Traffic Flow Forecast based on Vehicle Count

By Pavanee Weebadu Liyanage & K. P. G. C. D. Sucharitharathna

Curtin University Perth

Abstract- Real-time traffic predictions have now become a time-being need for efficient traffic management due to the exponentially increasing traffic congestion. In this paper, a more pragmatic traffic management system is introduced to address traffic congestion, especially in countries such as Sri Lanka where there is no proper traffic monitoring database. Here the real-time traffic monitoring is performed using TFmini Plus light detection and ranging (LiDAR) sensor and vehicle count for next five minutes will be predicted by feeding consecutively collected data into the LSTM neural network. More than ten separate prediction models were trained, varying both window size and the volume of input data delivered to train the models. Since the accuracy results of all prediction models were above 70%, it demonstrates that this system can produce accurate predictions even if it is trained using less input data collection. Similarly the sensor accuracy test also resulted in 89.7% accuracy.

Keywords: traffic monitoring, lstm neural network traffic predictions, vehicle count, traffic flow forecast, real- time traffic monitoring, lidar sensor traffic monitoring.

GJCST-D Classification: LCC: TE175-178

TRAFFICFLOWFORECASTBASEDONVEHICLECOUNT

Strictly as per the compliance and regulations of:

Traffic Flow Forecast based on Vehicle Count

Pavanee Weebadu Liyanage [α] & K. P. G. C. D. Sucharitharathna [σ]

Abstract- Real-time traffic predictions have now become a time-being need for efficient traffic management due to the exponentially increasing traffic congestion. In this paper, a more pragmatic traffic management system is introduced to address traffic congestion, especially in countries such as Sri Lanka where there is no proper traffic monitoring database. Here the real-time traffic monitoring is performed using TFmini Plus light detection and ranging (LiDAR) sensor and vehicle count for next five minutes will be predicted by feeding consecutively collected data into the LSTM neural network. More than ten separate prediction models were trained, varying both window size and the volume of input data delivered to train the models. Since the accuracy results of all prediction models were above 70%, it demonstrates that this system can produce accurate predictions even if it is trained using less input data collection. Similarly the sensor accuracy test also resulted in 89.7% accuracy.

Keywords: traffic monitoring, lstm neural network traffic predictions, vehicle count, traffic flow forecast, real- time traffic monitoring, lidar sensor traffic monitoring.

I. Introduction

Vehicle usage has risen significantly throughout the world in recent years. As a result, road traffic has gotten more complex and chaotic. Traffic congestion is becoming a more highlighted issue in cities throughout the world, owing to continuous urbanization and population increase. According to the Victoria Transport Policy Institute's Congestion Costing Critique (CCC) Critical Evaluation of the "Urban Mobility Report" (UMR), which was published on 1 September 2021 [1], congestion cost is estimated to cost between $130 to $500 per capita annually, particularly in comparison to $2,000 in crash damages, $3,000 in vehicle ownership costs and $1,800, $600, $400 respectively in parking, pollution damage, and roadway costs. Similarly, the UMR estimates that by 2025, congestion cost will have risento $200 billion [1].

According to the statistical records by The National Highway Traffic Safety Administration (NHTSA) [2], in 2013, more than 2.3 million injuries and 32,719 deaths were recorded due to vehicular accidents. Among those recorded fatalities majority were aged between 4–27 [2]. According to the NHTSA report, the direct economic cost and the social impact that occurs due to vehicular accidents are accounted for $871 billion per year, with an average of 5.8 million collisions each year only on the United State's highways [2]. NHTSA statistical records also emphasize that most of road accidents occur as a result of traffic congestion [2]. Further almost 5.5 billion hours are lost recorded due to traffic congestion, resulting in 2.9 billion gallons of fuel loss [2]. It is also estimated that automobile tailpipes produce around 31% (or 56 billion pounds) of $CO2$ due to vehicle traffic congestion each year [3].

The Global Traffic Scoreboard for 2021, published by data analytics company INRIX [5], provides insights into how people move throughout the world and provides more detailed information on congested traffic and commute times during the peak and off-peak hours in over 1000 urban cities across 50 countries. It also examines the per capita time spent on traffic congestion [5]. According to research [5], the most congested and crowded cities are "either older or developing cities". It also exclaimed that many of those developing cities where the significant population expansion is met, yet have inadequate infrastructures [5]. Therefore, the failure to develop adequate transportation infrastructures and roads to meet the increasing demand has led to traffic congestion.

Similarly, when global traffic mitigation strategies are considered, there can be seen several efficient approaches, such as intelligent traffic management systems. However, most of these mitigation methods have been limited to developed countries. Most of the remaining countries do not even have adequate traffic sensing infrastructures to monitor the traffic flow [5]. Therefore, there is no way to have an acknowledgment or a method to analyze the existing congestion having a confirmation of whether the implemented or existing congestion mitigation projects have achieved the desired progress. Since the traffic is always random, which means the vehicular traffic on the roads varies from time to time depending on the number of vehicles on the road, it is important to have a clear dataset of the number of vehicles on the road at a considerable very short intervals and effective feedback on-road operations are essential in order to carry out effective traffic congestion management.

Both vehicle detection and surveillance play an integral role in both effective traffic congestion management and intelligent traffic management systems (ITMS). Since ITMS plays a critical role in national traffic management systems (TMS), the quality of provided data and the geographical arrangement of traffic sensors are also important factors for ITS success

Author α: School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University Perth, Australia.
e- mail: p.weebadul@student.curtin.edu.au
Author σ: Department of Electrical and Electronic Engineering Sri Lanka Institute of Information Technology Sri Lanka. e-mail: charith.s@sliit.lk

[6]. Similarly, vehicle detection and surveillance functions are always subject to constantly being improved in order to increase vehicle detection and monitoring, observing traveling direction, counting their speed, and categorizing vehicles [6].

There is no doubt that vehicle identification and monitoring are vital components of both successful traffic management and ITMS. Similarly more effective traffic management may be achieved if future traffic volume estimates can be made through predictions. It performs a vital service for traffic management. Because the traffic density prediction methods depend on several factors, such as current and historic traffic measurements [6]. Accurate calculations are always important for accurate predictions. Therefore, the overall success predictions depend on the precise and efficient operation of the utilized medium to monitor traffic behavior.

The remainder of this paper is organized as follows. In section II, the research problem is discussed under the context of Sri Lanka. In section III, background and earlier studies carried out on real-time traffic monitoring and forecasting are presented, evaluating the research gaps. Moreover, part A and B of section III provide an extensive evaluation of each traffic monitoring and forecasting method, including tables that compare the performances of each technology. Section IV expresses the methodology, whereas section V includes the results and validation. Section VI summarizes all the facts as a conclusion and introduces future work.

II. Research Problem Definition in the Context of Sri lanka

Traffic congestion is one of the major issues in Sri Lanka. Initially, this problem was found only in the urban areas, including Colombo, but now with the increasing number of vehicles being added to the road day by day, the problem is further enhanced in other suburban areas as well [7], becoming traffic congestion is a prominent issue.

Excessive fuel consumption due to prolonged travel time results in economic loss and frequent acceleration and breakage, leading to frequent repairs and replacements, which create a significant loss to the national economy [8]. The cost of traffic congestion in Sri Lanka's Western Province exceeds Rs. 20,000 million per year, which is approximately 2% of regional GDP [9]. Total greenhouse gas (GHG) and CO_2 emissions in the Sri Lankan transport sector are accounted for 25% of GHG and 47% of CO_2 which indicates that traffic congestion has a substantial environmental impact in Sri Lanka when compared to global averages [10].

Nowadays Sri Lankan government is grappling with a large fiscal deficit, with the depreciation of the Sri Lankan rupee against major currencies and high debt. Traffic congestion is also a factor affecting a country's economy. And especially in a country like Sri Lanka, which is suffering from fuel and energy crises along with recent economic losses, this traffic congestion problem now has become an issue that can no longer be overlooked.

Several remedial strategies have been taken by the government to avoid rising traffic jams and to prevent it being worsening the congestion level. Short-term strategies such as adding new transport infrastructures: new roads, and expressways, and improving existing infrastructure capacities to be compatible with increasing road and transport capacity have been taken. Similarly, as with other possible alternative actions, the number of lane miles has been expanded, and alternate routes have been created to increase capacity and enhance the efficiency of existing infrastructure [4]. Long- term strategies such as introducing new vehicle ownership and public transport strategies and policies to be compatible with existing road and transport capacity have been taken to overcome this issue [9]. But the existing traffic congestion proves that those actions are insufficient to create an effective impact on traffic management.

Figure 1: Traffic Behaviour of Countries Such as Australlia, Russia and European Countries

Figure 2: Traffic Behaviour of Sri Lanka

Despite the above-mentioned facts, compared to some other countries where the drivers' disciplines are very high, in Sri Lanka, vehicles do not move systematically in a lane following one after the other while maintaining adequate space. Figures 1 and 2 which are presented above, illustrate the traffic behavior of Sri Lanka in comparison to other countries such as Russia, Australia, and other European countries. In the first figure, there can be seen adequate spaces between the vehicles and all of them are moving forward following one lane. But according to Figure 2, which illustrates the traffic behavior in Sri Lanka, vehicles move in closer to each other without following any lane.

This traffic behavior in Sri Lanka is the main issue, especially behind the difficulty in implementing image processing-based traffic monitoring systems since contours cannot be properly generated with this irregular traffic behavior in Sri Lanka. And it is harder to identify vehicles, especially in real-time traffic monitoring systems. Even though color lights are used to reduce traffic congestion since they are not intelligent and programmed for a certain period and there can be frequently seen several instances where the green color is still in "on" state even when there are no vehicles on the road. These phenomena can be commonly seen which causes traffic jams by lacking the chances for the vehicles on other lane's access to the road. Similarly, a large number of road accidents has also significantly increased due to the traffic jams in Sri Lanka, which now has become an adverse social impact [10].

Moreover, there can be seen poor technical and digital literacy in Sri Lanka compared to other countries. Even though many traffic surveys are conducted annually, there cannot be seen any proper traffic database. Similarly, there is no any automated system to retrieve past data from that analysis to do comparisons or other estimations such as traffic predictions [11].

III. Literature Review

As early as the 1970s, an autoregressive integrated moving average (ARIMA) model for short-term highway traffic flow forecasting was introduced which has been recorded as the foremost approach under ITMS strategies [12]. Since then, scholars and academics from a variety of fields, including transportation engineering, electronic and mechatronic engineering, statistics, economics as well as machine learning have proposed a wide range of models for forecasting traffic flow. Similarly, a considerable number of surveys and extensive analyses have also been carried out on traffic management in recent years, concentrating on various traffic factors.

A massive project was carried out to enhance traffic management and control of Hong Kong's Road network, which is known as one of the busiest roads in the world [13]. In this project [12], an ITMS was built beginning in 2001 and was successfully finished in 2010. Traffic monitoring, control operations, data collection, and analysis are among the primary platforms featured in this project to manage traffic congestion. By tracking all major highways and road tunnels, this initiative has ensured effective traffic control [12].

A low-cost sensor-based network instrument for traffic monitoring was developed and tested to be used in a work zone [14]. In that project, the entire sensor network system was used to collect data from that work zone [14]. Then the data was sent for post-facto analysis and uploaded to the internet. Al-Holou et al.[15] developed a multi-dimensional model to estimate the influence of vehicles on the environment, traffic congestion, and traffic safety.

This chapter includes a comprehensive literature review on similar traffic monitoring and prediction approaches and effective traffic monitoring and short-term forecasting techniques. First of all, the overview of real-time traffic monitoring and forecasting projects, research and surveys that have been carried out worldwide and their approaches are evaluated using the above paragraphs with the research gaps identified. Then the literature analysis is further extended, covering the traffic detection and prediction methods in separate sections.

a)   Traffic Detection Methods

Vehicle detection and surveillance play an integral role in both effective traffic congestion management and ITMS. Since ITMS plays a critical role in national traffic management systems (TMS), the quality of provided data and the geographical arrangement of traffic sensors are also important factors for ITMS success [6]. Similarly, both the vehicle detection and surveillance functions are always subject to constantly being improved in order to increase vehicle detection and monitoring, counting their direction headway, and speed, and categorizing vehicles. Further the real-time traffic information such as the number of vehicles and various sorts of road users and the vehicle types are useful to enhance the performances of the traffic management system.

Vehicle detection technologies that are widely used can be classified into three groups: intrusive, non intrusive, and off- roadway sensors [16]. The inductive loops, magnetic detectors, piezoelectric sensors, weight in-motion sensors and pneumatic road tubes are considered invasive sensors according to the above classification. These are usually embedded in the road surface after saw-cutting the surface or adding roadway holes. The detection methods such as vision- based systems such as image processing traffic monitoring systems, infrared sensors, microwave radar and ultrasonic detectors are categorized as non-intrusive sensors which can be installed atop roadway or roadside surfaces or mounted overhead. Remote sensing by airplane or satellite, as well as probe vehicles equipped with GPS receivers, are examples of off roadway sensors that do not require installation on highways [16]. More detailed descriptions of these technologies are available in [17], [18]. Consequently, these sensors are not suitable for large-scale integration or temporary installation. They are exclusively stationed in strategic areas and operate independently of one another.

Video Image Processor is a very common traffic monitoring method since now Image processing has become a tendency and the most prominent traffic monitoring system in the world [6]. Video Image Processor (VIP) systems normally consist of a camera, a processor-based workstation for analyzing the images, and software for understanding the images and transforming them into data. This can also be operated in multiple lanes. Image processing systems provide live images of real-time traffic status, which covers multiple detection zones. So that it offers broad area detection [19]. In addition to that, wide-area detection can also perform by gathering information generated from cameras located at different locations. Vehicular detection of the image processing systems is performed with the assistance of the contours drawn in the snapshots taken in constant time intervals thereby the vehicular presence is identified. It offers occupancy, classification, and count of vehicles, as usual in most other sensors.

Moreover, in the literature, several disadvantages of image processing have also been discussed. Being sensitive to weather conditions, vehicle shadows, and dust on the camera lens is notable. Lane closure requirements for installation and maintenance, specific camera mounting height requirements for better vehicle presence detection and speed measurement, higher installation, and maintenance costs are also significant drawbacks of this camera-based traffic monitoring system [19]. In spite of the weaknesses like costly equipment for the transfer of real-time video-image data, separating algorithms required for day and night traffic detections, possibilities of discrepancies appearing during traffic data transition, and performance prone to obscurants and heavy atmospheric conditions cannot be ignored [19]. Besides image processing, there are a variety of technologies for traffic monitoring that use various types of electronic sensors.

Light Detection and Ranging (LiDAR) technology is a novel technology in which research and investigations have been performed in recent years [24], [25]. LiDAR is a remote sensing method that uses light in the form of a pulsed laser to measure ranges of variable distances. The point cloud of LiDAR data is made up of thousands of points in X, Y, and Z coordinates. A point cloud depicts the environment in three dimensions. However, point cloud data is huge and contains duplicate information [25]. Downsampling, noise reduction, ground removal, object grouping, distant irrelevant object rejection, and ultimately vehicle recognition utilizing point cloud data are all part of this architecture.

The authors of [26] have proposed a technique for removing backdrops and detecting lanes from a point cloud based on roadside LiDAR data. Several articles and researchers have proposed many approaches for vehicle detection using LiDAR technology [26]. The authors propose an L-shape fitting model for automobile identification. It makes use of the fact that a vehicle's top view in LiDAR data resembles an L shape. A segment-based technique for recognizing automobiles using mobile LiDAR has also been suggested [25], [26]. Geometric characteristics such as size, form, and height are retrieved for categorization in this method. The distance traveled by the segments and the direction of movement is utilized to locate moving autos [25]. The strengths and weaknesses of each traffic monitoring technique are compared in Table 1 below.

Table I: Strengths and Weakness Comparison of Detector Types[6]

Type of the Detector	Detector Methods Comparison		
	Working Mechanism	Strengths	Weaknesses
Inductive loop	Detects the vehicle by sensing the loop	• Flexible design to fulfill a great variety of applications. • Unresponsive to bad weather. • Offers accurate count data.	• Traffic disruptions while installation and maintenance • Prone to damage by heavy vehicular movements and during road repairs h
Microwave radar	Transmits signals in the recognition regions and captures the echoed signals from vehicles	• Unresponsive to bad weather. • Speed is measured directly. • Multiple lane operation.	• Continuous-wave Doppler sensors are incapable of sensing immobile vehicles.
Acoustic	Detect audible sounds produced by vehicular traffic. Using them vehicle presence and speed are measured.	• Unresponsive to precipitation. • Multiple lane operation.	• Vehicle count accuracy may be affected by cold temperature.
Magnetic	Detects the presence of a vehicle by measuring the perturbation in the Earth's magnetic field	• Applicable where loops are not likely. • Insensitive to bad weather. • Installation of some models does not require a pavement cut. • Less prone than inductive loops to pressures of vehicles.	• Installation needs a boring under the road. • Incapable of sensing immobile vehicles.
Ultrasonic	Transmits ultrasonic waves and again collect the echoed waves from an object.	• Monitors multiple lanes. • Proficient in detecting over - height vehicles.	• Performance is affected by environmental circumstances. • Occupancy measurement on freeways may be degraded with large pulse repetition periods.
VIP (Video image processor)	This system normally consists of a camera, processor, and software.	• Monitors multiple lanes. • Simple to add and change detection areas. • Offers broad-area detection.	• Installation and maintenance require lane closure. • Performance is sensitive to bad weather, vehicle shadows, and dust on the camera lens.

Type of the Detector	Detector Methods Comparison		
	Working Mechanism	Strengths	Weaknesses
			• Requires specific camera mounting height for finest vehicle presence detection and speed measurement.
Light Detection and Ranging (LIDAR)	This is a remote sensing method that uses a pulsed laser to measure ranges of variable distances.	• Monitors multiple lanes. • Simple to add and change detection areas. • Insensitive to bad weather. • Offers broad-area detection.	• Traffic disruptions while installation and maintenance.

b) Prediction Methods

Traffic flow is a real-time, totally non-linear, high-dimensional, and non-stationary random process. Reviewing the literature emphasizes that vehicle forecasting is a common research topic and that more research has been done while many approaches have been taken. Most of the research articles have a special focus on these traffic predictions due to the uncertainty and random non-linearity of the traffic flow. Similarly, most of them have introduced Traffic predicting and forecasting traffic strategies which assist to minimize these increasing traffic jams [27], [28]. According to the reference [28], traffic forecasting is classified into two basic categories: long-term prediction and short-term prediction. The projections for the near future are known as short-term predictions typically these short-term predictions are for the next 5 to 10 minutes in immediate future. Short-term predictions are a common prediction method since to the changing weather conditions, cultural or political occasions, and events, road accidents may cause sudden changes in traffic conditions [28].

There can be seen several research articles which evaluate the accuracy and the efficiency of several prediction models comparing one another. Among them [27], [28] reference emphasizes that both data-driven and experimental traffic flow prediction approaches can be classified as parametric, non-parametric, or hybrid, each having its own set of benefits and drawbacks [28]. The majority of traffic flow prediction research has been conducted under regular traffic conditions, while unusual traffic variables such as climate, the presence of noise in the data, and highway disruptions have seldom been addressed. Providing a complete detailed analysis of these three prediction methods, above mentioned reference emphasizes linear regression, maximum likelihood (ML), Historical Mean Average, exponential smoothing method, and time series model as parametric prediction approaches [28]. Moreover, it claimed that the parametric prediction method is more accurate than the other two methods, still its poor functionality amid the noise and other disturbances is an encountered major drawback [28].

Extensive research by Van Lint and Hoogendoorn which provides an overview of predictive models for short-term traffic forecasting also emphasizes that forecasting methods can be divided into three major groups: parametric, non-parametric, and hybrid (naive) [29].

i. Non-Parametric Models

Non-parametric models imply that the number of parameters which assigned to a model is flexible. This means it is not fixed, in which the model structure and parameters should be developed based on the available data. Usually, the amount of data should be significant in contrast to the other two approaches. These models have the benefit of allowing for the discovery of intricate non-linear correlations between traffic factors. However, the disadvantage of these models is that while the model's structure is derived from the data, unanticipated events and outliers may have an impact [29]. The intricacy of these models, as well as their reliance on vast amounts of data, are also other significant drawbacks. Non-parametric prediction approaches use current models rather than traditional models to estimate traffic flow in proportion to road conditions. Neural networks, such as the multilayer perceptron (MLP), time-delay neural network (TDNN), and radial basis function (RBF) are the most popular and prominent non-parametric approaches. Besides these neural networks, Fuzzy [30], Bayesian networks k-nearest neighbor (KNN)[31], support vector machine, and wavelet are other non-parametric methods used for predictions [28].

Neural networks are the most common and popular model that is used for traffic predictions. They consist of the ability to model and simulate complicated non-linear relationships [29]. There can be seen different neural network types such as Feed forward artificial neural networks, Convolutional Neural Network (CNN), Recurrent neural networks (RNN) as well as Long Short-Term Memory networks (LSTM) based on their training procedure, internal structure, methods of preprocessing input data and their models including spatial or temporal patterns. Among them, the feed forward artificial neural network is the simplest neural network category while the LSTM is the most powerful model to process sequential data. The type of the neural network varies depending upon the task being performed with it. Typically, a neural network comprises an input layer, several hidden layers and an output layer [29]. According to the research paper by Van Hinsbergen, Van Lint, et al., a typical neural network might deliver reliable findings in terms of extensions required for higher accuracy in traffic predictions [29].

ii. Parametric Models

Parametric models imply that the model's structure and the number of parameters are tightly established, and the model's parameters must be derived using data. The advantage of such models is that unseen cases and incidents can be captured. Another advantage is the necessity of less data. Some models are capable of offering higher accuracy even with less computational work [29].

The hybrid model is the combined version of both parametric and non-parametric prediction models where the accuracy of the prediction is higher than the other two types. Most hybrid prediction models are combinations of neural networks and other parametric and non-parametric models such as ARIMA models, MLP, and fuzzy. According to the reference [28] the Neural network-based MLP model is the most suitable prediction model for data-driven traffic forecasting

systems combined with image processing compared to all other models.

Despite the facts that were included on these parametric, non-parametric and hybrid prediction models, in the literature section, there can be seen a vast number of projects and research approaches subject to predictions and forecasting models which evaluate the existing gaps. Further, when the literature on traffic prediction methods are referred to, there can be seen an outstanding tendency toward using LSTM neural networks for traffic predictions, especially in recently published research articles. Similarly, novel projects and research have given great attention to these LSTM neural networks. The reference [33] evaluates the recent rise in the usage of LSTM algorithms for traffic prediction. It reveals that now the LSTM model has become common and prominent. According to another research carried out on LSTM[34], most prediction approaches have concentrated on accuracy rather than immediacy.

LSTM neural network is an improved version of RNN where the vanishing gradient problem and long-term dependency of recurrent neural networks are successfully overcome. Input, output, and forget gates are three gates that are introduced in LSTM. However, the reference [35] claims that even though most of the novel traffic management approaches used Long short-term memory (LSTM) models, those existing projects and their models have failed to address the issue of massive traffic flow data being processed simultaneously with parallel to computing and distributed data storage. But it also emphasizes that the LSTM model is a better prediction model for more random and time-varying predictions such as traffic flow [35].

According to reference [40], traffic prediction research should focus on the data-intensive era, which is now missing. The existing traffic categorization algorithms are ineffective in low-light situations [41]. Also, there can be seen a lack of studies focusing on the time series for the Internet of Things (IoT) traffic forecast [33]. Existing research has not properly simulated or created to be compatible with the dynamic traffic patterns of irregular locations [33].

Traffic prediction studies are controversial due to a lack of more efficient computational methodologies and algorithms, as well as excellent quality data. Research article [43] indicated that the performance of the Convolutional Neural Network (CNN) for traffic prediction was somewhat disappointing based on the implementations of prior research. Table II compares the strengths and weaknesses of each prediction model.

Table II: Prediction Models Comparison

Prediction Model	Prediction Models Comparison	
	Strengths	*Weaknesses*
Mean Average model	• Low Prediction Error.	• Poor functionality in the presence of noise. • Average of all inputs are needed for the predictions. High dependency on recorded data.
Linear Regression Method	• Low Prediction Error. • Predicts the next variable online using real data.	• Poor functionality amid the noise and other disturbances.
Maximum Likelihood (ML)	• Low Prediction Error. • Robust for sensor failures and rapidly changing conditions.	• High dependency on recorded data.
Exponential Smoothing Method	• Low Prediction Error.	• Poor functionality amid the noise and other disturbances. • Difficulties to determine constant coverage.
ARIMA Model	• Low Prediction Error. Simplicity. • More mathematical model • Obtain the relationship between past and future data.	• Poor functionality amid the noise and other disturbances. • High dependency on recorded data.
MLP Model	• High Accuracy.	• High dependency on recorded data.

Prediction Model	Prediction Models Comparison	
	Strengths	*Weaknesses*
	• Predict traffic flow in proportion • to road conditions.	
Fuzzy Model	• Low Prediction Error. Simplicity. • High Accuracy.	• Poor functionality amid the noise and other disturbances. • High dependency on recorded data.
KNN Model	• High Accuracy. • Predict traffic flow in proportion to road	• Poor functionality amid the noise and other disturbances. • High

IV. METHODOLOGY

The methodology is introduced by being split into two main sections: the traffic monitoring stage and the forecasting stage using neural networks emphasizing the main milestones of the project.

a) Traffic monitoring stage

Vehicle detection was accomplished using a TFMini Plus LiDAR sensor installed by the side of the road, to monitor the vehicles directed to one side, determining the desired tilt angle and height. Sensor positioning is illustrated in below figure 3.

TFMini Plus sensor's Time of Flight (ToF) principle is used to detect the presence of vehicles. The sensor's periodically emitted modulation waves are used to detect and calculate the proximity to the object and its time-of-flight is estimated by measuring the round-trip phase difference of its reflection when it contacts an object [49]. These periodic modulation waves are always set to be directed towards the road and to be released at a frequency of 16.667 Hz with 60-millisecond intervals. Universal Asynchronous Receiver / Transmitter (UART) communication was used to communicate with the microcontroller. Distance limitations for the sensor were configured as the distance should be between the range of 800 to 1220 meters considering white lines which indicate the margins and use to separate traffic traveling in the same direction, assuring the sensor is not triggered for pedestrian movements or the vehicles in the other direction.

Additionally, distance measurements must persist longer than 50 milliseconds and return to zero which is the initial distance value. The presence of a vehicle will therefore be counted as present if both prerequisites are met, where the count will be one. Likewise, the vehicle count is taken and increased within a 5-minute interval. And the vehicle count is programmed to be zero after every 5 minutes (5 $\times 1000 \times 60$ milliseconds) intervals and the latest value which is recorded as the vehicle count inside that interval is to be delivered to the ThingSpeak IoT platform as well as to an SD card. Here the ThingSpeak online IoT platform was used as the source to create the database. A database was developed by recording consecutively collecting data. ESP8266 (NodeMCU) module was used as the microcontroller which can enable further improvements with IoT connecting several nodes.
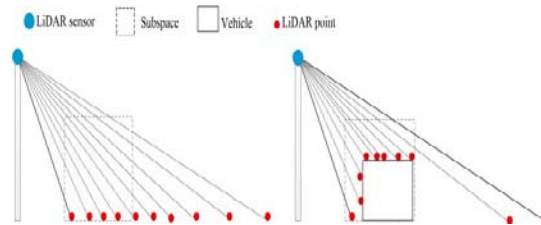

Figure 3: An Illustration of a Subspace Without (Left) Or With (Right) a Vehicle [50]
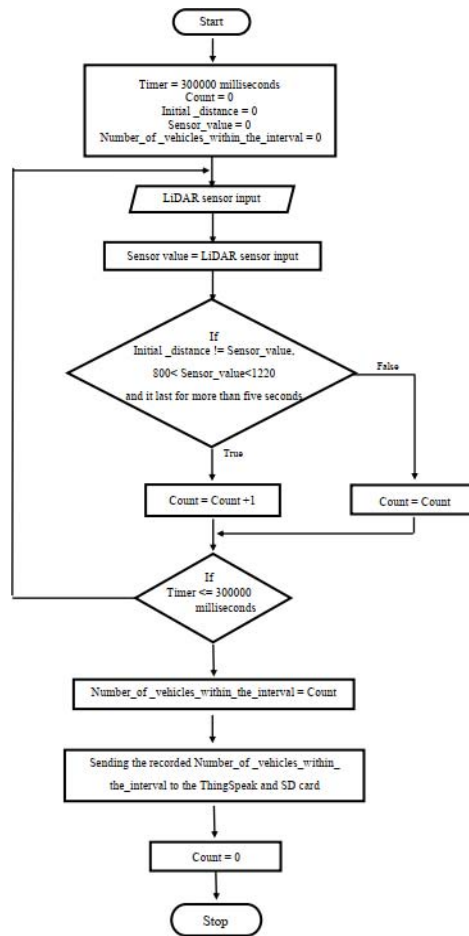
Figure 4: Flow Chart of the Vehicle Detection Algorithm

Then, sensor accuracy was tested by plotting the actual vehicle counts vs sensor detected vehicle counts graph considering 200 data points of that data entries. Similarly, the accuracy of the sensor was also measured as a percentage to ensure the sensor's performance.

b) Neural Network Training and Forecasting Stage

This stage includes training LSTM neural network model. To accomplish this task, data preparation and segmentation were carried out before sending to the neural network model for training. The data gathered to Thing Speak was exported as. csv files to train the LSTM model. During the data consolidation process, .csv files gathered from Thing Speak were modified separating date and time into two columns, removing the entry ID column, and adding missing entries. Similarly, the values which imply significant deviations were removed from the data set. Those modified .csv files were fed into the system as inputs by dividing them into four segments. And the input data volume that is taken to train the model and window size were used as variable parameters to find the best case of the prediction model.

Since the vehicle count is taken in five minutes intervals,

In one hour: 12 data points
In one day: 12 × 24 = 288 data points
In a week: 12 × 24 × 7 = 2016 data points
|In a month: 12 × 24 × 30 = 8640 data points

Likewise, data gathered over five weeks by monitoring the vehicular traffic were employed to create the forecasting models.

Then the entire data collection was separated into two sets for training purposes and testing purposes. Eight thousand sixty-four data points (data gathered over a duration of a month) were employed as the training data set while the remaining set of 2016 data points (data gathered over a week) were used as test data for the model testing purposes. The entire training data set was again split into four segments as week 1, week 2, week 3 and 4.

Then obtained data patterns over a day and week were observed to identify the patterns of the traffic behavior. After that more than ten separate models were tested to find the best traffic prediction model. Initially, the first set of models were created by setting the window size to 12 while changing the input data volume from 2016 data points to 8064 data points. When the 2016 input data set is considered, it includes the data gathered

during the last week of the month, which is fed to the system for training. Accordingly, the 4032 data points indicates the data gathered over the last two weeks and 6048 indicates the collection of data over three weeks and 8064 data points represents the data gather over a month. (i.e., a separate data set with 2016 data points collected over another week was employed as test data to test the models.)

In the second model set, depending upon the results of the first four models, input data volume was set to 8064 data points and the window size was changed to 6, 12, 24, 60 and 288. Window size 6 means 6 data points are fed into the system at a time, which means 30-minute (1/2 hour) intervals. 12 means 1-hour data is fed at a time. Likewise, three other models were trained for 2 hours, 5 hours and 24 hours. Depending upon the results of those five models, the desired window size was determined. And then in the third model set, the window size was set to 24 and once again the input data volume was changed from 2016 data points to 8064 data points.

After obtaining all of these prediction model sets, all the models that have been trained varying both window size and the volume of input data delivered to train the models were critically evaluated in terms of accuracy in order to identify the most suitable prediction method. Accuracy formulas provide a vital contribution to projects since precise calculations are always helpful to critically evaluate the resulted outcomes to determine the most suitable solutions to apply. Accordingly, all accuracy-related calculations of this project (sensor accuracy calculations, prediction models' accuracy calculations) have been done using the following two equations which are used to measure the world-accepted accuracy in accordance with IEEE standards [51]. The accuracy formula provides accuracy as a difference of error rate from 100%. To find accuracy, first it is needed to calculate the error rate. And the error rate is the percentage value of the difference between the observed and the actual value, divided by the actual value [51].

Therefore, below equation 1 was utilized to determine the error rate.

$$\text{Error Rate} = \frac{|Observed\ Value - Actual\ Value|}{\text{Actual Value}} \times 100$$

Once the error rate is calculated the accuracy is determined using the below equation 2 as the difference of error rate from 100%.

$$\text{Accuracy} = 100\% - \text{Error Rate}$$

V. Results and Validation

a) Results of the Sensor Accuracy Test

First of all, the accuracy of the sensor was tested considering the data gathered during a random day. Actual vehicular counts were taken by counting the actual vehicle count and the sensor detected vehicle count was directly obtained from the sensor records. Then the actual vehicle count vs sensor-detected vehicle count was plotted and the accuracy of the sensor has been tested considering 200 data points.

Obtained the accuracy results of the sensor was 89.86%. Figure 5 illustrates the behavioral differences between actual vehicle count and sensor-detected vehicle count. This emphasizes that sensor detected vehicle count is almost closer actual vehicle count. Similarly, this sensor accuracy plot is also helpful to understand the traffic behavior of a day, starting from 0.00hr to 23.59hr. Accordingly, it was understood that the traffic behavior of a day more takes normal distribution behavior, in which the higher values cluster in the middle of the range and the rest taper off symmetrically toward either extreme.
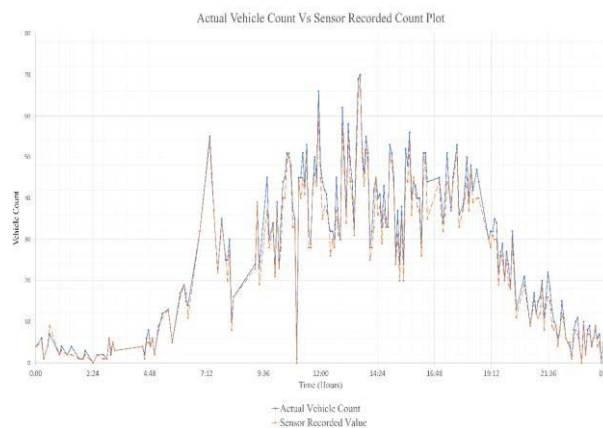

Figure 5: Actual Vehicle Count vs Sensor-Recorded Vehicle Count Plot

b) Traffic Behavior Analysis

After observing the traffic behavior of a day, the traffic behavior of a week was also plotted. There also the traffic behavior of each day takes normal distribution behavior. Therefore, the traffic behavior of a week was a combination of seven normal distributions. Additionally, after gathering traffic data over a month, the traffic behavior of each week was also plotted to identify whether the traffic flow remains the same following a pattern every week or is there any significant increase or decrease in the traffic behavior. Belowfigure 6 represents the overall traffic behavior of the entire month. These plots which were plotted observing the traffic behavior of a day, a week and entire month were helpful in understanding the traffic pattern as well as in deciding the window size that should be taken to train the LSTM neural network since there was not any other source to be observed to identify the traffic patterns.

c) Prediction Results

When obtained data patterns were observed, it was recognized that there could be significant deviations in traffic flow after intervals of 60 minutes which means during an hour, there could be seen slight differences but from hour to hour there can be seen significant increasements or decrements in traffic behavior. Therefore, first it was identified it is better totake 12 data points at a time to train the prediction model. Then the desired window size was kept to twelve and using data gathered over a month the very first LSTM neural network model was trained. Below figure 7 illustrates the resulted predicted traffic behavior vs test data plot and figure 8 illustrates the entire plot of the trained data, test data and resulted predicted traffic data.

Finding the best traffic prediction model is one major research objective of this project. Therefore, more than ten LSTM neural network models were created and evaluated in terms of accuracy to identify the most suitable forecasting model. Here all the model sets that were tested able to result accuracy more than 70%. And no model exhibits large deviations in predicted traffic behavior vs test data plot proving that fact. Further, it seems that the strategy that has been carried out by changing the window size and the volume of the input data as the variable parameters to figure out the best model was successful because there can be seen a noticeable pattern in the prediction accuracies.
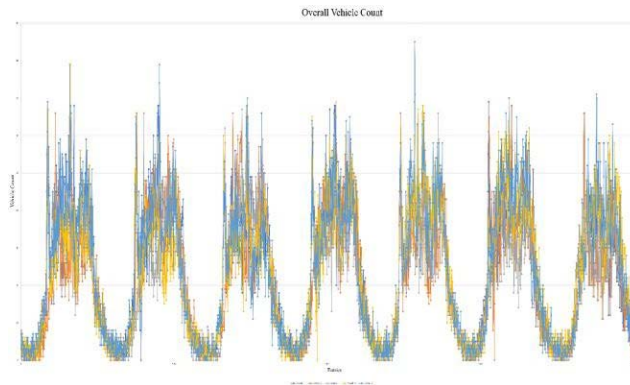

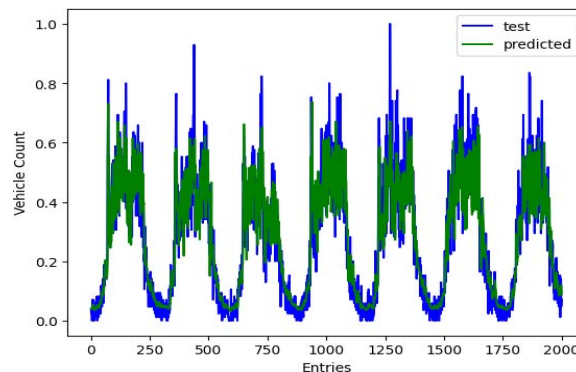*Figure 6:* Overall Traffic Behavior of Each Week

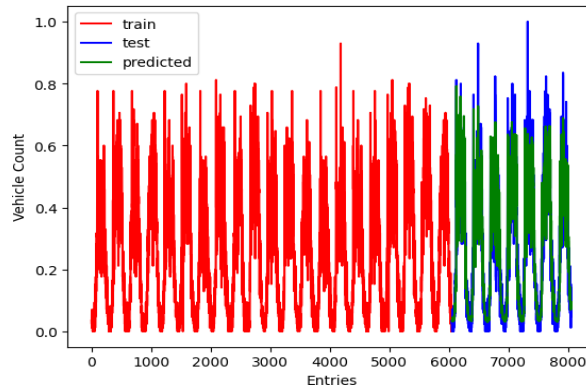*Figure 7:* Predicted Traffic Behavior vs Test Data Plot

Figure 8: Train, Test and Predictions Results of the Most Precise Forecasting Model

The accuracy results for each prediction mode that was developed employing both window size and the volume of input data delivered into the model to be trained as variable parameters are shown in Table III, IV, and Table V below.

Table III: Accuracy Table of the Models Trained using Window Size_12

Volume of input (Data Points)	Window size (Data Points)	Accuracy %
2016	12	72.87
4032	12	72.91
6048	12	73.42
8064	12	74.59

Table IV: Accuracy Table of the Models Trained using Variable Window Sizes

Volume of input (Data Points)	Window size (Data Points)	Accuracy %
8064	6	74.49
8064	12	74.59
8064	24	75.34
8064	60	73.58
8064	288	72.07

Table V: Accuracy Table of the Models Trained using Window Size_24

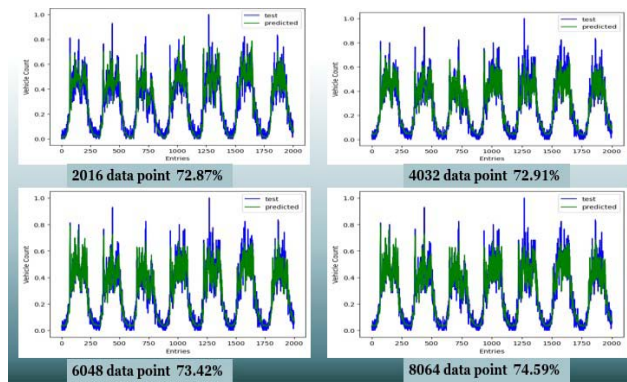| Volume of input (Data Points)	Window size (Data Points)	Accuracy %
2016	24	74.20
4032	24	73.78
6048	24	73.96
8064	24	75.34

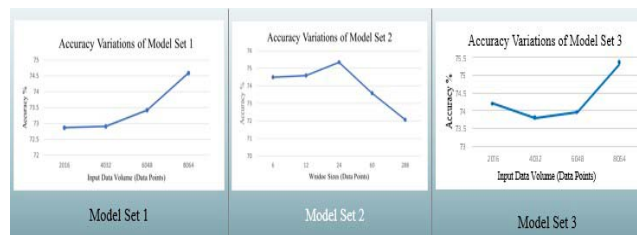*Figure 9:* Predicted Traffic Behavior vs Test Data Plots

Figure 10: Accuracy Variations of the Prediction Model Sets

Further, the accuracy results of these test models which are presented above shown in Table III, IV, and Table V emphasize that both window size and the volume of the input data creates a significant impact on the prediction accuracy of the prediction models. The accuracy test results of all the 3 model sets show that all of them were above 70%.

Above Table III includes accuracy results of the training model set 1 which was created by setting the window size to 12 while changing the input data volume from 2016 data points to 8064 data points. As per described in the methodology section, when 2016 input data set is considered, it includes the data gathered during the last week of the month and that data was fed to the system for training. Accordingly, the other models were tested using increasing input data points while keeping the window size constant. In model set 1, there can be seen a gradual increase in accuracy when the input data volume gets larger.

In the second model set (represented in Table IV), the input data volume was set to 8064 data points and the window size was changed to 6, 12, 24, 60 and 288 (explained in the methodology section). Even though the accuracy has increased in the first three models, there can be seen a noticeable decrease in accuracy with larger window sizes. This implies that the most suitable window size is 24.

Third model set in which results are presented in Table V, the window size was set to 24 and once again the input data volume was changed from 2016 data points to 8064 data points. There also can be seen a gradual increase in accuracy when the input data volume is getting larger.

Above figure 9 illustrates several predicted traffic behavior vs test data plots. All of them are quite similar and both predicted vs actual traffic behavior lies on top of each other showing a closer behavior. Similarly, there are no larger deviations. Figure 10 includes the graphical representation of the above Tables III, IV and V. It also provides graphical comparison on the accuracy results that were obtained from each model set.

As illustrated, there can be seen noticeable changes when varying both window size and the volume of input data. A gradual increase in accuracy can be seen when increasing the input data volume. Meanwhile, it implies that the window size should remain under a fixed value, and increasing it enhances the accuracy but after a certain value, the accuracy may decrease if it is further increased.

Therefore, according to the obtained results it was identified that the highest prediction accuracy could be obtained by setting the window size to 24 windows in which the data gathered for two hours will be fed into the system to be trained at a time and increasing the input data volume to the greatest extent possible. Similarly, resulted accuracy values also indicate that the model should be trained using 80% data and tested using 20% data from the gathered data in order to obtain higher precision.

VI. Conclusion

The main objective of this project was to introduce A better traffic management system and traffic forecasting system to reduce traffic congestion,

especially in the context of Sri Lanka. The results of both the sensor accuracy tests and the results obtained from the prediction models emphasize that the aforementioned objectives have been successfully achieved and the strategies that were utilized in this project to address the issues have been able to effectively contribute to the successful completion of the entire project by accomplishing their individual tasks. Here the efficiency of the methodology has been successfully illustrated by the practical implementation and using the obtained accuracy test results.

Further, there can be seen several benefits in this project. The traffic monitoring and forecasting system, which is introduced in this project, is low-cost as well as it can be easily implemented. This system can be implemented in any country, under any circumstances, since it has the ability to adapt to any type of traffic behavior. It is a well-known fact that monitoring vehicular movement is a really challenging task, especially in Sri Lanka, where the traffic behavior is messy and chaotic since vehicles do not move consistently following a lane, maintaining adequate distance one after another. However, the traffic monitoring method that has been utilized in this project which employs TFmini plus LiDAR sensor to monitor the traffic, is able to provide 89.86% accurate readings, avoiding difficult circumstances, even under an unsystematic traffic condition like Sri Lanka. Therefore, since this sensor system is able to address the traffic issue in Sri Lanka successfully, it proves the adaptability of implementing it anywhere under any type of traffic behavior.

Moreover, even though image processing-based video traffic monitoring systems are commonly used in many other countries, it appears that those image processing-based systems require advanced technological feasibility as well as considerable capital to implement them. Therefore, when considering the existing technical feasibilities and the traffic infrastructures in Sri Lanka, this system is more appropriate as a low-cost as well as a system that can be implemented with low technical capabilities. Especially its resistance to dust and rain makes it possible to place it anywhere in a country like Sri Lanka.

The other thing is most of the traffic forecasting models need a sizable traffic database to train their prediction models. Prediction models based on image processing need a previous database to identify the vehicles by drawing the contours. And it takes significant time to train the model. But especially in the countries like Sri Lanka, there is no traffic database gathered over a considerable time period. Therefore, the prediction model which is introduced in this project is more suitable for such countries because it does not require a large volume of data to be trained. Obtained prediction accuracy results depict that this prediction model can produce predictions with an accuracy rate of 74.20% even if it is trained using the data gathered over one week. Therefore, this implies that the absence of a proper traffic database would no longer be an issue for implementing traffic prediction models with this system.

Therefore, it is clear that this traffic monitoring and forecasting system is a more practical approach to solving the traffic congestion issue effectively, especially in Sri Lanka, where there is no valid database regarding the traffic behaviors and congested areas, where the traffic behavior is messy and chaotic and where there are less advanced technical feasibilities. And the other important fact is that the accuracy testing results of the prediction models reveal that when the volume of the input data is increased it significantly increases the accuracy of the prediction model. Therefore, at the beginner level, this traffic monitoring and forecasting system can be implemented with the data gathered over a short period like one month, and later this can be developed further, with better forecasting accuracies by including data collected by training the system.

References Références Referencias

1. T. Litman, "Congestion Costing Critique Victoria Transport Policy Institute 2", [Online]. Available: www.vtpi.orgInfo@vtpi.org.

2. National Highway Traffic Safety Administration and others, "Traffic safety facts, 2012 data: pedestrians," vol. 65, pp. 1–452, 2015.

3. "TTI's 2012 Urban Mobility Report: Powered by INRIX Traffic Data Google Search." https://www.google.com/search?q=TTI%27s+2012+Urban+Mobility+Report%3A+Powered+by+INRIX+Traffic+Data&oq=TTI%27s+2012+Urban+Mobility+Report%3A+Powered+by+INRIX+Traffic+Data&aqs=chrome..69i57.551j0j15&sourceid=chrome&ie=UTF-8 (accessed Apr. 28, 2022).

4. "Traffic surveillance by wireless sensor networks : Final report | CiNii Research all 検索 ." https://cir.ii.ac.jp/all?q=Traffic%20surveillance%20by%20wireless%20sensor%20networks%20:%20Final%20report (accessed Apr. 28, 2022).

5. "INRIX 2021 Global Traffic Scorecard: As lockdowns ease UK city centres show signs of return to 2019 levels of congestion - INRIX." https://inrix.com/press-releases/2021-traffic-scorecard-uk/ (accessed Apr. 26, 2022).

6. K. Nellore and G. P. Hancke, "A Survey on Urban Traffic Management System Using Wireless Sensor Networks," *Sensors 2016, Vol. 16, Page 157*, vol. 16, no. 2, p. 157, Jan. 2016, doi: 10.3390/S16020157.

7. "Registered Motor Vehicles | Lanka Statistics." https://lankastatistics.com/social/registered-motor-vehicles.html (accessed Mar. 31, 2022).

8. L. N. v Alwis and N. Amarasingha, "Estimating the fuel loss during idling of vehicles at signalized intersections in colombo," in *2017 6th National Conference on Technology and Management*

(NCTM), 2017, pp. 132–137. doi: 10.1109/ NCTM.2017.7872841.

9. A. S. Kumarage, "URBAN TRAFFIC CONGESTION: THE PROBLEM & SOLUTIONS Paper Published in the Economic Review, Sri Lanka," Moratuwa, Sri Lanka, 2004. Accessed: Mar. 28, 2022. [Online]. Available: https://kumarage.files.wordpress.com/20 15 /03/2004-p-06-ut- urban-traffic-congestion-the-problems-and-solutions-economic- review-sri-lanka-2004.pdf

10. "Kanishka Werawella Presents Solutions to Traffic Congestion in Sri Lanka | Institute of policy studies Sri Lanka." https://www.ips.lk/kanishka-werawella-presents-solutions-to- traffic-congestion-in-sri-lanka/ (accessed Apr. 03, 2022).

11. "Traffic information system for Sri Lanka." http://dl. lib.uom.lk/handle/123/10529 (accessed Mar. 31, 2022).

12. ."Hong Kong Intelligent Transport System (ITS) - Verdict Traffic." https://www.roadtraffic-technology. com/projects/hong-kong/ (accessed May 03, 2022).

13. "A Distributed Instrument for Measuring Traffic in Short-Term Work Zones." https://trid.trb.org/view/ 1229740 (accessed May 03, 2022).

14. "TS-45A Multi dimensional Model for Vehicle Impact on Traffic Safety, Congestion, and Environment | Michigan-Ohio (MIOH) University Transportation Center (UTC)." https://mioh- utc.udmercy.edu/ research/ts-45/index.htm (accessed May 03, 2022).

15. W. Wen, "A dynamic and automatic traffic light control expert system for solving the road congestion problem," *Expert Systems with Applications: An International Journal*, vol. 34, no. 4, pp. 2370–2381, May 2008, doi: 10.1016/J.ESWA. 2007.03.007.

16. M. Bernas, B. Płaczek, W. Korski, P. Loska, J. Smyła, and P. Szymała, "A Survey and Comparison of Low-Cost Sensing Technologies for Road Traffic Monitoring," *Sensors 2018, Vol. 18, Page 3243*, vol. 18, no. 10, p. 3243, Sep. 2018, doi: 10.3390/ S18103243.

17. X. Mao, S. Tang, J. Wang, and X. Y. Li, "ILight: Device-free passive tracking using wireless sensor networks," *IEEE Sensors Journal*, vol. 13, no. 10, pp. 3785–3792, 2013, doi: 10.1109/JSEN.2013.22679 59.

18. R. Gade and T. B. Moeslund, "Thermal cameras and applications: a survey," *Machine Vision and Applications 2013 25:1*, vol. 25, no. 1, pp. 245–262, Nov. 2013, doi: 10.1007/S00138-013-0570-5.

19. I. Gulati and R. Srinivasan, "Image processing in intelligent traffic management," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2 Special Issue 4, pp. 213–218, Jul. 2019, doi: 10.35940/IJRTE.B1040.0782S419.

20. M. R. Islam, N. I. Shahid, D. T. Ul Karim, A. al Mamun, and M.

21. K. Rhaman, "An efficient algorithm for detecting traffic congestion and a framework for smart traffic control system," *International Conference on Advanced Communication Technology, ICACT*, vol. 2016-March, pp. 802–807, Mar. 2016, doi: 10.1109/ICACT.2016.7423566.

22. G. Padmavathi, D. Shanmugapriya, M. Kalaivani, G. Padmavathi, D. Shanmugapriya, and M. Kalaivani, "A Study on Vehicle Detection and Tracking Using Wireless Sensor Networks," *Wireless Sensor Network*, vol. 2, no. 2, pp. 173–185, Mar. 2010, doi: 10.4236/WSN.2010.22023.

23. D. Guilbert, C. le Bastard, S. S. Ieng, and Y. Wang, "State Machine for Detecting Vehicles by Magnetometer Sensors," *IEEE Sensors Journal*, vol. 16, no. 13, pp. 5127–5128, Jul. 2016, doi: 10.1109/JSEN.2016.2560903.

24. M. Tomic, P. T. Sullivan, and V. K. Mcdonald, "Wireless, acoustically linked, undersea, magne tometer sensor network," *MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges, OCEANS 2009*, 2009, doi: 10.23919/ OCEANS.2009.5422465.

25. T. W. Yeh, S. Y. Lin, H. Y. Lin, S. W. Chan, C. T. Lin, and Y. Y. Lin, "Traffic Light Detection using Convolutional Neural Networks and Lidar Data," *Proceedings - 2019 International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS 2019*, Dec. 2019, doi: 10.1109 /ISPACS48206.2019.8986310.

26. B. Anand, V. Barsaiyan, M. Senapati, and P. Rajalakshmi, "Region of Interest and Car Detection using LiDAR data for Advanced Traffic Management System," *IEEE World Forum on Internet of Things, WF-IoT 2020 - Symposium Proceedings*, Jun. 2020, doi: 10.1109/WF-IOT48130.2020.9221354.

27. J. Wu, H. Xu, and J. Zheng, "Automatic background filtering and lane identification with roadside LiDAR data," *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, vol. 2018-March, pp. 1–6, Mar. 2018, doi: 10.1109/ITSC.2017.8317723.

28. J. Barros, M. Araujo, and R. J. F. Rossetti, "Short-term real-time traffic prediction methods: A survey," in *2015 International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 2015, pp. 132–139. doi: 10.1109/MTITS. 2015.7223248.

29. S. M. H. Hosseini, B. Moshiri, A. Rahimi-Kian, and B. N. Araabi, "Traffic Flow Prediction Using MI Algorithm and Considering Noisy and Data Loss Conditions: An Application to Minnesota Traffic Flow Prediction," *Promet-traffic & Transportation*, vol. 26, pp. 393–403, 2014.

30. D. Zeng, J. Xu, J. Gu, L. Liu, and G. Xu, "Short term traffic flow prediction using hybrid ARIMA and ANN models," *Proceedings 2008 Workshop on Power Electronics and Intelligent Transportation System,*

52

PEITS 2008, pp. 621–625, 2008, doi: 10.1109/PEITS.2008.135.

31. A. Stathopoulos, L. Dimitriou, and T. Tsekeris, "Fuzzy Modeling Approach for Combined Forecasting of Urban Traffic Flow," *Computer-Aided Civil and Infrastructure Engineering*, vol. 23, no. 7, pp. 521–535, Oct. 2008, doi: 10.1111/J.1467-8667.2008.00558.X.

32. M. Dixit, R. Sharma, S. Shaikh, and K. Muley, "Internet traffic detection using naïve bayes and K-Nearest neighbors (KNN) algorithm," *2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019*, pp. 1153–1157, May 2019, doi: 10.1109/ICCS45141.2019.9065655.

33. Y. Tu, S. Lin, J. Qiao, and B. Liu, "Deep traffic congestion prediction model based on road segment grouping," *Applied Intelligence 2021 51:11*, vol. 51, no. 11, pp. 8519–8541, Apr. 2021, doi: 10.1007/S10489-020-02152-X.

34. H. Nicholson and C. D. Swann, "The prediction of traffic flow volumes based on spectral analysis," *Transp Res*, vol. 8, no. 6, pp. 533–538, 1974, doi: 10.1016/0041-1647(74)90030-6.

35. Y. Zhang, Y. Yang, W. Zhou, H. Wang, and X. Ouyang, "Multi- city traffic flow forecasting via multi-task learning," *Applied Intelligence 2021 51:10*, vol. 51, no. 10, pp. 6895–6913, Feb. 2021, doi: 10.1007/S10489-020-02074-8.

36. "Congestion Pattern Prediction for a Busy Traffic Zone Based on the Hidden Markov Model | IEEE Journals & Magazine | IEEE Xplore." https://ieeexplore.ieee.org/document/9308921 (accessed May 04, 2022).

37. L. Romo, J. Zhang, K. Eastin, and C. Xue, "Short-Term Traffic Speed Prediction via Machine Learning," *Communications in Computer and Information Science*, vol. 1311, pp. 31–42, 2020, doi: 10.1007/978-981-33-4532-4_3.

38. M. A. Abdelwahab, M. Abdel-Nasser, and R. ichiro Taniguchi, "Efficient and Fast Traffic Congestion Classification Based on Video Dynamics and Deep Residual Network," *Communications in Computer and Information Science*, vol. 1212 CCIS, pp. 3–17, 2020, doi: 10.1007/978-981-15-4818-5_1.

39. N. A. M. Razali, N. Shamsaimon, K. K. Ishak, S. Ramli, M. F. M. Amran, and S. Sukardi, "Gap, techniques and evaluation: traffic flow prediction using machine learning and deep learning," *Journal of Big Data*, vol. 8, no. 1, pp. 1–25, Dec. 2021, doi: 10.1186/S40537-021-00542-7/TABLES/4.

40. D. Xia *et al.*, "A distributed WND-LSTM model on MapReduce for short-term traffic flow prediction," *Neural Computing and Applications 2020 33:7*, vol. 33, no. 7, pp. 2393–2410, Jul. 2020, doi: 10.1007/S00521-020-05076-2.

41. D. H. Shin, K. Chung, and R. C. Park, "Prediction of Traffic Congestion Based on LSTM through Correction of Missing Temporal and Spatial Data," *IEEE Access*, vol. 8, pp. 150784– 150796, 2020, doi: 10.1109/ACCESS.2020.3016469.

42. Z. Wang, R. Chu, M. Zhang, X. Wang, and S. Luan, "An Improved Selective Ensemble Learning Method for Highway Traffic Flow State Identification," *IEEE Access*, vol. 8, pp. 212623–212634, 2020, doi: 10.1109/ACCESS.2020.3038801.

43. D. Liu, S. Hui, L. Li, Z. Liu, and Z. Zhang, "A Method for Short- Term Traffic Flow Forecasting Based on GCN-LSTM," *Proceedings - 2020 International Conference on Computer Vision, Image and Deep Learning, CVIDL 2020*, pp. 364–368, Jul. 2020, doi: 10.1109/CVIDL51233.2020.00-70.

44. Z. Bartlett, L. Han, T. T. Nguyen, and P. Johnson, "A Machine Learning Based Approach for the Prediction of Road Traffic Flow on Urbanised Arterial Roads," *Proceedings - 20th International Conference on High Performance Computing and Communications, 16th International Conference on Smart City and 4th International Conference on Data Science and Systems, HPCC/SmartCity/DSS 2018*, pp. 1285–1292, Jan. 2019, doi: 10.1109/HPCC/SMARTCITY/DSS.2018.00215.

45. Y. Zhang, H. Xu, and J. Wu, "An Automatic Background Filtering Method for Detection of Road Users in Heavy Traffics Using Roadside 3-D LiDAR Sensors with Noises," IEEE Sens J, vol. 20, no. 12, pp. 6596–6604, Jun. 2020, doi:10.1109/JSEN.2020.2976663

46. "12m IP65 distance sensor."http://en.benewake.com/product/ detail/5c345cd0e5b3a844c472329b (accessed Oct. 22, 2022).

47. "ToF LiDAR Distance Sensor|3D LiDAR|Laser Ranging Module|Solid-state Lidar Products|LiDAR with Mass-Poductive Price-BENEWAKE." http://en.benewake.com/?gclid=Cj0KCQjwhsmaBhCvARIsAIbE bH5179aXornxMS8TUaAyzXMqPfCFfAU4Y9Hs GJhiKwvi4 W5uRUzKzkaAnLoEALw_wcB (accessed Oct. 22, 2022).

48. "MIT Deep Learning and Artificial Intelligence Lectures | Lex Fridman." https://deeplearning.mit.edu/ (accessed Oct. 24, 2022).

49. "Theoretical and Advanced Machine Learning | TensorFlow." https://www.tensorflow.org/resources/learn-ml/theoretical-and- advanced-machine- learning?gclid=CjwKCAjwzNOaBhAcEiwAD7Tb6IC9_SQ 58aa JAt6H3o5mOePbVaHo03q9Sb_1azPy1OEB-0LTz18SHBoCVtMQAvD_BwE (accessed Oct. 24, 2022).

50. A. M. Schaefer, S. Udluft, and H. G. Zimmermann, "Learning long-term dependencies with recurrent neural networks," Neurocomputing, vol. 71, no. 13–15, pp. 2481–2488, Aug. 2008, doi: 10.1016/J.NEUCOM.2007.12.036.

51. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput, vol. 9, no. 8, pp.

1735–1780, Nov. 1997, doi: 10.1162/NECO.1997. 9.8.1735.
52. "IEEE SA - IEEE C57.13-2008." https://standards. ieee.org/ieee/ C57.13/4217/ (accessed Oct. 24, 2022).

List of Abbreviations

Abbreviation	Definition
LiDAR	Light Detection And Ranging
LSTM	Long Short-Term Memory
CCC	Congestion Costing Critique
UMR	Urban Mobility Report
NHTSA	National Highway Traffic SafetyAdministration
CO2	Carbon Dioxide
ITMS	Intelligent Traffic ManagementSystems
TMS	Traffic Management Systems
GDP	Gross Domestic Product
GHG	Greenhouse Gas
ARIMA	Autoregressive Integrated MovingAverage
GPS	Global Positioning System
VIP	Video Image Processing
ML	Machine Learning
MLP	Multilayer Perceptron
TDNN	Time Delay Neural Network
RBF	Radial Basis Functions
KNN	K-Nearest Neural Network
CNN	Conventional Neural Networks
RNN	Recurrent Neural Network
UART	Universal Asynchronous Receiver/Transmitter
IEEE	Institute of Electrical and ElectronicsEngineers