IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Integrated Real Time Traffic Analysis System

Prof. S. N. Chaudhari
Akshay Chavan, Shivam Chavan, Aditya Parkhe and Abhishek Patil
K. K. Wagh Institute of Engineering Education and Research, Nashik, India

Abstract:
The rapid boom of urbanization and transportation demands has considerably elevated complexity of coping with visitors gliding on highways. This mission introduces a progressive real-time machine that inte- grates car counting, classification, vehicle type, and emblem detection with the usage of deep getting-to-know methodologies. The motive of this mission is to decorate highway site visitors' control by imparting compre- hensive insights into vehicle motion patterns. Employing cutting-edge convolutional neural networks which include YOLO(You Only Look Once), the gadget achieves precise and adaptable automobile detection across various environmental situations. In addition to accurate vehicle detection and counting, this mission uniquely extends its competencies to perceive automobile kinds and brands. By leveraging the electricity of deep studying, the machine can distinguish between exceptional car categories or even understand unique brands, contributing to richer visitor insights. The real-time nature of the system guarantees timely records shipping for informed choice-making in visitor management and infrastructure planning. Extensive experimentation with real-international toll road video information validates the system's first-rate accuracy, processing veloc- ity, and effectiveness as compared to traditional techniques. This mission represents a full-size development in highway traffic control, offering a comprehensive, real-time, vision-primarily based answer that encompasses detection, counting, and type of vehicle kinds and brands.

Keywords: Vehicle dataset, Image segmentation, Vehicle detection, Vehicle counting, Highway management

1 INTRODUCTION

The document discusses the significance of automo- bile detection and monitoring systems in diverse civil- ian packages which includes site visitors analysis and infrastructure research. Traditional strategies rely heavily on human effort for statistics collection and evaluation, main to time-ingesting approaches and excessive error charges. Automated systems using computer imaginative and prescient algorithms offer a more efficient solution.

These structures hit upon and song cars by means of extracting facts from pix or video facts us- ing computer-primarily based algorithms. Different methods encompass using visible traits or motion of cars for detection. Recent advancements in ma- chine studying algorithms have improved accuracy via combining important features and education on huge datasets.

The proposed gadget, Autonomous Vehicle De- tection and Tracking (AVDT), integrates multiple pc vision algorithms to stumble on, classify, and tune motors in real-time. It makes use of pretrained Convolutional Neural Networks (CNN) for detection, SURF for function extraction, and CSRT for tracking. The device's overall performance is evaluated based on recollect rate and achieves high accuracy without the want for added hardware or software sources.

The document compares the proposed AVDT ma- chine with present methods, highlighting its ability to deal with various environmental conditions and with- stand changes in lighting fixtures with out requiring great image processing. Furthermore, it discusses comparisons with commercial software program pro- grams for traffic extent data collection, emphasizing the blessings of AVDT's flexibility and simplicity ofuse.

Overall, the AVDT machine presents a singular technique to car detection and monitoring, imparting real-time abilities and excessive accuracy throughout unique avenue environments.

2 RELATED WORKS

At gift, vision-primarily based vehicle item detection is divided into traditional device imaginative and pre- scient methods and complicated deep studying strate- gies. Traditional machine imaginative and prescient methods use the movement of a automobile to split it from a set heritage photograph. This technique can be divided into 3 classes (Al-Smadi, 2016): the ap- proach of using history subtraction (Radhakrishnan,
), the technique of using non-stop video body dis- tinction (Qiu-Lin, 2011), and the approach of using optical glide (iu, 2014). Using the video body dif- ference technique, the variance is calculated in line with the pixel values of two or 3 consecutive video frames. Moreover, the transferring foreground area is separated by means of the edge (Qiu-Lin, 2011). By the use of this method and suppressing noise, the stopping of the automobile also can be detected (Park,2007). When the historical past photo in the video is fixed, the heritage records is used to set up the her- itage version (Park, 2007). Then, every body photo is as compared with the heritage version, and the shift- ing item can also be segmented. The method of us- ing optical flow can stumble on the movement region within the video. The generated optical waft area rep- resents each pixel's path of movement and pixel pace (iu, 2014). Vehicle

detection strategies using vehi- cle features, which include the Scale Invariant Fea- ture Transform (SIFT) and Speeded Up Robust Fea- tures (SURF) techniques, have been widely used. Forexample, 3-d models have been used to finish car de-tection and class duties (Ferryman, 1995). Using the correlation curves of 3-D ridges at the outer surface ofthe car (Han, 2005), the automobiles are divided into3 categories: cars, SUVs, and minibuses.

3 VEHICLE DATASET

Surveillance cameras in roads were broadly set up worldwide however site visitors snap shots are hardlyever launched publicly because of copyright, privacy, and protection troubles. From the photo acquisition point of view, the visitors photo dataset may be di- vided into three categories: images taken by the autodigicam, photos taken through the surveillance digi- cam, and photos taken by way of non-tracking cam- eras (Luo, 2018). The KITTI benchmark dataset (Geiger, 2012) carries pictures of motorway scenes and normal street scenes used for computerized au- tomobile using and might remedy issues which in- clude three-D item detection and monitoring. The Tsinghua-Tencent TrafficSign Dataset (Zhe, 2016) has 100,000 photographs from car cameras protect- ing numerous lighting situations and weather situa- tions, despite the fact that no cars are marked. The Stanford Car Dataset (Krause, 2014) is a automobile dataset taken by means of non-tracking cameras with a vibrant vehicle appearance. This dataset consists of 19,618 classes of vehicles masking the brands, fash- ions, and production years of the vehicles.

4 METHODS

4.1 Road surface segmentation

This phase describes the technique of motorway road surface extraction and segmentation. We carried out floor extraction and segmentation using photo pro- cessing strategies, consisting of Gaussian combina- tion modelling, which allows better vehicle detection results whilst using the deep getting to know item detection method. The motorway surveillance video picture has a huge discipline of view. The car is the focal point of attention on this study, and the region of hobby in the image is therefore the motorway street floor area. At the identical time, in line with the digi-cam's taking pictures attitude, the road floor place is focused in a specific variety of the image. With this selection, we ought to extract the toll road avenue sur- face regions within the video.

As shown in Fig. 1, to cast off the impact of mo- tors on the road location segmentation, we used the Gaussian aggregate modeling method to extract the history in the first 500 frames of the video. The price of the pixel in the picture is Gaussian round a pos- itive significant value in a positive time range, and every pixel in every body of the picture is counted.If the pixel is far from the centre, the pixel belongs to the foreground. If the cost of the pixel factor de- viates from the centre price inside a sure variance, the pixel point is taken into consideration to belong to the historical past. The combined Gaussian version is mainly useful in images wherein historical past pixels have multi-top traits consisting of the dual carriage- way surveillance pics used on this examine.

After extraction, the background photograph is smoothed with the aid of a Gaussian filter with a 3*three kernel. The MeanShift algorithm is used to smooth the colour of the input picture, neutralize the shade with a similar coloration distribution, and erode the coloration region with a smaller vicinity. On this foundation, the flooding filling set of rules is used to split the road floor place. The flooding filling set of rules selects a point in the road surface region as a seed factor and fills the adjacent non-stop street sur- face areas with the pixel price of the seed point. The pixel fee of the adjacent continuous avenue surface re- gions is close to the seed factor pixel fee. Finally, the hole filling and morphological growth operations are accomplished to greater absolutely extract the road surface. We segmented the road surface vicinity to provide accurate input for subsequent automobile de-tection. For the extracted road floor image, a mini- mum circumscribed rectangle is generated for the im-age without rotation. The processed photo is divided into 5 identical elements, the 1/five vicinity adjacent to the foundation of the coordinate axis is defined be-cause the near far off location of the road floor, and the ultimate four/five region is defined because the near proximal area of the street floor. The close to proximal place and the near far off vicinity overlap by means of 100 pixels (as proven within the crimson a part of Fig. 2) to address the problem that the car inside the photograph may be divided into two ele- ments by using the above manner. The pixel values of the near proximal area and the close to remote region are searched column by column. If the pixel values in the column are all zero, the image of the column is all black and isn't always the street floor region; it is then deleted. After the now not-avenue-surface regions are excluded, the reserved regions are referred to as far off regions and proximal regions of the street floor.

4.2 Vehicle detection using YOLOv3

This segment describes the item detection strate- gies used in this look at. The implementation ofthe highway automobile detection framework used the YOLOv3 community. The YOLOv3 algorithm maintains the primary idea of the first generations of YOLO algorithms. The convolutional neural net- work is used to extract the capabilities of the input photo. According to the size of the function map, such as 13*thirteen, the enter image is divided into thirteen*thirteen grids. The centre of the object la- bel container is in a grid unit, and the grid unit is li- able for predicting the item. The network shape fol- lowed by YOLOv3 is known as Darknet-fifty three. This structure adopts the overall convolution method and replaces the previous version of the direct-related convolutional neural network with the residual shape. The department is used to without delay connect the input to the deep layer of the community. Direct gaining knowledge of of residuals guarantees the in- tegrity of picture feature facts, simplifies the complex- ity of training, and improves the overall detection ac- curacy of the network. In YOLOv3, each grid unit will have three bounding packing containers of vari- ous scales for one object. The candidate container that has the largest overlapping region with the annotatedbox may be the final prediction end result. Addition- ally, the YOLOv3 community has 3 output scales, andthe three scale branches are ultimately merged. Shal-low features are used to locate small items, and deep features are used to detect big objects; the communitycan as a consequence detect objects with scale adjust- ments. The detection pace is rapid, and the detection accuracy is excessive. Traffic scenes taken through toll road surveillance video have precise adaptabil- ity to the YOLOv3 network. The network will sub- sequently output the coordinates, self assurance, and category of the object.

When the usage of YOLO detection, photos are resized to the equal size, along with 416*416, while they're despatched to the network. Since the pho- tograph is segmented, the dimensions of the remote road floor will become deformed and larger. There- fore, extra feature points of a small vehicle object canbe acquired to avoid the loss of some object capa- bilities due to the automobile item being too small. The dataset supplied in "Vehicle dataset" section is positioned into the YOLOv3 community for educa- tion, and the car object detection version is acquired. The vehicle object detection model can hit upon 3 va-rieties of cars: motors, buses, and vans. Because there are few motorcycles on the highway, they were now not covered in our detection. The faraway region and proximal region of the street surface are despatched to the community for detection. The detected vehiclebox positions of the two regions are mapped again to the unique picture, and the proper object function is acquired inside the authentic picture. Using the auto- mobile item detection method for obtaining the class and location of the automobile can offer necessary in- formation for object monitoring. The above records is enough for vehicle counting, and the car detection ap- proach as a consequence does now not discover the unique characteristics of the automobile or the cir- cumstance of the car.

4.3 Multi-object tracking

This section describes how more than one items are tracked based totally on the object container detected in "Vehicle detection the usage of YOLOv3" section. In this take a look at, the ORB set of rules was used to extract the capabilities of the detected motors, and properly results had been acquired. The ORB al- gorithm shows advanced performance in phrases of computational performance and matching expenses. This algorithm is an exquisite alternative to the SIFT and SURF photo description algorithms. The ORB al- gorithm uses the Features From Accelerated Segment Test (FAST) to detect function points after which makes use of the Harris operator to carry out nook de-tection. After obtaining the characteristic factors, thedescriptor is calculated using the BRIEF algorithm. The coordinate device is established by way of takingthe function factor as the centre of the circle and usingthe centroid of the factor location as the x-axis of the coordinate machine. Therefore, when the photo is cir-cled, the coordinate device can be rotated in keeping with the rotation of the Picture, and the characteris- tic point descriptor therefore has rotation consistency. When the picture perspective is changed, a constant point also can be proposed. After obtaining the binary feature factor descriptor, the XOR operation is used to in shape the characteristic points, which improves the matching performance.

The tracking system is proven in Fig. 2. When the range of matching points received is greater than the set threshold, the point is taken into consideration to be correctly matched and the matching container of the item is drawn. The source of the prediction box is as follows: characteristic point purification is exe- cuted the usage of the RANSAC set of rules, which can exclude the wrong noise points of the matching errors, and the homography matrix is envisioned. Ac- cording to the anticipated homography matrix and the function of the unique object detection container, a attitude transformation is achieved to achieve a cor- responding prediction box. We used the ORB set of rules to extract function factors in the item detection container acquired through the vehicle detection set of rules. The object function extraction isn't always car- ried out from the complete avenue surface area, which dramatically Song European Transport Research Re-view (2019) eleven:fifty one Page eleven of 16 re- duces the quantity of calculation. In item tracking, theprediction field of the item in the next frame is drawn for the reason that alternate of the car item within the non-stop frame of the video is diffused consistent with the ORB characteristic extracted inside the item box. If the prediction container and the detection box of thesubsequent frame meet the shortest distance require- ment of the centre point, the same item effectively suits among the 2 frames. We outline a threshold T that refers to the maximum pixel distance of the de- tected centre factor of the vehicle item field, which movements between adjoining video frames. The po- sitional motion of the equal automobile inside the ad-jacent frames is less than the edge T. Therefore, whilst the centre point of the car item field movements over T within the adjacent frames, the motors in the two frames aren't the identical, and the statistics associ- ation fails. Considering the scale change throughout the motion of the automobile, the price of the thresh- old T is related to the size of the car object container. Different vehicle object containers have extraordinarythresholds. This definition can meet the needs of car movement and exclusive input video sizes. T is cal- culated by Eq. 1, wherein field peak is the peak of theautomobile object box.

$$T = containerheight/0.25$$

We delete the trajectory that isn't updated forten consecutive frames, that's suitable for the digi- tal camera scene with a extensive-angle of picture se-ries on the highway beneath observe. In this formof scene, the road floor captured by way of the cam- era is remote. In ten consecutive video frames, the car will flow farther away. Therefore, while the tra- jectory is not updated for ten frames, the trajectoryis deleted. At the same time, the automobile trajec- tory and the detection line will simplest intersect as soon as, and the brink putting consequently does no longer have an effect on the final counting end result. If the prediction field fails to in shape in consecutive frames, the item is considered to be absent from the video scene, and the prediction container is deleted. From the above manner, the worldwide item detec- tion outcomes and monitoring trajectories from the entire dual carriageway monitoring video angle are received.

5 RESULTS AND ANALYSIS

In this phase, we describe the performance trying out of the methods provided in "Methods" phase. We ex- perimented with the vehicle item dataset established in "Vehicle dataset" segment. Our test used high defi-nition dual carriageway motion pictures for three one-of-a-kind scenes, as shown in Fig. 1.

Figure 1: Vehicle detection.

5.1 Network training and vehicledetection

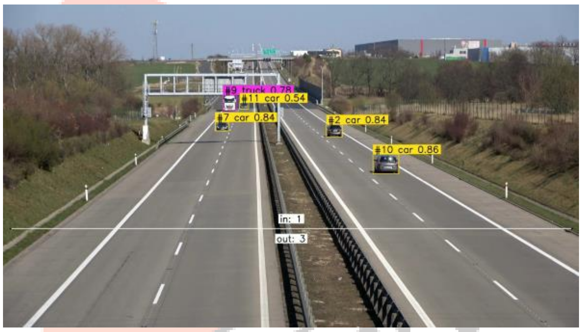
We used the YOLOv3 community for vehicle ob- ject detection and our set up dataset for community schooling. In community training, there's no ideal an- swer for the dataset department. Our dataset dividing method follows the standard utilization. We cut up the dataset into an 8020set images, and the check set pics are randomly decided on from the dataset. Due to a huge quantity of dataset photos, the charge of the check set and training set is sufficient to gain the model. To obtain an accurate version, the rate of the education set should be excessive. The education set has 8,904 pics, and severa automobile samples may be skilled to achieve correct fashions for detecting auto- mobiles, buses, and truck targets. The check set has 2225 photos with automobile targets that are abso- lutely unique from the education set, which is suffi- cient to test the accuracy of the model that has been skilled. We used a batch length of 32 and set the weight attenuation to zero.0005 and the momentum value to 0.Nine for the maximum number of train- ing iterations of 50,200. We used a gaining knowl- edge of price of zero.01 for the first 20,000 itera- tions, which modified to zero.001 after 20,000 iter- ations. This method made the gradient fall moder- ately and made the loss cost lower. To make the de- fault anchor field greater suitable for the dataset an- notation container to be annotated, we used the ok-way technique to make changes. The schooling set of our dataset calculated the default anchor containerlength at the community resolution of 832*832, and we received 9 sets of values: [13.2597, 21.4638], [24.1990, 40.4070], [39.4995, 63.8636], [61.4175, 96.3153], [86.6880, 137.2218], [99.3636, 189.9996], [125.6843, 260.8647], [179.7127 , 198.8155], [189.3695, 342.4765], with an average IOU of 71.20 percent. To enhance the detection impact of small ob- jects, we did now not discard samples with less than 1-pixel cost all through education however put them into the community for schooling. We output the end result of splicing the characteristic map of the pre- ceding layer of the routing layer before the remaining yolo layer of Darknet-fifty three and the 11th layer OfDarknet-53. We set the step size to four within the up- sampling layer before the ultimate yolo layer. When we set the image input to the community, the network decision turned into 832*832 rather than the default 416*416 resolution. After the input decision is ex- panded, whilst the community is output in the yolo layer, it can have a correspondingly larger decision and might for this reason enhance the accuracy of theitem detection.

A continuous 3000 frames of photos have been used for vehicle detection in numerous motorway scenes with the aid of the usage of our trained version. We extracted and divided the road surface location and positioned it into the community for automobile detection. Then, we in comparison our approach withthe detection of photographs with 1920*1080 resolu-tion into the community (with out dividing the street floor). We in comparison the quantity of item detec- tions underneath one of a kind strategies with the ac-tual quantity of vehicles.

Figure 2: Vehicle Counting.

5.2 Tracking and counting

After obtaining the object container, we achieved ve-hicle monitoring based on the ORB feature factor matching technique and carried out trajectory evalu- ation. In the experiment, when the matching factor of every object became greater than ten, the correspond- ing ORB prediction function changed into generated. Based at the path in which the monitoring trajectory became



generated, we used the detection line to judgethe direction of movement of the automobile and clas- sify it for counting. We carried out experiments on the different 3 films which can be similar to the scene in "Network training and vehicle detection" segment however with a unique variety of frames. We used the real time price to evaluate the velocity of the machine proposed on this paper, which is described as the ratioof the time required for the gadget to manner a videoto that of the unique video performed. In Eq. Four, the system going for walks time is the time required for the gadget to manner a video, and the video strollingtime is the time required for the unique video played.The smaller the actual time charge value is, the fasterthe device performs the calculations. When the cost of the actual time charge is less than or same to 1, the enter video can be processed in real time.

The consequences are proven in Table five. The outcomes show that the common accuracies of vehi- cle using route and vehicle counting are 92.Threetoll

road monitoring video, the auto magnificence has a small item and is easily blocked by using big motors. At the identical time, there could be multiple vehi- cles in parallel, to be able to affect the accuracy of the song counting. Our authentic video runs at 30 frames in step with second. From the calculation of the pace, it could be found that the vehicle monitoring set of rules primarily based at the ORB feature is speedy. The device processing velocity is related to the rangeof vehicles in the scene. The more the range of vehi- cles, the extra functions need to be extracted, and thedevice processing time will thus grow to be longer. In popular, the automobile counting machine proposed on this manuscript is very near real-time processing.

6 CONCLUSIONS

The examine brought a novel technique for vehi- cle detection and tracking in motorway surveillance video scenes, leveraging the YOLOv3 algorithm and ORB characteristic extraction. By defining awe-some road areas and reading vehicle trajectories, the approach validated effective overall performance in shooting driving direction, car type, and remem- ber, with practicality and fee-efficiency as compared to conventional hardware-primarily based monitor- ing structures. Importantly, the applicability of this method extends to European nations like Germany, France, the United Kingdom, and the Netherlands, where similar motorway surveillance camera setups exist, suggesting its capacity as a valuable reference for transportation research in Europe.

7 REFERENCES

1. Al-Smadi, M., A. K.-S. R. (2016). Traffic surveillance: A review of vision based vehicle detection, recognition and tracking. *International Journal of Applied Engi- neering Research*.
2. Ferryman, J.M., W. A. S.-G. B. K. (1995). *A generic de- formable model for vehicle recognition, In Proced- ings of the British Machine Vision Conference 1995.* British Machine Vision Association.
3. Geiger, A. (2012). *Are we ready for autonomous driving? the kitti vision IEEE Conference on Computer Visionand Pattern*
4. *Han, D., L. M. C.-D. M. J. (2005).* Vehicle class recognition from video-based on 3d curve probes. *IEEE.*
5. *iu, Y., Y. L. S.-Q. D. J. (2014).* Optical flow based urban road IEEE.
6. Krause, J., S. M. D.-J. L. F. (2014). *3d object representa- tions for IEEE.*
7. *Luo, Z. (2018).* Traffic analysis of low and ultra-low frame- rate videos, Doctoral dissertation. *Universite´ de Sher-brooke.*
8. *Park, K., L. D. P.-Y. (2007).* Video-based detection of street- parking IEEE.
9. Qiu-Lin, L.I., . J.-F. H. (2011). *Video object extraction byusing background Computer Engineering.*
10. *Radhakrishnan, M.* Video object extraction by using back-ground Digital Image Processing.
11. Zhe, Z., L. D. Z.-S. H. X. H. S. (2016). *Traffic-sign detec-tion IEEE.*