Cipher: Integration of Computer Vision Technologies in Monitoring of Overstaying Buses in EDSA Carousel Bus Stations

Patrick Lawrence Simpao[1], Jester Darylhans Ramos[1], Christian Michael Tan[1], Ricson Andrei Vergara[1], Mary Jane Arcilla[1], Christine Diane Ramos[1],and Joel Ilao[2],
[1] *Information Technology Department College of Computer Studies*
De La Salle University
[2] *Computer Technology Department College of Computer Studies*
De La Salle University
**Corresponding Author: joel.ilao@dlsu.edu.ph*

Abstract: "Cipher: A Bus Overstaying Monitoring System" is developed for the Metropolitan Manila Development Authority (MMDA) in monitoring and managing the EDSA carousel. It aims to automate the detection of bus overstaying violations through the integration of computer vision technologies and develop a dashboard for the visualization of data gathered, to allow for easier insights generation on the side of MMDA. Cipher has three modules: first is the data preparation module for handling video inputs, second is the detection module for bus detection, bus tracking, and bus timing, and lastly, the visualization module where data is displayed in a web dashboard consisting of charts and tables. It was tested through a comparison between the data acquired through manual monitoring of overstaying buses and the data generated by the detection module, yielding a 0.53% difference between the system-generated and manually generated data. Future studies aim mainly to focus on extending the system to other terminals as well as discussing the possibility of utilizing the system for other vehicle types.

Key Words: Computer Vision; EDSA Carousel; Bus Monitoring System

1. INTRODUCTION

In the context of the Philippines, public transportation is seen as a rudimentary aspect of every citizen's day-to-day life. Having said this, certain government branches are assigned responsibilities to ensure that every commuter's experience would be as seamless as possible. With the dawn of technological breakthroughs upon us, it is only fitting to integrate the existence of such innovative tools into our everyday tasks, including governmental processes, in order to reinforce the way they are executed. Given these opportunities, this study aims to utilize existing Computer Vision Technologies to streamline the apprehension of overstaying bus violations along the stations of the EDSA Bus Carousel. Through the integration of Computer Vision Technology and Data Visualization, involved organizations may be able to reduce the manpower required in the field to monitor violations and maximize the capabilities of the existing

CCTV cameras along the target locations, and even acquire the capability to produce significant insights based on the raw and processed data collected by the system.

To efficiently manage the entry and exit of buses in the stations, MMDA traffic enforcers implement the *bus-staying time rule*, which uses the staying time of a particular bus in the bay as the basis for sanctioning a violation: when a bus stays at the loading bay for more than 1 minute, it will be flagged as overstaying. Currently, this policy is being monitored manually by officers deployed on the field, which often leads to undetected violations and inconsistent violation detections by the personnel on duty. With this, automated bus monitoring and overstaying violation detection are recommended, as it will reduce the probability of human errors occurring and improve the effectiveness of the CCTV monitoring systems.

Automation technology can be used to address some of the issues and limitations of manual enforcement listed above; hence, the Cipher project was started. The Cipher project's goal is to integrate computer vision technology in order to improve MMDA's CCTV monitoring at the EDSA Carousel and provide an interface that contains the data produced in order to assist in producing valuable insights regarding the possible results of the implementation of the existing policies and aid the organization itself determine the final policy that would be imposed. The system integrated two computer vision algorithms, You Only Look Once (YOLOv4) and Simple Online and Real-time Tracking with Deep Association Metric (Deep SORT). YOLOv4 allows applications to detect objects from a given image by setting bounding boxes around the object and identifying the object class (Redmon et al., 2016). Deep SORT is a practical method for multiple object tracking that emphasizes simple algorithms (Wojke et al., 2017, p. 3645), and is utilized because it is able to track objects longer over an extended period of occlusion (Wojke et al., 2017, p. 3648).

2. METHODOLOGY

The designed system has three main modules, namely the data preparation module, the detection module, and the visualization module. By developing the system in modular outputs, it becomes more adaptive to possible changes in the future. Figure 1 illustrates the system architecture.

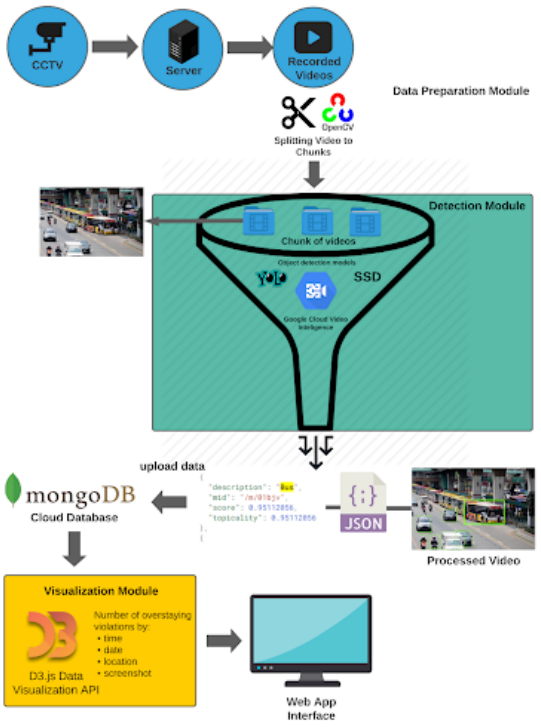

Fig. 1. Cipher: Design Architecture

The overall accuracy of the system will not be improved in the development of the project as it will be heavily dependent on the accuracy of the specific computer vision technology to be used which is the YOLOv4 algorithm and Deep SORT (Bochkovskiy et al., 2020; Wojke et al., 2017). The system will also analyze pre-recorded videos captured by the CCTVs, thus the quality of data analysis will also be based on the quality of the videos. The project is primarily focused on integrating existing computer vision technologies in guiding policy-making with regard to overstaying policies in the EDSA carousel. The following subsections describe each of the main modules further.

2.1 Data Preparation Module

The system includes the option to split input videos into five-minute chunks. The input is prepared using OpenCV, a computer vision library, and FFMPEG (FFmpeg, n.d.), a library that has the function to split large videos into smaller videos. The CCTV footage collected from the agency will be cut into 5-minute shorter videos in order to lessen the resources needed to complete the entire process and is then forwarded to the

detection module.

2.2 Detection Module

For detecting and tracking buses, several computer vision techniques were used. YOLOv4 was mainly used for identifying the bus vehicle objects in the frames. Each frame is analyzed by the YOLOv4 algorithm. Bus objects that are detected are being tracked by the system through Deep SORT, which takes into account both the bounding box parameters of the detected object and the information about the appearance of the tracked objects and associates the detections in a new frame with the previously tracked objects (Wojke et al., 2017). Buses that enter the loading bay are monitored for their staying times and these would be computed once they leave the frame; the staying time of a bus is computed as time duration from when it entered the loading bay until it leaves the frame. The staying time of each bus object is also displayed just above the object's bounding box, together with a status indicating whether the bus is overstaying or cleared.

The overall accuracy of the system mainly depends on the existing technologies used in the project and wasn't optimized to have higher accuracy as the main goal of the project is to integrate existing technologies. Detecting buses at night is feasible assuming that the area has enough lighting. Multiple occlusions might affect detecting multiple busses but as long as the main features of a bus are visible, Cipher can still detect it.

2.3 Visualization Module

The users can access the data gathered through a user interface. With the use of a visualization tool, Charts.js, the information in the cloud database will be converted into visual data that the users can easily comprehend (Kurkela, 2020). Reports, tables, and graphs will be provided in the web user interface in different formats. Reports can be seen as a summary report or by individual stations in EDSA Carousel. The reports will include time and date, location, and a total count of violations per data or location, depending on the preference of the user. Descriptive analytics will also be possibly implemented, depending on the number of CCTV footage the proponents will be able to gather and process. The user interface will also include screenshots of violators to serve as evidence if the agency decides to penalize them.

3. RESULTS AND DISCUSSION

The detection module was tested on 3 whole-day videos from the Main Avenue station of the EDSA Carousel. In order for the system to only count buses passing in the carousel, a region of interest was defined, in order to ignore other buses that could be seen from other areas of the video.

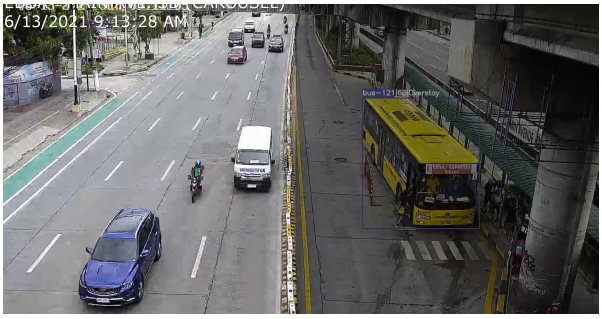

Fig. 2. Region of interest, shown in the darker shade at the right half of the frame, in the detection module

In Fig. 2, an example can be seen of the detection module detecting and tracking buses. Buses are detected once they have entered the region of interest within the carousel and will be tracked up until the time it leaves the frame. The figure above shows the shaded area representing the region. Data such as date, time entered, time left, staying time, and status of all detected buses are saved into the database. Buses that stay for more than a minute will have their status changed to overstaying.

While the system is able to detect and flag overstaying buses, there are some limitations that may come into play regarding the computation of staying time and accuracy. These limitations may be due to misdetections that can be caused by some challenges such as viewpoint variation, occlusion, illumination conditions, and cluttered or textured backgrounds (Pokhrel, 2020).

The data collected from the video processing will be displayed in a web dashboard composed of several data visualization charts, cards, and tables to represent the data collected and give visual insights for the MMDA. Figure 3, 4, & 5 shows the tabular presentation of data, visualization of data through charts, and summary cards.

Buses Detected

Minimum staying time: Maximum staying time:

Show 10 entries Copy CSV Print

id	Station	Date	Start Time	End Time	Staying Time
1	Main Ave.	06/11/2021	08:00:00 AM	08:00:52 AM	52
2	Main Ave.	06/11/2021	08:00:00 AM	08:00:20 AM	20
4	Main Ave.	06/11/2021	08:00:20 AM	08:00:29 AM	9
6	Main Ave.	06/11/2021	08:00:39 AM	08:01:28 AM	49
7	Main Ave.	06/11/2021	08:00:52 AM	08:01:21 AM	29
9	Main Ave.	06/11/2021	08:01:11 AM	08:02:07 AM	56
12	Main Ave.	06/11/2021	08:01:54 AM	08:01:58 AM	4
13	Main Ave.	06/11/2021	08:02:09 AM	08:02:19 AM	10
14	Main Ave.	06/11/2021	08:02:18 AM	08:03:02 AM	44
16	Main Ave.	06/11/2021	08:02:22 AM	08:02:52 AM	30

Showing 1 to 10 of 3,005 entries

Fig. 3. Tabular Presentation of Data

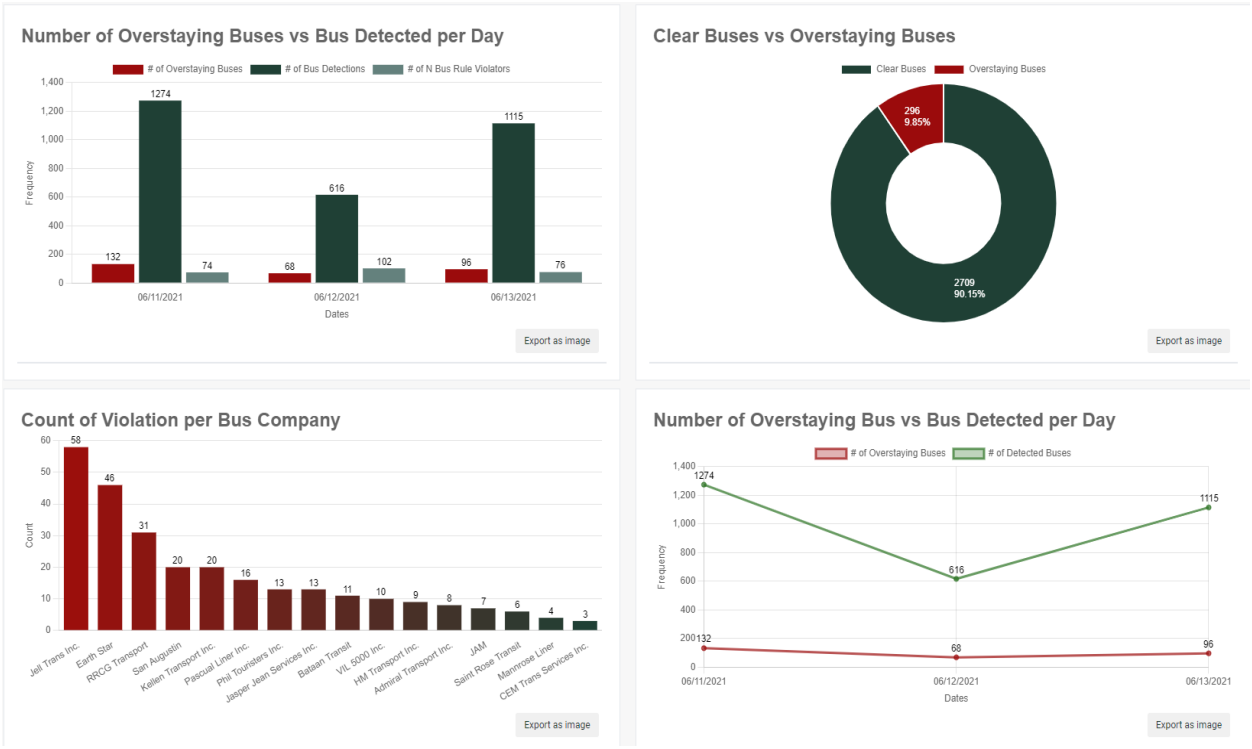

Fig. 4. Visualization of Data through charts

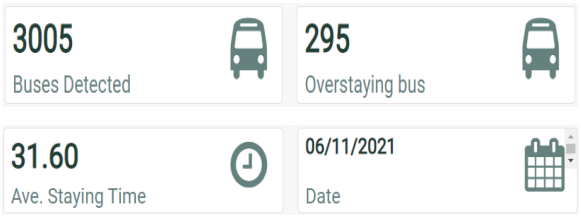

Fig. 5. Summary cards

In order to validate the results of the detection module, two histograms were prepared that show results from a manual counting of bus staying times and the automated counting from the bus detection module. For manual counting, the researchers manually tallied the number of buses that passed through the loading station as well as their staying time and recorded them in a spreadsheet. On the other hand, automated counting refers to the count of buses detected by the system in the bus detection module. The data generated was based on 3 whole-day videos from the EDSA Carousel.

In Fig. 6, the manual count resulted in around 2989 buses. In the manual counting, buses were only counted once they had been fully visible in the EDSA Carousel station. The histogram also shows that the detection module counted 3005 buses. Buses were counted once they had entered the region of interest specified by the system. In regards to the mismatching results of the manual and automated count, the cause for the imperfect result may be due to the misdetections that were processed by the system.

In Fig. 7, based on the threshold of 60 seconds for overstaying violations, the manual counting of violations totaled 310 overstaying violations while automated counting by the system resulted in 295 overstaying violations. This amounts to a 4.96% difference between the two methods.

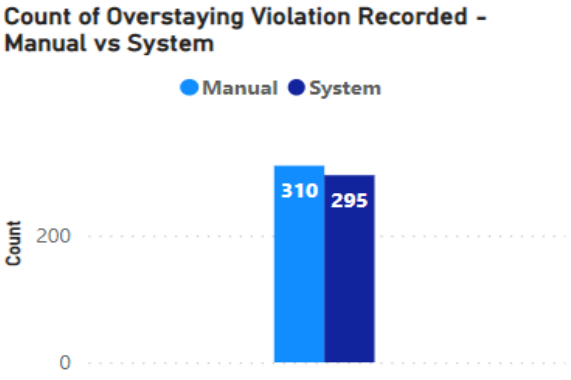

Fig. 7. Summary cards

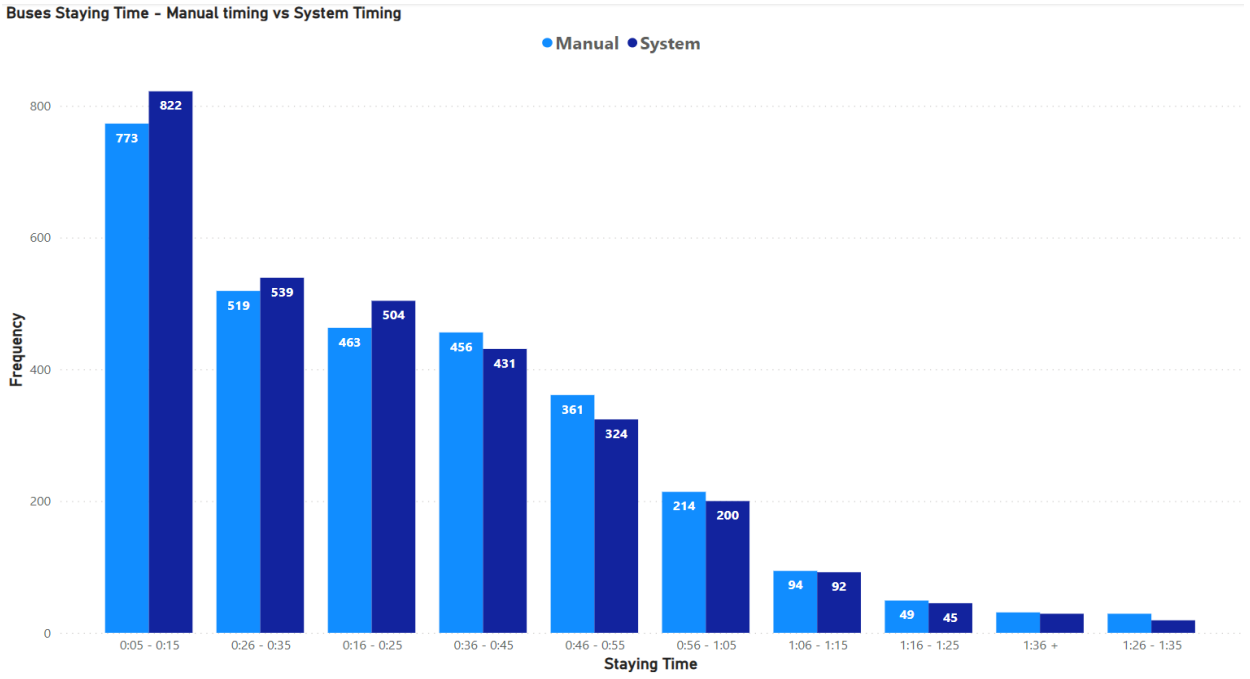

Fig. 6. Histogram of manual bus count vs. automated bus count

4. CONCLUSIONS

In this paper, the Cipher bus monitoring system is described as a tool that can automate and aid in the detection of bus staying violations through the lens of the staying time rule, through the utilization of object detection and tracking. This indicates that the system is capable of counting the time a particular bus is within the loading bay and flagging it as an overstaying violator when it reaches a certain threshold of time.

To test the accuracy, the data produced by the system was compared to the manually processed data. This yielded a 0.53% difference between the system-generated and manually generated data.

The data provided by the system can help the MMDA improve in its management of the EDSA Carousel. The system will be able to provide on-the-ground traffic information for crafting better traffic management policies. By utilizing the web dashboard, the MMDA will be able to gain valuable insights regarding the day-to-day operations of the carousel. In the future, the system may also be extended to cover other vehicles and terminals, for as long as surveillance cameras have already been installed.

5. ACKNOWLEDGMENTS

The authors greatly acknowledge the MMDA Metrobase personnel for providing recorded CCTV Videos that were used in the study and for being part of the interviews. Also, the authors would like to give thanks to the DOTr personnel and TITAN team for giving insights and suggestions to the project.

6. REFERENCES

An act creating the metropolitan manila development authority, defining its powers and functions providing funding therefor and for other purposes 1994 (RA) s. 2-5 (Phils). https://mmda.gov.ph/mmda-about-us.html

Bochkovskiy, A., Wang, C., & Mark, H. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv:2004.10934. https://arxiv.org/abs/2004.10934

Dela Cruz, R. (2020, August 20). 15 stations complete, 284 buses running at EDSA Busway: DOTr. Philippine News Agency. https://www.pna.gov.ph/articles/1112939

DOTr. (2020, October 17). EDSA transformation. Department of Transportation. https://dotr.gov.ph/55dotrnews/2525-edsa-transformation.html

FFmpeg. (n.d.). FFmpeg. https://www.ffmpeg.org

Kurkela, J. (2020, March 25). Charts.js. https://www.chartjs.org/

Metrobase Citizens charter 2020. (2020). Metropolitan Manila Development Authority. https://mmda.gov.ph/10-transparency/4310-citizens-charter-2020.html

Pokhrel, S. (2020, February 17). 6 Obstacles to Robust Object Detection - Towards Data Science. Medium. https://towardsdatascience.com/6-obstacles-to-robust-object-detection-6802140302ef

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-time Object Detection. Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788). https://arxiv.org/pdf/1506.02640v5.pdf

Wojke, N., Bewley, A., & Paulus, D. (2017). Simple online and real-time tracking with a deep association metric. 2017 IEEE International Conference on Image Processing (ICIP), 3645-3649. https://doi.org/10.1109/icip.2017.8296962