Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231

e-ISSN : 2541-2019
p-ISSN : 2541-044X

# Detection and tracking different type of cars with YOLO model combination and deep sort algorithm based on computer vision of traffic controlling

**Nisma Novita Hasibuan[1)*], Muhammad Zarlis[2)], Syahril Efendi[3)]**
[1)2)3)] Magister of Informatics Program, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Medan, Indonesia
[1)] nismanisma23@gmail.com

**Submitted** : Dec 9, 2021 | **Accepted** : Dec 23, 2021 | **Published** : Dec 23, 2021

**Abstract:** The application of CCTV cameras for traffic surveillance and monitoring is one effective solution to address urban traffic problems, as the number of vehicles that continue to increase rapidly but the area of the road remains the same will cause congestion. However, the problem in traffic surveillance and monitoring is not just focusing on vehicle detection based on category inference on video sequence data sourced from CCTV cameras alone, another important, challenging task is to combine calculations, classification and tracking of different vehicle movements in urban traffic control systems. The study expanded on previous research by breaking down the problem into different sub-tasks using the YOLOv4 approach combined with the Deep Sort algorithm for the detection and tracking of objects directly on CCTV footage of vehicle activity on the city's three-stop highway. Based on the results of YOLOv4 testing resulted in a detection accuracy rate with mAP of 87.98% where the combination of YOLOv4 with the Deep Sort algorithm can detect, track and calculate 13 types of vehicles.

**Keywords:** Traffic surveillance, Urban traffic problems, YOLOv4, Deep Sort algorithm, traffic light controlling

## INTRODUCTION

Intelligent Transportation System (ITS) is one of the effective solutions to overcome urban traffic problems, because the number of vehicles that continue to increase rapidly but the road area remains the same will cause congestion. Determining vehicle density on the road and video surveillance analysis help a lot in traffic management systems (Sang et al. 2018). One of the implementation applications of ITS in various major cities in Indonesia such as Medan City is to apply CCTV cameras for traffic surveillance and monitoring. However, actual traffic monitoring still relies on human operators to detect, track, warn, and initiate responses to various traffic incidents on the highway network, so stored video data has not automatically detected various entities useful in decision making such as vehicle detection, pedestrians, vehicle density, congestion and traffic violations.

Advances in deep learning-based computer vision of the past decade have been a leading method for high-quality general object detection, including vehicle detection as a real-time traffic control system with different objectives, such as vehicle detection (Hsu, Huang, and Chuang 2018; Kim 2019; Meng et al. 2020), vehicle calculation (Asha and Narasimhadhan 2018; Nam Bui, Yi, and Cho 2020; Oltean et al. 2019), vehicle tracking (Gu et al. 2017; He et al. 2019), vehicle classification (Schumann et al. 2018; Watkins, Pears, and Manandhar 2018) and others. Various techniques have been proposed for intelligent traffic control systems, but deep learning techniques are more accurate than traditional techniques (Li et al. 2018) because they have the ability to learn image or video features and classification and regression tasks (Song et al. 2019).

Today's popular object detection method consists of a one-stage and two-stage approach. R-CNN, Fast R-CNN, Faster R-CNN and Spatial Pyramid Pooling Network (SSP-Net) are two-stage approaches by optimizing the Selective Search to Regional Proposal Network (RPN) function to generate candidate object boxes through various algorithms, then classify objects with Convolutional neural networks that produce high accuracy (Husein

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

et al. 2020). But too computationally intensive, the bound boxes are too slow for real time or close to real time (Zhang, Li, and Yang 2019).

One-stage detection method approaches such as Single Shot Multibox Detector (SSD), You Only Look Once (YOLO) implement convolution networks to directly predict the class and location of objects trained end to end faster than the two-stage method. SSDs focus on detection of objects of different scales so that they perform well in speed and accuracy, but ignore the relationships between the various layers of the pyramid so that they perform poorly in the detection of small objects (Zhao, Ni, and Jia 2017) and have lower accuracy than YOLO for vehicle detection on traffic monitoring videos (Husein et al. 2020). Other works such as RefinaDet (Zhang et al. 2018), CenterNet (Duan et al. 2019), RetinaNet (Lin et al. 2020), EfficientNet (Tan and Le 2019), EfficentDet (Tan, Pang, and Le 2020), FreeAnchor (Zhang et al. 2021) achieved a better trade-off between speed and accuracy.

YOLO network architecture is one of the most popular models of researchers' interest in building traffic control systems (Huang et al. 2020; Mahto et al. 2020; Ouyang and Wang 2019; Wang et al. 2020), several performance improvements and proposed new models such as LittleYOLO-SPP (Li et al. 2018), Gaussian-YOLOv3 (Choi et al. 2019), MME-YOLO (Zhu et al. 2021) and SF-YOLO (Han et al. 2020). Furthermore, in recent years, YOLO has performed well in many areas, such as helmet detection (Wu et al. 2019), tiger detection (Wei, He, and Lu 2020), opium detection (Zhou et al. 2019), pedestrian detection (Lai, Sun, and Liu 2020), traffic sign detection (Arcos-García, Álvarez-García, and Soria-Morillo 2018), vehicle plate detection (Laroca et al. 2019) and others. Recently, Alexey Bochkovskiy released a new network architecture YOLOv4 (Bochkovskiy, Wang, and Liao 2020) which is an improvement from YOLOv3 to detect objects in some detection areas quickly and accurately. YOLOv4 uses Cross-Stage-Partial-Connections (CSP) as a new backbone that can improve CNN's learning capabilities faster than other one-stage approaches.

As one of the most representative networks, yolov4 network achieves good performance in speed and accuracy and has high potential in the field of intelligent transportation. But the problem in its field is not only focusing on vehicle detection based on category inference on video sequence data sourced from CCTV cameras, vehicle calculations and tracking the movements of different vehicles is an important but challenging task in urban traffic control systems. Therefore, the main contribution of this study is to expand on previous research by breaking down the problem into different sub-tasks such as automatic classification, detection, tracking and calculation of vehicles. The proposed framework applies YOLOv4's deep learning techniques to vehicle detection, then vehicle tracking is adopted with the Deep Sort algorithm. In particular the main objective of this research is to research and analyze the application of computer vision with deep learning techniques to be applied as a solution in the construction of traffic control systems based on CCTV video of highway traffic.

## LITERATURE REVIEW

In this section will be outlined review literature that contains information about the development of intelligent transportation system research with the application of computer vision and deep learning techniques consisting of several sub-chapters.

### Vehicle detection

Detection of computer vision-based vehicles with deep learning techniques is one of the areas of research that has attracted many researchers' interest in recent years, especially in the application of the field of intelligent transportation systems. Various new models with deep learning techniques are reported in many studies with promising performance results, for example (Xiang, Lv, et al. 2018) proposing a Faster-RCNN framework for identifying vehicles of different types of vehicle objects, Faster-RCNN models are modified to adjust the position of roi layers, then convolution layers are added to improve detection accuracy whereas (Hsu et al. 2018) eliminate regional network proposal functions. (RPN) fast RCNN network for vehicle detection in virginia tech transportation institute (VTTI) dataset shrp 2 NDS.

Furthermore, (Li et al. 2018) adopted the YOLOv2 network architecture by building a new YOLO-vocRV network to detect multi-target detection of highway vehicles specifically recorded as method testing materials. Hoanh Nguyen (Nguyen 2019) proposes a Faster-RCCN approach to rapid vehicle detection on kitti and LSVH dataset sets. The paper adopts the MobileNet architecture at the faster-RCNN base layer as well as applying NMS to duplication problems resulting in a 4% accuracy improvement in kitti datasets and LSVH datasets of 24.5%. The proposed new model framework of DP-SSD (Zhang et al. 2019) is based on the SSD network architecture. Developed to improve the detection of different types of vehicles in real-time, this model was tested on the KITTI and UA-DETRAC datasets which produced an accuracy rate comparable to other models except YOLOv3. Adaptive Perceive-SSD (AP-SSD) was proposed (Wang et al. 2018) based on an improved SSD object detection framework for multiobjection accuracy and detection speed in traffic scenes and (Biswas et al. 2019) compared SSD performance with MobileNet-SSD for traffic density estimation where SSD models produce better performance than MobileNet-SSD.

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

**Vehicle Tracking**

Tracking vehicles in intelligent transportation systems is one of the important problems in traffic control systems. Most multi-object tracking methods use Detection-Based chase (DBT) and Detection-Free chase (DFT) for object initialization. The DBT method uses background modeling to detect moving objects in a video frame before tracking. The DFT method needs to initialize tracking objects but cannot handle the addition of new objects and the departure of old objects. Dual Object Tracking algorithms need to consider the similarity of intra-frame objects and inter-frame object related issues. The similarity of intra-frame objects can use normalized cross-correlation (NCC). Bhattacharyya distance is used to calculate the distance of the color chart bar between objects. When inter-frame objects are linked, it should be determined that an object can only appear in one path and one path can only be related to one object. Currently, detection rate exceptions or track-level exceptions can solve this problem. To solve problems caused by scale changes and changes in the illumination of moving objects, (Mu, Hui, and Zhao 2016) used the SIFT points feature for vehicle object tracking, although this was slow. The ORB feature point detection algorithm proposed by (Rublee et al. 2011) obtains the extraction feature point better with a much higher principle speed than sift.

Another method was proposed by Bolchinski (Bochinski, Eiselein, and Sikora 2017) using simple Intersection over Union (IoU) based object matching by extracting position information from overlapping frames. This model results in very fast tracking. However, accuracy will be reduced when used on complex objects or in difficult landscapes. The Simple Online and Real-time Tracking (SORT) algorithm proposed by (Bewley et al. 2016) is a framework for tracking multiple visual objects based on basic data associations and status estimation techniques. This method generates object identity quickly allowing the tracker to simultaneously learn the features of the object while performing the tracking task. SORT's accuracy and precision outperform traditional IoU-based methods but have a tendency to generate more false positives and Deep SORT solves some of this problem by introducing reliance on detection results and limiting box coordinates (Wojke, Bewley, and Paulus 2018). Therefore, the Deep Sort algorithm approach was adopted in this study for multi-vehicle tracking.

**Vehicle Calculation**

Automatic vehicle detection and counting is considered important in improving traffic control and management; some of the work in the literature reports new vehicle counting methods based on the video sequence of toll roads reported by (Meng et al. 2020) where SSD models are applied to vehicle classification and detection. Furthermore, the work (Kim 2019) proposes an aggregated channel features (ACFs) algorithm approach and adaboost algorithm for autonomous vehicle detection. Some researchers adopted the YOLO network architecture combined with other algorithms, such as (Asha and Narasimhadhan 2018) applying YOLO with correlation filter (CF) algorithm, (Oltean et al. 2019) using the Fast Motion Estimation (FTE) algorithm. Generally, visual speed estimation is done by tracking objects through sequential frames. Estimated speed of the object being tracked is often considered a sub-task when compared to object detection and tracking as this only involves converting the speed of the limiting box across the frame to the inertial frame.

Recently, many studies have proposed a video-based vehicle counting framework. For example, Xiang et al. (Xiang, Zhai, et al. 2018) present a new framework for calculating vehicles using aerial video. Specifically, the object can be detected following two cases such as a static background to detect and a moving background to estimate the movement of the vehicle. In the case of highway scenarios, Song et al. introduced a counting system using YOLOv3 to detect the type and location of vehicles. Then, Song et al. (Song et al. 2019) applied the ORB algorithm (Rublee et al. 2011) to vehicle trajectories based on video, while (Dai et al. 2019) proposed a vehicle counting framework with matching method algorithms, based on detection output to analyze traffic flows in complex areas (e.g., intersections) with different types of vehicles. The YOLOv3 network architecture was adopted for vehicle detection directly on video. In the case, proposed a vehicle counting framework with the YOLOv3 algorithm was adopted for vehicle detection. Furthermore (Liu, Zeng, and Jiang 2017) proposed adaptive patterns based on virtual loops and detection line methods to improve vehicle counting performance at highway intersections.

Based on the description of related research, there are many models and techniques proposed in much literature with promising performance results to be applied as intelligent traffic control system solutions. But it is difficult to identify and apply optimal methods to be exposed, this is because in many previous research works using different stages, data sets and testing conditions performed differently. Therefore, in this study the authors identified and tested the most commonly used deep learning techniques and reported the results of comparative experimental results on the highway CCTV video dataset, then combined with the Deep Sort algorithm for vehicle tracking and calculation.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

## METHOD

The model architecture proposed in this study is based on deep learning techniques by applying the YOLOv4 model for vehicle classification and detection. The model will be analyzed for performance for vehicle detection with the aim of getting the most optimal model applied when implementing a traffic control system in Medan City. Furthermore, for direct vehicle tracking, the authors adopt the Deep Sort algorithm which is one of the most optimal algorithms currently used for multi-object tracking compared to the ORB and Kalman Filter algorithms, lastly for vehicle calculation needs a background reduction approach is proposed. More clearly the proposed model can be seen in Fig 1.

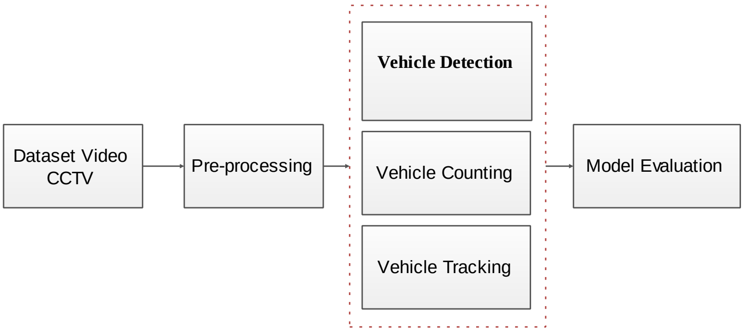



Fig 1 Model architecture proposed

## RESULT

The experimenting process and dataset training is carried out on the specifications of core i5 10 T Gen laptop devices, 8 GB RAM, WIN 10 with Nvidia GeForce MX350 GPU. Based on the results of the evaluation of the CCTV dataset used in this study, the next step is to conduct training on yolov4 models to detect vehicle types. The first process is carried out labeling vehicle datasets sourced from CCTV videos to be used, in this study the author uses labeling from Roboflow (https://app.roboflow.com/) which provides labeling applications. The video testing data used will be extraction into images with a limit of 500 images on each video, then labeling data with a division of 70% of training data, 10% valid and 20% test data.

**Table 1** Details of the Training Dataset

| Picture | Training Data | Data *Valid* | Test Data |
|---|---|---|---|
| 25,000 | 17,500 | 2,500 | 5,000 |

Overall the total images used for labeling amounted to 25,000 images with details of 17,500 for training data, 2,500 valid data and 5,000 test data. The training process with an iteration value (epoch) of 1000-8000 for each class with model performance evaluation using the mAP (mean Average Precision) approach where the training results will be stored under the format name cctv_iterasi.weights. During the training process, the application will store two automatic files, namely cctv_best.weights and cctv_last.weights. In the cctv_best.weights file will store the best accuracy value among the number of iterations, while cctv_last.weights will store the accuracy value of the last training, it will be very useful to repeat the training if there is a problem during training that concerns the application stopped or forced to be issued. In each iteration will be stored with multiples of 1000, in this study the authors conducted iterations of training up to multiples of 8000 Training results on each iteration presented in table 2.

**Table 2** Results of Training Evaluation

| **Epoch** | **mAP** | **precision** | **Recall** | **F1-score** | **averager IoU** |
|---|---|---|---|---|---|
| 1000 | 59.79 | 0.41 | 0.56 | 0.47 | 28.79 |
| 2000 | 66.49 | 0.62 | 0.77 | 0.68 | 46.15 |
| 3000 | 73.63 | 0.66 | 0.77 | 0.71 | 50.84 |
| 4000 | 79.92 | 0.67 | 0.78 | 0.72 | 51.47 |
| 5000 | 80.11 | 0.78 | 0.76 | 0.72 | 53.36 |
| 6000 | 84.66 | 0.78 | 0.79 | 0.74 | 55.05 |
| 7000 | 87.98 | 0.82 | 0.81 | 0.74 | 63.43 |

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

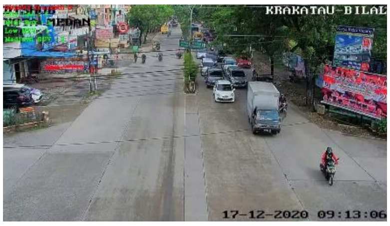
| | | | | | |
|---|---|---|---|---|---|
| 8000 | 85.36 | 0.69 | 0.78 | 0.74 | 61.65 |
| best | 87.98 | 0.70 | 0.77 | 0.73 | 63. 43 |

The testing process is carried out on a video testing dataset that has been evaluated on vehicle activity recordings from December 17, 2020 to December 23, 2020 for 7 days. Further, the detection results will be evaluated and combined with the deep sort method for vehicle tracking, finally the entire model will be built applications for detection, tracking and calculating the number of types of vehicles passing through which will be made a line (ROI) barrier. All vehicle detection, tracking and calculation activities will be stored in a .csv format file for each video testing.

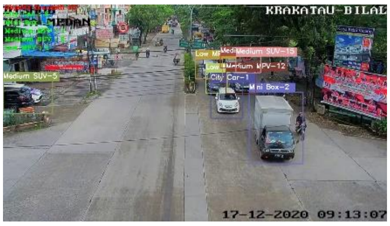
**Table 3** Vehicle Detection


Based on table 3 it can be seen that the proposed yolov4 model can detect several types of vehicles at once in CCTV footage video of highway activity, in the first frame the proposed model detects 5 types of vehicles namely City Car 1, Mini Box 1, Low MPV 1 and Medium SUV 1, but for the position of the second Medium SUV class in a stop position. In the frame of 13 class Medium SUVs that are parallel to the Low MPV class so that it is not detected, but the car behind is detected as a Mini Bus class, this is likely to happen that the model detects the medium SUV and the City Car becomes one class so that it becomes a box becomes large.

**Table 4** Vehicle Tracking


Frame 0


Frame 13

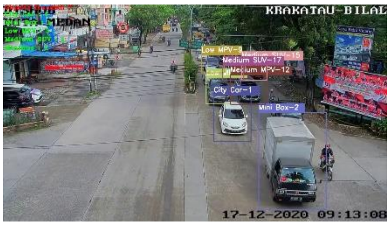

Frame 26

*name of corresponding author

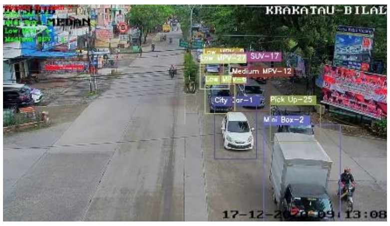
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231

e-ISSN : 2541-2019
p-ISSN : 2541-044X

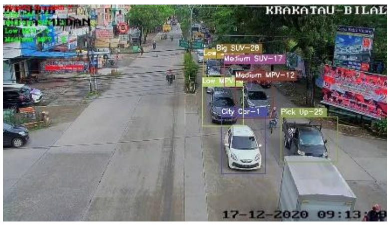
Frame 39



Frame 52

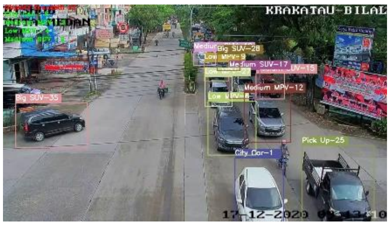


Frame 65



Frame 78

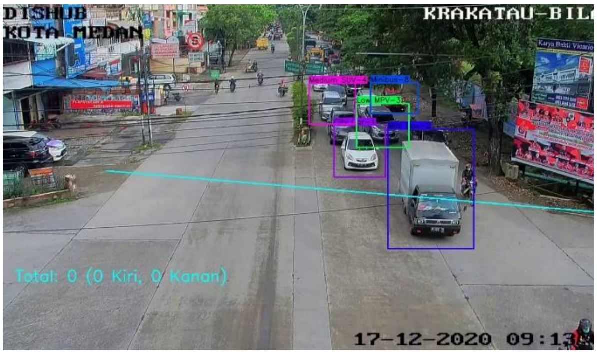
Frame 91



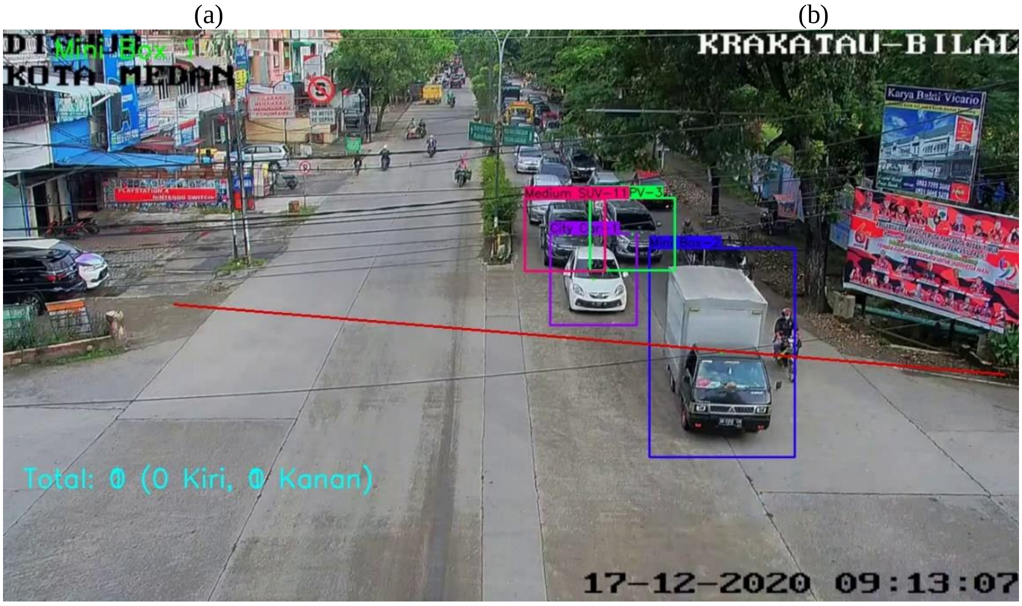
Frame 104

In figure 4 describes the illustration of vehicle tracking using the deep sort algorithm, in frame 0 all objects have not been detected, frame 13 class City Car detected first resulting in ID 1, and class Mobil Box with ID 2. In the City Car class with ID 1 and Car Box ID 2 class can be seen in frames 0, 13, 26, 39, 52, 65, 78, 91. While ID 3 is in the Low MPV class name behind the city car, ID 4 Mini Bus which is a detection error between the Low MPV type car and the car behind it (red) while the ID 5 in the Medium SUV class with a stop position in the right corner of the camera. In frame 104 of the car box vehicle type has not been detected while City Car ID 1 is still detected. Overall, each class of vehicle type that has been detected will be created a virtual ID as a tracking marker and its results will be stored into the file.

(a)



(b)




(c)

Fig 2 Vehicle Calculation

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

In the initial stage, a line will be created that can automatically be created directly on video image 2 (a). This aims because the camera position is not fixed on each video, then the application will record activities by detecting vehicle types and tracking vehicles in figure 1 (b), after the vehicle crosses the line will be calculated by storing vehicle ID and class information. The results of the calculation will be stored on the file and displayed on the video as seen in figure 2 (c) of the left corner. In table 5 is the overall result of total vehicles passing and fig 3 is the frequency of vehicle activity in intersection, namely simpang krakatau-bilal, simpang glugur dan simpang lapangan merdeka.

**Table 5** Total Vehicles Passing By

| Date | Save | | |
|---|---|---|---|
| | **Krakatoa Bilal** | **Glugur** | **Wipe. Independent** |
| 17/12/2020 | 245 | 285 | 480 |
| 18/12/2020 | 290 | 356 | 392 |
| 19/12/2020 | 283 | 76 | 414 |
| 20/12/2020 | 187 | 244 | 315 |
| 21/12/2020 | 240 | 211 | 529 |
| 22/12/2020 | 251 | 317 | 504 |
| 23/12/2020 | - | 209 | 500 |
| 24/12/2020 | - | 341 | 446 |

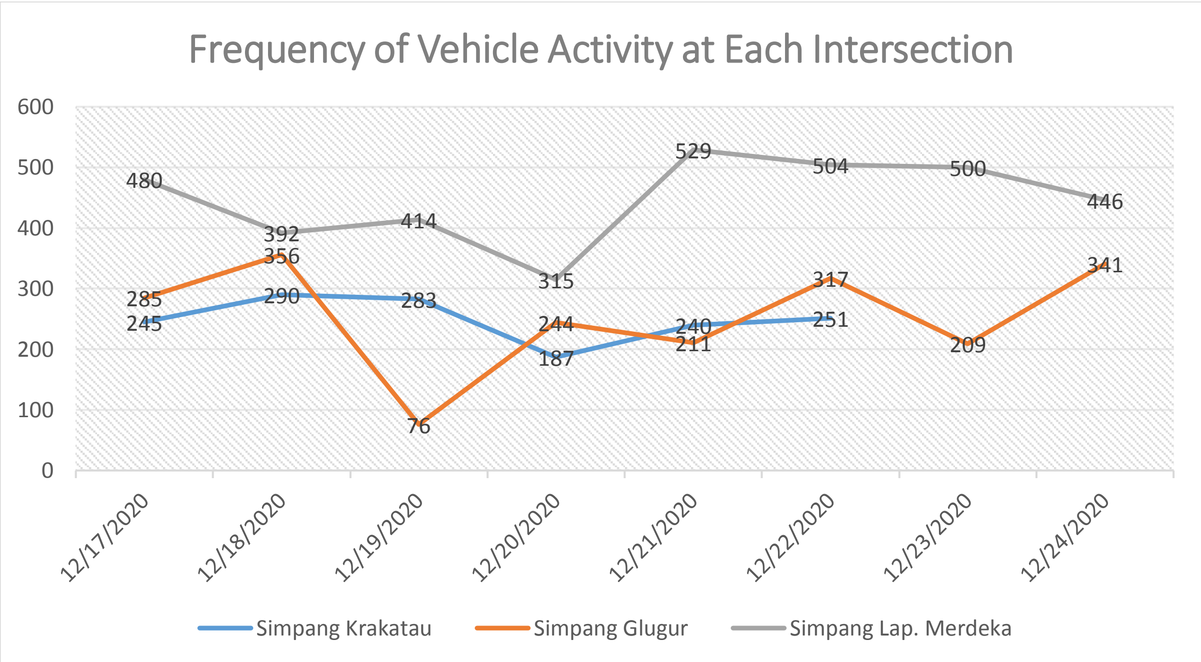



**Fig 2** Frequency of Vehicle Activity in Each Store

In fig 2 it is seen that the highest activity occurs in intersection lapangan merdeka which occurred on 21/12/2020 on Monday which is the day of start of work, the decrease of activity on this road occurred on 20/12/2020 which is on Sunday day which is a holiday. The type of vehicle that will be counted is a vehicle that passes towards the city hall road. Different results occurred in glugur reservoir, where the most crowded activity occurred on 18/12/2020 on Friday with a total of 356 vehicles passing through, and on the day of19/12/2020 there was a drastic decrease with the number of vehicles as many as 76. On the road junction krakatau bilal activity the number of vehicles passing looks normal, where the most number of vehicles occurred on 19/12/2020 day Saturday with the number of 283 and at least occurred on the day Sunday on 20/12/2020 with the number of 187 vehicles. From the total results of vehicles calculated in each intersection is still limited to one direction, while the other direction is difficult to calculate, this is because of the limited data recording of vehicle activity on the highway that can not be collected by the author, in addition to the condition of the CTTV installed currently adjusted to the

*name of corresponding author

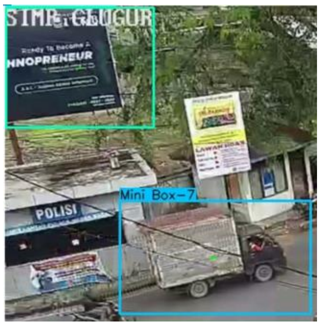
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

travel needs of officials such as governors, mayors, regional police officers. or other important guests to facilitate the *controlling of* traffic lights.

## DISCUSSIONS

Based on the results of application testing developed for the detection, tracking and calculation of vehicle types in Krakatau Bilal, Glugur and Merdeka Square in general has been successfully implemented, but there are still some problems in the results of vehicle type detection as shown in fig 3.

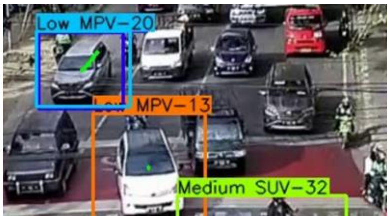
(a)



(b)

**Fig 3** Detection Error

In fig 3 it can be seen that there are some errors in the detection of vehicle type objects applied directly on CCTV such as fig 3 (a) ID 331 detected with the type of Mini Box vehicle while in fig 3 (b) the billboard detected the anchor box but the type of vehicle does not appear. In addition, the scene of the object recorded on CCTV cameras still has an influence on visibility as seen in figure 4.



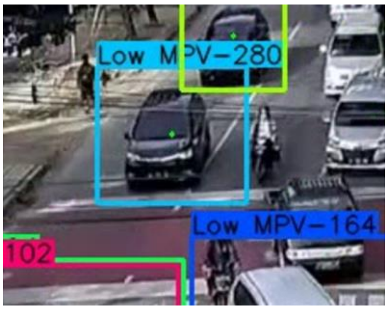
(a) Frame 120

(b) Frame 220

(c) Frame 260

**Fig 4** Vehicle Type Detection Error

In figure 4 is an illustration of vehicle type detection error where (a) object is detected with *Low* MPV *class* with ID 20, but at the time *frame* 220 object can not be detected so tracking ID does not exist. At the time *frame* 260 of the object was detected with pick up *class* and ID 15, this indicates that the change in dimensions of the object in *frame* 260 resulted in an error so it still needs to be further examined. Object detection errors are very influential at the time of labeling where the size of the anchor box applied to indicate the *class* of vehicle type done manually needs to be evaluated, but the tracking results applied in this study can be a good solution to overcome repeated detection problems so that the calculation of passing objects is more accurate. Illustration of vehicle tracking shown in figure 5.



Frame 620

Frame 640

Frame 660

**Fig 5** Tracking Results

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

In figure 5 show the tracking work applied to improve the detection results and calculation of vehicle types, from the illustration can be seen in frame 620 a *low* MPV vehicle type object with ID 280 detected, then in frame 640 the object can not be detected so that the ID also does not exist, then in frame 660 objects detected with *class* the same so that the same ID on frame 620 will reappear. In this case it can prove that tracking with the approach *deep sort* can work as expected.

## CONCLUSION

In this study, the authors developed the YOLOv4 approach combined with a deep sort algorithm for the detection and tracking of objects directly on CCTV footage of vehicle activity on medan city highways at kratau bilal, glugur and merdeka fields, in addition to being able to calculate the type of vehicles passing by. From the results of the tests that have been done, then the author can draw several conclusions, namely:

1. The implementation of YOLOv4 and Deep Sort Methods can be accurately combined into one application that serves for vehicle type detection and tracking.
2. YOLOv4 method produces detection accuracy rate with mAP of 87.98%
3. Deep Sort models can perform vehicle tracking directly thus optimizing yolov4 models for object detection and avoiding miscalculating the same type of vehicle.
4. The application developed can be applied as a traffic control system on medan city highways by generating total vehicle information and vehicle types.

## REFERENCES

Arcos-García, Álvaro, Juan A. Álvarez-García, and Luis M. Soria-Morillo. 2018. “Evaluation of Deep Neural Networks for Traffic Sign Detection Systems.” *Neurocomputing* 316:332–44. doi: 10.1016/j.neucom.2018.08.009.

Asha, C. S., and A. V Narasimhadhan. 2018. “Vehicle Counting for Traffic Management System Using YOLO and Correlation Filter.” Pp. 1–6 in *2018 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*. IEEE.

Bewley, Alex, Zongyuan Ge, Lionel Ott, Fabio Ramos, and Ben Upcroft. 2016. “Simple Online and Realtime Tracking.” Pp. 3464–68 in *Proceedings - International Conference on Image Processing, ICIP*. Vols. 2016-Augus. IEEE Computer Society.

Biswas, Debojit, Hongbo Su, Chengyi Wang, Aleksandar Stevanovic, and Weimin Wang. 2019. “An Automatic Traffic Density Estimation Using Single Shot Detection (SSD)and MobileNet-SSD.” *Physics and Chemistry of the Earth* 110(January 2019):176–84. doi: 10.1016/j.pce.2018.12.001.

Bochinski, Erik, Volker Eiselein, and Thomas Sikora. 2017. “High-Speed Tracking-by-Detection without Using Image Information.” *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2017* (August). doi: 10.1109/AVSS.2017.8078516.

Bochkovskiy, Alexey, Chien Yao Wang, and Hong Yuan Mark Liao. 2020. “YOLOv4: Optimal Speed and Accuracy of Object Detection.” *ArXiv*.

Choi, Jiwoong, Dayoung Chun, Hyun Kim, and Hyuk Jae Lee. 2019. “Gaussian YOLOv3: An Accurate and Fast Object Detector Using Localization Uncertainty for Autonomous Driving.” *Proceedings of the IEEE International Conference on Computer Vision* 2019-Octob:502–11. doi: 10.1109/ICCV.2019.00059.

Dai, Zhe, Huansheng Song, Xuan Wang, Yong Fang, Xu Yun, Zhaoyang Zhang, and Huaiyu Li. 2019. “Video-Based Vehicle Counting Framework.” *IEEE Access* 7:64460–70. doi: 10.1109/ACCESS.2019.2914254.

Duan, Kaiwen, Song Bai, Lingxi Xie, Honggang Qi, Qingming Huang, and Qi Tian. 2019. “CenterNet: Keypoint Triplets for Object Detection.” *Proceedings of the IEEE International Conference on Computer Vision* 2019-Octob:6568–77. doi: 10.1109/ICCV.2019.00667.

Gu, Xiao Feng, Zi Wei Chen, Ting Song Ma, Fan Li, and Long Yan. 2017. “Real-Time Vehicle Detection and Tracking Using Deep Neural Networks.” *2016 13th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP 2017* 167–70. doi: 10.1109/ICCWAMTIP.2016.8079830.

Han, Byung Gil, Joon Goo Lee, Kil Taek Lim, and Doo Hyun Choi. 2020. “Design of a Scalable and Fast Yolo for Edge-Computing Devices.” *Sensors (Switzerland)* 20(23):1–15. doi: 10.3390/s20236779.

He, Zhiqun, Yu Lei, Shuai Bai, and Wei Wu. 2019. “Multi-Camera Vehicle Tracking with Powerful Visual Features and Spatial-Temporal Cue.” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* 10.

Hsu, Shih Chung, Chung Lin Huang, and Cheng Hung Chuang. 2018. “Vehicle Detection Using Simplified Fast R-CNN.” *2018 International Workshop on Advanced Image Technology, IWAIT 2018* 1–3. doi: 10.1109/IWAIT.2018.8369767.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Huang, Yi Qi, Jia Chun Zheng, Shi Dan Sun, Cheng Fu Yang, and Jing Liu. 2020. "Optimized YOLOv3 Algorithm and Its Application in Traffic Flow Detections." *Applied Sciences (Switzerland)* 10(9):3079. doi: 10.3390/app10093079.

Husein, Amir Mahmud, Christopher Christopher, Andy Gracia, Rio Brandlee, and Muhammad Haris Hasibuan. 2020. "Deep Neural Networks Approach for Monitoring Vehicles on the Highway." *SinkrOn* 4(2):163. doi: 10.33395/sinkron.v4i2.10553.

Kim, Jong Bae. 2019. "Efficient Vehicle Detection and Distance Estimation Based on Aggregated Channel Features and Inverse Perspective Mapping from a Single Camera." *Symmetry* 11(10):8–10. doi: 10.3390/sym11101205.

Lai, Yanyu, Fuchun Sun, and Huaping Liu. 2020. "Small Object Detection Base on YOLOv3 for Pedestrian Recognition." Pp. 235–41 in *2020 5th International Conference on Control and Robotics Engineering, ICCRE 2020*.

Laroca, Rayson, Luiz A. Zanlorensi, Gabriel R. Gonçalves, Eduardo Todt, William Robson Schwartz, and David Menotti. 2019. "An Efficient and Layout-Independent Automatic License Plate Recognition System Based on the YOLO Detector." *ArXiv*.

Law, Hei, and Jia Deng. 2020. "CornerNet: Detecting Objects as Paired Keypoints." *International Journal of Computer Vision* 128(3):642–56. doi: 10.1007/s11263-019-01204-1.

Li, Xun, Yao Liu, Zhengfan Zhao, Yue Zhang, and Li He. 2018. "A Deep Learning Approach of Vehicle Multitarget Detection from Traffic Video." *Journal of Advanced Transportation* 2018. doi: 10.1155/2018/7075814.

Lin, Tsung Yi, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. 2020. "Focal Loss for Dense Object Detection." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42(2):318–27. doi: 10.1109/TPAMI.2018.2858826.

Liu, Fei, Zhiyuan Zeng, and Rong Jiang. 2017. "A Video-Based Real-Time Adaptive Vehicle-Counting System for Urban Roads" edited by X. Hu. *PLoS ONE* 12(11):1–16. doi: 10.1371/journal.pone.0186098.

Mahto, Pooja, Priyamm Garg, Pranav Seth, and J. Panda. 2020. "Refining Yolov4 for Vehicle Detection." *International Journal of Advanced Research in Engineering and Technology* 11(5):409–19. doi: 10.34218/IJARET.11.5.2020.043.

Meng, Qiao, Huansheng Song, Yu'An Zhang, Xiangqing Zhang, Gang Li, and Yanni Yang. 2020. "Video-Based Vehicle Counting for Expressway: A Novel Approach Based on Vehicle Detection and Correlation-Matched Tracking Using Image Data from PTZ Cameras." *Mathematical Problems in Engineering* 2020. doi: 10.1155/2020/1969408.

Mu, Kenan, Fei Hui, and Xiangmo Zhao. 2016. "Multiple Vehicle Detection and Tracking in Highway Traffic Surveillance Video Based on Sift Feature Matching." *Journal of Information Processing Systems* 12(2):183–95.

Nam Bui, Khac Hoai, Hongsuk Yi, and Jiho Cho. 2020. "A Multi-Class Multi-Movement Vehicle Counting Framework for Traffic Analysis in Complex Areas Using CCTV Systems." *Energies* 13(8). doi: 10.3390/en13082036.

Nguyen, Hoanh. 2019. "Improving Faster R-CNN Framework for Fast Vehicle Detection." *Mathematical Problems in Engineering* 2019. doi: 10.1155/2019/3808064.

Oltean, Gabriel, Camelia Florea, Radu Orghidan, and Victor Oltean. 2019. "Towards Real Time Vehicle Counting Using YOLO-Tiny and Fast Motion Estimation." *SIITME 2019 - 2019 IEEE 25th International Symposium for Design and Technology in Electronic Packaging, Proceedings* 240–43. doi: 10.1109/SIITME47687.2019.8990708.

Ouyang, Lecheng, and Huali Wang. 2019. "Vehicle Target Detection in Complex Scenes Based on YOLOv3 Algorithm." *IOP Conference Series: Materials Science and Engineering* 569(5). doi: 10.1088/1757-899X/569/5/052018.

Rublee, Ethan, Vincent Rabaud, Kurt Konolige, and Gary Bradski. 2011. "ORB: An Efficient Alternative to SIFT or SURF." Pp. 2564–71 in *2011 International Conference on Computer Vision*. Vol. 2008. IEEE.

Sang, Jun, Zhongyuan Wu, Pei Guo, Haibo Hu, Hong Xiang, Qian Zhang, and Bin Cai. 2018. "An Improved YOLOv2 for Vehicle Detection." *Sensors (Switzerland)* 18(12). doi: 10.3390/s18124272.

Schumann, Arne, Lars Sommer, Krassimir Valev, and Jurgen Beyerer. 2018. "A Systematic Evaluation of Recent Deep Learning Architectures for Fine-Grained Vehicle Classification." P. 1 in *Pattern Recognition and Tracking XXIX*, edited by M. S. Alam. SPIE.

Song, Huansheng, Haoxiang Liang, Huaiyu Li, Zhe Dai, and Xu Yun. 2019. "Vision-Based Vehicle Detection and Counting System Using Deep Learning in Highway Scenes." *European Transport Research Review* 11(1). doi: 10.1186/s12544-019-0390-4.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 6, Number 1, October 2021
DOI : https://doi.org/10.33395/sinkron.v6i1.11231
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Tan, Mingxing, and Quoc V. Le. 2019. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." *36th International Conference on Machine Learning, ICML 2019* 2019-June:10691–700.

Tan, Mingxing, Ruoming Pang, and Quoc V. Le. 2020. "EfficientDet: Scalable and Efficient Object Detection." Pp. 10778–87 in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society.

Wang, Xiaolan, Shuo Wang, Jiaqi Cao, and Yansong Wang. 2020. "Data-Driven Based Tiny-YOLOv3 Method for Front Vehicle Detection Inducing SPP-Net." *IEEE Access* 8:110227–36. doi: 10.1109/ACCESS.2020.3001279.

Wang, Xinqing, Xia Hua, Feng Xiao, Yuyang Li, Xiaodong Hu, and Pengyu Sun. 2018. "Multi-Object Detection in Traffic Scenes Based on Improved SSD." *Electronics (Switzerland)* 7(11). doi: 10.3390/electronics7110302.

Watkins, Rohan, Nick Pears, and Suresh Manandhar. 2018. "Vehicle Classification Using ResNets, Localisation and Spatially-Weighted Pooling." *ArXiv*.

Wei, Runchen, Ning He, and Ke Lu. 2020. "YOLO-Mini-Tiger: Amur Tiger Detection." Pp. 517–24 in *ICMR 2020 - Proceedings of the 2020 International Conference on Multimedia Retrieval*.

Wojke, Nicolai, Alex Bewley, and Dietrich Paulus. 2018. "Simple Online and Realtime Tracking with a Deep Association Metric." Pp. 3645–49 in *Proceedings - International Conference on Image Processing, ICIP*. Vols. 2017-Septe.

Wu, Fan, Guoqing Jin, Mingyu Gao, Zhiwei He, and Yuxiang Yang. 2019. "Helmet Detection Based on Improved YOLO V3 Deep Model." *Proceedings of the 2019 IEEE 16th International Conference on Networking, Sensing and Control, ICNSC 2019* 363–68. doi: 10.1109/ICNSC.2019.8743246.

Xiang, Xuezhi, Ning Lv, Xinli Guo, Shuai Wang, and Abdulmotaleb El Saddik. 2018. "Engineering Vehicles Detection Based on Modified Faster R-CNN for Power Grid Surveillance." *Sensors (Switzerland)* 18(7). doi: 10.3390/s18072258.

Xiang, Xuezhi, Mingliang Zhai, Ning Lv, and Abdulmotaleb El Saddik. 2018. "Vehicle Counting Based on Vehicle Detection and Tracking from Aerial Videos." *Sensors (Switzerland)* 18(8):1–17. doi: 10.3390/s18082560.

Zhang, Fukai, Ce Li, and Feng Yang. 2019. "Vehicle Detection in Urban Traffic Surveillance Images Based on Convolutional Neural Networks with Feature Concatenation." *Sensors (Switzerland)* 19(3). doi: 10.3390/s19030594.

Zhang, Shifeng, Longyin Wen, Xiao Bian, Zhen Lei, and Stan Z. Li. 2018. "Single-Shot Refinement Neural Network for Object Detection." Pp. 4203–12 in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Zhang, Xiaosong, Fang Wan, Chang Liu, Xiangyang Ji, and Qixiang Ye. 2021. "Learning to Match Anchors for Visual Object Detection." *IEEE Transactions on Pattern Analysis and Machine Intelligence*. doi: 10.1109/TPAMI.2021.3050494.

Zhao, Xia, Yingting Ni, and Haihang Jia. 2017. "Modified Object Detection Method Based on YOLO." Pp. 233–44 in.

Zhou, Jun, Yichen Tian, Chao Yuan, Kai Yin, Guang Yang, and Meiping Wen. 2019. "Improved UAV Opium Poppy Detection Using an Updated YOLOV3 Model." *Sensors (Switzerland)* 19(22):1–23. doi: 10.3390/s19224851.

Zhu, Jianxiao, Xu Li, Peng Jin, Qimin Xu, Zhengliang Sun, and Xiang Song. 2021. "MME-YOLO: Multi-Sensor Multi-Level Enhanced Yolo for Robust Vehicle Detection in Traffic Surveillance." *Sensors (Switzerland)* 21(1):1–17. doi: 10.3390/s21010027.

*name of corresponding author