ISSN 2394-739X

IJRSET FEBRUARY Volume 12 Issue 2

International Journal for Research in Science Engineering & Technology (IJRSET)

https://www.doi.org/10.5281/zenodo.14810392

Vehicle Detection Network using CNN and instance Segmentation

[1] Priyanka, [2] Mr. Kameshwar Rao
[1, 2] Assistant Professor,
[1, 2] Shri Rawatpura Sarkar University Raipur (C.G.), 492015, India.

Abstract:- These days, when an accident happens, people either hide from the emergency services or cause a lot of disturbance while reporting it, or the accident goes unreported and by the time help arrives, it's too late. A comprehensive system has been developed to actively detect all types of traffic accidents and notify the appropriate parties. In the event of an accident, this includes the closest police station, hospitals, general ambulances, the registered owner of the vehicle involved in the accident, and their emergency contacts. In the event of a hit-and-run, the police can obtain the vehicle number of the accused vehicle. Numerous applications exist for vehicle detection in aerial pictures, and most vehicle detection techniques employ the bounding-box approach for localization. Using a private dataset, a convolutional neural network is proposed in this letter to detect cars. Based on experimental data, CNN performs better, with an accuracy rate of 94.54%.

Keywords: [Convolutional Neural Network (CNN), Data Preprocessing, Deep Learning (DL), Feature Engineering (FE), Vehicle Detection (VD).]

1. INTRODUCTION

The number of motor vehicles on the road has significantly increased due to the recent decades' population boom. Because more cars and people are using the roads, urban traffic has become a significant issue for growing populations. [1] [2]. Due to the economy's rapid growth and the continuous improvement in people's quality of life, which has led to an increase in the number of families owning private cars, traffic management faces significant challenges [3]. Efficient surveillance is essential for alleviating the persistently growing traffic issues. An essential technology for an autonomous vehicle to sense its surroundings is object detection and applications for object detection in various fields include surveillance, automatic emergency braking systems, and other applications. Even though object detection research is very important and popular, it is also one of the hardest tasks in computer vision. The research field in vehicle detection for intelligent traffic surveillance systems is essential to modern computer vision [4]. In addition to identifying objects that differ considerably from one another, there is an increasing need to recognize the kind of comparable objects [5]. There are several restrictions on collecting visual data from stationary cameras, which are typically placed at crossings. These restrictions depend on the coverage area, camera angles, etc. Additional ad-hoc drone-based traffic surveying devices can be very helpful in a variety of scenarios to improve the real-time traffic statistics data along the traffic area in the absence of fixed cameras. Drone operator overflies area of interest; aerial view video data is transferred to a cloud-based machine learning server for analysis and extraction of statistical data, including total number of cars, direction, and congestion for both general traffic and vehicle type-specific traffic [6].

There are numerous Deep Learning approaches that can be used to generate situations that are suitable for our solutions. Detection, tracking, and trajectory processing are all integrated into the computer vision-based speed detection and vehicle counting approach [7]. A vision-based traffic monitoring and surveillance system consists of three primary steps: tracking, categorization, and vehicle detection. The great majority of vehicle identification techniques use picture difference to recognize moving objects from a fixed, predetermined camera. However, tracking systems play a major role in target identification and vision-based surveillance. The combined objective of tracking and vehicle identification has so advanced to include extensive monitoring. The main goal of vision-based object tracking is to locate and identify an object in a series of consecutive frames [8] [9].

2. LITERATURE REVIEW

Many studies have been done on this subject, and some of the literatures that have been produced include:

Using a transfer learning strategy, ApoorvaOjha et al. [10] provide a state-of-the-art implementation of Mask R-CNN for vehicle detection via instance-wise segmentation, which simultaneously generates bounding box and object mask. A segmentation-based technique is used because autonomous systems require accurate and faultless vehicle recognition. The Tensorflow and Keras frameworks, along with an online GPU and cloud services from Google Colab, were used to complete this study. Using a combination of benchmark datasets yields a mAP of 90.27% and mAR of 92.38%.

Fulian Li et al. [11] state that YOLOv5, a one stage target identification technique, satisfies the auto drive system's vehicle recognition requirements due to its high precision and quick detection speed. In order to provide a baseline model that will be useful for future auto drive system research, this article builds a custom dataset and educates model networks with varying depth and width using the YOLOv5 framework.

Using on the YOLOv5 algorithm, Alavikunhu Panthakkan et al. [12] devised a vehicle recognition approach using UAV aerial photos. This technique trains YOLO using the VEDAI dataset, which is accessible to the public. With a 93 percent hit rate, 96 percent accuracy, and 98 percent precision, the suggested approach performs well. The study's findings unequivocally demonstrate that the technique performs better than the current techniques.

In order to introduce the SRBD-CDSO algorithm, R. Anitha et al. [5] propose our model of dataset design (C-DSO) and combine the YOLO (You Only Look Once) V3 algorithm with SRBD.

This method detects the clever way of integrating the pertaining of the image, labeling of the vehicle image, image smoothing, and annotation.

Additionally, the dataset has been updated, new vehicle picture categories have been included, and the experimental findings demonstrate an improved to identification rate and decreased validation loss in the dataset.

In order to determine which Yolo version is best for vehicle detection and counting, author Ayush Dodia et.al [3] implemented three different versions of the Yolo algorithm: YOLO-v3, YOLO-v5, and YOLO-v7. The results showed that YOLO- v7 had the highest accuracy for vehicle detection and counting, at 93.67%. The newest iteration of the YOLO algorithm is called YOLO v7. This survey is beneficial in determining which algorithm is best for a vehicle detection system.

Shuiqiang Zhang et al. [13] present the Night vision GAN method to improve vehicle features at night by improving the feature information of the vehicle and the contrast between the vehicle and background. The issue of missing tiny targets is resolved and the average detection accuracy is increased by 12.05% after cascading with target detection algorithms. Pruning and knowledge distillation techniques are combined to efficiently minimize the model' s computation and parameter size, doubling the detection speed at a sacrifice of 2.04% in AP. This solves the problem of sluggish detection speed.

In aerial photography situations, Jiaquan Shen et al. [14] have developed an anchor-free multilayer attentional depth aerial vehicle object detection method that is capable of efficiently identifying small targets of vehicles. Two attention mechanism models the convolution alattention technique and the triplet attention technique are constructed on the feature extraction network for the purpose of feature fusion. In the Munich dataset and the created dataset, the proposed detection algorithm achieves 89.1% and 92.6% mAP, respectively, and its detection times for each image are 1.21s and 0.036s, correspondingly. This suggested detection algorithm is evaluated on both public aerial dataset and our gathered aerial dataset.

By employing stacked pictures from wavelet transform as input to the FCNN, Monika To desk's et.al [15]suggested the method resolves this problem and produces fewer FLOPs (floating-point operations) and faster inference times compared to conventional CNN-based techniques. Furthermore, employing a dataset of about 150.000 photos of both small and large vehicles to train the neural network yields encouraging accuracy rates for the suggested approach. The photos were taken from video records that were acquired from security camera systems that were put up on city streets and highways. All things considered, this method is a major advancement towards the creation of real-time object detection systems.

For vehicle recognition in aerial images, JiaquanShen et al. [16] developed a lightweight backbone network with a context information module and an attention mechanism module. This design enables the feature extraction network to make better use of prominent regions and contextual information additionally, to predict the bounding box in the detection model, we employ adaptive anchor-free. On the German Aerospace Centre (DLR) 3K dataset and the generated dataset, the suggested detection algorithm achieves 89.7% and 94.1% mean average precision (mAP), while the detection times for each image are 1.66 and 0.049 s, respectively.

Thesuggestedmodelby Mrs .Jayashree Metal. [17] shown excellent accuracy and efficiency in identifying a legitimate accident from live CCTV footage by utilizing TensorFlow, Canny edge detection, and Convolutional Neural Networks (CNN).Then, to prevent chaos, notify all surrounding emergency services of the accident and provide them with the necessary information to respond promptly. Additionally, inform the emergency contacts of the registered vehicles involved in the incident first responders the information they need to locate a car in the event of a hit-and-run, enabling them to take fast action.

A Fully Convolutional Lightweight Pyramid Network (FCLPN) is proposed by Qingsong Du et al. [18] to identify cars in visible-spectrum aerial photos. In contrast to the bounding-box method, FCLPN carries out localization and classification at the pixel level. To verify the generalization strengthofFCLPN,whichwastrainedontheDLR-3Kdataset, it is directly evaluated on the DLR-3K, VEDAI, COWC, The VAID dataset and experimental results show that FCLPN outperforms state-of-the-art techniques for aerial vehicle detection in terms of precision, F1 score, and mAP, achieving a precision value of 95.75.

Deep learning is used by Prasoon Bharat Mishra et al. [19] to assess research on object recognition. The study begins with an overview of deep learning and convolution neural networks (CNNs), which are among its main techniques. Next the focus will be on several popular generic object detection techniques, along with useful enhancements and approaches for enhancing detection overall. Next, some frequent genealogy patterns for object recognition will be discussed, along with some adjustments and useful techniques to improve detection, greater efficiency additionally, because they exhibit a variety of attributes, it briefly examines several specific detection tasks, including face detection, pedestrian identification, and the precognition of noteworthy objects. There are also experimental analyses available to evaluate other approaches and obtain some valuable insights. There are many different potential directions and goals covered by the recommendations for additional research in both the object and object-oriented disciplines.

ISSN 2394-739X
IJRSET FEBRUARY Volume 12 Issue 2

3. METHODOLOGY

DataCollection:We employ a camera to record a video in the vicinity of the local traffic signal, and Mat lab is used to extract pictures from the video. Once the photos have been extracted from the movie, we save them in a different folder and give them the following labels: 0 for rikshaws1 for automobiles, 2 for bikes, 4 for trucks and buses, and 3 for non-vehicles. With regard to this instance, the dataset included 951 photographs, of which 127 were labelled as 0, 409 as 1, 195 as 2, 219 as 3, and the remaining images as 4.

2) Data Pre-processing:Data preprocessing is the straightforward conversion of unprocessed data into a format that can be understood. There are times when noisy, redundant, inconsistent, and incomplete data are present in the real world. Data preparation involves a variety of procedures to convert unprocessed data into a format that is processed and useful [20]. Using a camera, we record a video close to the local traffic signal. Matlab is then used to extract the images from the movie. We save the photos in a different folder and name them as follows: 0 for rikshaws, 1 for automobiles, 2 for bikes, 3 for non-vehicles, and 4 for trucks/buses after extracting the images from the film using the cropdata.m file.

Figure.1. Input Images from dataset

Training Data: At the data division stage, the input data set needs to be divided into a test set and a training set. If a classification model has been developed for the test set of data and the classifier's output accurately predicts the test set, then the training set of data ought to be utilized for training. Data splitting creates two collections of data that are mutually exclusive. There is nothing that two pieces of data that are mutually exclusive can have in common. There are contents unique to each data collection. In our sample, the ratio of training to testing data is 8:2.

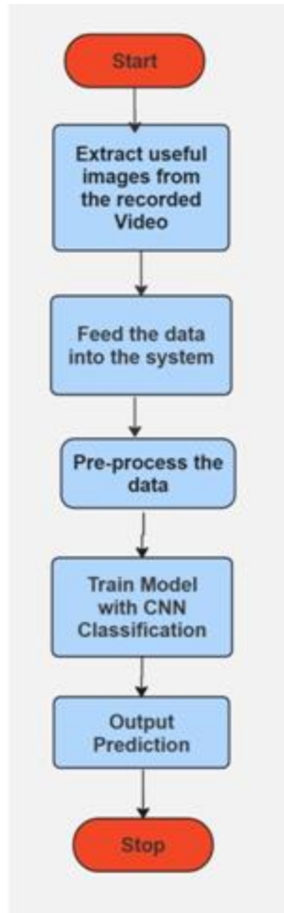

Figure.2. Proposed Work Flow Diagram

4) Proposed System: After removing images from the recoded movie, the model builds its dataset. Subsequently, the pre-processed images are fed into the system to train the CNN classifier model. In reference to the dataset under consideration, it comprises 951 photos, of which 127 were labelled as 0, 409 as 1, 195 as 2, 219 as 3, and the residue as 4. The objective of this research is to use a sliding window technique to detect different vehicle types in a picture by training a CNN model on a variety of vehicle photographs.

5) Applied Algorithm: Rather than using IOT devices inside the car, which can produce false positives, we propose an efficient method of accident detection from CCTV live footage using convolutional neural network (CNN) technology. Tensor flow is used to train a model, which can identify an accident given the correct training dataset. The convolutional neural

network represents the most modern approach to picture classification and recognition. It contains a stack of convolutional, ReLu, and pooling layers. This is thoroughly discussed in [21]. Convolutional layers are used in feature extraction techniques to extract various aspects from images. These features can range from more fundamental ones like edges to more specific ones like objects and colours. By utilizing the most appropriate type of layer pooling, such as average or maximum pooling, the dimensionality is reduced, which lowers the number of parameters to learn and the amount of computation done by the network. The features in a section of the feature map produced by a convolution layer are summarized by the pooling layer. A large image collection, a lot of computing power, and time are required to train such a network [17] [22].Layers such as the input data shape of [100*100,3], the conversion of 64 3x3 filters with ReLU activation, Pooling using a 2x2 filter, Conv: 32 3x3 filters activated by ReLU Pooling: using a 2x2 filter; Conversion: 32 3x3 filters activated by ReLUPooling: using a 2x2 filter Completely Connected: including 256 neurons, ReLU initiation, and a 0.75 likelihood of dropout, Fully Connected: with probability 0.75, 256 neurons, ReLU activation, and dropout Fully Connected output layer: containing a SoftMax classifier and 5 neurons (equivalent to the number of classifications).

4. RESULT ANALYSIS

The suggested technique uses a sliding window approach to identify distinct vehicle kinds in the image by training a CNN model on photos of various automobiles. To collect the data, we use a camera to take a video at the nearby traffic signal. The images from the movie are then extracted using MATLAB. We give the pictures the following names and save them in a different folder: After using the cropdata.m file to extract the images from the film, the numbers correspond to: 0 for rikshaws, 1 for cars, 2 for bikes, 3 for non-vehicles, and 4 for trucks/buses. Finally, use the Adam optimizer to run the CNN model for 20 epox, yielding an accuracy of 0.9454 and a validation accuracy of 0.8586 with a learning date of 0.001. Comparison of existing algorithm with CNN is given below:

1) Cross Validation Analysis: Cross-validation is a statistical approach that is used to evaluate and compare learning systems. It entails dividing the data into two sets: one for training and the other for model validation. The sets used for training and validation must intersect in successive rounds of a conventional cross-validation in order to guarantee that every data point has an opportunity to be validated against. The most basic kind of cross-validation is called k-fold cross-validation. Some cross-validation methods repeat k-fold cross-validation or update the k-fold cross-validation methodology. The data is first split into k segments or folds of the same (or nearly similar) size for k-fold cross-validation. Subsequently, k iterations of training and validation are carried out, with a distinct fold of the data held aside for validation and the remaining k-1 folds used for learning. We used fresh data that had not been used for training to assess our models' performance using the 10-fold cross validation method.

S.No.	Model	Obtain Accuracy
1.	LightweightBackbone Network	94.1%
2.	YOLO-V5	93.67%
3.	Fully Convolutional Light Weight Network	95.75%
4.	Convolutional Neural Network	85.86%

TABLE – I 10-FOLD CROSS VALIDATION ACCURACY OF THE CNN MODEL

Theaverage cross validation score for the employed approaches is displayed in Table II. Using the CNN architecture, the best outcome in 10-fold cross validation was 85.86%.

The above Table-I shows the obtain 10-Fold Cross validation techniques accuracy of proposed CNN network using private dataset.

CONCLUSION

The goal of this effort is to use the CNN network to recognize automobiles on unseen data. So far, the challenge has been to record a minimum of 25 minutes of video so that a minimum of 1000 photos can be cropped from the video. It was also difficult because the CNN network produces varying accuracies while utilizing 5 and 20 epox. On the other hand, we can experiment with different learning rates and raise the epox to get better results (greater accuracy).

REFERENCES

[1]. S. M. John, F. A. Kareem, S. G. Paul, A. G. M, S. Al Mansoori, and A. Panthakkan, "Enhanced yolov7 model for accurate vehicle detection from uav imagery," in 2023 International Conference on Innovations in Engineering and Technology (ICIET), 2023, pp. 1–4.

[2]. W. Yang, "Multi-type vehicle detection algorithm based on improved yolov5," in 2023 3rd International Conference on Frontiers of Electronics, Information and Computation Technologies (ICFEICT), 2023, pp. 599–603.

[3]. A. Dodia and S. Kumar, "A comparison of yolo based vehicle detection algorithms," in 2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1), 2023, pp. 1–6.

[4]. X. Li, G. Tian, Z. Lu, and G. Zhang, "Yolo v5-max: A multi-object detection algorithm in complex scenes," in 2023 IEEE 6th International Conference on Industrial Cyber-Physical Systems (ICPS), 2023, pp. 1–6.

[5]. R. Anitha and P. Prabakaran, "Vehicle detection and classification based on c-dso dataset using yolo v3 with srbd method for intelligent transportation applications," in 2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), 2023, pp. 1–5.

[6]. V. Singh, P. Kaewprapha, and C. Boonmee, "Ad-hoc aerial-view vehicle detection and tracking for real-time traffic monitoring using yolov7," in 2023 International Electrical

ISSN 2394-739X

Engineering Congress (iEECON), 2023, pp. 327–331.
[7]. F. Ahmad, M. Z. Ansari, S. Hamid, and M. Saad, "A computer vision-based vehicle counting and speed detection system," in 2023 International Conference on Recent Advances in Electrical, Electronics Digital Healthcare Technologies (REEDCON), 2023, pp. 487–492.
[8]. P. K. Vishwakarma and N. Jain, "Deep learning-based methods in image analytics for vehicle detection: A review," in 2023 6thInternational Conference on Information Systems and Computer Networks (ISCON). IEEE, 2023, pp. 1–6.
[9]. C. Bae, K. Kang, G. Liu, and Y. Y. Chung, "A novel real time video tracking framework using adaptive discrete swarm optimization," Expert Systems with Applications, vol. 64, pp. 385–399, 2016.
[10]. A. Ojha, S. P. Sahu, and D. Dewangan, "Vehicle detection through instance segmentation using mask r-cnn for intelligent vehicle system," 05 2021, pp. 954–959.
[11]. F. Li, Z. He, and Y. Yu, "Vehicle target detection in day-night mode based on yolov5," in 2023 IEEE International Conference on Image Processing and Computer Applications (ICIPCA), 2023, pp. 641–646.
[12]. A. Panthakkan, N. Valappil, S. Al-Mansoori, and H. Al-Ahmad, "Ai based automatic vehicle detection from unmanned aerial vehicles (uav) using yolov5 model," in 2022 IEEE 5th International Conference on Image Processing Applications and Systems (IPAS), vol. Five, 2022, pp. 1–5.
[13]. S. Zhang and Y. Tong, "Nighttime vehicle detection algorithm enhanced by nightvisiongan," in 2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2023, pp. 672–676.
[14]. J. Shen, W. Zhou, N. Liu, H. Sun, D. Li, and Y. Zhang, "An anchor free lightweight deep convolutional network for vehicle detection in aerial images," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 12, pp. 24 330–24 342, 2022.
[15]. M. Todevska, A. Zlatkova, D. Taskovski, and Z. Ivanovski, "Real–time vehicle detection based on wavelet decomposition and cnn," in 2023 30th International Conference on Systems, Signals and Image Processing (IWSSIP), 2023, pp. 1–5.
[16]. J. Shen, N. Liu, H. Sun, D. Li, and Y. Zhang, "Lightweight deep network with context information and attention mechanism for vehicle detection in aerial image," IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1–5, 2022.
[17]. J. M, R. P, A. Kunjumon, M. Thamban, and A. Roy, "Convolutional neural networks (cnn)-based vehicle crash detection and alert system," in 2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), 2023, pp. 161–164.
[18]. Q. Du, T. Celik, Q. Wang, and H.-C. Li, "Fully convolutional lightweight pyramid network for vehicle detection in aerial images," IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1–5, 2022.
[19]. P. B. Mishra, A. Malik, M. Safa, G. Saranaya, and D. Arun, "Enhanced object detection with deep convolutional neural networks for vehicle detection," in 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), 2022, pp. 1–7.
[20]. V. Agarwal, "Research on data preprocessing and categorization technique for smartphone review analysis," International Journal of Computer Applications, vol. 131, no. 4, pp. 30–36, 2015.
[21]. L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al Shamma, J. Santamar´ıa, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions," Journal of big Data, vol. 8, pp. 1–74, 2021.
[22]. A. Pathak and A. Dhole,"Lungs cancer identification by deep learning 3d cnn architecture." Grenze International Journal of Engineering & Technology (GIJET), vol. 8, no. 1, 2022.