VEHICLE DENSITY ESTIMATION IN QUEZON CITY USING OBJECT-BASED FEATURE EXTRACTION ON SATELLITE IMAGES

H. B. Banaag , M. S. Litana , R. V. Ramos

Department of Geodetic Engineering, University of the Philippines, Diliman
Quezon City 1101, Philippines– (hbbanaag, mslitana, rvramos) @up.edu.ph

KEYWORDS: Transportation, vehicle count, object-based, feature extraction, Google Earth imagery

ABSTRACT:

Manual vehicle counting is often tedious, expensive, and time-consuming. While automatic counting from CCTV allows for annual average daily traffic estimation, CCTV files in the Philippines are not available to the public and do not fully cover all road extents. In this study, Remote Sensing and Geographic Information Systems (GIS) techniques are employed to use readily available satellite images to obtain vehicle count in selected road segments in the Central Business Districts of Quezon City before and after the COVID-19 lockdown. Using the existing Google Earth Images, a segmentation algorithm using ENVI Feature Classification was developed to allow remote counting of vehicles from the earliest image in 2018. The devised algorithm was able to delineate, identify, and classify according to the types of vehicles that are visible on the image. An average error rate of 12.24% was found by comparison of automated counts and manual counts on the images, while a regression analysis yielded a value of $R^2 = 0.9227$ that denoted a strong relationship between automated and manual counts. Vehicle density was calculated, and percent differences were obtained to determine the relative differences of the vehicle counts from the vehicle count of the earliest image taken in 2018. It was found that the vehicle density declined by at least 81% by March 25, 2020. The methodological framework presented in this study provides estimates of vehicle counts and vehicle density. It can be further improved if vehicle counts, on the same location and period, from field validation surveys are available.

1. INTRODUCTION

1.1 Background of Study

The Philippines declared its first case of the novel coronavirus on January 30, 2020, followed suit by the proclamation of the World Health Organization (WHO) that COVID-19 is a global pandemic. On March 15, 2020, the National Capital Region (NCR) and other key regions were placed on strict community quarantine (CNN Philippines, 2020). One of the quarantine measures imposed by the local government units in NCR is the adoption of work-from-home arrangement for schools, offices, and businesses, which made private and public transport very limited for healthcare workers and delivery of essential goods and services. With that, the transportation sector has been significantly affected. Recent studies involve assessing the changes in various sectors, including the transport sector, brought about by the COVID-19 pandemic.

Transportation is a crucial indicator of economic growth and is used by the government and investors to make big decisions regarding business, location, and development (Casey et al, 2001). In their study outlining transport policy implications of pandemics and lockdowns, Hasselwander et al (2021) stated that while many studies produce generalized findings, there is a need for separate investigations for megacities due to their distinctive features. By studying changes in transport behavior, we can quantify the adverse effects of inefficient transportation, determine the sector and individuals most affected, and form plans, and policy changes specifically catered to specific individuals and occasions to minimize its impact (Monschauer et al, 2020). Different variables can be considered in evaluating the changes in transport behavior of the residents such as vehicle counts (Brown et al, 2016), number of person trips and other indicators describing how people move from their homes to their workplace and vice versa.

In 2019, an Asian Development Bank (ADB) study reported that Metro Manila is the "most congested city" of 278 cities in Asia. According to the Metro Manila Development Authority (MMDA), the daily average number of vehicles that passed EDSA in 2019 (excluding other NCR roads) was 405,882. The number of cars exceeded the total number of jeepneys, UV express, and taxis by almost tenfold (255, 732 compared to 27, 364). As public transit and people's mobility were restricted, the Philippines, by 2020, has entered an economic recession (Dela Cruz, 2020). In order to revive the country's economy, transport restrictions were eased. The change in modes of quarantine restrictions would have resulted in changes in transport behavior, which may be quantified by the number of people on the streets, changes in transport routes, or the volume of vehicles passing through the Metro Manila roads.

1.2 Research Objectives

The study aims to create a methodological framework for delineation, counting, and classification of vehicles using publicly available satellite images and object-based feature extraction algorithm, to calculate vehicle density on selected roads in Quezon City and to determine whether these values changed before and during the lockdown.

1.3. Scope and Limitations

This study is mainly focused on using object-based image feature extraction on satellite images to obtain a traffic count of Quezon City CBDs (i.e., Triangle Park, Eastwood City, Araneta Center). The study concentrated only on specific sites/road segments within the three CBDs (refer to Figures 1 to 3) for the period 2018 - 2021. The datasets and imagery used for this study are limited on the available Google Earth Pro images. The researchers only used the ENVI software, so the workflow of this study is based on the features and parameters that are built into the software. A limitation of this study is that no ground truth data could be used to determine the accuracy of the vehicle counts.

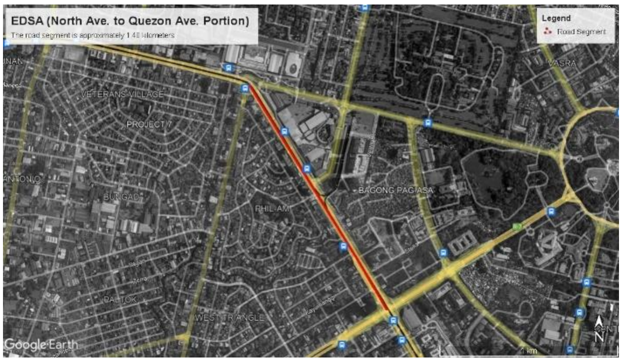

Figure 1. Map of EDSA (North Avenue to Quezon Avenue portion) Segment

Figure 2. Figure 3.3. Map of EDSA Segment (Cubao to White Plains Ave. portion)

Figure 3. E. Map of Rodriguez Jr. Avenue

1.4 Review of Related Literature

1.4.1. Traffic Counting: When the lockdown was imposed, Filipinos in NCR noted cleaner air, decongested roads, and faster travel time (Conde, 2020). These may be attributed to fewer people going to work or roaming the streets and the imposed ban on public transportation earlier this year. Hasselwander et al. (2021), in their study of Metro Manila as "a characteristic megacity that experienced one of the most stringent lockdowns worldwide," wrote how the COVID-19 pandemic has "affected human mobility via lockdowns, social distancing rules, home quarantines, and the full or partial suspension of transportation." In their paper, studies were cited which prove that transport contributes to the geographical spread of the disease, and it highlights having an increased risk of infections in public transport.

Traffic Data Collection is necessary for road planning and management schemes. Traffic data is essential knowledge in drawing up a rational transport policy for the movement of passengers and goods by both government and the private sectors. Obtaining traffic data is performed through traffic counting, of which there are two types: Manual counting and Automatic counting. Manual counting uses people to man specific junctions and intersections on the road and counts the number of vehicles as they pass by. In contrast, automatic counting counts vehicles in a non-intrusive fashion using electromagnetic spectra and wireless communication media (Roads Department of Botswana, 2004). For this research, the authors used automatic counting to obtain traffic data.

ADT (Annual Daily Traffic) is the most used measure of traffic. The ADT of a hiTrghway may be visualized as if someone counts the number of vehicles passing a given point in that highway for 24 hours a day, 365 days a year, and then dividing all vehicles counted by 365. As this may be tedious, ADTs may be annualized by obtaining a count for shorter durations such as a week, then applying adjustment factors from nearby count stations. These adjustments are performed to overcome biases and effects of seasonal fluctuations to short-duration counts. The adjusted values are called the AADT (Annual Average Daily Traffic) wherein it is used to describe traffic volume characteristics for transportation planning (Schroeder, 2016).

1.4.2. Feature Classification: In a study conducted by Chabot et al., 2018, the authors developed an approach for detecting and counting birds from aerial imagery using object-based image analysis software. The images were existing digital imagery taken from aerial surveys of breeding colonies throughout the Canadian Arctic. During the initial review of the images, it was judged by the authors that images with resolutions coarser than 6 cm are not suitable for computer-automated detection of geese because of the decreasing accuracy by which the birds could be identified. Hence, the authors used images with 3 cm to 6 cm resolution. First, the images were opened in Quantum GIS (QGIS), and a manual identification was performed by marking each goose with a point. The points were saved in a shapefile. The authors noted that although the geese were easy to identify, they wanted to account for possible omissions and false identifications.

The second part of their methodology employed ENVI's feature extraction method, which uses a watershed segmentation algorithm. A set of standard segmentation parameters suitable to all images was selected so the rest of the images may be batch-processed. Image processing produced a polygon shapefile output for each image "delineating all segmented objects and an associated attribute table containing the spatial, spectral, and texture attribute values for each object." A bird object classification rule set was developed so that by looking at the text files, only those who meet all the criteria in the ruleset will be classified as birds. To sequentially process all image files, an Interactive Data Language (IDL) script was written. Effective segmentation was found on 4cm and 5 cm image resolution. The study showed that the developed automated analysis routine was more efficient than manual counting, although the manual counting was helpful to countercheck the results obtained from the automated analysis.

This research employed the same manner in counting the cars from the roads. Since cars are much larger than geese, we may be fine with using images with larger resolutions. However, Chabot et al. were careful to note that the background can skew the identification of the geese; hence the researchers minimized the unnecessary elements in the photos used in this research. The software used in Chabot et al.'s research is ENVI, a software for processing and analyzing geospatial imagery.

The 2018 study by Guirad et al. used Google Earth images for their method in whale detection: a robust and generalizable CNN-based (Convolutional Neural Networks) system. The dataset for testing and validating the whole procedure consists of RGB satellite images downloaded from Google Earth in 14,148 cells of 71 m x 71 m distributed worldwide. Since the Google Earth images are free as opposed to costly surveys and expensive production of high-resolution orthoimages, they deemed the method as cost-effective with a performance of (F1-measure) of 84% in detecting and 97% in counting 80 whales.

The goal of the devised workflow for this research is to minimize the errors as much as possible and to have an error that is within the acceptable range. To be able to ensure that the results obtained in the automated workflows are accurate and conform to the actual values, there should be a way to validate the results.

Mishuk, in his 2020 Master's Thesis "An Approach to Counting Vehicles from Pre-Recorded Video Using Computer Algorithms," validated his obtained results by comparing the automated counts to the manual counts. The manual counts were obtained by counting the number of vehicles that passed through the video frame in a given period. The difference between the two values was computed by calculating the percentage difference. The author mentioned how "there is no actual ground value to estimate errors of manual counts," but there is a way to estimate errors indirectly, and that is to have a repeat count. This is similar to how Chabot et al, in their above-mentioned study, validated their results. The obtained values from the ENVI workflow were compared to the values obtained through manual count in QGIS. The percent difference between the two values was also computed to determine if the automated count conforms to the manual count.

In order to properly delineate and classify vehicles, information regarding the spatial and spectral properties of such should be outlined. Spatial properties refer to the physical attributes of the vehicles, which are not only helpful in identifying if an object is a car but also as to its type and classification (motorcycle, car, van, bus, truck). These properties include but are not limited to length, width, height, area, and color. Spectral properties provide the observable signature for hyperspectral remote sensors and are measurable quantities. These include spectral reflectance, transmittance, and emissivity.

2. METHODOLOGY

2.1 Datasets and Materials

The test sites in this study include Quezon City's three (3) Central Business Districts (CBD), namely Triangle Park, Araneta Center, and Eastwood, which represent the city at a micro-level. Roads around the three CBDs were selected as the test site according to the following criteria: proximity to establishments, road classification, and the number of intersections with other roads and avenues.

Available raw images from Google Earth Pro covering the specified CBD areas, a total of 220 images with the highest resolution (4800 x 2782), were utilized in this study. Images obtained are within the span of January 18, 2018, until February 25, 2021. All downloaded images were in JPEG format (120 dpi).

2.2 Image Preprocessing

Clipping the roads from the images reduces the processing time of the ENVI software. The downloaded images from Google Earth Pro contain buildings, trees, and other objects that are not needed for this study and may be eliminated to reduce the processing time of the ENVI software.

In order to clip the roads, a polygon buffer is needed that will cover the whole stretch and width of the roads. This is to provide the extent of the image to be clipped. However, the road buffers cannot be decided arbitrarily nor subjectively chosen as this can affect the integrity of the results. It is a must that the buffers amply cover the width of the road and at the same time allow room for error due to vehicles going out of lane on the outermost width of the road. Therefore, the buffers were created using standard values for road lanes, adding an allowance for error.

The road buffers were created using the buffer tool in ArcMap 10.3. Buffer values were calculated by multiplying the number of lanes to the standard national road width, equivalent to 3.4 meters, and adding it to the widths of train tracks (approximately 3 to 4 m) and allowances for roadsides (refer to Table 1).

Roads	Number of Lanes	Buffer Diameter (m)
EDSA (Cubao to White Plains)	14	54.4
EDSA (North Ave. to Quezon Ave)	14	54.4
E. Rodriguez Jr. Ave.	10	40.8

Table 1. Number of Lanes and Buffer Diameter Values of Road Segments

2.3 ENVI Feature Extraction Workflow

2.3.1 Establishment of Image Segmentation Parameters: Feature classification involves determining objects on satellite images using defined segmentation parameters. The object-based approach employs the concept of a segment, which is described as a group of pixels with similar spectral, spatial and/or texture attributes (Chabot et al, 2018). Two primary image segmentation parameters were analyzed in this study, and these are (1) segment value, which indicates the degree of grouping neighboring pixels with a common value, and (2) merge value, which indicates the degree of combining adjacent segments with similar spectral attribute. Different settings and level values were tested for their accuracy in delineating objects.

2.3.2. Creation of Ruleset for Vehicle Identification and Classification: To further enhance the accuracy of delineating vehicles, a rule set based on vehicle lengths was set to distinguish vehicles from other ground features such as road lanes, buildings, road barriers, and other objects near the vicinity of the roads. The spectral attributes of vehicles are not yet established, so the researchers relied on the vehicles' spatial attributes that are apparent on a bird's eye view (i.e., vehicles' dimensions and color) to create the ruleset.

The researchers used the vehicles' lengths as the basis for image classification. That is, a vehicle of a specific length would be classified as a car, a van/jeep, a bus, or a truck, depending on the length delineated by the algorithm. The standard lengths (ground image length) of representative vehicles as taken from the websites of their respective companies are converted image object lengths. The algorithm calculates and gives the image object length as its output as opposed to calculating and giving the ground object length; hence, the image object length is used as the condition for the created ruleset (refer to Table 2).

Vehicle types were segregated from the output vector classes and were counted using table operations in ArcGIS.

Vehicle	Range of Length for Ruleset
Cars	20.00000 - 50.00000
Vans/Jeepneys	50.00001 - 70.00000
Buses	70.00001 - 90.00000
Trucks	90.00001 - 200.00000

Table 2. Range of Length for Ruleset per Vehicle Type

2.4 Accuracy Assessment

The downloaded images were opened in QGIS, and a manual identification was performed by marking each vehicle with a point. The total number of placemarks in a single image is taken as the manual count of cars for an image. The automated count is compared to the manual count to determine the percent error.

Errors may be attributed to image quality and misclassification of features such as train tracks, parking lot facilities, and buildings into vehicles.

3. RESULTS AND DISCUSSION

3.1 Vehicle Delineation, Classification, and Counting

The image segmentation parameters that are observed to be optimal in the feature classification workflow consist of a low segment value equivalent to 20 and a high merge value equivalent to 90, resulting to 9.71% relative accuracy.

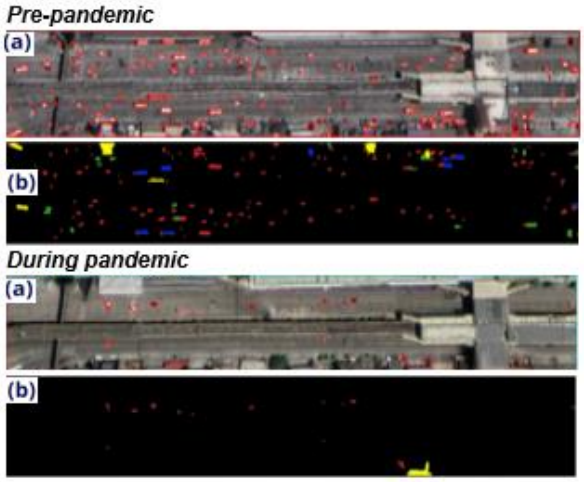

Figure 4. Vehicles in Triangle Park delineated in (a) Google Earth image and (b) segmentation output image

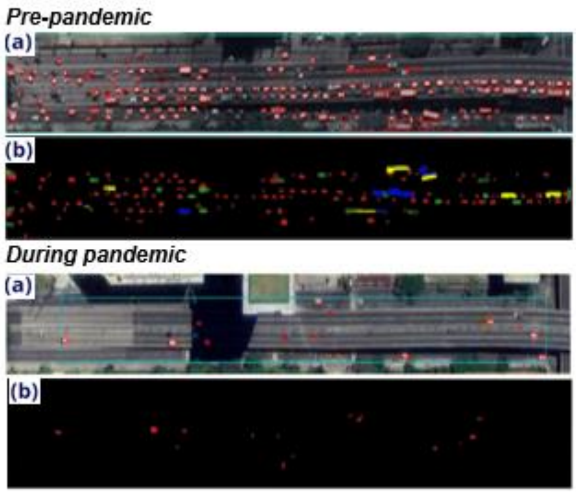

Figure 5. Vehicles in Eastwood delineated in (a) Google Earth image and (b) segmentation output image

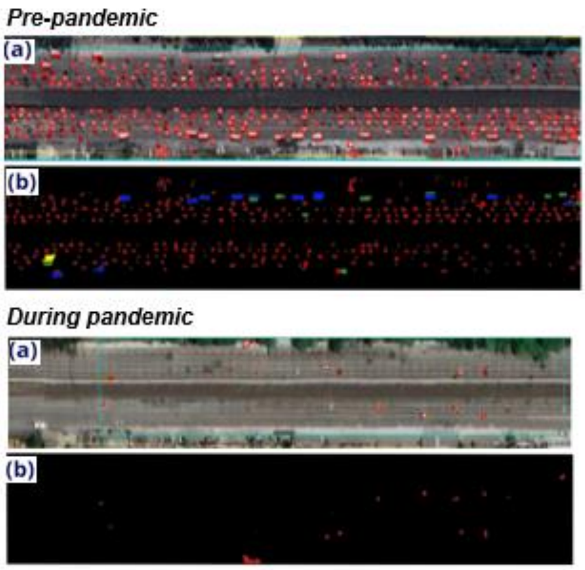

Figure 6. Vehicles in Araneta Center delineated in (a) Google Earth image and (b) segmentation output image

The results indicate the estimated extent and types of vehicles delineated in each site before and during the lockdown (refer to Figures 4 to 6). Vehicles are represented as polygons categorized into cars (in yellow), vans/jeepneys (in green), buses (in blue), and trucks (in yellow).

The overall error rate of the object-based feature extraction algorithm ranges from 0.57% to 22.22%, with an average error of 12.24%. The time it took to process an image to obtain vehicle count is 40 seconds, while manual counting takes around one to seven minutes depending on the number of vehicles in an image.

The feature extraction results show that there are a total of 1,044 vehicles delineated in Triangle Park, 1,883 vehicles in Araneta Center, and 2,217 vehicles in Eastwood during the lockdown, which are relatively lower than the number of vehicles observed from the year 2018 prior to the pandemic (refer to Table 3).

The figures (refer to Figures 7 to 9) illustrate that vans/jeepneys have the highest decline, a relative decrease of 92% between the days representing the pandemic period (March 25, 2020) and pre-pandemic (April 17, 2018).

It is worth noting that the vehicle count in the figures are not the average vehicle counts of the inclusive years and the numbers are only from a portion of a road segment (i.e. EDSA Triangle Park and EDSA-Araneta Center)

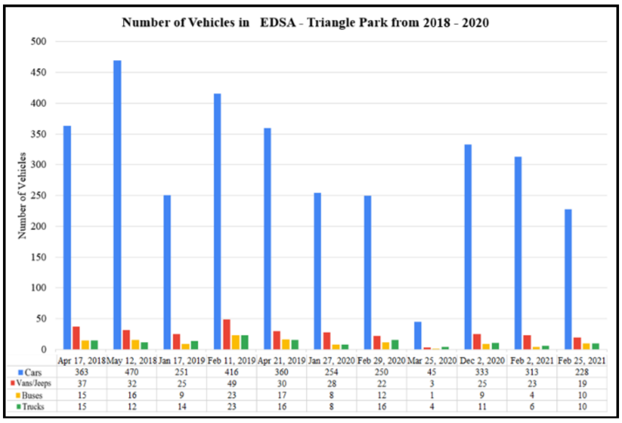

Figure 7. Number of Vehicles per Classification in EDSA Triangle Park, Eastwood City from 2018-2021

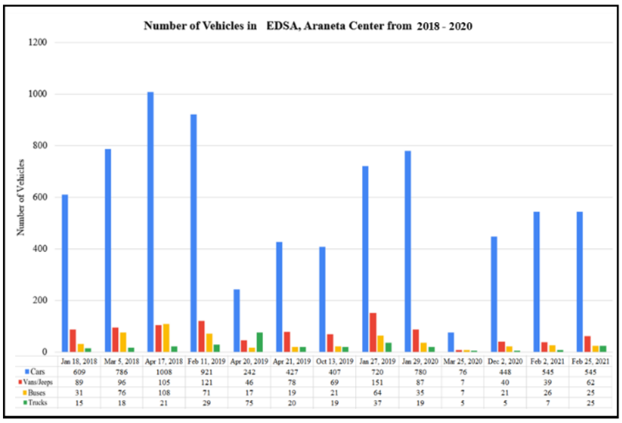

Figure 8. Number of Vehicles per Classification in E. Rodriguez Avenue, Eastwood City from 2018-2021

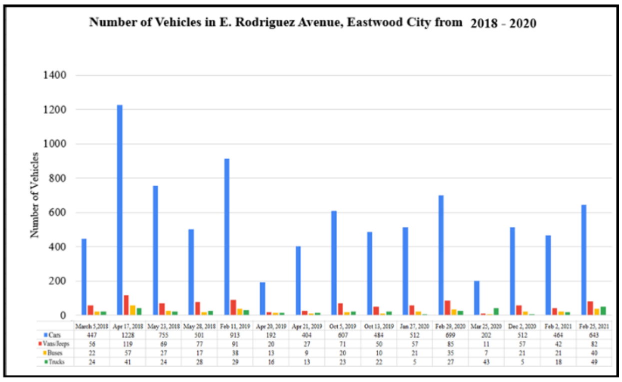

Figure 9. Number of Vehicles per Classification in E. Rodriguez Avenue, Eastwood City from 2018-2021

Image Date	Triangle Park	Eastwood City	Araneta Center
Jan. 18, 2018	--	--	744
Mar. 5, 2018	--	549	976
Apr 17, 2018	430	1445	1242
May 12, 2018	530	--	--
May 23, 2018	--	875	--
Jan. 17, 2019	299	--	--
Feb. 11, 2019	511	1071	1142
April 20, 2019	--	241	380
Apr. 21, 2019	423	453	544
Oct. 5, 2019	--	721	--
Oct. 13, 2019	--	566	516
Jan. 27, 2020	298	595	972
Jan. 29, 2020	--	--	921
Feb. 29, 2020	300	846	--
Mar. 25, 2020	53	263	95
Dec. 2, 2020	378	595	514
Feb. 2, 2021	346	545	617
Feb. 25, 2021	267	814	657

Table 3. Total Vehicle count from 2018 to 2020 in the three sites

3.2 Vehicle Density

The data from April 17, 2018 was used as the baseline for calculating the percent differences in the data from all dates. All the study areas had a decline in vehicle density ranging from 81% to 92% decline (refer to Tables 4 to 6). This goes to show that the restrictions in transportation imposed effectively reduced the number of vehicles on the road during the time of lockdown.

Date of Image	Vehicle Density (Vehicles/Km)	Percent Difference
2018 Apr 17	26.029	0
2019 Feb 11	30.932	18.8366821
2019 Apr 21	25.605	-1.62895232
2020 Jan 27	18.039	-30.6965307
2020 Feb 29	23.184	-10.9301164
2020 Mar 25	3.208	-87.6752852
2020 Dec 2	22.88	-12.0980444
2021 Feb 2	20.944	-19.5359022
2021 Feb 25	16.16	-37.9154020

Table 4. Percent Differences of Vehicle Density from EDSA-Triangle Park

Date of Image	Vehicle Density (Vehicles/Km)	Percent Difference
2018 Apr 17	49.503	0
2019 Feb 11	36.691	-25.88125972
2019 Apr 21	15.519	-68.65038483
2020 Jan 27	23.09	-53.35636224
2020 Feb 29	28.983	-41.45203321
2020 Mar 25	9.01	-81.79908288
2020 Dec 2	20.387	-58.81663738
2021 Feb 2	18.671	-62.28309395
2021 Feb 25	27.886	-43.66806052

Table 5. Percent Differences of Vehicle Density from E. Rodriguez Ave.

Date of Image	Vehicle Density (Vehicles/Km)	Percent Difference
2018 Apr 17	42.636	0
2019 Feb 11	39.203	-8.051881039
2019 Apr 21	18.332	-57.00347124
2020 Jan 27	33.368	-21.73749883
2020 Feb 29	31.617	-25.84435688
2020 Mar 25	3.261	-92.35153392
2020 Dec 2	17.645	-58.61478563
2021 Feb 2	21.181	-50.3213247
2021 Feb 25	22.554	-47.10104137

Table 6. Percent Differences of Vehicle Density from E. Rodriguez Ave.

In all study areas, the vehicle density increased on December 2, 2020. This may be attributed to the easing of quarantine measures in NCR. However, this decline is not equal for all study areas. Based on the ranking, E. Rod Avenue - Eastwood City showed the most significant decline in vehicle count during the pandemic, followed by EDSA - Araneta Center and then by EDSA- Triangle Park.

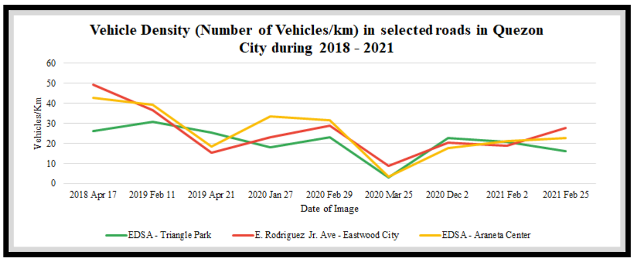

Figure 10. R Vehicle Density (Number of Vehicles/Km) in select roads in Quezon City during 2018 - 2021

The trend (refer to Figure 10) as to whether the vehicle density continued to increase or decrease during the dates succeeding December 2, 2020, varies. As mentioned above, in EDSA-Triangle Park, the vehicle density continuously declined after March 25, 2020, in EDSA- Araneta Center, the vehicle density increased after March 25, 2020; in E. Rodriguez Avenue - Eastwood City, the vehicle density fluctuated from increasing to decreasing and increasing again.

4. CONCLUSION

The segmentation algorithm can delineate, identify, and classify according to the types of vehicles that are visible on the image. Prior to processing, the image was cropped down to the road extent to ensure that the objects classified were most likely vehicles. Using the optimal values for the segment and merge parameters and incorporating a ruleset to filter the classified objects further have enhanced the ability of the algorithm to perform the said objective.

A limitation of this study is that no ground truth data could be used to determine the accuracy of the algorithm. If there was a published vehicle count on that specific date and time on the exact location, an error computation might be performed. However, the researchers have found no such data.

There are gaps in the data that the authors could not have sufficiently provided an answer to as not all days nor months are fully accounted for. However, what they can establish are the relative differences between the vehicle count before and after the lockdown. As shown in the results, there was a significant decline in vehicle volume on March 25, 2020 (during lockdown) from February 29, 2020 (before lockdown). This is true for all study areas. This shows that the lockdown and the imposed restrictions in transportation have significantly reduced the number of vehicles on the road.

Also mentioned are the trends of increase or decrease in vehicle count after March 25, 2020. There were periods where quarantine restrictions were eased or raised, so this would account for the increase in vehicle count, especially in the month of December where it was declared GCQ in NCR. However, the trends for the study areas come February 2021 differ between the study areas. The percentage of decline in vehicle count per study area also differs. This may be attributed to the nature of the location itself, such as the population, economic status of its residents, and the establishments surrounding the area.

REFERENCES

Brown, T. H., Cunningham, C.M., Findley. D.J., Schroeder, B., 2016: *Highway Engineering: Planning, Design and Operations.* Elsevier Inc.

Casey, J., Norwood, J., 2001: *Key Transportation Indicators.* NATIONAL ACADEMY PRESS. Washington, D.C., 22

Chabot, D., C. Dillon, and C. M. Francis., 2018: *An approach for using off-the-shelf object-based image analysis software to detect and count birds in large volumes of aerial imagery*. Avian Conservation and Ecology 13(1):15.

CNN Philippines, 2019: *Metro Manila is most congested city in developing Asia, study shows*. https://cnnphilippines.com/news/2019/9/26/Metro-Manila-most-congested-city-Asia-ADB.html

CNN Philippines, 2020: *Metro Manila to be placed on "lockdown" due to COVID-19*. https://www.cnnphilippines.com/news/2020/3/12/COVID-19-Metro-Manila-restrictions-Philippines.html

Conde, M., 2020: *As lockdown ends, Manila's dirty air is back. It doesn't have to stay*. https://news.mongabay.com/2020/06/as-lockdown-ends-manilas-dirty-air-is-back-it-doesnt-have-to-stay

Guirado, E., Tabik, S., L. Rivas, M., Alcaraz-Segura, D., Herrera, F., 2018: *Automatic whale counting in satellite images with deep learning*. 10.1101/443671.

Hasselwander, M., Tamagusko, T., Bigotte, J., Ferreira, A., Mejia, A.& Ferranti, Emma., 2021: *Building back better: The COVID-19 pandemic and transport policy implications for a developing megacity*. Sustainable Cities and Society. 69. 102864. 10.1016/j.scs.2021.102864.

Mishuk, M., 2020: *An Approach to Counting Vehicles from Pre-Recorded Video Using Computer Algorithms*. LSU Master's Theses. 5231. https://digitalcommons.lsu.edu/gradschool_theses/5231

Monschauer, Y., Sung, J., 2020: *Changes in transport behaviour during the Covid-19 crisis*. International Energy Agency.

Schroeder, B.J., 2016: *Part 2 - Transportation Planning*. Highway Engineering, Butterworth-Heinemann, 17-88. doi.org/10.1016/B978-0-12-801248-2.00002-2.