can continuously analyze streaming images, promptly identifying instances of individuals not wearing helmets. For batch processing, the model systematically evaluates a collection of images, providing efficient and accurate results more deliberately. The implementation leverages the model's learned features to discern the presence or absence of helmets, contributing to enhanced safety measures. This application finds practical utility in areas like surveillance, enforcing safety protocols, and fostering proactive intervention to mitigate risks associated with non compliance.

E. *Vehicle Number Plate Recognition*

Diverse license plate representations are ensured by curating a distinct dataset for vehicle number plate recognition. Convolutional layers are used in the CNN architecture to extract features, while pooling layers are used to capture important spatial information. Fully connected layers are employed for comprehensive pattern recognition, facilitating accurate identification of alphanumeric characters on license plates. The annotated dataset, which labels each image with the associated vehicle number plate information, is used to train the algorithm. Leveraging optimization techniques, such as gradient descent, the model refines its parameters during training. This comprehensive approach aims to develop a robust number plate recognition system, contributing to enhanced vehicle surveillance and management.

F. *Integration*

Integrating the helmet detection and number plate recognition models involves creating a unified system where the models run sequentially. Upon detecting a non wearing helmet, a trigger mechanism activates the number plate recognition model, enhancing surveillance and identification capabilities simultaneously. This synergistic integration optimizes safety and compliance measures in various scenarios. Establish a database to efficiently store identified number plates. Upon detecting a non- wearing helmet, the system dynamically records the associated vehicle number plate in the database. This seamless integration ensures accurate and retrievable documentation, providing valuable insights into non-compliance instances for subsequent analysis or enforcement purposes.

G. *Testing and Evaluation*

Conduct rigorous testing on an independent dataset to assess the integrated system's performance. Evaluation metrics such as recall, precision and F1 score to gauge recognition rate. If needed, fine-tune the helmet detection and number plate recognition models iteratively to enhance overall system efficacy and ensure reliable real-world application.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13830

IV. Feature Selection Techniques

A. SSD-MobileNet

The SSD-MobileNet model is perfect for actual time object identification on devices with restricted resources because it combines the effectiveness of Single Shot Multibox Detector (SSD) with MobileNet. Its use is not limited to important domains such as construction worker safety. With 67.95% of accidents attributed to improper helmet use, a real-time automatic system becomes essential. SSD-MobileNet, leveraging deep learning, offers promise in effectively monitoring helmet usage. While challenges exist in unclear images and complex backgrounds, this approach holds the potential for enhancing safety on construction sites.

B. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) stand as pivotal deep learning models tailored for structured grid data processing, primarily in the realm of image analysis. Widely applied in computer vision, CNNs leverage convolutional layers to autonomously learn hierarchical feature representations from input data. By utilizing filters, these layers adeptly identify patterns such as edges and textures, allowing the network to comprehend intricate visual hierarchies. CNN architectures typically include fully connected layers for classification and pooling layers for reduction of dimensionality. Celebrated for their proficiency in capturing spatial dependencies within data, CNNs have revolutionized image recognition tasks, playing a crucial role in facial recognition, object detection, and medical image analysis advancements. In the evolution of computer vision, from traditional methods like SIFT and HOG to Fukushima's "recognition" and LeCun's pioneering CNN, these models have simplified and enhanced image understanding, transforming modern image analysis with efficient solutions for recognition and segmentation tasks.

C. DH-SiamRPN

DH-SiamRPN, a potent visual object tracking algorithm, employs a fusion of deep features and correlation filters for robust performance. Ideal for real-time tracking in projects, it finds applications in surveillance, autonomous vehicles, and diverse computer vision tasks, offering both accuracy and efficiency. Its proficiency in obscured object tracking, dynamic parameter adaptation, and integration of Distance IoU enhances precision. With a sophisticated network architecture, it excels in image restoration, proving promising for real- world scenarios like fog or haze, where visual challenges abound.

D. You Only Look Once (YOLO)

You Only Look Once is a computer vision technique for an actual-time object detection algorithm, crafted by Joseph Redmon and Santosh Divvala. YOLO's unique approach involves processing images in a single pass, breaking them into a n number of grids, and assigning bounding boxes with class probabilities like x and y coordinates along with probability of having object for within each grid cell. This strategy delivers swift and efficient object detection, rendering YOLO ideal for real-time applications such as video analysis and autonomous vehicles. The algorithm strikes a balance between speed and accuracy, presenting a robust solution for diverse object detection tasks. YOLO's ability to simultaneously identify multiple objects using a singular forward pass through the neural network enhances its versatility. With its innovative regression-based approach and emphasis on minimizing false positives, YOLO continues to redefine object detection, achieving speeds like 155 frames per second with variants like Fast YOLO, making it a vital tool across various domains, including the dynamic landscape of autonomous driving.

E. Faster Region-based Convolutional Neural Network

Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun proposed Faster Region-based Convolutional Neural Network is a pivotal breakthrough in computer vision's object detection. Introducing a two-stage architecture, it uses a Region Proposal Network (RPN) within the model, enhancing both speed and accuracy in object detection. The RPN generates candidate object regions for final classification and bounding box prediction. This innovative approach finds extensive applications in image analysis, autonomous vehicles, and video surveillance. Faster R-CNN efficiently addresses computational bottlenecks, predicting object bounds and generating high-quality proposals. With a commendable 5fps frame rate on GPUs and state-of-the-art accuracy across datasets, it has become widely adopted in commercial systems, emphasizing its practical utility and establishing itself as a cornerstone in contemporary object detection frameworks.

F. DenseNet

The dense connectivity design of DenseNet, a deep learning architecture for image categorization, allows each layer to acquire input from all layers that come before it. In doing so, vanishing gradient problems are addressed and reuse of features, model compactness, and flow of gradients are encouraged. The architecture includes dense blocks, transition layers, and global average pooling, offering high accuracy with parameter efficiency. DenseNet's novel approach connects each layer in a feed forward manner, significantly improving information flow and reducing parameters. The

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024
DOI: 10.17148/IJARCCE.2024.13830

architecture's scalability to hundreds of layers ensures consistent accuracy improvement.

G. Support Vector Machine

Encouragement Strongly suited for both classification and regression, Vector Machine is a powerful supervised machine learning technique with a wide range of applications. SVM endeavors to discover the optimal hyperplane differentiating data points in feature space, maximizing the margin between classes. Support vectors, proximate to the decision boundary, crucially influence this determination. With proficiency in handling high dimensional data, SVM's effectiveness extends to nonlinear classification through kernel functions. Its versatility finds applications in image classification, text categorization, and bioinformatics, facilitating both linear and non- linear pattern recognition tasks. SVM's adaptability is evident in regression tasks, where it captures a margin around data points. In practical instances, such as robust visual object recognition, SVM's linear variant, when coupled with Histograms of Oriented Gradient descriptors, excels in human detection, emphasizing fine scale gradients and spatial binning to mitigate false positives. This showcases SVM's relevance in addressing complexities within datasets, affirming its status as a valuable tool across varied domains.

V. CONCLUSION

In conclusion, the introduced advanced system, utilizing state-of-the-art deep learning technology, emerges as a robust solution for the real-time identification of individuals without helmets at signal stops in road traffic. With a primary focus on enhancing compliance with helmet regulations, the system demonstrates exceptional accuracy in detecting non-compliant riders during stationary conditions at intersections. Its adaptability to diverse environmental factors, coupled with real-time processing capabilities, underscores its practicality for comprehensive road safety measures. The secure database storage of the system ensures meticulous record-keeping, facilitating detailed documentation and reporting. By emphasizing the importance of helmet usage and vigilantly monitoring non compliance, this innovative technology significantly contributes to fortifying safety standards on the roads. The ultimate goal is to create a safer environment for all road users, mitigating risks associated with riding without helmets and promoting adherence to critical safety regulations. As a vigilant guardian of road safety, this system provides a valuable tool for authorities to enforce and monitor helmet usage, fostering a culture of responsible and safe riding.

REFERENCES

[1]. Y. Li, H. Wei, Z. Han, J. Huang, and W. Wang, “Deep learning-based safety helmet detection in engineering management based on convolutional neural networks”, Adv. Civil Eng., vol. 2020, pp. 1_10, Sep. 2020.
[2]. R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 580_587.
[3]. S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards realtime object detection with region proposal networks,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137_1149, Jun. 2017.
[4]. Y. Gu, S. Xu, Y. Wang, and L. Shi, “An advanced deep learning approach for safety helmet wearing detection," in Proc. Int. Conf. Internet Things (iThings), IEEE Green Comput. Commun. (GreenCom), IEEE Cyber, Phys.Social Comput. (CPSCom), IEEE Smart Data (SmartData), Jul. 2019, pp. 669_674.
[5]. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in Proc. IEEE Conf. Comput. Vis.Pattern Recognit. (CVPR), Jul. 2017, pp. 2261_2269.
[6]. N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Jun. 2005, vol. 1, no. 1, pp. 886_893.
[7]. K. Han, J. Peng, Q. Yang, and W. Tian, “An end-to-end dehazing Siamese region proposal network for high robustness object tracking," IEEE Access, vol. 9, pp. 91983_91994, 2021.
[8]. M. Li, Z. Zhang, L. Lei, X.Wang, and X. Guo, “Agricultural greenhouses detection in high-resolution satellite images based on convolutional neural networks: Comparison of faster R-CNN, YOLOv3 and SSD," Sensors, vol. 20, no. 17, p. 4938, Aug. 2020.
[9]. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ``You only look once: Uni_ed, realtime object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 779_788.
[10]. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W.Wang, T. Weyand, M. Andreetto, and H. Adam, “MobileNets: Ef_cient convolutional neural networks for mobile vision applications,” 2017, arXiv:1704.04861.