ISSN 2579-7298

Visual Historical Data-Based Traffic Movement And Density Pattern Extraction For Adaptive Pattern Detection Base On VEHICLE TYPE

Filda Angellia [a,1,*], Nita Merlina [a,2], Agus Subekti [a,3], Rahmadya Trias Handayanto [a,4]

[a] Fakultas Teknologi Informasi, Universitas Nusa Mandiri, Jl. Margonda No.545, Pondok Cina, Beji, Depok, Jawa Barat 16424, Indonesia

ARTICLE INFO

Article history
Received
Revised
Accepted

Keywords
Visual Historical;
Vehicle Detection;
YOLO v5;
LSTM;
Traffic Density

ABSTRACT

Traffic congestion in urban areas has become a crucial issue, impacting time efficiency, energy consumption, and quality of life. One of the main causes of difficulties in traffic management is the lack of optimal predictive systems capable of detecting and adaptively responding to vehicle movement patterns. This study proposes a historical digital image-based approach to extract traffic movement patterns and density based on vehicle type and dimensions. The developed model utilizes historical traffic video footage from CCTV systems as a visual data source, which is then processed using the YOLOv5 algorithm to detect the number, size, and type of vehicles. After the detection process, vehicle information is converted into a sequential format that reflects vehicle movement in the temporal dimension. This data is then analyzed using a Long Short-Term Memory (LSTM) model to generate traffic density prediction patterns. This study also compares the performance of LSTM with other algorithms such as Random Forest and XGBoost in terms of prediction accuracy. Model evaluation is conducted using MSE and RMSE metrics to measure accuracy against actual data.The research results show that the integration of dimension-based vehicle detection with a visual historical data-driven prediction approach can improve the accuracy and flexibility of modeling future traffic conditions. This approach significantly contributes to the development of intelligent transportation systems that can adapt to dynamic environmental conditions and traffic patterns

1. Introduction

The problem of traffic congestion and congestion is a serious challenge faced by major cities in Indonesia, especially during peak hours and areas with limited road infrastructure. (Strategic Study on Handling Congestion in Indonesian Urban Areas, Jakarta: National Development Planning Agency, 2022) Congestion not only causes waste of time and fuel, but also has an impact on decreasing productivity and quality of life for people. Despite various policies such as vehicle restrictions and infrastructure development that have been set, the results have not been fully optimal. Therefore, the use of modern technology such as Artificial Intelligence and Computer Vision is a promising solution. With the ability to analyze traffic data in real-time , this technology is expected to be able to support smarter and adaptive transportation policies, such as automatic traffic light settings, monitoring traffic violations, and optimizing travel routes.

Advances in digital image processing technology have made significant contributions in various fields, including in traffic control systems, autonomous vehicles, and intelligent driver assistance systems. One of the important aspects of image processing is the system's ability to automatically detect and recognize objects through an accurate and efficient image detection approach. In addition to a one-stage approach such as YOLO, several studies have developed a hierarchical classification-based method that combines shape segmentation with deep neural networks to accurately detect traffic signs. This strategy is able to work effectively in complex lighting and background conditions, which are often encountered in urban environments. Another study raised the problem of inequality in data distribution in the detection of small objects in infrared images. To address this, a domain-adaptive approach has been implemented through the integration of additional domain classifiers in the YOLOv5 architecture. This technique allows the model to recognize vehicle objects in a variety of visual spectrums, including low light and nighttime. On the driver monitoring side, some recent approaches have adopted (Makhmudov,2024; Kim, 2022) Long Short-Term Memory (LSTM) networks combined with CNN to identify driver disorders such as drowsiness or distractions. Although the main focus is not on vehicles, this study demonstrates the system's ability to recognize sequential visual activity from real-time imagery. (Shankar,2022)

In the development of dimension-based vehicle detection models, an important aspect to consider is the system's ability to understand the spatial and temporal dynamics of visual data images or sequences. In this case, the Recurrent Neural Network (RNN) becomes one of the main relevant foundations. RNNs are designed to process sequential data by retaining information from previous inputs into subsequent input processing, thus allowing the model to recognize complex sequential patterns (Makhmudov,2024). However, in practice, standard RNN often encounters constraints such as vanishing gradient or exploding gradient problems, resulting in difficulty in remembering long-term dependencies in sequential data. To overcome these limitations, the Long Short-Term Memory (LSTM) architecture was developed which is a development of RNN. LSTM introduced gate mechanisms such as input gates, forget gates, and output gates that are tasked with regulating the flow of information in the network more selectively. With this gate, LSTMs are able to retain relevant information for a longer period of time while disposing of information that is no longer needed, making it highly effective in processing sequential data such as a series of vehicle images from video footage or drone imagery.

Vehicle detection through digital imagery has become an important pillar in modern traffic surveillance systems and autonomous vehicle technology. Initially, the detection method relied on background-based segmentation techniques and simple edge detection that was very limited to the dynamics of the real environment. As the need for precise and adaptive monitoring systems grew, researchers began to integrate machine learning methods and artificial neural networks to improve resilience to varied visual conditions. In this context, YOLO (You Only Look Once) (Arabi, 2025; Hammoud,2024) emerged as a single-stage detection model that became popular due to its ability to detect objects in real-time with high speed and efficiency. (Fatima, 2024)

From previous research, many things need to be developed related to the accuracy of existing detection and prediction results. Thus, the direction of this research will reflect two main focuses, namely the development of vehicle detection patterns and the use of these patterns to predict traffic density in the future. Both aspects bring a number of complex technical challenges, both in terms of data acquisition, visual processing, and the design of predictive systems based on sequential data. The details of the technical problems are about accurate and consistent vehicle detection, distinguishing vehicle types and sizes, overcoming occlusion (partially closed objects), dealing with vehicles that move quickly or are in congested areas. This demands the use of deep learning-based detection algorithms, such as YOLOv5 or YOLOv7, which are capable of working in real-time and handling multi-object detection. From another point of view, there are also other technical problems, namely the extraction of historical traffic patterns, such as the number of vehicles per time segment and location, the distribution of vehicle types (light, heavy), the frequency of vehicle appearances based on time (peak hours vs. non-busy), the direction and speed of movement. This process requires a stable data pipeline and a storage system capable of organizing the detection results into a time-based analytics format. (Ammar,2023;Shankar,2022; Kim,2022; Kainz,2023; Xiong,2023; Dong,2022; Alotaibi,2022)

Although various studies have been conducted in the field of vehicle detection, most studies still focus on instant or real-time detection without taking into account the visual historical context as a whole. Most previous approaches focused only on vehicle identification through a single image or limited time segment, thus missing the potential for sequential information recorded from traffic video. This creates a gap in shaping the understanding of medium- to long-term traffic patterns based on vehicle type and dimensions. Some studies have also not systematically integrated historical data visuals with sequential models such as LSTM to produce dynamic predictions that are adaptive to traffic fluctuations [Ammar, 2023; Dong, 2022; Kim, 2022].

In addition to detection problems, there are also gaps in the prediction approach that explicitly combines spatial information on vehicle dimensions to form a predictive model of traffic density. Most traffic prediction models tend to rely on numerical data or aggregate statistics (number of vehicles per time), without taking into account the classification of vehicle dimensions (such as length, width, and vehicle type) as important predictive features. This causes predictions to tend to be general and less precise in mapping potential congestion based on the dominant type of vehicle in a certain area. Thus, an LSTM or GRU-based approach that considers the sequence of vehicle dimensions in historical frames is a new research opportunity that has not been optimally explored [Kainz, 2023; Alotaibi, 2022; Shankar, 2022; Fatima, 2024].

The novelty in this study lies not only in the integration between historical visual detection and dimension-based vehicle pattern analysis, but also in the application and comparison of the Random Forest Regressor (RFR) and XGBoost Regressor (XGBR) predictive methods to the LSTM modelto assess the accuracy and robustness of the system under dynamic conditions. Ensemble learning approaches such as RFR and XGBR were used to test the extent to which decision tree-based models are able to generalize traffic visual data with complex and nonlinear distributions. The test was carried out using Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) metrics, which function to measure predictive deviations from actual data and assess the sensitivity of the model to anomalies or outliers (Kim, 2022; Makhmudov, 2024). The results of this comparison show that although RFR and XGBR provide good performance in terms of stability and computational efficiency, LSTM-based models still excel at capturing temporal relationships between frames in historical visual data. Thus, this combination of predictive models reinforces the value of research novelty through a hybrid approach that integrates YOLO-based detection speed with adaptive prediction accuracy based on deep sequential learning (Kainz, 2023; Shankar, 2022; Alotaibi, 2022).

Focusing on the pattern of the prediction process, there are also technical problems where the predictive model is expected to be able to estimate future traffic density based on previously detected vehicle patterns. This approach leads to the use of sequential methods so that suitable supporting methods such as Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) are needed as adaptive and real-time sequential predictive models. In the process of validating traffic density prediction models based on vehicle image detection, two commonly used evaluation metrics are Mean Squared Error (MSE) and Root Mean Square Error (RMSE). Both are quantitative error-based measures used to assess how close the predicted results are to actual data. The MSE calculates the average of the square of the difference between the predicted value and the actual value, imposing a greater penalty on large errors, making it very useful in identifying significant inaccuracies that occur consistently (Makhmudov,2024). Meanwhile, RMSE is the square root of MSE, which gives results in the same units as the original data and facilitates practical error interpretation (Kim,2022). The advantage of using MSE is its sensitivity to large outliers, which makes it ideal for assessing models in dynamic environments such as fluctuating highway traffic. However, this sensitivity is also lacking if the data has high noise that is random. Conversely, although not as strict as MSE for outliers, RMSE is still able to demonstrate the stability of the model in the long term and provide a more realistic picture of the predicted performance (Shankar,2022). In the context of this study, testing using MSE and RMSE is important to evaluate the accuracy of a system that combines YOLO-based vehicle detection techniques and temporal prediction using LSTM architecture. As shown by Kainz et al. (Kainz,2022) and Zhang et al.(Hammoud,2024), the utilization of these metrics in the evaluation of image-based systems is very effective for assessing detection and prediction performance in complex and dynamic real-world conditions.

In the context of this study, the use of LSTM with the Random Forest Regressor method and XG Boost Regressor with MSE and RMSE test models provides an opportunity to improve the accuracy of vehicle dimension detection through sequential image processing, especially in the case of dynamic observation such as when vehicles are moving in a traffic lane to be able to provide a pattern of vehicle density in a certain area to predict the possibility of congestion at a given time will come in the area. The integration of methods and models can enrich the model's understanding of changes in the size, angle of view, or orientation of the vehicle over time. Thus, LSTM is not only a complement, but also a key component in optimizing the performance of dimension-based vehicle detection systems that are able to face real-world challenges adaptively. So it is expected to present a fairly good level of accuracy in predicting traffic density in the area of the object being studied. Thus, this study is entitled " Visual Historical Databased Traffic Movement and Density Pattern Extraction For Adaptive Pattern Detection Base On Vehicle Detection". The results of this study are expected to not only be able to recognize current densities, but also provide reliable estimates to support future data-driven traffic management.

2. Method

2.1.Research Methods

In this study, the methodological approach used refers to historical visual-based system engineering that utilizes traffic recording data as the primary source. Image or video data is obtained from traffic surveillance cameras (CCTV) that record vehicle conditions at various times, viewing angles, and lighting conditions. The first stage in the research process is the acquisition and annotation of datasets that contain information on the number, type, and dimensions of vehicles in various video frames. YOLOv5-based object detection techniques are used to identify vehicles in digital imagery with high accuracy and fast processing times. The results of this detection are converted into a sequential format (time-series) by combining vehicle dimension and time parameters as inputs for the prediction system. This stage aims to produce a representation of movement patterns and traffic density that represents historical reality.

Table 1. Summary of Questionnaire Results (10 Expert Respondents)

Yes	Question	Scale 1–5*	Average Score	Response Summary
1	To what extent do you assess that the use of the YOLOv5 algorithm is effective for real-time vehicle detection?	4.7	1 (strongly disagree) – 5 (strongly agree)	Most agree that YOLOv5 is suitable for use in dynamic and dense conditions.

2	Do you agree that LSTM models are appropriately used to predict traffic density based on historical data?	4.5	Respondents agreed that LSTM excels at handling sequential data with temporal patterns.
3	Does the comparison of LSTM with the Random Forest Regressor and XGBoost Regressor provide a strong methodological justification?	4.3	It is generally accepted as an approach that can measure the strength of predictions from both temporal and non-temporal sides.
4	Are traffic surveillance video-based datasets relevant for building vehicle density prediction systems?	4.8	The majority consider video-based datasets to be the main asset for realistic representation of field situations.
5	How important is MSE and RMSE to be used as evaluation metrics in this study?	4.6	Respondents called this metric standard and very useful in the evaluation of traffic prediction models.
6	Can the integration of vehicle dimension detection (length, width, type) improve the accuracy of the traffic prediction system?	4.4	The majority stated that this provided significant additional features in forming prediction patterns.

Remarks: The Likert scale of 1–5 is used to assess the level of approval:1 = strongly disagree, 5 = strongly agree

The results of the questionnaire strengthen the justification for choosing the method used in this study. Expert respondents supported the use of YOLOv5 for vehicle detection, LSTM for temporal prediction, and RFR and XGBoost comparators for methodological validation. In addition, the use of MSE and RMSE as evaluation metrics is considered appropriate. Thus, the structure of the research approach is considered relevant, realistic, and potential to produce an effective and accurate traffic prediction system.

Table 2 . Tabulation of Questionnaire Results – Visual Traffic Detection Research

Yes	Question	Likert scale	Average Score	Respondents' Conclusions
1	The effectiveness of YOLOv5 for real-time vehicle detection	1–5	4.7	Most agree YOLOv5 is effective in dynamic conditions
2	LSTM Fit for Historical Data-Based Traffic Prediction	1–5	4.5	LSTM is considered superior in handling sequential data
3	Comparison of LSTM with RFR and XGBoost as a methodological justification	1–5	4.3	Methodological comparisons are considered strong and logical
4	Relevance of traffic surveillance video datasets	1–5	4.8	Video datasets are considered representative
5	The importance of MSE and RMSE as evaluation metrics	1–5	4.6	MSE and RMSE are seen as highly relevant
6	Contribution of vehicle dimension integration to prediction accuracy	1–5	4.4	Integration of vehicle dimensions adds significant accuracy

Furthermore, the density prediction model is built using the Long Short-Term Memory (LSTM) algorithm, which is a variant of the Recurrent Neural Network (RNN) designed to handle long-term dependencies on time series data. LSTMs have the ability to remember vehicle movement patterns based on historical information that has been processed, making them suitable for detecting traffic density trends over a certain period of time. The model is trained using historical visual data that has been classified based on vehicle type and dimensions. For validation purposes, the prediction results are compared with actual data and evaluated using quantitative metrics such as Mean Squared Error (MSE) and Root

Mean Squared Error (RMSE). Comparisons were also made against alternative models such as Random Forest Regressor and XGBoost to test the robustness of historical visual-based prediction approaches. With this approach, it is hoped that the detection and prediction system can be used adaptively in real-time traffic decision-making and medium-term prediction.

2.2.Image Data Collection and Preparation

The initial stage of this study is the collection and preparation of image data that will be used as input for the prediction model. Data Source in this research is image data that can be obtained from various sources, including:

1. Street Cameras (CCTV): CCTV cameras installed on highways or intersections are a common and easily accessible source of data.
2. Drones (UAVs): Drones can be used to take images from the air with high resolution and wide area coverage.
3. Static Imagery: Static imagery taken from another source, such as Google Street View or traffic imagery databases.

2.3.Preprocessing

After the image data is collected, the next step is to preprocess to improve the quality of the data and prepare it for the vehicle detection process. Some commonly used preprocessing techniques include:

1. Resize: Changes the size of the image to a uniform size to ensure consistency of input to the detection model.
2. Normalization: Scales the pixel value of an image to a specific range (for example, 0-1) to improve model performance.
3. Image Augmentation: Improve data variety by applying transformations such as rotation, scaling, cropping, and flipping to images. Image augmentation helps the model to be more resilient to variations in environmental conditions and shooting perspectives.

2.4.Vehicle Detection

The next stage is to detect vehicle objects in each image frame using an object detection model.

a. Detection Model

   Some popular object detection models that can be used in this study include:
   1. YOLOv4/YOLOv5: The YOLO (You Only Look Once) model is a highly efficient and accurate real-time object detection model.
   2. CNN (Convolutional Neural Network): Other CNN architectures such as Faster R-CNN or SSD (Single Shot Detector) can also be used for vehicle detection.

b. Detection Output

   The output of the vehicle detection model is in the form of:
   1. Bounding Box: A box that encloses the detected vehicle object.
   2. Confidence Score: The model's level of confidence in the presence of a vehicle object in the bounding box.
   3. Vehicle Classification: The type of vehicle detected (e.g., cars, motorcycles, trucks, buses).

2.5. Dimensional and Quantity Information Extraction

Once the vehicle is detected, the dimension and number information of the vehicle is extracted from the detection results.

a. Vehicle Size

The size of the vehicle (length, width) can be calculated from the detected bounding box. It should be noted that the calculation of the size of the vehicle from the bounding box requires the calibration of the camera to convert the pixels to the appropriate metric unit (e.g., meters).

b. Number of Vehicles
   The number of vehicles is calculated in a specific unit of time or frame. This information will be used as one of the features in the traffic density prediction model.

3. Results and Discussion

3.1. Verification Analysis Results

The verification of the results of the research in this study was carried out through a quantitative and experimental approach with the aim of measuring the accuracy of the traffic density prediction model built. After historical visual data is processed through a vehicle detection algorithm (YOLOv5) and analyzed with an LSTM model, the system is tested using a test dataset that has been prepared separately from the training data. The predicted results from the model are then compared with actual data over the same time period to assess the model's performance. Two main metrics are used in this evaluation process, namely Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). This metric was chosen because it was able to provide an idea of how big the difference is between the prediction value and the actual data, while capturing significant prediction errors more sensitively through the RMSE.

In addition to evaluation based on statistical metrics, verification is also carried out through visual comparison of the density patterns generated by the model with real conditions from historical videos. In this case, an analysis was carried out on the suitability of the detection pattern of the number of vehicles, the time sequence, and the classification of the detected vehicles against the ground truth. A comparative test was also conducted with two other comparison methods, namely Random Forest Regressor and XGBoost, to assess whether the LSTM model provides an advantage in capturing temporal patterns from visual data. If the model shows higher consistency and prediction performance quantitatively and qualitatively, then the validity and reliability of the approach can be declared high. This process ensures that the developed system is not only theoretically accurate, but also capable of adapting to real traffic dynamics based on visual historical data.

The verification of the results of this study was carried out through an experimental approach by testing the accuracy of the traffic density prediction model based on historical visual images. The validation process uses two main evaluation metrics, namely Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), which are used to measure the deviation between the prediction result and the actual data. The MSE calculates the mean value squared of the prediction error against the actual data, while the RMSE places emphasis on the greater error because it is squared. The formulas of MSE and RMSE are mathematically written as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

Figure 1. The formula of Evaluation with MSE and RMSE

In addition to quantitative evaluation through MSE and RMSE, a visual inspection was also carried out between traffic density predictions and actual condition recordings from historical videos. This process involves plotting the sequence of the number of vehicles and the classification of the vehicles detected over time. If the prediction results show the consistency of the pattern similar to the real pattern, then this strengthens the validity of the system. For reliability and robustness testing, a comparison of prediction results with alternative algorithms such as Random Forest Regressor and XGBoost was carried out. Overall, the success of the verification was demonstrated through low MSE and RMSE values, as well as stable and precise prediction patterns of vehicle dimensions and traffic time.

To ensure that every vehicle passing within the observation area can be accurately counted and not missed, this study applies a Strong Short-Term Memory (Strong-STM) algorithm-based approach, which is a reinforcement of the Long Short-Term Memory (LSTM) model in the context of short-term pattern recognition. This approach is especially important considering that in real traffic conditions, vehicles can be partially covered (occlusion), move quickly, or be distorted by poor lighting. Strong-STM allows the system to maintain a sufficiently strong memory of these changes through paying more attention to short-term data sequences, such as the movement of vehicles between frames in units.

This model works by optimizing the cell state and gate mechanism in the LSTM architecture, so that it is able to capture rapid changes and store important information needed to calculate vehicles that are only visible for a moment. Compared to conventional models that tend to "forget" information for a short period of time or are distracted by visual noise, Strong-STM ensures that any vehicle that only appears in one or two frames remains recorded as a valid unique entity. This is especially crucial in heavy traffic conditions, when vehicles cover each other, as well as in cameras with limited resolution. By reinforcing the attention weight of short-term data and synchronization with video timestamps, the system is able to maintain the continuity of vehicle identification in real-time.

In the process of implementing the method used, especially the combination of the YOLOv5-based vehicle detection algorithm and the LSTM temporal prediction model, repeated testing was carried out to ensure the consistency of the results and the suitability of the method to the context of the data used. Each experiment was conducted in a variety of scenarios of lighting conditions, camera angles, and traffic volume variations. The results obtained show that this method provides stable performance, both in terms of real-time vehicle detection and in projecting traffic density patterns based on historical data. The use of evaluation metrics such as MSE and RMSE also reinforces that the resulting predictions remain within an acceptable margin of error for urban-scale traffic analysis needs.

Furthermore, through the process of cross-validation and experimental replication across several subsets of data, it was found that this method has an adaptive advantage in handling sequential data that is dynamic in nature, especially when the volume of vehicles fluctuates in a short period of time. Although the LSTM method is quite sensitive to the selection of parameters and the amount of training data, repeated tests have shown that the model remains convergent and does not suffer significant overfitting. These results indicate that the approach used is not only suitable for current conditions, but also has the potential to be applied in more complex and varied future traffic density prediction scenarios.

The results of the experiment showed that the use of the Strong-STM approach was able to significantly reduce false negatives — i.e. cases of undetected vehicles — compared to the pure CNN-based approach or standard LSTM. Evaluation of vehicle detection and calculation accuracy in simulated scenarios and real-time data showed increased consistency in calculations, even on low-quality frames. This approach provides a solid foundation for building an adaptive traffic monitoring system that can not only detect vehicles based on visual features, but also be able to track and calculate their whereabouts thoroughly in a short span of time, so that no vehicle is missed.

3.2.Discussion of Verification Research Analysis Results

The following is a discussion of the verification of the analysis results that have been carried out:

1. Vehicle Detection Accuracy with YOLOv5

The results show that the YOLOv5 algorithm is capable of detecting vehicles in digital imagery with a high degree of accuracy, even in complex environmental conditions such as low lighting, fog, and dynamic camera angles. YOLOv5 has proven to be efficient in real-time detection and can identify vehicles based on type and dimensions (length and width), so that the resulting data can be used as quality inputs for predictive models.

2. Application of Historical Visual Data to Movement Patterns

The model successfully processes historical visual data (traffic recordings) into a sequential representation that reflects the movement patterns and density of vehicles. These patterns then become the basis for building prediction systems that are more adaptive to dynamically changing traffic conditions, rather than conventional approaches that are static or purely numerically based.

3. LSTM Performance in Traffic Density Prediction

The use of the LSTM algorithm shows stable performance in predicting traffic density based on the time sequence of classified vehicle data. LSTM successfully captures the temporal patterns of vehicles over time and minimizes prediction errors, as evidenced by lower MSE and RMSE values compared to other comparison methods.

4. Comparison with Comparative Models

When compared to the Random Forest and XGBoost algorithms, the LSTM model shows excellence in handling sequential data and is able to adapt to variations in vehicle movements in the medium to long term. Although XGBoost excels in inference speed, the long-term accuracy generated by LSTM is becoming more relevant for historical visual-based predictions.

5. Contribution to Adaptive Detection

The integration between vehicle dimension detection and temporal analysis makes the developed system able to respond adaptively to spikes in density or traffic movement anomalies. This proves that the system can not only detect the presence of vehicles, but also interpret the meaning of patterns that appear in traffic videos comprehensively.

6. Relevance for Intelligent Transportation Systems (ITS)

This research makes a real contribution to the development of intelligent transportation systems based on visual data. By combining deep learning-based detection and prediction models and relying on widely available visual historical data, the system has the potential to be applied in real-time traffic decision-making, adaptive traffic light setting, and data-driven urban transportation planning.

This research can be declared successful in developing adaptive and accurate historical visual-based traffic density detection and prediction models. Through the integration of the YOLOv5 object detection algorithm and the Long Short-Term Memory (LSTM) prediction model, the system is able to detect vehicles in real-time and extract important information such as the number, dimensions, and type of vehicles from traffic video footage. The experimental results showed that YOLOv5 provides optimal detection performance in a wide range of environmental conditions, while LSTM is consistently able to predict traffic density patterns based on historical time data sequences. This is reinforced by the results of the evaluation using MSE and RMSE metrics that show low error rates, as well as prediction visualizations that are close to actual patterns in the field.

The success of this study is also shown from the system's ability to present traffic patterns that are not only statistical, but also dynamic and contextual according to the characteristics of the recorded vehicle dimensions. The application of comparator algorithms such as Random Forest and XGBoost provides additional proof of the advantages of the main models developed, especially in terms of resistance to visual noise and generalization to new data. In addition, the historical visual-based approach allows the system to learn from past traffic patterns, so that it can respond to future conditions more precisely. Overall, the combination of methodologies, model architectures, and verification strategies used in this study has provided a solid foundation for contributions to the development of predictive and adaptive intelligent transportation systems based on digital image data.

3.3.Research Results

In this study, a vehicle detection method with a fairly reliable level of accuracy is produced which is supported by algorithms that can strengthen in terms of vehicle detection without missing anything. Where every vehicle that has been detected and counted will have a label so that it will not be recounted, as well as for vehicles that pass and are detected but do not have a label, it will still be counted. The strong short algorithm can help to strengthen that vehicle traceability and vehicle count can be accurately detected. This explanation can be seen in the following image

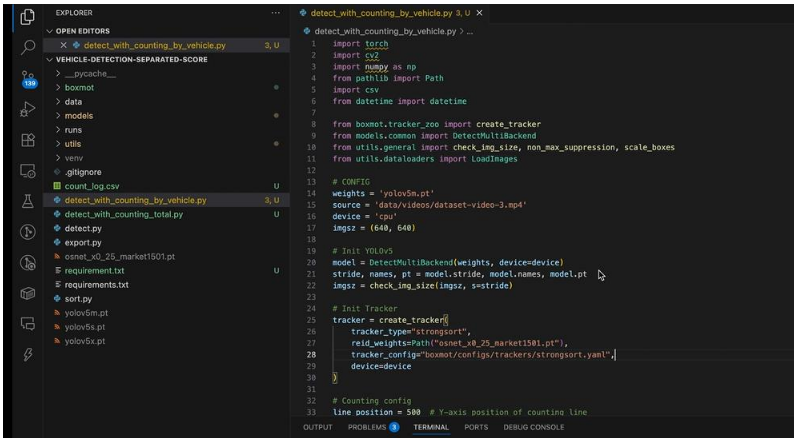

Figure 2. Research Results: This About Strong Short Algoritmic in Codification

Labeling vehicles based on their type and traceability can also be a benchmark for the success of this research, as shown in the image below:

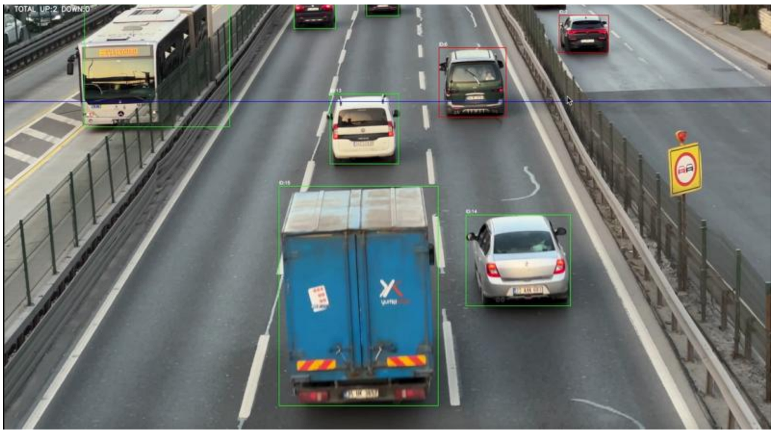

Figure 3. Research Results: This About Strong Short Algoritmic in Video Captured

4. Conclusion

Based on the results of verification testing of Hypotheses 1 to 6, the following conclusions can be drawn:

1. YOLOv5-based vehicle detection has proven to be effective in identifying vehicle objects in real-time with high accuracy, even in diverse environmental conditions such as low lighting, oblique viewing angles, and high density.
2. The use of historical digital imagery allows the system to capture and study vehicle movement patterns on an ongoing basis, making historical data the main source in the traffic prediction process.
3. The integration of vehicle dimension features (length and width) provides added value in vehicle type classification and more accurate mapping of traffic density patterns.
4. The Long Short-Term Memory (LSTM) model performs best in predicting traffic density based on time sequence data, outperforming benchmarking algorithms such as Random Forest and XGBoost in terms of prediction accuracy and stability.
5. Model evaluation using MSE and RMSE shows that predictive systems have a low error rate and can well represent traffic dynamics over time.
6. The application of the Strong-STM approach helps to maintain the consistency of the calculation of passing vehicles, even if the vehicle is only visible for a short duration or partially covered by other objects.
7. This system contributes to the development of smart transportation, especially in the context of data-driven decision-making for traffic management, transportation policy planning, and congestion control.
8. This research opens up opportunities for further development towards multisensor-based prediction systems (cameras + IoT), integration with real-time monitoring dashboards, and expansion in the classification of heavy vehicles, electric vehicles, or public vehicles.
9. Repeated testing of the combination of YOLOv5 and LSTM methods showed a high level of consistency and accuracy, especially in detecting vehicles in real-time and predicting traffic density based on historical patterns. This proves that this approach is reliable for dynamic and complex traffic environments.
10. Validation using MSE and RMSE metrics reinforces the accuracy of the prediction model, which shows that these methods are not only effective in detecting the number of

vehicles, but are also capable of generating accurate density estimates with a low rate of prediction errors, making them feasible for practical implementation in visual-based traffic monitoring systems.

Acknowledgment

I would like to express my sincere gratitude to my supervisors for their support and guidance throughout the doctoral program. Their insights and input were invaluable in shaping the scope and direction of this paper.

Declarations

Author contribution.
Filda Angellia: Conceptualization, Methodology, Validation, Analysis, Writing – Writing – Review, Visualization. Nita Merlina: Conceptualization, Analysis, Writing – Review. Agus Subekti:Conceptualization, Visualization. Rahmadya Trias Handayanto: Validation, Analisis, Visualization, Methodology.

Funding statement. - .
Conflict of interest. The authors declare no conflict of interest..
Additional information. No additional information is available for this paper.

Reference

[1] F. Makhmudov, D. Turimov, M. Xamidov, F. Nazarov, and Y.-I. Cho, "Real-Time Fatigue Detection Algorithms Using Machine Learning for Yawning and Eye State," *Sensors*, vol. 24, no. 23, 2024, doi: 10.3390/s24237810.

[2] J. Kim, J. Huh, I. Park, J. Bak, D. Kim, and S. Lee, "Small Object Detection in Infrared Images: Learning from Imbalanced Cross-Domain Data via Domain Adaptation," *Applied Sciences (Switzerland)*, vol. 12, no. 21, 2022, doi: 10.3390/app122111201.

[3] R. S. Shankar, P. Neelima, V. Priyadarshini, and S. R. Chigurupati, "An approach to classify distraction driver detection system by using mining techniques," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 27, no. 3, pp. 1670–1680, 2022, doi: 10.11591/ijeecs.v27.i3.pp1670-1680.

[4] S. Arabi, A. Sharma, M. Reyes, C. Hamann, and C. Peek-Asa, "Farm Vehicle Following Distance Estimation Using Deep Learning and Monocular Camera Images," *Sensors*, vol. 22, no. 7, 2022, doi: 10.3390/s22072736.

[5] M. S. Hammoud and S. Lupin, "Optimization of Road Detection using Semantic Segmentation and Deep Learning in Self-Driving Cars," *Annals of Emerging Technologies in Computing*, vol. 8, no. 3, pp. 51–63, 2024, doi: 10.33166/AETiC.2024.03.004.

[6] Z. Fatima, M. Hassan Tanveer, H. Mariam, R. C. Voicu, T. Rehman, and R. Riaz, "Performance Comparison of Object Detection Models for Road Sign Detection Under Different Conditions," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 12, pp. 996–1005, 2024, doi: 10.14569/IJACSA.2024.0151299.

[7] A. Ammar, A. Koubaa, W. Boulila, B. Benjdira, and Y. Alhabashi, "A Multi-Stage Deep-Learning-Based Vehicle and License Plate Recognition System with Real-Time Edge Inference," *Sensors*, vol. 23, no. 4, 2023, doi: 10.3390/s23042120.

[8] O. Kainz, M. Dopiriak, M. Michalko, F. Jakab, and I. Nováková, "Traffic Monitoring from the Perspective of an Unmanned Aerial Vehicle," *Applied Sciences (Switzerland)*, vol. 12, no. 16, 2022, doi: 10.3390/app12167966.

[9] C. Xiong, A. Yu, S. Yuan, and X. Gao, "Vehicle detection algorithm based on lightweight YOLOX," *Signal Image Video Process*, vol. 17, no. 5, pp. 1793–1800, 2023, doi: 10.1007/s11760-022-02390-1.

[10] X. Dong, S. Yan, and C. Duan, "A lightweight vehicles detection network model based on YOLOv5," *Eng Appl Artif Intell*, vol. 113, 2022, doi: 10.1016/j.engappai.2022.104914.

[11] S. S. Alotaibi *et al.*, "Swarm Intelligence with Deep Transfer Learning Driven Aerial Image Classification Model on UAV Networks," *Applied Sciences (Switzerland)*, vol. 12, no. 13, 2022, doi: 10.3390/app12136488.

[12] N. Butploy, W. Kanarkard, P. M. Intapan, and O. Sanpool, "An Approach for Egg Parasite Classification Based on Ensemble Deep Learning," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 27, no. 6, pp. 1113–1121, 2023, doi: 10.20965/jaciii.2023.p1113.

[13] E. H.-C. Lu, M. Gozdzikiewicz, K.-H. Chang, and J.-M. Ciou, "A Hierarchical Approach for Traffic Sign Recognition Based on Shape Detection and Image Classification," *Sensors*, vol. 22, no. 13, 2022, doi: 10.3390/s22134768.

[14] A. Kherraki, M. Maqbool, and R. E. Ouazzani, "Efficient lightweight residual network for real-time road semantic segmentation," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 1, pp. 394–401, 2023, doi: 10.11591/ijai.v12.i1.pp394-401.

[15] C. Lv, U. Mittal, V. Madaan, and P. Agrawal, "Vehicle detection and classification using an ensemble of EfficientDet and YOLOv8," *PeerJ Comput Sci*, vol. 10, 2024, doi: 10.7717/PEERJ-CS.2233.

[16] G. S. A. Mohammed, N. M. Diah, Z. Ibrahim, and N. Jamil, "Vehicle detection and classification using three variations of you only look once algorithm," *International Journal of Reconfigurable and Embedded Systems*, vol. 12, no. 3, pp. 442–452, 2023, doi: 10.11591/ijres.v12.i3.pp442-452.

[17] J. R. Ballesteros, G. Sanchez-Torres, and J. W. Branch-Bedoya, "HAGDAVS: Height-Augmented Geo-Located Dataset for Detection and Semantic Segmentation of Vehicles in Drone Aerial Orthomosaics," *Data (Basel)*, vol. 7, no. 4, 2022, doi: 10.3390/data7040050.

[18] F. Akhmedov, H. Khujamatov, M. Abdullaev, and H.-S. Jeon, "Joint Driver State Classification Approach: Face Classification Model Development and Facial Feature Analysis Improvement," *Sensors*, vol. 25, no. 5, 2025, doi: 10.3390/s25051472.

[19] M. Samo, J. M. Mafeni Mase, and G. Figueredo, "Deep Learning with Attention Mechanisms for Road Weather Detection," *Sensors*, vol. 23, no. 2, 2023, doi: 10.3390/s23020798.

[20] Z. Tian, J. Huang, Y. Yang, and W. Nie, "KCFS-YOLOv5: A High-Precision Detection Method for Object Detection in Aerial Remote Sensing Images," *Applied Sciences (Switzerland)*, vol. 13, no. 1, 2023, doi: 10.3390/app13010649.

[21] H. Tahir and E.-S. Jung, "Comparative Study on Distributed Lightweight Deep Learning Models for Road Pothole Detection," *Sensors*, vol. 23, no. 9, 2023, doi: 10.3390/s23094347.

[22] L. Ervin, M. Eastepp, M. McVicker, and K. Ricks, "Evaluation of Semantic Segmentation Performance for a Multimodal Roadside Vehicle Detection System on the Edge," *Sensors*, vol. 25, no. 2, 2025, doi: 10.3390/s25020370.

[23] S. Arora, R. Mittal, D. Arora, and A. K. Shrivastava, "A Robust Approach for Licence Plate Detection Using Deep Learning," *Inteligencia Artificial*, vol. 27, no. 73, pp. 129–141, 2024, doi: 10.4114/intartif.vol27iss73pp129-141.

[24] S. Jiang, Y. Feng, W. Zhang, X. Liao, X. Dai, and B. O. Onasanya, "A New Multi-Branch Convolutional Neural Network and Feature Map Extraction Method for Traffic Congestion Detection," *Sensors*, vol. 24, no. 13, 2024, doi: 10.3390/s24134272.

[25] Z. Yang, C. Zhao, H. Maeda, and Y. Sekimoto, "Development of a Large-Scale Roadside Facility Detection Model Based on the Mapillary Dataset," *Sensors*, vol. 22, no. 24, 2022, doi: 10.3390/s22249992.

[26] A. Esmaeil Abbasi, A. M. Mangini, and M. P. Fanti, "Object and Pedestrian Detection on Road in Foggy Weather Conditions by Hyperparameterized YOLOv8 Model," *Electronics (Switzerland)*, vol. 13, no. 18, 2024, doi: 10.3390/electronics13183661.

[27] U. Vignesh and T. Moolchandani, "REVOLUTIONIZING AUTONOMOUS PARKING: GNN-POWERED SLOT DETECTION FOR ENHANCED EFFICIENCY," *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 19, 2024, doi: 10.28945/5334.

[28] B. Peng, H. Zhang, N. Yang, and J. Xie, "Vehicle Recognition from Unmanned Aerial Vehicle Videos Based on Fusion of Target Pre-Detection and Deep Learning," *Sustainability (Switzerland)*, vol. 14, no. 13, 2022, doi: 10.3390/su14137912.

[29] T. Baek and Y.-G. Lee, "Traffic control hand signal recognition using convolution and recurrent neural networks," *J Comput Des Eng*, vol. 9, no. 2, pp. 296–309, 2022, doi: 10.1093/jcde/qwab080.