## Project Background

### General Background

For most people, losing their automobile is a very distressing situation. These situations are often exacerbated when the victims face the frustration of long investigation periods and, as in most cases, the law enforcement authorities being unable to retrieve the stolen vehicle.

Motor vehicle theft, also known as car theft and grand theft auto, refers to the criminal act of stealing or attempting to steal a motor vehicle [1]. Statistically speaking, there were 352 cases of “Missing Motor Vehicle” cases in Hong Kong during the first six months of 2022 [2]. This is a common phenomenon in urban spaces; in Europe, motor vehicle theft accounted for approximately 7.4 billion Dollars worth of damages in 2020 [1].

The prevalence of this issue can mostly be attributed to the fact that most countries still rely on old methods for identifying and tracking down stolen vehicles, for example in the United Kingdom, law enforcement relies on methods such as random street checks, license plate checking of stolen vehicles and tips from crowd-sourced websites [1].

### Research Gap

Despite recent advances in technology, the current method of identifying and spotting stolen vehicles relies heavily on manual labor. Law enforcement agencies (LEA) watch several hundreds of hours of surveillance videos to try and identify potential suspects of stolen vehicles, and they invest even more time and energy investigating each potential suspect.

The danger of this method is that, due to its time-consuming nature, by the time the stolen vehicle is identified it might be too late to efficiently intercept it. This issue is compounded when it is taken into account the fact that there are several cases of motor vehicle theft each year and many of these cases have to be resolved concurrently to one another.

This project tries to address this problem by introducing a new method of identifying and tracking stolen motor vehicles. The project aims to automate the process of watching surveillance footage by creating an Artificial intelligence (AI) model which can identify cars and classify their features such as color, type, and license plate number. By comparing passing vehicles against a list of stolen vehicles law enforcers can be alerted whenever a potential suspect is identified, thus saving both time and energy.

### Project Objective, Deliverable, and Scope

The goal of this project is to create a Machine Learning (ML) Model that can be integrated into Hong Kong’s existing network of surveillance cameras. The model aims at matching each vehicle it sees against the description of stolen vehicles, notifying the LEA in the event of a potential suspect.

The project can be defined by three distinct deliverables: an ML model that analyzes surveillance videos, a database server to store descriptions of stolen vehicles, and a mobile app interface for authorities to view the results of the model’s analysis.

This three-part system allows great efficiency in deploying the project. It also makes for easy maintenance and upgrades in case of developments in the future; only the part which has to be upgraded can be turned offline and replaced while the rest of the system remains functional.

It also provides LEA with real-time information and updates as the notification system removes the need to manually request analysis results.

The mobile app interface enables the LEA to view the results of the ML model and make the active decision of whether or not the potential suspect is correctly identified. This interface allows for efficient human intervention to prevent false positives from wasting precious time and resources.

### Justification of Work

This project would greatly help in reducing, and possibly resolving, the issues faced by LEA in cases related to motor vehicle thefts.

Firstly, since the ML model is used to handle the analysis of surveillance videos, it becomes a simple task to run multiple instances of the ML model. In this manner, multiple video feeds are analyzed simultaneously allowing LEA to save on labor. costs. It also helps save time as the model can be left running continuously without any breaks.

Secondly, the model can be optimized to help LEA reduce the number of false positives; this helps prevent innocent civilians from being mistakenly labeled as potential suspects. This is highly beneficial as the reduced number of potential suspects translates to lesser amounts of time and labor required to resolve a case.

Finally, the model outputs can be analyzed to trace the path of stolen vehicles. Since all the surveillance camera’s video feed is being watched simultaneously, the moments a potential suspect passes through a camera are time-stamped. Multiple time stamps could then be plotted on a map to trace the path taken by potential suspects.

### Outline

## Project Methodology

### Overview

### Design Principles

### Hardware Architecture

### Software Architecture

### Mobile App Interface

### Database Server

## Project Status

### Evaluation of Progress

### Expected Outcomes

### Project Schedule

### Remaining Work

### Difficulties Encountered

## Conclusion

## References

## Appendix

List of Figures

Fig 1.

List of Tables

Tab. 1 5

## 3. Project Status

With the above setup of hardware and software architecture, a system with live cameras programmed with a Convolution Neural Network model can be delivered to detect the stolen vehicles and alert the authorities. Although the AI model is expected to have a low recall and precision on recognizing the license plate number, the error can be addressed with the help of combining all features of a vehicle on determining the matching rate.

3.1 Evaluation of Progress

The project has completed its data collection and labeling stage. In the data collection stage, The training dataset of vehicles was collected with smartphone fixing on a point to do recording for a few hours. The smartphone was placed on different spots like parking lot, open streets, and traffic lights to increase the variety of vehicles. For the data labeling stage, the collected videos were framed and processed with augmentation first to increase the size of the dataset to 5000 photos. Then the photos were annotated with classes of vehicle and position of license plate. Finally, the processed dataset was split into 80% for training and 20% for validation.

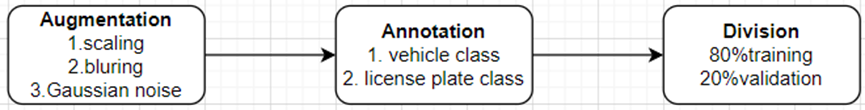


Fig.2 steps of data processing

3.2 Outcomes

Runtime, recall and precision are three main factors for evaluating the algorithm. Runtime is defined by the time needed to process one frame of the video by the algorithm; Recall is defined by dividing the true positive number by the sum of all instances; Precision is defined by dividing the true positive number by the sum of correctly and incorrectly predicted instances.

Precision = TP/(TP + FP) (5)

Recall = TP/(TP + FN) (6)

True positive (TP)—total number of correctly identified stolen vehicles.

False positive (FP)—total number of incorrectly identified stolen vehicles.

False negative (FP)—total number of stolen vehicles not being identified.

AI model can recognize most of the vehicles with the average accuracy of 65.4%. It has a precision and recall of 0.612 and 0.622 respectively. However, the precision and recall on recognizing the license plate number are 0.521 and 0.517, which are comparatively low . It is expected the Al model based on CNN algorithm will yield a high precision and recall on recognizing the vehicle object if the size of input dataset is large enough. However, the precision and recall will be undesirably low on recognizing the license plate number due to the small size of the license plate when compared to the vehicle. The distance between the vehicle and the camera will also affect the performance of the model. The AI model will not be able to identify the vehicles until they reach a relatively close distance.

3.3 Projects Schedule

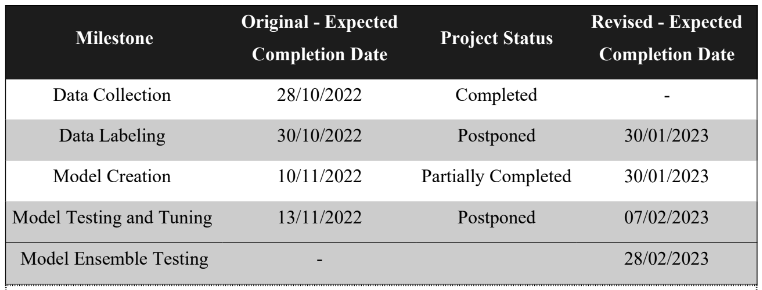


Table : Milestones for the first phase of model development and their expected completion dates.

Due to the increased workload in the project teams’ academic courses and scheduling conflicts, the project had to be postponed. In reflection of this, the deadlines had to be extended to accommodate the academic requirements of the team.

3.4 Remaining Work

The project has completed data collection, data labeling and preliminary model creation stages. Although the model can recognize the presence of a vehicle and license plate, extra effort and time needed to be put to tuning the model to recognize the numbers of the license plate in order to distinguish the stolen vehicle. The next step of the project would be the final

testing and fine-tuning of the model. At this stage, a video feed collected from the street camera will be fed to the model to examine its performance in a simulated environment.

Any incorrect parameters will be fine-tuned before finally testing the model against a live feed of a street camera.

3.5 Difficulties Encountered

The project has faced three main obstacles up til the writing of this paper.

Firstly, the project was suffering from label imbalance in the data collected. Label imbalance refers to the problem where a certain label occurs much more frequently than other labels. In this particular context, since the data collection relies on using image frames from traffic videos (see section 2.4.a for more details), there were more images of empty streets than there were of motor vehicles. This difference in label frequencies can harm the accuracy of the model.

To address this issue the project used data augmentation to synthesize images of motor vehicles to reduce the difference in label frequencies. Data augmentation refers to techniques where simple changes can be applied to images to create two different images. These changes may include flipping the image horizontally or skewing the image at an angle to create different images of the same view. While this measure is usually simple to implement, it has to be carefully calibrated to prevent the model from over-fitting. Over-fitting refers to a problem in ML where a model becomes too accustomed to a set of data during training that it performs poorly on new data provided during testing.

The second problem the project faced was data mismatch for the naive model implementation. Since the project relies on the naive model to automate the data labeling process a data mismatch problem would impair the accuracy of the naive model and require more human intervention. Data mismatch is caused when a model is trained by data taken using different hardware configurations.

Since the open-sourced data used to create the naive model was taken using a high-resolution

camera, the naive model suffered when processing data taken using the project's low-resolution cameras. This problem was solved by editing the open-sourced training data and blurring the images, thus reducing their resolution.

A third problem faced by the project was the occurrence of polluted data in the collected datasets. Since the camera is left to record videos outdoors, it is not uncommon for birds to perch on the camera stand causing the camera to tilt and change the orientation of the recorded video. A change in video orientation can greatly affect the accuracy of the model and might influence the projects viability after deployment. This problem was resolved by having a team member monitor and protect the camera while it was recording.

4. Conclusion