Jurnal Kejuruteraan 35(3) 2023: 779-787
https://doi.org/10.17576/jkukm-2023-35(3)-25

Implementation of A.I. Vehicle Detection for Traffic Analysis Using In-situ Surveillance Infrastructure

Saadullah Hyder*, Marjan Gul, Sadiq Hussain, Syed Ilyas Ahmed, Aamir Nazeer & Faheem Ahmed

Department of Civil Engineering, Balochistan University of Information Technology, Engineering and Management Sciences, Quetta, Pakistan

**Corresponding author: saad.shahwani@gmail.com*

Received 1 March 2022, Received in revised form 29 November 2022
Accepted 30 December 2022, Available online 30 May 2023

ABSTRACT

Traffic flow parameters are required for optimizing traffic operations, design of pavements, and future planning of traffic networks. Unfortunately, due to the unique characteristics and variety of vehicles in the sub-continent i.e., size and design, the accuracy of results for a vision-based system is challenged, since most thorough datasets are based on European and American traffic. This paper proposes a solution by developing a detection model ground-up using a dataset created from the local traffic surveillance footage, and creating a python pipeline for vehicle speed detection and classification. The vehicle classification model is developed using the state-of-the-art YOLO object detector which significantly reduces the computation time required to maintain the efficiency of the proposed solution. Furthermore, a computer-vision script is developed to track the movement of vehicles in the footage and record the speeds in a spreadsheet. The technique used eliminates the video calibration, including distance and angle, required for detecting accurate speeds. Finally, the real-time traffic data is analyzed to derive the fundamental traffic flow parameters and discuss the relation between flow and density. To ascertain the validity of this survey technique, the results are compared to the following renowned traffic flow models: The Modified Greenberg model, Eddie's model, and The Two-regime model. The results are found to closely follow the models in all three cases.

Keywords: Traffic analysis; vehicle classification; vehicle detection; YOLO-v3

INTRODUCTION

The study of traffic is an evolving discipline, and many techniques have been developed to make the flow of traffic as smooth as possible. Besides the complete automation of cars, there has never been a complete solution devised for traffic management with the smoothest possible flow. But complete automation is still a very young and upcoming concept, and it is not applicable in underdeveloped regions of the world. No degree of management can eliminate traffic issues, but with the proper application of engineering principles, they can be reduced to a minimum.

The demand for robust traffic management arises as a result of high demographic growth and hyper-urbanization, as seen especially in Asian countries. The lack of safe transport, poor traffic management, and crowded road networks significantly worsen the situation. Proper utilization of investments in this sector requires comprehensive traffic studies, to eliminate future bottlenecks in the road transport system.

Thus, it is critical to adopt a suitable methodology for conducting road traffic surveys that are technically and scientifically sound. It necessitates the requirement of exceptional data which closely mimics the reality of the situation on the roads. This could help increase the efficiency of urban planning designs and traffic management; reducing congestion to a minimum. This makes data collection the most important aspect of any traffic engineering process, and it must be made sure that these three objectives are met in every traffic survey: A rich and reliable survey, conducted without much human effort, and an analysis platform for the data.

The traditional methods for traffic data collection do not satisfy the three objectives all at once. Moreover, the accuracy achieved by manual methods fluctuates heavily between 70% — 95% depending on the effort invested, and is greatly affected by the length of the survey, time of desired delivery of results, fatigue of the staff, etc. The error rate is, therefore, changing over time, says Stofan (2018). The current data collection methods on traffic and transportation infrastructure are generally inefficient. They are either resource-intensive or time-consuming; usually providing single-purpose data only, on top of being expensive and complicated to perform.

Alternatively, the use of artificial intelligence for video analytics can produce compelling results while maintaining the accuracy and precision of results. Stofan (2019) showed that the accuracy of results from computer vision analysis can be maintained above 90%, constantly. The amount of footage processed for traffic analysis is far beyond human capacity to do so. Artificial intelligence provides a much simpler solution, with much more accurate results. Enabling researchers to analyze more data than ever before, improving the understanding of a city's traffic, and helping the decision-makers to make more informed decisions.

This paper identifies the gap in Quetta City's traffic infrastructure and outlines the development of a computer-vision data collection system and compares its performance against renowned traffic models. The video footage is extracted from the existing surveillance matrix of an arterial road in Quetta city in Balochistan, Pakistan. The unique requirements of the city's traffic were kept in mind while developing the vision-based algorithm.

LITERATURE REVIEW

RELATED WORK ON VEHICLE DETECTION

Vehicle detection through artificial intelligence has been recently gaining popularity in free-flowing intelligent transportation systems (ITS) domains. The methodologies are rapidly evolving with time, improving detection, speed, and reliability. Some of the prominent approaches for traffic surveys used, including the development of artificial intelligence for vehicle detection, are discussed as follows:

A computer vision-based nonlinear system-based filter which reduces the computational complexity was proposed by Lei, Xue-Fei, and Yin-ping (2008). It is was an iterative approach to detecting vehicles with the use of a filter that subtracted the background of the vehicle and then accurately detected the vehicle through mathematical morphology and wavelet transform. Researchers have used the same approach as Zheng Lei, but also further added a vehicle classification system to it. A Curvelet coefficient method for vehicle recognition was developed by Rahati et al. (2008) by applying standard deviation and achieving feature vectors. The curvelet coefficient was used because it offered optimal sparseness for images. The number of features of the image decreased from 1894105 through the curvelet coefficient to 81 through the standard deviation of different scales and orientation of the curvelet coefficient. The greater the scale the greater the feature numbers and greater accuracy. Li, Liang, and Zhang (2014) proposed a new and comprehensive method of object detection by applying a Gaussian filter to the video to smooth the frame and convert all RGB (standard color format) images to grayscale. The software then differentiated between the grayscale image background and real frame image and gave back a foreground for the moving vehicle. Then binarization of the image was used to extract the foreground and convert it into a binary image. Otsu's algorithm was then applied to calculate the optimal threshold value. It is the best method to get a threshold value in the image segmentation algorithm. The shadow removal algorithm was used to remove shadows. Then a combination of the virtual detector and vehicle detector was programmed to count and detect every car entering ROI (Region of Interest). Li, Liang, and Zhang (2014) gave a comprehensive method in which the margin of error is minimal, and Tian et al. (2014) further used the Kalman filter to predict the trajectory of the vehicle. The Kalman is a recursive estimation method. However, this method was based on a high-resolution camera, limiting its application and making it ineffective to use with a midrange camera such as a typical video surveillance camera.

Anandhalli and Baligar (2018) showed that the software and hardware can be brought together to get economic results by using a Raspberry PI unit for object detection. In this process, the video in RGB was converted to HSV (Hue meaning color, S defines saturation, and V for luminosity), and then chain codes were used to change the vehicle image into white blobs. Kalman filter was then used to track the centroid of the vehicle. The advantage of this system is that it is the most portable and easy-to-process method, as it does not need a workstation for processing. Moreover, the proposed algorithm was compared to the rear-view vehicle detection system of Tian et al. (2014) and was found more accurate for vehicle detection. Their novel approach not only developed the software but also proposed the use of cheap hardware to transmit the video. By assigning a static IP address, the Raspberry Pi could be remotely accessed and the camera feed from the Pi transmitted to a remote computer, and a background subtraction technique was used to detect approaching vehicles. The background and foreground of the video were crucial perimeters for the detection. Vehicles were classified based on the area they cover.

DISCUSSION ON THE TRAFFIC FLOW MODELS

Traffic analysis is a complex process, so models were developed to fully understand and describe it. The models discussed below are abstractions of the real world that are based on mathematical equations and physical intuition. Over the years, researchers tried overcoming the shortcomings of previous models by introducing new parameters. Speed, flow, and density are the fundamental variables that are necessary to describe these models.

A comparative analysis of traffic models was conducted by Jabeena (2013), showing that the Greenshield model proposed a linear relationship between speed and density, which was an empirical approach for traffic analysis. The shortcoming of this model was that there is almost no linear relationship between speed and density on the field, so the results of this model are often questioned. However, the logarithmic model put forth by Greenberg (1959) assumed that the traffic flows like a continuous fluid. Although the model showed good results than the Greenshields model, it received criticism for its inability to predict speeds at lower densities, it was because when the density approaches zero the speed increase to infinity which violates the boundary conditions. The Underwood exponential model for speed-density relationships tried to cover the flaws of the Greenberg model. This model tends to predict lower speeds than the real ones in the free-flow region and speeds a bit higher than the real ones in the congested portion. The drawback of this model is when speed becomes zero, the density becomes infinite, so it could not be used for the prediction of speed at high density. Pipes (1967) introduced a new parameter "n" in his model. The parameter n can be negative, zero, or positive; in any of the cases, different curves are achieved. If n<1 curve of type A is represented, then it means that the speed of vehicles is fast until the concentration approach maximum. If n = 0 it would give a linear relation like Greenshield, this is a curve type B. If n>1, the vehicles travel slowly at low concentrations, this type of behavior can be seen while driving at night or in a tunnel. Rakha and Crowther (2002) proposed a fourth parameter to cover the shortcomings of traffic stream models by capturing the macroscopic and microscopic behavior of the steady traffic behavior. Van Aerde's single model regime is a combination of the Greenshield and Pipes models, and it aims to overcome the main flaws of these models. Van Aerde assumes that the speed of a vehicle in the uncongested regime is not dependent on traffic density. Ardekani and Ghandehari (2008) said that there are some vehicles on the road even in very light traffic, so they modified the original Greenberg model and introduced k_0 as the minimum density in the model.

METHODOLOGY

Videos from multiple locations with varying light conditions and camera angles are collected for a broader range of perspectives. Each video is manually analyzed to extract photos for training the dataset. Six discrete classes of vehicles are identified based on their PCU (Passenger Car Unit) following recommendations from Adnan (2014).

TABLE 1. Passenger car equivalent factors in heterogenous traffic environment

Class	Categories	PCU
A	Cars	1.00
B	Motorcycles	0.25
C	Rickshaw	0.50
D	Vans	1.00
E	Buses	2.50
F	Trucks	3

A YOLO model is trained using over 12,000 unique photographs of local vehicles in varying lighting conditions. For expanding the training dataset multiple-folds, the Image Augmentation through Jupyter Notebook is utilized to artificially create training images by combining multiple processes such as rotation, shifts, shear, and flips on the existing dataset.

FIGURE 1. Original image (Top), Augmented dataset (Bottom)

The open-source annotation tool Label Img is used to Classify each photograph according to its class, creating a text file accompanying the image with the information about the coordinates of the bounding box. This spatial information is used to train the model in Google Colab, by executing the python pipeline for dataset training using YOLO-V3 architecture.

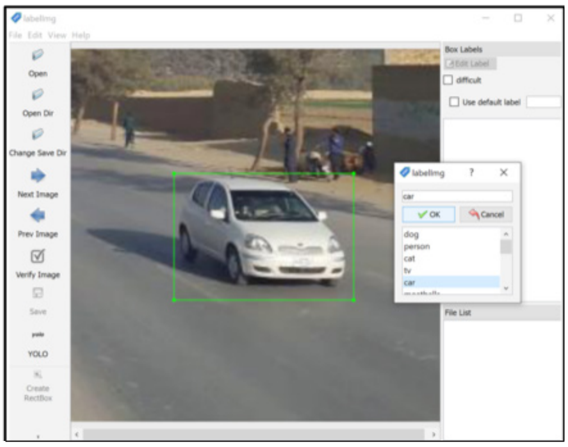

FIGURE 2. Labelling vehicles respective of its type for training YOLO model

Firstly, an object tracker is developed for enabling the computer to count and detect speed. The centroid of each detected object is located using the python script. Treating video frames as a series of images, the object tracker calculates the trajectory of the centroid using the Kalman filter. Secondly, the python script estimates the speed of the object, in pixels per second, on screen by calculating the number of pixels the centroid of each object takes to cross the frame of the video. A horizontal line placed at the center of the frame is programmed to record the speed and category of the object in a spreadsheet.

FIGURE 3. A detected car with speed in pixels per second, with 99% confidence

To convert the pixels into meters, various benchmarks, such as horizontal, vertical, and oblique measurements of road marking and width of the section, from the site are collected. These spatial parameters gathered from the site are used to calibrate the conversion factor. This eliminates the need for a predetermined distance for the ideal location of the camera – thus the existing framework of surveillance cameras in the city can be used for collecting data. Additionally, multiple passes of vehicles with known original speeds are also compared to the detected speed to verify the accuracy of the results. The speed of the object is converted into km/hr using the equation:

$$\frac{\text{pixels}}{\text{second}} * \frac{\text{meters}}{\text{pixels}} * \frac{\text{kilometers}}{\text{meter}} * \frac{\text{seconds}}{\text{hour}} \tag{1}$$

RESULTS AND DISCUSSION

It should be noted that although the traffic flow patterns appear random and influenced by complex factors, they follow a random distribution that can be drawn into defined patterns and later classified and analyzed. Sometimes even the automatic methods for traffic count are not capable enough to grasp the entire traffic behaviors of a section as discussed by The Ministry of Works and Transport Roads Department (2003).

The speed and count are used to calculate the macroscopic parameters i.e., density and flow. This comprehensive survey of data is collected in morning and evening peak traffic. It is reported and reset at every 5-minute interval for an hour each i.e., after counting for 5 minutes, the data is stored in the spreadsheet and the counter starts from zero. This data is used to study the influence of the macroscopic parameters on one another.

FUNDAMENTAL FLOW PARAMETER TRENDS

Speed Trends

Statistically analyzing the data presented in Figure 4, it is observed that the variation in average speed in the morning, at every 5-minute interval, has a correlation coefficient of 0.302, which translates to a weak uphill positive linear relation between speed and time, as discussed by Rumsey (2016). This weak relation indicates that there are factors at play that are affecting the variation in speed which may include human behavior, road condition, visibility, and serviceability of the road.

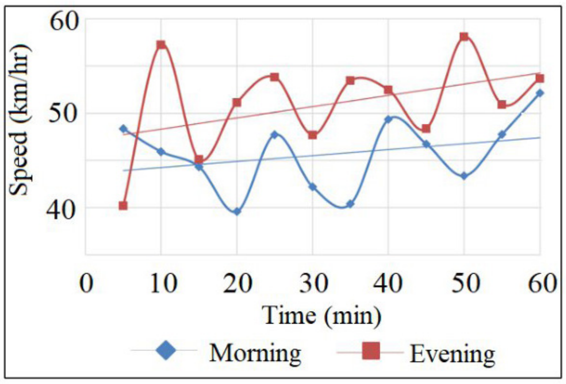

FIGURE 4. Comparison of speed-time scatter for morning and evening peak hours

Similarly, analysis of the evening data shows that the correlation coefficient is 0.423, indicating a moderate uphill linear relation between speed and time. It is then postulated that the speed of the vehicles in the evening is more time dependent, whereas, in comparison, the speed of the vehicles in the morning is not.

TABLE 2. Correlation coefficients of speed trends

		Time	Speed
Correlation for morning peak hour			
Time	Pearson Correlation	1	.302
	N	12	12
Speed	Pearson Correlation	.302	1
	N	12	12
Correlation for evening peak hour			
Time	Pearson Correlation	1	.423
	N	12	12
Speed	Pearson Correlation	.423	1
	N	12	12

Comparing both data sets, it is clear that the speed in the evening is much faster than that in the morning. The evening trend shows a steeper increase in speed with respect to time as compared to the morning. This indicates the possibility of human behavior, including physiological and psychological factors, impacting the variation in speeds. It can be deduced that the drivers in the morning are fresh as compared to those in the evening. Since the evening traffic comprises of people returning from work, and they tend to drive fast to reach their destination as soon as possible to rest.

Flow Trends

In this study, flow is defined as the number of vehicles (n_t) that pass a point during an interval of five minutes. The flow (q) is then given by one of the equations for the fundamental parameters of traffic flow:

$$q = \frac{n_t}{t} \tag{2}$$

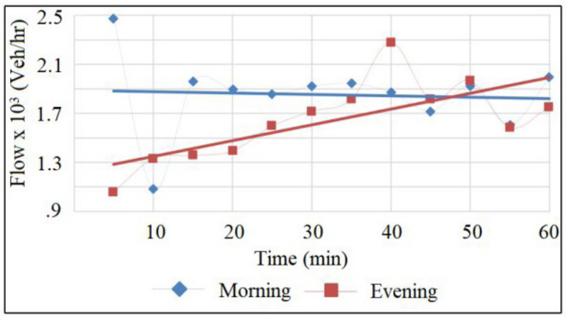

FIGURE 5. Comparison of flow-time scatter for morning and evening peak hours

From Figure 5, it is observed that the flow of morning traffic is decreasing as time passes. The correlation coefficient for morning data is -0.064 and it is showing no evident linear relationship. It is clear that time had no impact on the flow rate in the morning, indicating that there are other factors impacting the speeds, which are out of the scope of this paper. The data for evening flow shows an increase in flow with respect to time. The correlation coefficient for this scatter plot is 0.711, which is concrete evidence of a strong uphill linear relation between flow and time, which means that flow is rapidly increasing with time. This rapid increase in flow is due to the high speeds as discussed previously.

TABLE 3. Correlation coefficient of flow trends

Correlation for morning peak hour			
		Time	Flow
Time	Pearson Correlation	1	-.064
	N	12	12
Flow	Pearson Correlation	-.064	1
	N	12	12
Correlation for evening peak hour			
Time	Pearson Correlation	1	.711
	N	12	12
Flow	Pearson Correlation	.711	1
	N	12	12

Density Trends

Density (k) in this study is defined as the number of vehicles (n_l) present on a given length of a road section of length l:

$$k = \frac{n_l}{l} \tag{3}$$

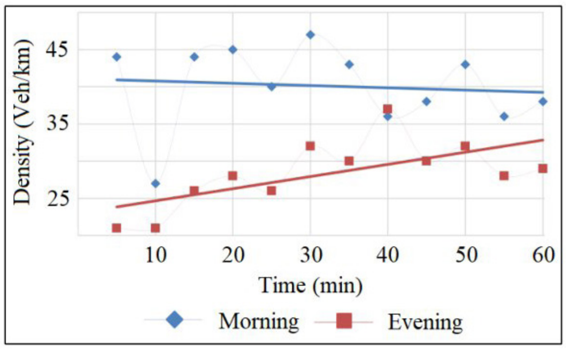

FIGURE 6. Comparison of density-time scatter for morning and evening peak hours

Figure 6 shows a slight decrease in density for morning traffic, which is backed by a correlation coefficient of -0.099. The negative sign shows that there is a downhill trend, but it is very close to 0, which means the impact of the variables on each other is negligible. This indicates that there is no relation between density and time in the morning.

The density calculated for evening traffic shows a rapid increase with time. This is backed by the correlation coefficient test of the two variables, which turns out to be 0.650, representing a strong uphill linear relation between density and time. This means that not only the speeds were higher in the evening as compared to the morning traffic, but the density was also higher. There is evidence that the density in the evening is dependent on time.

TABLE 4. Correlation coefficient of Density trends

Correlation for morning peak hour			
		Time	Density
Time	Pearson Correlation	1	-.099
	N	12	12
Density	Pearson Correlation	-.099	1
	N	12	12
Correlation for evening peak hour			
Time	Pearson Correlation	1	.650
	N	12	12
Density	Pearson Correlation	.650	1
	N	12	12

Flow and Density Relation

Figure 7 shows that the morning scatter plot varies in flow as density changes, which is clear from the uphill gradient of the graph. In the real world, it translates to a great change in flow due to density.

Evening data also shows a good relationship between density and flow, but the correlation coefficient value turns out to be 0.923, which is nearly a perfect positive relationship between the two variables. This means that the evening traffic was almost entirely dependent on the density of the traffic.

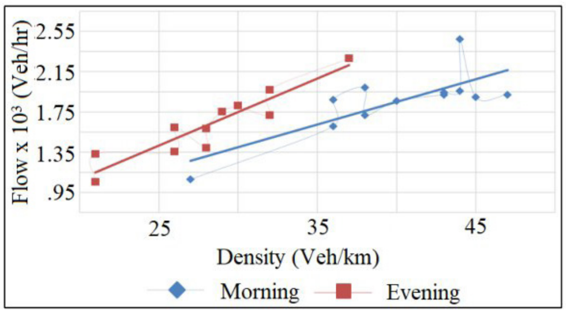

FIGURE 7. Comparison of Flow and density for morning and evening peak hours

When both peak hours are compared in, it is seen that the flow rate in the evening is higher compared to the morning traffic. There is also a greater change in flow with respect to density in the evening, in comparison to the morning traffic. Jabeena (2013) discussed a similar case which enables us to conclude that the increase in flow with density indicates that the traffic was in free flow condition in both graphs.

TABLE 5. Correlation of Flow and Density at morning and evening peaks

		Density	Flow
Correlation for morning peak hour			
		Density	Flow
Density	Pearson Correlation	1	.780**
	N	12	12
Flow	Pearson Correlation	.780**	1
	N	12	12
Correlation for evening peak hour			
Density	Pearson Correlation	1	.923**
	N	12	12
Flow	Pearson Correlation	.923**	1
	N	12	12

Speed and Density Models

In the following sections, the validity of the data is measured through graphs, by analyzing how closely the empirical data fits the ideal traffic models.

The Eddie Model

In Figure 8, the empirical data and the Eddie model are compared, and the similarities and differences are analyzed by plotting trend lines. The most dominant feature observed through the graph is the downhill relation of speed and density and a similar slope of both trend lines.

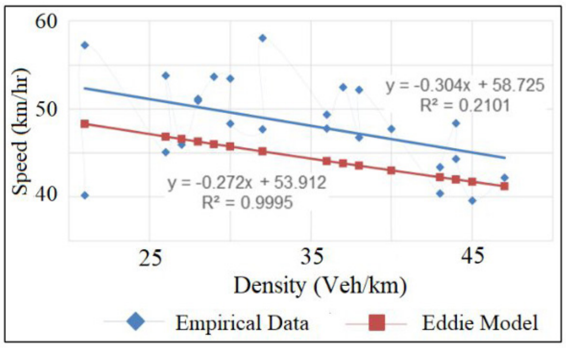

FIGURE 8. Comparison of the empirical speed-density relation against the Eddie's Model of traffic flow

The observation from the graph validates the empirical data. To further solidify this claim, both graphs are statistically compared in Table 6. Carefully studying the table model summary of empirical data and the Eddie model, it is evident that there is almost a 50% difference in r values, and the R^2 value for the empirical data is 5 times smaller than that of the Eddie model. The Eddie model is an ideal relation between speed and density, and the speed is completely dependent on density, whereas, in real life, many unaccounted variables impact this relation.

TABLE 6. Summaries – Empirical model v Eddie model

Model Summary – Empirical Data			
R	R^2	Adjusted R^2	Std. Error
.458[a]	.210	.174	4.68037
a. Predictors: (Constant), Density			
Model Summary – Eddie Model			
R	R^2	Adjusted R^2	Std. Error
1.000[a]	1.000	1.000	.04670
a. Predictors: (Constant), Density			

For free flow for $k \leq 50$ Wang et al. (2011):

$$V = 54.9e^{\frac{k}{163.9}} \tag{4}$$

The empirical data shows a lot of variation because many drivers intuitively slow down before an upcoming U-turn, while others do not. This completely depends on the driving experience and safety precautions. It is clear that human behavior plays a vital role in this speed variation. This data was collected near the U-turn having a speed limit sign of 70 kmph, which is 30 kmph higher than the recommended speed at any other U-turn in the area. This sign may have very well convinced many drivers to drive at higher speeds, while others may have been cautious.

The small R^2 value is an indication of human behavior interfering with the data, as discussed above. Rumsey (2016) shows that statistical data, with a factor of human behavior involved, result in an R^2 value less than 0.50 since the behaviors are not homogenous.

The study of the correlation between the model and the empirical data shows much more similarity as observed in the graphs. The linear regressions in Figure 8 for the model and the data are slightly different i.e., both the constants of the line equations, 53.912 and 58.725, show that the lines started very close on the y-axis. Moreover, the values of the coefficient of empirical data and the Eddie model were -0.304 and -0.272, respectively, which are statically very close to each other. This shows that both the empirical data and the model are behaving similarly. This validates the results and shows that the empirical data is valid enough to be accepted.

The Two-Regime Model

This two-regime model is based on a three-year extensive traffic survey, and it shows promising results. The empirical data was compared to this model to see if the surveying technique used for the collection of the empirical data stands up to the standards of such a robust model.

Comparing the empirical data and the model, it is observed that although both trend lines show a downhill relation, the empirical data is gradually moving away from the model.

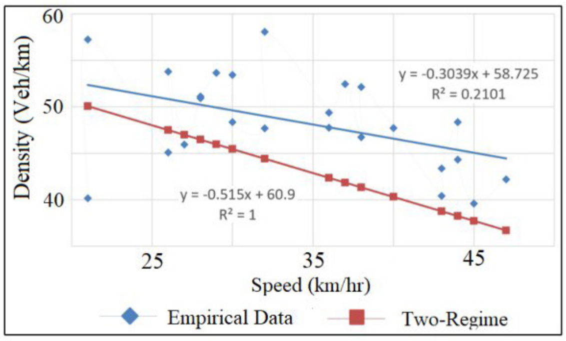

FIGURE 9. Comparison of the empirical speed-density relation against the Two-Regime traffic flow model

Based on the data of the model summary, the correlation coefficient R of the empirical data and two regime model were 0.458 and 1 respectively. The variance of both graphs was 0.210 and 1. The reasons for the smaller R^2 for the empirical data were discussed in the explanation for the Eddie model, and they remain the same for the Two-Regime model. The perfect R and R^2 values of the Two- Regime model is a result of the equation for speed, which is only dependent on density, thus the linear relationship between the two.

TABLE 7. Summaries – Empirical model v Two-regime model

Model Summary – Empirical Data			
R	R^2	Adjusted R^2	Std. Error
.458[a]	.210	.174	4.68037
a. Predictors: (Constant), Density			
Model Summary – Two-Regime Model			
R	R^2	Adjusted R^2	Std. Error
1.000[a]	1.000	1.000	.00000
a. Predictors: (Constant), Density			

The equation for two regime model for free flow when $k \leq 65$ Wang et al. (2011):

$$V = 60.9 - 0.515k \tag{5}$$

After analyzing the coefficient data, it is seen that the empirical data is drifting further away from the model with the increase in speeds. This behavior is due to the coefficient value of the Two-regime model, which is almost twice as compared to the empirical data. The constants for both equations were 58.725 and 60.9 respectively, showing that the graphs started from points much closer to each other on the y-axis, this difference increased with increasing speeds, because the two-regime model was built on an extensive survey of three years. It must be noted that the two regime model data sets do not represent the data sets in near capacity conditions, and therefore the empirical data is drifting away from the model trend.

The Modified Greenberg Model

The Modified Greenberg model was an approach to overcome the drawbacks of the outdated Greenberg model. The empirical data had to be broken into two sets in order to compare it with the free flow and congested flow of the Modified Greenberg model.

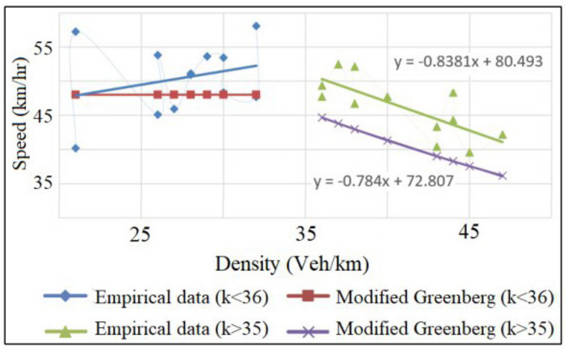

FIGURE 10. Comparison of the empirical speed-density relation against the Modified Greenberg model of traffic flow

Firstly, the data obtained through the software, up to the density of 35 Veh/km (Vehicles per kilometer), was compared to the model. The empirical data showed an uphill relation while the Modified Greenberg model showed no relation between the density and the speed. Further investigation includes statistical analysis of the graphs.

TABLE 8. Coefficients of Empirical Data (k<36)

Coefficients of empirical data					
	Unstandardized		Standardized		
	B	Std-Err[1]	Beta	T	Sig.
Constant	39.57	12.170		3.252	.009
Density	.396	.439	.274	.901	.389
a. Dependent Variable: Empirical Speed (k<36)					

* 1. Standard Error

Upon comparing the graphs, it is seen that the Greenberg model is constant at 48 km/hr as the density increases up to 35 veh/km, but at the same time, empirical data has a moderate uphill change with a change in density. Based on the practical conditions, vehicles cannot maintain a constant speed at a macro level. On the other hand, the speed of empirical data increases with an increase in density because the traffic is in a free flow.

TABLE 9. Model summaries when k>35

Model Summary – Empirical Data			
R	R^2	Adjusted R^2	Std. Error
.758[a]	.575	.532	2.91090
Model Summary – Modified Greenberg			
R	R^2	Adjusted R^2	Std. Error
.999[a]	.999	.999	.10276

Looking at the table 9, it is seen that there is a difference in the R and R^2 data of empirical data and the model. The R values are 0.758 and 0.999 for empirical data and the modified Greenberg model respectively. Similarly, the values of R^2 of empirical data were 0.575 and 0.999 for the modified Greenberg model in congested flow.

The data used for the modified Greenberg model for congested flow originated from the equation of speed solely dependent on density, therefore the speed value changes with the change in density and it is plotted linearly on the scatterplot, this explains the reason for higher R and R^2 values. Equation of speed for modified Greenberg model when $k \geq 35$ Wang et al. (2011):

$$V = 32 l_n\left(\frac{145.5}{k}\right) \tag{6}$$

TABLE 10. Coefficients of Empirical and Modified Greenberg (k>35)

Dependent Variable: Empirical Speed (k>35)					
	Unstandardized		Standardized	T	Sig.
	B	Std. Error	Beta		
(Constant)	80.493	9.369		8.591	.000
Density	-.838	.228	-.758	-3.675	.004
Dependent Variable: Modified Greenberg (k>35)					
	Unstandardized		Standardized	T	Sig.
	B	Std. Error	Beta		
(Constant)	72.807	.331		220.130	.000
Density	-.784	.008	-.999	-97.387	.000

CONCLUSION

The technique used for surveying includes the use of state-of-the-art artificial intelligence to identify vehicles and detect their speeds. This technique eliminates human error involved with manual traffic surveys. The data acquired through A.I. helped simulate the traffic conditions very close to the actual traffic condition on the road. The comparison of results to renowned traffic models, namely, The Eddie model, Two-Regime model, and Modified Greenberg model, verified the data collected using computer-vision. These models represent ideal traffic conditions, and the survey data follows the trends of all these models close enough to hold statistical significance, signifying that the computer-vison with a dedicated dataset can produce very accurate results, and is a valid technique for traffic surveying in underdeveloped regions of the world.

Any difference in the models and survey data is due to human behavior – the models represent the ideal conditions. This deviation can be explained by two well-known psychological theories: Social Cognitive Theory and Experiential Learning Theory. This theory postulates that learning is a by-product of multiple factors, including people's past experiences. Bruneel, Yli-Renko, and Clarysse (2010) discussed this theory and how it directly influences how an individual maintains a behavior. These behaviors are further reinforced due to engagement in a specific behavior and the reasons why a person engages in that behavior. The second theory was proposed by Kolb, known as the Experiential Theory. This theory takes into account the environmental factors and emotions that influence the learning process. Kolb defines this behavioral learning as "The process whereby knowledge is created through the transformation of experience. Knowledge results from the combinations of grasping and transforming the experience"

ROTH and JORNET (2014). These theories help explain why there is variation in average speeds over five-minute intervals in the recorded survey. The people who have experienced and witnessed accidents tend to brake before the junction, and the ones without such experiences have reinforced the idea of total control over the vehicle and maintaining their speeds, which are high in the case of this arterial road.

The purpose of the study was to investigate the possibility of using computer vision for traffic studies in an underdeveloped city by conducting a short-term traffic analysis. The results show that it can successfully identify and store traffic data with limited computing resources. Future work includes the use of this technique for long-term traffic analysis and the environmental impact of vehicles in the city.

ACKNOWLEDGEMENT

We would like to thank the Chairperson of Department of Civil Engineering, BUITEMS, and our supervisors for their help during our research. We highly appreciate the reviewers for their useful feedback to improve the quality of our work. The corresponding author would especially like to thank Zain Tareen for his technical support to make this project possible.

DECLARATION OF COMPETING INTEREST

None

REFERENCES

Adnan, Muhammad. 2014. Passenger car equivalent factors in heterogenous traffic environment-are we using the right numbers? *Procedia Engineering* 77:106–113. DOI: 10.1016/j.proeng.2014.07.004.

Anandhalli, Mallikarjun, & Vishwanath P. Baligar. 2018. A novel approach in real-time vehicle detection and tracking using Raspberry Pi. *Alexandria Engineering Journal* 57 (3):1597-1607. DOI: https://doi.org/10.1016/j.aej.2017.06.008.

Ardekani, Siamak A. & Mostafa Ghandehari. 2008. A Modified Greenberg Speed-flow Traffic Model.

Bruneel, Johan, Helena Yli-Renko & Bart Clarysse. 2010. Learning from experience and learning from others: How congenital and interorganizational learning substitute for experiential learning in young firm internationalization. *Entrepreneurship J* 4: 164-182. DOI: 10.1002/sej.89.

Greenberg, Harold. 1959. An Analysis of Traffic Flow. *Operations Research* 7 (1):79-85.

Jabeena, M. 2013. Comparative study of traffic flow models and data retrieval methods from video graphs.

Lei, Z., Z. Xue-fei & L. Yin-ping. 2008. Research of the Real-Time Detection of Traffic Flow Based on OpenCV. 2008. International Conference on Computer Science and Software Engineering, 12-14 Dec. 2008.

Li, Da, Bodong Liang & Weigang Zhang. 2014. Real-time moving vehicle detection, tracking, and counting system implemented with OpenCV. *2014 4th IEEE International Conference on Information Science and Technology*: 631-634.

Ministry of Works and Transport Roads Department, Gaborone, Botswana. 2003. Traffic Data Collection and Analysis. In.

Pipes, Louis Albert. 1967. Car following models and the fundamental diagram of road traffic. *Transportation Research* 1:21-29.

Rahati, S., R. Moravejian, E. M. Kazemi & F. M. Kazemi. 2008. Vehicle recognition using contourlet transform and svm. Fifth International Conference on Information Technology: New Generations (itng 2008), 7-9 April 2008.

Rakha, Hesham, & Brent Crowther. 2002. Comparison of Greenshields, Pipes, and Van Aerde Car-Following and Traffic Stream Models. *Transportation Research Record* 1802 (1):248-262. DOI: 10.3141/1802-28.

Roth, Wolff-Michael, & Alfredo Jornet. 2014. Toward a Theory of Experience. *Science Education* 98 (1):106-126. DOI: https://doi.org/10.1002/sce.21085.

Rumsey, Deborah J. 2016. How to interpret a correlation coefficient r. *Statistics For Dummies*.

Stofan, Daniel. 2018. The development of traffic data collection methods. https://medium.com/goodvision/the-development-of-traffic-data-collection-cd87cc65aaab.

Stofan, Daniel. 2019. The Guide on Making a Suitable Video Footage for Video-analytic AI.https://medium.com/@danielstofan/the-guide-on-making-a-suitable-video-footage-for-video-analytic-ai-820b7b013a52.

Tian, Bin, Ye Li, Bo Li, & Ding Wen. 2014. Rear-view vehicle detection and tracking by combining multiple parts for complex urban surveillance. *Intelligent Transportation Systems, IEEE Transactions on* 15:597-606. DOI: 10.1109/TITS.2013.2283302.

Wang, Haizhong, Jia Li, Qian-Yong Chen, & Daiheng Ni. 2011. Logistic modeling of the equilibrium speed–density relationship. *Transportation research part A: policy and practice* 45 (6):554-566.