Comparative analysis of optimizers and transfer learning methods for headgear detection using convolutional neural network

Sri Sai Teja M S
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
srisaitejms2002@gmail.com

Ullas P
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
ullasputtaraju@gmail.com

Vikas K R
Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
vikaskr8055@gmail.com

Kumar Swamy S
Professor & Head, Department of Computer Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
Skswamy@gat.ac.in

Ashwini Kodipalli
Associate Professor & Head, Department of Artificial Intelligence and Data Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
dr.ashwini.k@gat.ac.in

Trupthi Rao
Assistant Professor, Department of Artificial Intelligence and Data Science and Engineering
Global Academy of Technology
Bangalore, Karnataka, India
trupthirao@gat.ac.in

***Abstract*— In untapped nations, cars are quite uncommon because motorcycles have always dominated the transportation industry. Over the past few years, there has been an increase in motorbike accidents. Motorcyclists who do not use fluorescent helmets because they don't think they offer enough protection are frequently engaged in traffic accidents. Today, it is against the law to operate a motorbike without a helmet, which increases accidents and fatalities in India. Every motorcyclist must wear a helmet when operating a motorcycle by law for the protection of the public. However, the majority of people continue to disprove this by driving without wearing helmets. In order to solve this problem, we provide a method in this work that uses real-time surveillance footage to automatically identify motorcycle riders without helmets and identify license plates. When the traffic police observe motorbike riders riding without helmets in a group at intersections, they also use CCTV footage to seize control of the drivers of those cars and fine the helmet-less riders. But only human initiative and dedication can make it happen. Because of its great accuracy, CNNs are utilized for picture categorization and recognition. The CNN uses a hierarchical model that builds a network in the shape of a funnel and eventually produces a fully connected layer where all the neurons are connected to one another and the output is processed.**

Keywords—Helmet Detection, Traffic Surveillance, Deep Learning, Convolutional Neural Network.

I. INTRODUCTION

Among the main factors that contributes to human fatalities is traffic accidents. Due to the fact that most motorcyclists do not wear helmets, there is a rapid growth in motorbike accidents, making it a constant hazard. The majority of accidents recent years have seen the introduction on by head injuries. As a result, traffic laws make wearing a helmet mandatory. However, the majority of motorcyclists never abide by the law. A surveillance network is used in several places to keep an eye on motorcyclists who disobey helmet regulations. However, such a system will require human input.

According to current polls, human interventions are inefficient due to an increase in monitoring time as well as human monitoring errors. There are various ways to identify motorcyclists who do not wear helmets.

Due to occlusion, illumination, poor video quality, and other factors, It is challenging to identify the real percentage of motorcyclists who do not wear helmets. According to current polls, human interventions are inefficient due to an increase in monitoring time as well as human monitoring errors. There are various ways to identify motorcyclists who do not wear helmets. Here are a few methods that helmet detection can benefit society: Injury prevention, increased awareness of traffic safety, behavioural modification, legal enforcement, and data gathering for research. Numerous studies have examined the effects of motorcycle helmet rules on rates of helmet use, head injuries, and fatalities. Helmet wearing rates have been reported to rise to 90% or higher when mandatory helmet regulations are enforced; when such laws are abolished, wearing rates often decline to less than 60%.

Our project's main aim is to ensure the safety of motorbike riders. The detection of drivers without helmets. Other advantages include the capacity to motor vehicle department to lower accident rates, the potential of software to identify and punish defaulters in real time, the enforcement of law and order in society. Motorcycle riders must take additional steps to protect their bodies in addition to wearing helmets, which are designed to lessen the effect of a force or contact to the head in the event of an accident. Every motorcyclist is required by law to wear a helmet when operating a motorcycle.

The likelihood of survival is increased for those who wear helmets. We provide a technique for real-time, automatic detection of bike riders without helmets utilising feed from current security cameras in consideration of these difficulties and desired features. The remainder of this essay is structured as follows:

Section II evaluates the connected works, highlighting their advantages and disadvantages. Section III presents the suggested strategy. All the experimental information, findings, and analyses are presented in Section IV. The paper is summed up in the final portion.

II. Existing work

Methods based on deep learning have a lot of potential for identifying risky behaviour in humans. An answer to this problem has been offered by numerous earlier investigations. Studies that are noteworthy include the following: The results showed that the model can precisely detect safe and unsafe actions carried out by workers on-site. Ding et al. [1] developed a hybrid deep learning model that integrates a convolution neural network (CNN) and long short-term memory (LSTM) that automatically recognises workers' unsafe actions.

However, due to a lack of data, the tiny training sample size, and the few risky behaviors that were taken into account, some behaviors cannot be recognized.

A unique deep learning-based system was proposed by Fang et al. [2] to determine whether a site worker is adhering to the conditions of their certification. The framework incorporates worker competency assessment, trade recognition, and crucial video clip extraction. Results show that the suggested framework provides an efficient and workable way to identify noncertified activity. However, some employees cannot be identified when their faces are barely visible or blocked by safety helmets or other gear. Additionally, the worker who was right in front of the camera went unnoticed.

The broad topic of anomaly detection in surveillance footage includes the automatic identification of bike riders without helmets. Effective automatic surveillance systems typically involve the duties of environment modelling, item detection, tracking, and categorization [3, as stated]. Chiverton presented a method in [4] that makes use of the geometrical contours of the helmet and illumination variation at various points on the helmet. It makes use of the Hough transform-based circle arc detection method. The main drawback of this method is that it tries to locate the helmet in the entire frame, which is computationally expensive, and that it frequently misidentifies other items with similar shapes as helmets. Additionally, it ignores the reality that a helmet is only necessary for bike riders.

Convolutional neural network (CNN) technology was used by Boonsirisumpun et al. [5] to identify bikers riding without a helmet. Utilizing cameras, the input was captured. For training, 493 photos from the dataset were used.
Four CNN-based models were utilized by the system: Google Net, MobileNet, VGG19, and VGG16. The accuracy rating given by MobileNet was the highest (85.19%). Based on a deep learning technique, Raj et al.'s work [6] assisted in identifying bikers who had disregarded the helmet wearing regulations. Utilizing HOG, the work of motorbike detection is completed after choosing the area of interest. They used CNN technology to recognize licence plate numbers and spot bicyclists riding without helmets. It has been utilized the self-generated dataset from many sources. They cited a 94.70% accuracy rate.

To identify bicyclists riding without a helmet, Wu et al. [7] utilized the YOLOv3 and YOLO-dense models. They gathered data from the Internet and two self-generated sources. According to the experimental findings, they were able to reach 95.15% mAP for YOLOv3 and 97.59% for the YOLO-dense model. CNN was used by Mistry et al. [8] to find bikers riding without a helmet. In two levels, YOLOv2 was utilised. First, the system employed YOLOv2 to identify various objects and helmetless riders. For training, the COCO dataset has been utilised. The experimental result shows a 92.87% accuracy. Faster R-CNN was employed by Afzal et al. [9] to identify bikers who were not wearing helmets. A dataset that was created by the system itself was used to train it. The trial findings indicated a 97.26% accuracy.

A technique for identifying motorbike riders who are not wearing helmets was introduced by Kharade et al. [10] and is based on deep learning algorithms and the YOLOv4 model. Comparing the suggested model to the present CNN-based algorithms, the proposed model shows actual performance in traffic motion photos.

III. Proposed work

It is advisable to go with a new CNN as the current CNN is unable to control the difficulties in real-world helmet wear analysis, such as the detection of motorcycle helmets. This is due to the need to find a reliable and accurate model for classification.

A. Datasets Used

The 764 photos in the helmet detection dataset that is accessible on Kaggle have bounding box annotations in the PASCAL VOC format. For the purposes of helmet detection, the dataset focuses on two different classes:

Those of helmet-wearing riders and those of who did not. Images of people wearing helmets are represented by the class "with helmet".

Images of people without helmets are represented by the class "without helmet." The complete range of dataset can be seen in the below Figure 1:

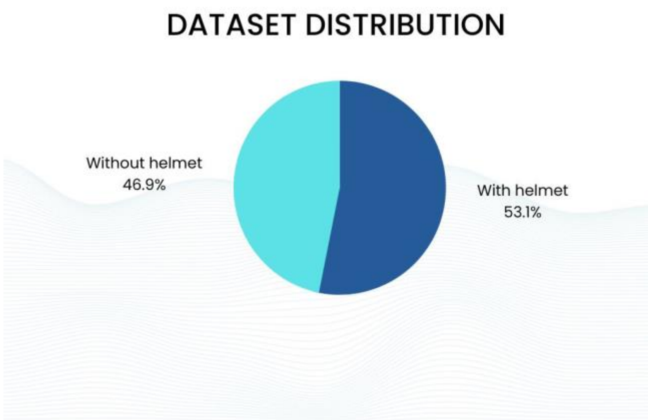

Fig.1. Total Distribution of images in the dataset

A training set and a testing set have been created from the dataset. The allocation of the photos is done in an 80:20 ratio, with roughly 80% going to the training set and the remaining 20% going to the testing set. The pie chart representation of each dataset elements are shown below in the Figure 2:

Fig.2. Percentage distribution of the images for two different sets

The training dataset contains 606 images, including both folders, and the other folder contains 158 for the testing datasets and the sample set of images can be seen in Figure 3 below:

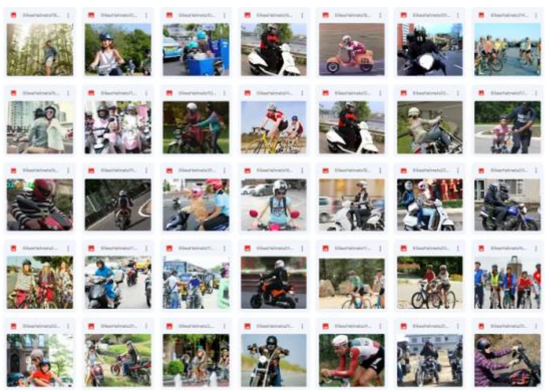

Fig. 3. Sample Images available in the dataset

B. Flowchart

The suggested method is based on a CNN deep learning model, which uses photos as an input to detect helmet violations and take appropriate action against offenders of traffic laws. The suggested system carries out several tasks sequentially. In addition, it divides riders into two groups: "With Helmet" and "Without Helmet." The elements of the suggested technique are separately described in the sections that follow:

3.1. Data Acquisition: Data acquisition is the first part of the suggested technique. To do this, an appropriate dataset for headgear detection must be gathered. A significant number of photos divided into the categories "helmet wearing" and "helmet not wearing" should make up the dataset. The pictures must to be varied and realistic of actual situations. The dataset can be obtained using a variety of methods, including gathering photographs from the internet, taking pictures using sensors or cameras, or using headgear detection databases that are already publicly available.

3.2. Pre-processing: Following the acquisition of the dataset, this phase is carried out. This entails carrying out the steps required to get the data ready for the CNN training. Among the pre-processing steps are:

Image resizing: In this instance, to guarantee uniformity in input dimensions, the photos are shrunk to a fixed size of (64, 64, 3). Normalisation is the process of setting an image's pixel values to a range between 0 and 1. This process aids in training convergence that occurs more quickly.

Data augmentation: Techniques for enhancing data can be used to expand the dataset and provide variability. The training data can be enhanced using methods such as rotation, flipping, zooming, and introducing noise to assist the model be more general.

C. Proposed CNN Model:

The proposed method for headgear detection relies heavily on the CNN architecture. The architecture is made up of interconnected layers that use input photos to learn hierarchical properties. Following is a description of the CNN's architecture shown in Figure 6:

Input layer: The input layer is where the images with the dimensions (64, 64, 3) are sent. (64, 64) refers to the image's width and height, while (3) denotes the number of RGB colour channels.

Convolutional Layer: To find regional patterns and characteristics, the convolutional layer applies a series of filters to the input images. A single convolutional layer with 32 filters is utilised in this architecture. The number of filters chosen depends on the difficulty of the task and the resources for computation that are available.

Pooling Layer: By downsampling the input, the pooling layer decreases the input's spatial dimensions. This architecture employs maxpooling with a pooling size of (4, 4).

Flatten layer: A 1-dimensional vector is created from the output of the preceding layer by the flatten layer. The multi-dimensional feature maps are transformed into a linear array and fed into the fully connected layers.

Fully Connected Layer (Hidden Layer): High-level representations of the input features are learned by the fully connected layer (also known as the hidden layer). A hidden layer with 156 neurons is utilised in this architecture. Rectified Linear Unit is the activation function utilised for the hidden layer. Non-linearity is introduced by ReLU, which aids in the learning of intricate correlations between features.

Output Layer: The CNN's final layer, known as the output layer, generates the desired output. In this instance, there is only one neuron with the sigmoid activation function in the output layer. The likelihood that the input image belongs to the "wearing helmet" class ranges from 0 to 1, according to the sigmoid activation function. The final prediction can be determined by setting a threshold.

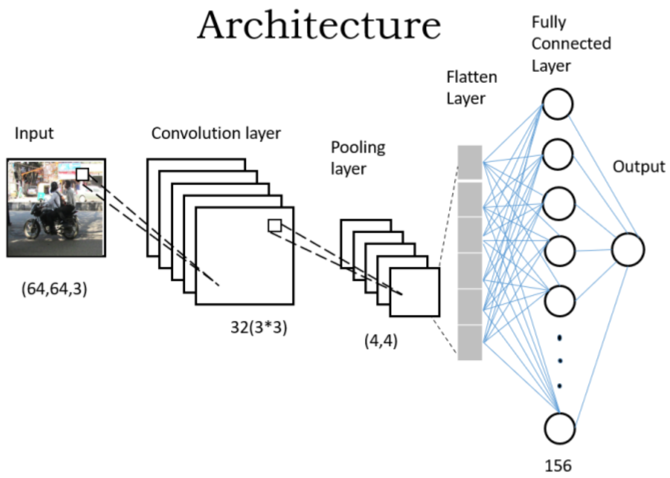

Fig. 6 CNN Architecture for the proposed model.

The fundamental components, the Conv layer and the pooling layer, are able to extract object features. The Conv layers can extract and amplify the object features. The pooling layers allow for the filtration of numerous features, the removal of extraneous features, and the compression of features. The activation layers successfully tackle nonlinear problems by using nonlinear activation functions to improve the neural network models' capacity for expression. The feature values are output by the FC layers, which combine the data features of objects. By doing this, the CNNs can layer by layer convert the original input images' pixel values to the final classification confidence.

Convolutional neural networks (CNNs) are being used in the project to do a comparative examination of optimizers and transfer learning techniques for headgear identification. The study intends to increase the precision of headgear detection models through the investigation of various optimizers and transfer learning approaches.

Optimizers: Deep learning model training, optimizers are essential. They assist in modifying the model's weights and biases in accordance with the inclinations of the loss function.

Adam (Adaptive Moment Estimation) is an approach for optimising the adaptive learning rate that integrates concepts from AdaGrad and RMSProp. According to the average of previous gradients and their squares, it modifies the learning rate for each parameter. The Adam's update rule's formula is as follows:

$$m\_t = \beta 1 * m\_\{t-1\} + (1 - \beta 1) * g\_t$$

$$v\_t = \beta 2 * v\_\{t-1\} + (1 - \beta 2) * (g\_t^2)$$

$$m\_t\_hat = m\_t / (1 - \beta 1^t)$$

$$v\_t\_hat = v\_t / (1 - \beta 2^t)$$

$$\theta\_\{t+1\} = \theta\_t - lr * m\_t\_hat / (\surd (v\_t\_hat) + \varepsilon)$$

In the equations above, t is the time step, t stands for the current model parameters, lr is the learning rate, and is a small constant to prevent division by zero. m t and v_t are the first and second moment estimates of the gradients, respectively, while 1 and 2 are the decay rates.

Transfer learning:

A pre-trained model on a sizable dataset is utilised as a starting point for a new task or dataset in the transfer learning technique. Transfer learning utilises the information and learnt features of a pre-trained model, which are then fine-tuned on the target dataset, rather than training a CNN from start. This method can assist in overcoming the limits of sparse training data, which is especially helpful when the target dataset is tiny.

As a way to enhance the model's ability to learn complicated patterns, more layers, such as multiple convolutional layers and extra hidden layers, can be added. The architecture shown above is a simplified representation. The suggested method makes use of this CNN architecture to train a headgear detection model, with the output indicating whether or not a person is wearing a helmet in an image

IV. Experiments and Results

The experiments involve the subsequent steps:

1) Evaluation of Optimizers in Comparison

This experiment compares the effectiveness of various optimizers for CNN-based helmet identification. Several optimizers, including Adam, SGD, RMSprop, Adadelta, and Adagrad, are used to train the CNN model. The learning rate, batch size, and 50 iterations are some of the hyperparameters that are specific to each optimizer. Following that, the models are assessed using the testing dataset to gauge their effectiveness in terms of testing accuracy and testing loss.

2) Dropout Research

The performance of the headgear detection model is being examined in this experiment in order to determine the effect of dropout regularisation. The CNN model is trained both with and without dropout regularisation using the optimizer found to be the most efficient in Comparative Analysis of Optimizers. The accuracy of the models' performances is contrasted.

3) Analysis of Transfer Learning

In this experiment, pre-trained CNN models for headgear detection are used in conjunction with transfer learning approaches. Particularly, pre-trained models are applied to well-known architectures as AlexNet, ResNet50, and VGG16. These models are polished using the headgear detection dataset, and the testing dataset is used to gauge how well they work. Accuracy is one of the parameters for evaluation.

4) Integrated Analysis

This experiment attempts to assess the performance of a combined strategy by building on the findings from Comparative Analysis of Optimizers and Transfer Learning Analysis. The best-performing architecture from the transfer learning analysis is used to implement transfer learning once the CNN model has been trained using the best optimizer found in the comparative analysis of optimizers.

5)Robustness Evaluation

Additional tests are run in this step to assess the models' durability. The CNN design has several hyperparameters, such as learning rate, batch size, and number of layers. On many dataset modifications, including various lighting conditions, angles, and resolutions, the models' performance is assessed.

A. Results and Discussion

The findings of the comparative evaluation of five alternative optimizers for headgear detection using a CNN as displayed in Table 1. Two sets of accuracy values—one with dropout regularisation and the other without—are included in the table. According to these findings, RMSprop outperforms the other investigated optimizers, obtaining excellent accuracy in both situations. While SGD and Adadelta display relatively lower accuracy, Adam and Adagrad both work admirably.

TABLE 1. EXPERIMENTAL RESULTS FOR OPTIMIZERS

OPTIMIZERS	WITH DROPOUT ACCURACY	WITHOUT DROPOUT ACCURACY
Adam	82.85	88.57
SGD	81.43	61.42
RMSprop	81.42	84.28
Adadelta	57.4	57.14
Adagrad	75.71	75.71

The below bar chart represents the testing accuracy of the model using various optimizers comparing by using with dropouts and without using dropouts in Figure 7:

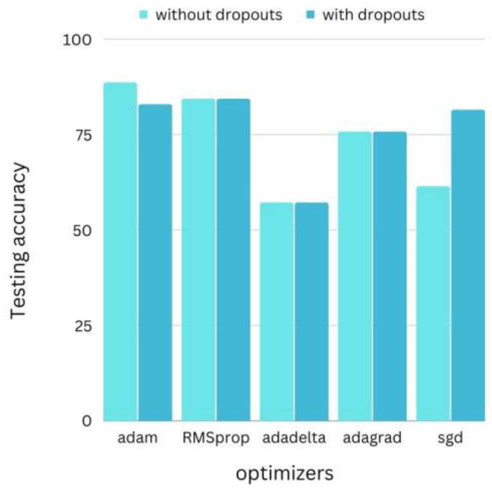

Fig. 7 Bar chart representation of the accuracy with and without dropouts

TABLE 2. EXPERIMENTAL RESULTS FOR DIFFERENT TRANSFER LEARNING MODELS

TRANSFER LEARNING	TESTING ACCURACY	TESTING LOSS
AlexNet	80.00	44.33
ResNet50	74.28	46.12
VGG16	42.85	70.77
Proposed CNN Model	88.57	27.81

According to these findings, our Proposed CNN model outperforms the other transfer learning models that were put to the test, reaching the highest level of accuracy of 88.57%. This can be seen in the below Figure 8 where there is bar chart representation of accuracy of different transfer learning models.

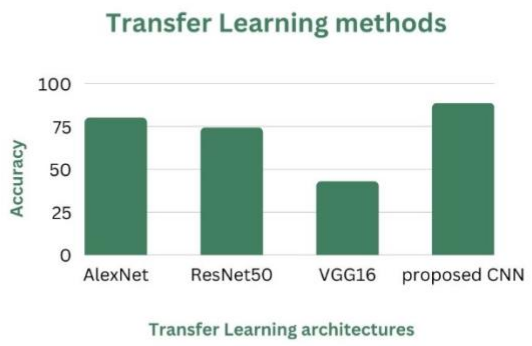

Fig. 8 Bar chart representation of the accuracy of various transfer learning models

How well the trained model performs on unobserved data can be inferred from the validation loss vs. validation accuracy graph. The validation loss is often biggest at the start of training since the model has not yet developed strong generalisation skills.

It's crucial to remember that this particular experiment uses 50 epochs, with 20 steps per epoch. This is shown in the below Figure 9:

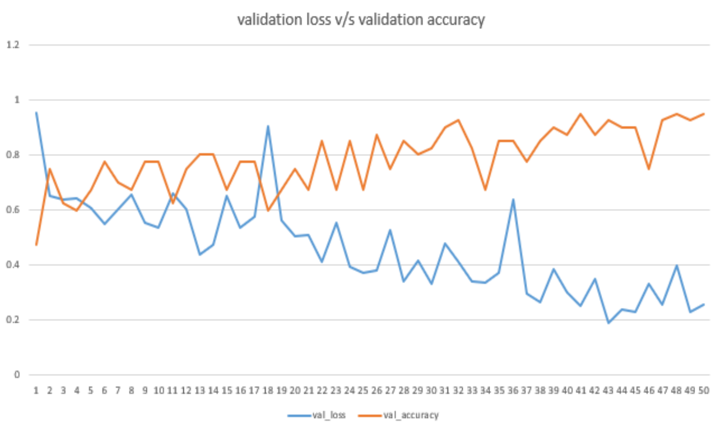

Fig. 9. Validation Loss v/s Validation Accuracy

The validation accuracy also rises linearly with time after each epoch. This suggests that the model learns to extract pertinent information and generate increasingly precise predictions as its performance on the validation dataset advances steadily.

To improve generalization and performance, it reveals if the model is overfitting or underfitting, enabling for changes to be made to the hyperparameters or the training procedure. The validation loss vs. validation accuracy graph aids in evaluating the trained model's performance on unobserved data and directs additional optimisation efforts.

V. Conculsion

In this research, we suggest a framework for the real-time detection of riders who disobey traffic laws by riding without a helmet. The proposed framework will help the traffic police find these offenders in unusual weather, such as a hot sun, etc.

Experimental results show that the detection of bike riders and lawbreakers has the highest accuracy, 88.57%. Additionally, the suggested architecture automatically adjusts to new cases as needed with minor tweaking. This system can be improved to track down and report violators' licence plates. Increasing the training data set and image quality will increase accuracy.

In this study, we looked into a useful and cutting-edge technique for detecting motorcycle riders wearing safety helmets, which can track riders' helmet use in real time. The current and next work will concentrate on enhancing the speed and accuracy of safety helmet wear detection. Sometimes the classifier may produce inaccurate results due to variations in light intensity.

References

[1] Ding, L., Fang, W., Luo, H., Love, P. E., Zhong, B., & Ouyang, X. (2018). A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Automation in construction*, *86*, 118-124.

[2] Fang, Q., Li, H., Luo, X., Ding, L., Rose, T. M., An, W., & Yu, Y. (2018). A deep learning-based method for detecting non-certified work on construction sites. *Advanced Engineering Informatics*, *35*, 56-68.

[3] Hu, W., Tan, T., Wang, L., & Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *34*(3), 334-352.

[4] Chiverton, J. (2012). Helmet presence classification with motorcycle detection and tracking. *IET Intelligent Transport Systems*, *6*(3), 259-269.

[5] Boonsirisumpun, N., Puarungroj, W., & Wairotchanaphuttha, P. (2018, November). Automatic detector for bikers with no helmet using deep learning. In *2018 22nd International Computer Science and Engineering Conference (ICSEC)* (pp. 1-4). IEEE.

[6] Raj, K. D., Chairat, A., Timtong, V., Dailey, M. N., & Ekpanyapong, M. (2018, January). Helmet violation processing using deep learning. In *2018 International Workshop on Advanced Image Technology (IWAIT)* (pp. 1-4). IEEE.

[7] Wu, F., Jin, G., Gao, M., Zhiwei, H. E., & Yang, Y. (2019, May). Helmet detection based on improved YOLO V3 deep model. In *2019 IEEE 16th International conference on networking, sensing and control (ICNSC)* (pp. 363-368). IEEE.

[8] Mistry, J., Misraa, A. K., Agarwal, M., Vyas, A., Chudasama, V. M., & Upla, K. P. (2017, November). An automatic detection of helmeted and non-helmeted motorcyclist with license plate extraction using convolutional neural network. In *2017 seventh international conference on image processing theory, tools and applications (IPTA)* (pp. 1-6). IEEE.

[9] Afzal, A., Draz, H. U., Khan, M. Z., & Khan, M. U. G. (2021, April). Automatic helmet violation detection of motorcyclists from surveillance videos using deep learning approaches of computer vision. In *2021 International Conference on Artificial Intelligence (ICAI)* (pp. 252-257). IEEE.

[10] Kharade, N., Mane, S., Raghav, J., Alle, N., Khatavkar, A., & Navale, G. (2021, September). Deep-learning based helmet violation detection system. In *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)* (pp. 1-4). IEEE.

[11] Dahiya, K., Singh, D., & Mohan, C. K. (2016, July). Automatic detection of bike-riders without helmet using surveillance videos in real-time. In *2016 International Joint Conference on Neural Networks (IJCNN)* (pp. 3046-3051). IEEE.

[12] Silva, R., Aires, K., Veras, R., Santos, T., Lima, K., & Soares, A. (2013). Automatic motorcycle detection on public roads. *CLEI Electronic Journal*, *16*(3), 4-4.

[13] Chiu, C. C., Ku, M. Y., & Chen, H. T. (2007, June). Motorcycle detection and tracking system with occlusion segmentation. In *Eighth International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'07)* (pp. 32-32). IEEE.

[14] Chiverton, J. (2012). Helmet presence classification with motorcycle detection and tracking. *IET Intelligent Transport Systems*, *6*(3), 259-269.

[15] Li, Y., Wei, H., Han, Z., Huang, J., & Wang, W. (2020). Deep learning-based safety helmet detection in engineering management based on convolutional neural networks. *Advances in Civil Engineering*, *2020*, 1-10.

[16] Waris, T., Asif, M., Ahmad, M. B., Mahmood, T., Zafar, S., Shah, M., & Ayaz, A. (2022). CNN-Based Automatic Helmet Violation Detection of Motorcyclists for an Intelligent Transportation System. *Mathematical Problems in Engineering*, *2022*.

[17] Vishnu, C., Singh, D., Mohan, C. K., & Babu, S. (2017, May). Detection of motorcyclists without helmet in videos using convolutional neural network. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 3036-3041). IEEE.

[18] Peter, A. V., Lakshmi, C. S., Prathyusha, V. N. L., Kumar, P. A., & Manikanta, C. H. (2023, May). Helmet detection and license plate recognition. In *AIP Conference Proceedings* (Vol. 2492, No. 1). AIP Publishing.

[19] Balpande, V. P., Parashar, K., Sakharkar, Y., Shingade, K., & Pankhaniya, A. HELMET DETECTION AND LICENSE PLATE RECOGNITION FOR TRAFFIC SURVEILLANCE.

[20] Shreenithi, S., & Ruby, J. AUTOMATED HELMET DETECTION FOR MULTIPLE MOTORCYCLE RIDERS USING CNN.

[21] Li, J., Liu, H., Wang, T., Jiang, M., Wang, S., Li, K., & Zhao, X. (2017, February). Safety helmet wearing detection based on image processing and machine learning. In *2017 Ninth international conference on advanced computational intelligence (ICACI)* (pp. 201-205). IEEE.

[22] Tran, D. N. N., Pham, L. H., Jeon, H. J., Nguyen, H. H., Jeon, H. M., Tran, T. H. P., & Jeon, J. W. (2023). Robust automatic motorcycle helmet violation detection for an intelligent transportation system. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5340-5348).

[23] Silva, R. R. V. E., Aires, K. R., & Veras, R. D. M. (2018). Detection of helmets on motorcyclists. *Multimedia Tools and Applications*, *77*, 5659-5683.