OENG118X/COSC250X Engineering & IT Capstone Project

Capstone Project Final Report

Develop AI Camera system for traffic control.

Student team: Ice Tea

Student name & ID: Ngô Quang Khải - s3836387

Trần Nguyễn Anh Khoa - s3863956

Nguyễn Hữu Đức - s3669698

Trần Vĩnh Khang – s3855823

Company: Gtel

Academic supervisor: Trần Nhật Quang

Nguyễn Ngọc Thành

Industry supervisor: Nguyễn Xuân Thọ

Võ Minh Tiến

Table of Content

Executive Summary

Vietnam's traffic issues are complex and challenging, requiring new AI technology to help identify any irregularities in street scenes. One of the main problems is detecting vehicle-related anomalies, such as sudden stops, lane changes, or collisions, that can disrupt the traffic flow and cause accidents. In this paper, we aim to develop a solution to this problem by incorporating object detection techniques with anomaly detection methods to enhance the traffic monitoring system. Using machine learning with camera technology, we can analyze the trajectories of vehicles and detect any abnormal behaviors in real time.

Detecting unusual traffic patterns in Vietnam's crowded and unpredictable roads is difficult. One of the difficulties is the need for more sufficient data for training and testing purposes. Creating dangerous situations on the road to generate more data is not a safe option, and the scarcity of abnormal events captured on camera makes it hard to gather enough data for practical training. Another area for improvement is the reliability of the existing methods for detecting unusual trajectories. Many current methods only focus on individual points that fail to consider their context in the scene, leading to false alarms or missed detections. Therefore, a more reliable approach is needed to consider the trajectories' spatial and temporal information. Furthermore, it can capture the subtle differences between normal and abnormal behaviors.

Our proposed solution to this problem is a computer vision-based approach that uses trajectory images generated by our multiple-object tracker. We used YOLOv8 to detect and track the objects in the scene and extract features such as the centroid coordinates of each object's bounding box. Unlike other methods that only consider individual points, our approach examines the relationship between points to detect unusual events. We measure angles between the points to determine whether they move in opposite or perpendicular directions. Based on this, we can effectively identify unusual trajectory coordinates by converting anomaly detection from points-based to visual-based detection by drawing the trajectory as an image. For instance, a normal trajectory image would show smooth and continuous curves, while an abnormal image would show sharp turns or breaks.

Additionally, this method allows for the quick generation of anomalous data by drawing a distinct pattern, such as a zigzag, a spiral, or a dot. Our testing has shown that this approach effectively identifies anomalous traffic trajectories. Our approach's advantages are its high accuracy, low computational cost, and easy scalability. However, some of the limitations are its sensitivity to occlusion and its dependence on the quality of the object detection.

Project background & problem statement

Figure 1 – Vietnam Traffic Scene (source: [26])

This project aims to address Vietnam's complex traffic challenges on its congested and unpredictable roadways. Detecting irregularities in street scenes, such as sudden stops, lane changes, and collisions, is a central problem that can disrupt traffic flow or lead to fatal accidents. Vietnam's major cities, including Hanoi and Ho Chi Minh City, are experiencing rapid population growth and urbanization driven by economic growth. This expansion has placed immense pressure on the existing road infrastructure, resulting in traffic congestion that costs the economy a staggering $6 billion annually in Ho Chi Minh City alone, according to local news source reported that in 2022[1]. This project's significance lies in its potential to mitigate traffic congestion's detrimental effects and improve the overall transportation system in Vietnam. By collecting valuable data and developing a more intelligent transportation system, we aim to significantly reduce economic losses, promote productivity, and attract further investment.

Despite an enormous advantage, this subject has yet to be tackled by many in the industry in Vietnam due to the uniqueness of Vietnamese traffic. The main objective of our project is an anomaly detection system to work in real-time (or near real-time), using a combination of traffic cameras and machine learning algorithms to detect and classify anomalous trajectories on the road. To achieve the objective, we proposed to develop a system for detecting anomalous trajectories on roads in Vietnam inspired by [25] to further enhance the data we can extract from the multi objects tracking model.

The motivation and significance of our project are as follows:

1. **Enhancing Traffic Safety:** The project seeks to develop an AI-driven solution to identify abnormal traffic patterns and behaviors in real time. This technology can significantly enhance road safety by quickly detecting potential hazards and alerting authorities, reducing the risk of accidents.
2. **Efficient Traffic Flow:** By accurately detecting anomalies such as sudden stops and illegal lane changes, the project aims to contribute to smoother traffic flow. This solution has the potential to help reduce congestion, minimize travel times, and improve overall transportation efficiency.

3. **Data-Driven Insights**: The project acknowledges the challenges of gathering sufficient data for training and testing AI models in the context of traffic monitoring. It proposes a solution that does not rely on creating dangerous situations on the real road but instead leverages existing camera technology and machine learning. This approach can provide valuable insights into traffic behavior, which can be applied in Vietnam and similar urban environments globally.
4. **Context-Aware Solutions**: The project recognizes the limitations of current anomaly detection methods that focus on isolated data points without considering the broader scene context. It aims to develop a more reliable and context-aware approach by incorporating spatial and temporal information from vehicle trajectories. This innovation has the potential to revolutionize traffic monitoring systems in Vietnam and other regions grappling with similar challenges.
5. **Generalized Application**: While the immediate focus is on Vietnam's traffic issues, the project's outcomes can have broader traffic management and surveillance applications. The solution architect developed here can serve as a blueprint for addressing traffic challenges in various urban settings.

We introduce a novel approach to trajectory anomaly detection, forming the cornerstone of our project's traffic monitoring and safety innovation. Our framework is grounded in the realm of computer vision, further harnessing cutting-edge technology of image classification to enhance anomaly detection capabilities.

In our approach, we pivot from traditional point-based methods and instead leverage the full potential of trajectory data. We achieve this by constructing trajectory images from the coordinates generated by our advanced multi-object tracker. Unlike previous methods that primarily rely on isolated data points or segments, our approach capitalizes on the rich context provided by the trajectory relationships between these points for event detection.

One of the key distinguishing features of our method is its transformation of anomaly detection from a numerical paradigm to a visual one. By rendering trajectories and saving them as images, we effectively convert anomaly detection from a numerical analysis to a visually driven process. This visual anomaly detection approach has been extensively explored and validated in various domains, underscoring its effectiveness in identifying deviations from established patterns.

Moreover, the adaptability of our method sets it apart. Anomalies can be effortlessly generated by drawing a different pattern within the trajectory data, making it a versatile tool for simulating abnormal traffic behaviors. This adaptability and ease of generating anomalous data contribute to the method's practicality and robustness.

Following rigorous testing and validation, our approach has demonstrated its effectiveness in identifying anomalous trajectories within complex traffic scenarios. It promises to be a valuable asset in our mission to enhance road safety, streamline traffic flow, and provide actionable insights for traffic management in Vietnam and beyond. This innovative trajectory anomaly detection method represents a significant stride toward our overarching goal of revolutionizing traffic monitoring systems and making urban environments safer and more efficient.

Literature review & market research

The anomaly detection market in the IT industry is experiencing significant growth [10] due to the increasing complexity of data and rising cybersecurity threats. Additionally, global anomaly detection market to reach $8.6 billion USD by the year 2026 [11]. Anomaly detection plays a crucial role in network security, system health monitoring, fraud detection, cloud security, IT operations management, and incident response. Key trends include a focus on network security, proactive system monitoring, fraud prevention, cloud security, IT process optimization, and incident response. Major players in the market include Cisco Systems Inc., WSO2 Inc., Microsoft Corporation, Verint Systems Inc., and Broadcom Inc. (Symantec Corporation). The market is moderate to highly concentrated, with notable developments like the launch of automated anomaly detection applications and advanced risk analytics. The outlook is positive, with continued growth expected as digitization and advancements in AI and cloud computing drive the demand for anomaly detection solutions. Anomaly detection remains a critical tool for securing digital infrastructure and addressing emerging challenges in the IT industry.

STUD [12] is a new unknown-aware object detection framework that distils unknown objects from videos in the wild and meaningfully regularizes the model’s decision boundary. This framework attempts to capitalize on the unknown ones for model regularization rather than known objects. Because of this framework, there was not any source code that we could try. However, it provides a framework on how we can further improve our object detector by using their framework. It also helps us to gain some more understanding on how people usually approach anomaly detection on road.

UNKAD [13] is a novel training strategy for object detection that can predict unknown objects without requiring any annotation of them, exploiting non-annotated objects that are already present in the background of training. This has the potential to help improve our model and provide another different direction in which we can train our model.

In the realm of trajectory analysis, this innovative approach has reshaped how we go about detecting anomalies. It adopts a unique two-stage method that leverages deep neural networks, self-supervised learning, and one-class classification. In the first phase, the approach trains a deep neural network using self-supervised learning, allowing it to discern key features within the input data without relying on explicit labels. These learned features become valuable assets in the second stage, where one-class classification techniques like OC-SVM and kernel density estimators come into play. These methods are adept at identifying data points that stand out significantly from the typical patterns, making them ideal for spotting anomalies. The fusion of self-supervised learning and one-class classification has propelled this approach to deliver exceptional performance across a range of benchmark datasets. However, it's important to be aware of its limitations, such as the need for a substantial amount of unlabeled data and its sensitivity to the choice of one-class classification techniques and hyperparameters.[27]

The Trajectory Outliner Detection approach represents a groundbreaking stride in the domain of trajectory analysis. Unlike traditional methods that broadly label entire trajectories as anomalies, this approach takes a more nuanced approach. It breaks down intricate trajectories into smaller, manageable sub-trajectories, allowing it to scrutinize data with a fine-tooth comb. What truly sets it apart is its knack for identifying not just complete outlier trajectories but the specific sub-trajectories within them that deviate significantly from the rest. Picture it as a detective carefully examining the pieces of a puzzle rather than just the picture on the box. This precision in pinpointing anomalies within trajectories is immensely valuable, particularly when the need arises to precisely locate anomalies amidst the complexity of real-world data. Nevertheless, it's important to acknowledge that, like any tool, this approach has its own set of challenges, such as potential limitations in handling large or noisy datasets and the need for parameter tuning in certain situations. Nonetheless, its unique strengths in anomaly localization make it a promising tool in trajectory analysis and anomaly detection.[28]

A Vision-based System for Traffic Anomaly Detection using Deep Learning and Decision Trees [14]. This paper proposes a vision-based system for traffic anomaly detection using deep learning. The system consists of three main components: a background estimator, a road mask extractor, and a decision tree for anomaly detection. The background estimator estimates the static background scene. The road mask extractor extracts the road region from the input video. The decision tree uses a combination of motion and appearance features to detect anomalies such as stopped vehicles, wrong-way driving, and pedestrians on the road. The proposed system achieved high performance on several benchmark datasets for traffic anomaly detection.

The TraClet method, introduced in Harnessing the power of computer vision for trajectory classification study [24], showcases an innovative approach that utilizes image representations of trajectories for intuitive human-like classification using computer vision techniques. This technique involves mapping low-dimensional features into a distinct space using a partition-wise activation function applied before linear transformation, with varying parameters for different partitions. This allows for the joint modeling of trajectories in a nonlinear hierarchical deep structure.

An additional information we gather from the previous paper is that we have found out this model was entered in a competition hosted by Nvidia called "Nvidia AI City Challenge" [15]. Digging deep into that competition, we have found several models and articles related to our project.

We have found that there were some teams from Vietnam that have develop something like our project in 2020 [4],[5]. The team did not focus exclusively on anomaly detection but used it with another model too. Their method was using a "GAN-based Forward prediction and Backward-

Tracking Refinement". Using a combination of backward vehicle tracing technique (taking average on position) with the GAN-based future frame prediction they were able to achieve S4 score of 0.9059 and F1 score of 0.9412 (5^{th} place).

More from the Nvidia Ai City Challenge 2020, we found the 2nd place model [17]. That model also used a simple architect object detector with an unsupervised anomaly detection model. This model's approach was very closely matched to our need because we do not have much and cannot find enough videos that contain anomalous objects. Moreover, since the system uses an unsupervised approach, they only need a little data to train it to get some good results. A drawback of this architecture is that we have to run two models simultaneously. The final S4 score computed using the model was 0.5763, with an F1 score of 0.5926. According to the author, external data is needed to improve the model further, indicating that the more data the model can train on, the better it is due to the unsupervised component. This project proposed a very interesting way to tackle anomaly object detection problem and is worth noting.

In other findings related to anomaly detection on the street of Vietnam, we have found two projects which aim to produce anomalous data from the street of Vietnam for other models to use. UIT-Anomaly: A Modern Vietnamese Video Dataset for Anomaly Detection [7] and [8] the "VNAnomaly" looks like it an exclusive dataset made for the student at that school and does not allow for commercial use. However, it is good to know that there are still projects related to ours. It is also confirmed that not many anomalous objects are detected on the street, primarily anomalous events caused by humans.

The paper is a foundation for our project, which presents a traffic anomaly detection system that employs a three-module pipeline to effectively detect anomalies on the road. The system [6] uses deep learning techniques to analyze the traffic flow and identify anomalies. It begins with the background modeling module, which analyzes the traffic flow to obtain accurate road segmentation results. Then, the perspective detection module comes into play, detecting vehicles in each frame of the video feed. The spatial-temporal matrix discriminating module plays a vital role by constructing a spatial-temporal information matrix based on the detection results. This matrix is used to determine whether an anomaly has occurred. To obtain an accurate timeline, the system backtracks and merges any detected anomalies. Deep learning techniques are used for vehicle detection and anomaly analysis in the system. By weighting and overlaying frame results, a frequency map of vehicle distribution is generated. The image is normalized, followed by a binarization process to obtain a segmentation map of the traffic flow.

Scope & Solution design

Scope of work & performance evaluation

Scope and milestones

Below is the scope for our completion plan in Capstone B:

Milestones:
Apart from that we also achieve these milestones quickly in the development phase:

1. Successfully recreate and customize the model we based on [6]: We want to fully understand the logic behind the model that we are building upon. So, we would know what and how to improve it in later stages.
   - Because of this achievement, we were able to come up with our own idea by building upon the code.
2. Finding the right dataset to train on: It is very important to find the right dataset to train our model to suit our needs. The videos/ data should be created after the year 2015. We hope to find at least 50 videos, each around 15 minutes. The video resolution must be in an acceptable range (800x 410) with a frame rate of 30fps. This instance, we would need to ask GTEL for assistance in finding these videos.
   - We have also successfully asked for NVIDIA dataset which the team use. GTEL have also provided us with the video from their camera. Luckily, in one of the videos, there was an anomaly event that gave rise to this architecture. That anomaly was a truck making an illegal U-turn.

Original completion plan scope:

1. Synthesize (mimic) anomalous data: To improve the model and validating purposes, we will need to create some video with anomalous object in it. It is even better if we could find some real footage from the video provided.
2. Successfully recreate and customize the model we based on [6]: We want to fully understand the logic behind the model that we are building upon. So, we would know what and how to improve it in later stages.
3. Achieve high accuracy object detection model: The model must be able to classify the normal trajectory of that road and the anomalous trajectory of that road. With F1 score > 80%.
4. Implement object detector with anomaly detection techniques: The project must be able to classify anomalous objects on the road. The system must be able to flag the video with anomalous object.

5. The system must have a low latency: We are aiming to achieve close to near real-time with our model for it to work with traffic cameras. We want to achieve this without sacrificing too much of the model's accuracy.
6. Produce a report/ document for the project: After developing the system, we want to produce documentation on the design of our project. Ex: how it works, what technology is being used.

The scopes highlighted in red have been changed from our original complete plan since our client's requirements have broadened a bit. Firstly, we have decided to change "Synthesize (mimic) anomalous data" to finding a way to create anomalous data easily —moreover, a way to consume those synthesized data in the solution.

In our original plan, we have moved the system to have low latency to be out of scope. However, after discussing this with our industry supervisors, this is one of the mandatory scopes of our project, with slight adjustments. Our industry supervisor added that we do not need to handle the inferred time of our model due to hardware constraints. Instead, we need to make it so that the solution's architecture can run in real-time with minimal latency, excluding the infer steps.

With that, we have also updated our performance metrics and evaluation method to fit the new solution. The evaluation method is now more straightforward than the original NVIDIA AI City Challenge proposed, raising a warning as soon as an anomaly trajectory is detected. However, we must still test our trained One-class SVM (OC-SVM) model to choose the correct combination. Because of this, we generated a mixed dataset of normal and anomaly trajectory images of that road to evaluate the OC-SVM. We decided to focus on the F1 score and Recalls of class anomaly since we want to maximize it.

Tech stack

Software:

- Python 3.10
- Matplotlib, numpy, tensorflow.
- OpenCV
- Ultralytics YOLOv8 (object detector) with BoT-SORT as tracker
- Pre-trained ResNet50 as feature extractor
- One-class SVM
- Pytorch/TensorFlow library
- Google Collab or Azure Databrick (if local hardware is not sufficient)

Evaluation method

In this document, we define a set of success metrics that are directly aligned with Capstone B's specific functions and features. These success metrics are justified as being quantifiable and observable, and a clear plan is in place for evaluating the project's success based on these metrics.

F1-score: F1-score measures the balance between precision and recall, incorporating the number of true positives (TP) and the avoidance of false positives (FP) and false negatives (FN). A higher F1 score signifies better anomaly detection performance. A true-positive (TP) detection will be considered the predicted anomaly within 10 seconds of the actual anomaly with the highest confidence score. Each anomaly will be counted as one TP. A false-positive (FP) is a predicted anomaly that is not a TP. Finally, a false-negative (FN) is a true anomaly that was not predicted.

The accuracy of the model for detect anomaly is based on three score:

- **Precision Score:** Precision gauges how accurate our model is when it predicts positive instances. It answers the question: "Of all the instances our model said were positive, how many were actually correct?" A high precision score means our model's positive predictions are reliable and have fewer false positives.
- **Recall:** Recall measures how well our model identifies all the actual positive instances. It answers the question: "Of all the actual positive instances, how many did our model manage to capture?" A high recall score indicates that our model is effective at minimizing false negatives by correctly capturing most of the relevant positives.
- **F1-Score:** The F1-score is a balanced measure that combines precision and recall into a single value. It helps us strike a balance between precision (accuracy of positive predictions) and recall (ability to capture all positive instances). It's particularly valuable when we need to consider both false positives and false negatives in our evaluation.

Alignment with Capstone B functions/features: The success metrics are aligned with the essential functions/features of the planned Capstone B project, which primarily focuses on anomaly detection. By evaluating the model's F1-score and detection time error, we can evaluate its effectiveness in accurately identifying anomalies.

Measurable and observable: The selected success metrics, F1-score, and Recall. The F1-score is calculated based on the number of TP, FP, and FN detections, while the recall score correctly predicted positive observations to the total actual positive observations. These metrics provide a quantifiable assessment of the model's performance.

Evaluation plan: Evaluate the project's success using the defined metrics; the following steps will be taken:

- Object Tracking with YOLOv5: We employed YOLOv5 for object tracking and to extract the coordinates of the tracked objects in the video.
- 2D Path Visualization: We visualized the paths of tracked objects on a flat 2D surface, creating a clear representation of their movement patterns.
- Dataset Splitting: The dataset was split into two categories: one for collecting normal paths, meticulously ensuring that it contains nearly 100% genuinely normal paths, and another for collecting both anomaly and normal paths.

- **Training with OC-SVM:** We trained our model using One-Class Support Vector Machine (OC-SVM) on the dataset containing only normal paths. This model aims to identify deviations from normal patterns.
- **Fine-Tuning for Optimal Benchmark Scores:** Through fine-tuning, we optimized the model to achieve high scores during benchmarking. Notably, the hyperparameter "nu" was set equal to the threshold used during testing, ensuring consistency.
- **Benchmarking with Precision, Recall, and F1-Score:** The benchmarking process provided accuracy metrics, including Precision, Recall, and F1-Score, which range from 0 to 1. These metrics help us evaluate the model's performance in detecting anomalies effectively.
- **Testing on Real Video Data:** Finally, we put the model to the test on real video data to assess its ability to capture anomalies. This practical evaluation will determine the model's real-world effectiveness in identifying deviations from normal object paths.

Solution Design

I. Overview

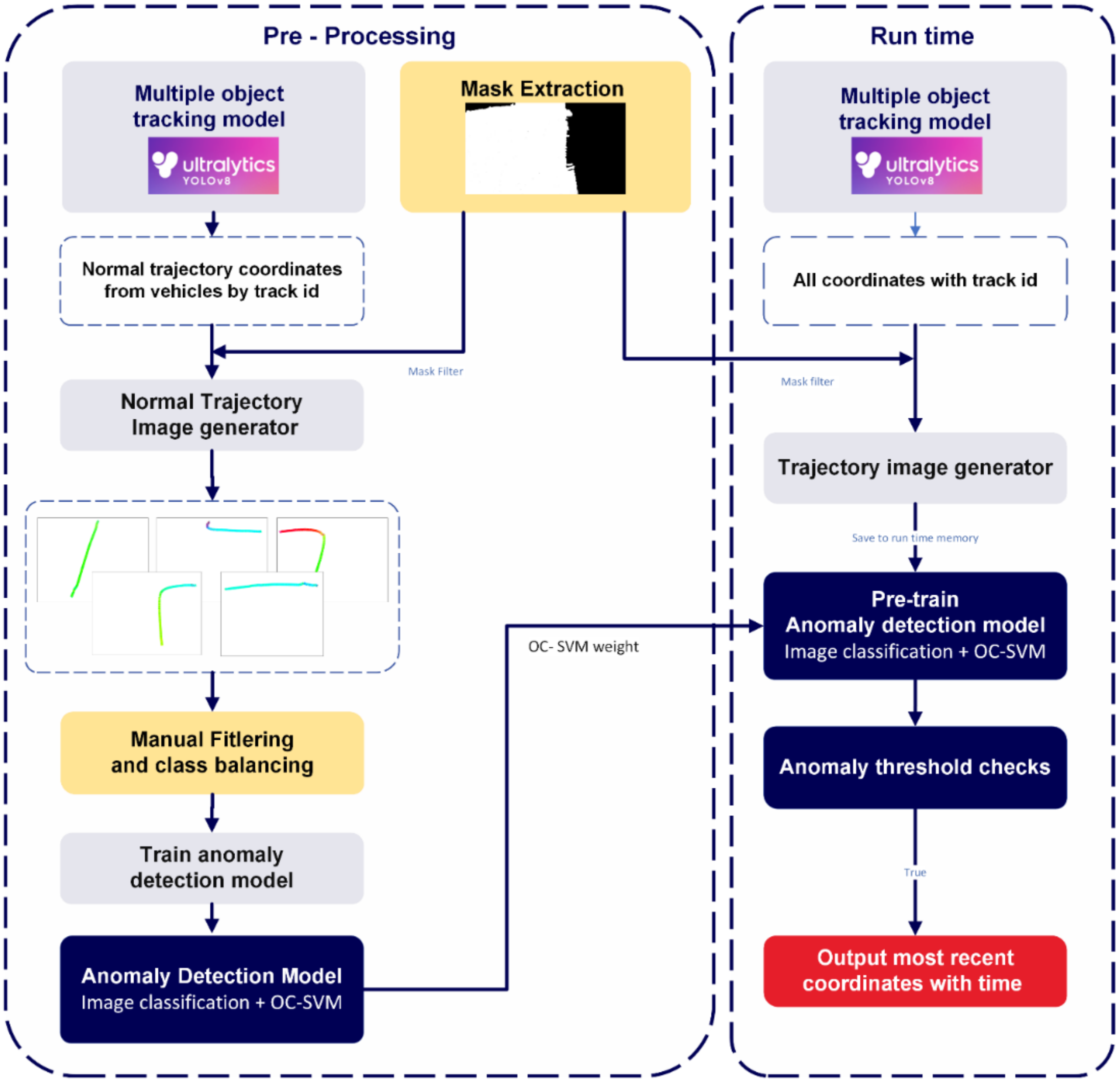


Figure 2 – Detail view of Box Level Tracking Branch (source: [6])

Our solution is divided into two main stages: pre-processing stage and run-time stage. The pre-processing stage is for generating artifacts that will assist in our runtime like the video mask and anomaly detection (OC-SVM) model weight as pickle file. The yellow boxes are external processes that run separately from the main flow in that stage. After running the pre-processing stage, the run-time stage will load all the artifacts to use in the detection pipeline. In the pipeline, if an image of the trajectory is classified as an anomaly over a set threshold, we will conclude that it is a true anomalous trajectory.

II. Pre-Processing

1. Extraction of Region of Interest Mask

Inspired by the method proposed in this paper [multi granularity], we use motion-based and trajectory-based methods to help us extract our region of interest more accurately. The reason is to prevent the solution from classifying vehicles stopping for a long time in parking lots or on side roads as an anomaly; we need to separate the possible abnormal areas automatically and extract our region of interest. This method is challenging with a segmentation model because the road scene is complex, and the anomalies sometimes go off to the side road. The two methods are as follows:

- **Motion-based mask**: We compare two consecutive frames and keep the areas with more changes than the difference threshold. These areas have moving objects. To combat shaky frames, we ignore the changes if the abnormal mask of a frame is more significant than another threshold M because a camera shake or scene change could cause it. We add all the change areas of a video to get a motion mask.
- **Trajectory-based mask**: We use BOT-SORT to track the vehicle movement. For each track, we ignore it if it is too short (less than n) or too small (less than d). These tracks are either false positives or on the side roads. We adjust the detection boxes for the remaining tracks according to the vehicle size, especially for large vehicles. We also remove small, connected regions that are noise-like side roads. We sum up the detection results of each track to get a trajectory-based mask.

We combine the two masks by taking their intersection. This way, we can get a better abnormal mask that covers motion and trajectory information.

2. Trajectory extraction and augmentation

During both stages, we utilized YOLOv8 with the BOT-SORT tracker as the initial feature extractor. This model helped us extract positional features of each vehicle based on their ID. We then added the x-centroid and y-centroid of each vehicle's bounding box to their corresponding trajectory. Next, we divided the trajectory into small, and connected segments using a simple algorithm. To better describe the motion of our trajectory, we transformed these segments into vectors with direction. We employed a formula, like the one utilized in [25], to encode the vector direction as a color in an image. We established a rule where each angle corresponds to a different color on the HSV color chart. To calculate the direction of each vector relative to the camera, we utilized the following formula:

$$V = arctan2(dy, dx) \tag{1}$$

Where dx and dy are the difference between the end and start point of each segment and V is the vector direction angle. After we calculate all the V using formula (1) for all small segments, we will use each segment points to project it on a two-dimensional space (x, y) along with the color to generate images that will be used for anomaly detection model training.

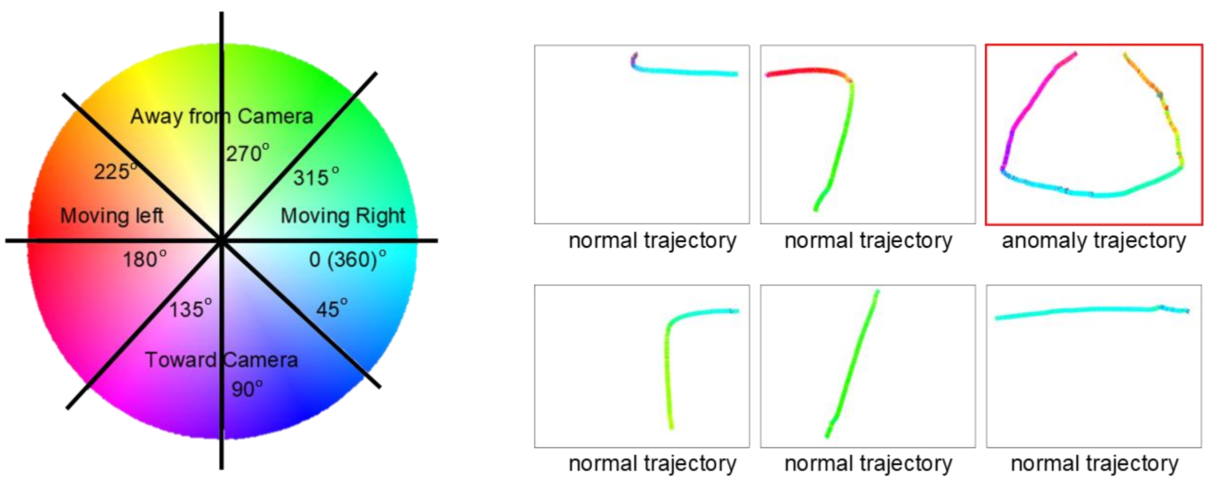


Figure 3 – Color meaning and sample trajectory images generated by this step.

After we have successfully enriched our trajectory, we can clearly see which one an anomalous trajectory easily. We then store all the trajectories images into a folder for further processing.

3. Anomaly classification model

Before we can train our model, we must manually filter out all the anomalous trajectory images or images that look anomalous. We want to ensure that our OC-SVM only learns to classify normal trajectories. This paper [29] has introduced this architecture to us, where a pre-trained Resnet50 will act as a deep feature extractor for OC-SVM to classify. Because we never train our OC-SVM model on any anomalous trajectory image, it will classify images with features significantly different from the majority class, which in this case is normal trajectory. We want to note that this architecture is not limited to Resnet50 with OC-SVM; we have tested some combination of image classification models with OC-SVM, which we will discuss further in the result section. We then test the model with mixed dataset and choose the best parameters for OC-SVM by focusing on F-1 score and Recall score.

Another point we want to make is that one trajectory image might be normal for that road section but anomalous for another. We would need to re-train our OC-SVM weight to fit that road better. However, this training process is not computationally expensive.

III. Run-time process

1. Anomaly trajectory detection pipeline

During this step, we can either run the model in real time or process the images produced by the image generator. The multiple object tracking model tracks IDs to generate coordinates in a real-time environment. These coordinates help us construct an image of the trajectory, which we feed into our pre-trained detection model from the previous step. For real-time processing, we avoid classifying premature trajectory images as anomalies by setting a minimum threshold of points that a trajectory must have before we generate that trajectory image to classify. Additionally, to prevent memory overflow, we set up a trajectory decay method. If a trajectory exceeds the maximum threshold for the number of points, the solution will delete the oldest point in that trajectory, shrinking it if it keeps appearing. To avoid classifying a car stopping at a traffic light as an anomaly, we set another anomaly threshold on how often that trajectory id has been labeled anomalous before concluding it as a true anomaly. If the solution determines that it is an actual anomaly trajectory, it will raise a warning and output all the trajectory information.

Result analyses & discussions

The objective of this project revolves around analyzing a video capturing the dynamic crossroads of Vietnam during both day and night. The video showcases many vehicles, including cars, trucks, and motorcycles. We undertook a rigorous data cleaning process to prepare the dataset for training and manually labeled trajectories as either normal or anomalous. It is worth noting that if there was any uncertainty during the labeling process, the trajectory was categorized as anomalous to ensure the utmost cleanliness of the dataset. This meticulous preparation was essential to achieve meaningful results in our subsequent model evaluation.

Our primary focus was to train a robust anomaly detection model to distinguish between normal and anomalous patterns within this video effectively. To achieve this, we implemented several architectural variations and fine-tuning approaches. One of the critical models we employed was the one-class Support Vector Machine (SVM).

In the case of the one-class SVM, we fine-tuned it with a hyperparameter, "nu," set to 0.00001. Nu is a crucial parameter in training one-class SVM models, and it plays a pivotal role in balancing the model's ability to classify data points as either outliers (anomalies) or inliers (normal data). Smaller nu values, closer to 0, make the one-class SVM more tolerant of outliers in the training data. This flexibility allows the model to consider more normal training points, which aligns with our objective of training on a normal pattern.

The same principle was applied during the testing phase, where we used a threshold of -0.00001 for the decision score to classify anomalies. Negative scores were associated with anomalies, while positive scores were linked to normal patterns. This approach was carefully chosen to ensure anomalies were effectively identified in the video.

Combination	F-1	Recall

ResNet50 + OC - SVM	0.96	0.91
EfficientNet + OC -SVM	0.72	0.68

Our one-class SVM model delivered impressive results, with an F1-score of 0.96 and a recall of 0.91 with the ResNet50 combination. In this project, our primary emphasis was on recall since our goal was to capture most anomalous objects, even in cases where some might appear normal. Recall, also known as sensitivity or the true positive rate, served as a critical metric for evaluating the model's performance. It measures the model's effectiveness in correctly identifying true anomalies among all actual anomalies in the dataset. A high recall score indicates that the model excels at capturing genuine anomalies while minimizing false negatives, showcasing its proficiency in identifying the majority of anomalies present in the dataset.

As evident from the comparison above, Resnet50 significantly outperformed EfficientNet. EfficientNet, on the other hand, is thought to perform and compile more quickly than Resnet50.

Summary & reflection

Reflecting on our journey, it's clear that our initial approach, inspired by pixel tracking [6], faced a significant setback when we presented it to our industry supervisor. They expressed a need for more generalized solutions and emphasized the importance of real-time functionality. This feedback prompted us to step back and reconsider our strategy, leading us to delve into further research to address the problem in a broader manner.

One example of experimentation that did not yield the expected results was our attempt at point-based trajectory clustering using various clustering algorithms [25]. However, with this failure, we also gain more insight to develop our new solution. This experiment is a valuable lesson, indicating the importance of flexibility and adaptability in our approach.

Communication challenges with the industry partner at Capstone A led to the abandonment of some ideas due to time constraints. This taught us the importance of clear and effective communication, ensuring that we align our efforts with the project's goals from the outset.

On a positive note, successfully running our sample project, "pixel tracking" provided us with a solid foundation and a deeper understanding of the problem. This accomplishment gave us the confidence to build upon this foundation and explore new avenues for solutions.

Rapid testing and experimentation became a crucial part of our development process. Regular testing allowed us to identify what was working and what wasn't early in the development cycle, enabling us to make necessary adjustments promptly.

Additionally, dividing tasks and assigning them to team members with the most relevant experience proved to be an effective strategy. This approach allowed us to develop different components of the project simultaneously, optimizing our progress and efficiency.

In summary, this project development is a series of trials and errors, but these experiences have shaped our approach and methodology. We've learned to be adaptable, communicate effectively, and value experimentation as a means of achieving success in our project.

References

[1] Tuoitrenews. Traffic congestion costs Ho Chi Minh City $6 billion each year. Tuoitrenews. Oct 2022.

Available: https://tuoitrenews.vn/news/society/20221001/traffic-congestion-costs-ho-chi-minh-city-6-billion-each-year/69343.html
[2] VietnamNet. Traffic congestion still a big problem for Hanoi. May 2021.

Available: https://vietnamnet.vn/en/traffic-congestion-still-a-big-problem-for-hanoi-649480.html

[3] Vietnamnet. "Parking space still big problem in Hanoi," Vietnamnet. Sep2021.

Available: https://vietnamnet.vn/en/parking-space-still-big-problem-in-hanoi-754522.html.

[4] Tran, H., Trinh, T., Nguyen, H., & Duong, T. iTASK - Intelligent Traffic Analysis Software Kit. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops ,pp. 1509-1516, June 2020.

Available: https://openaccess.thecvf.com/content_CVPRW_2020/papers/w35/Tran_iTASK_-_Intelligent_Traffic_Analysis_Software_Kit_CVPRW_2020_paper.pdf

[5] Github, Minh-Triet Tran and Tam V. Nguyen and Trung-Hieu Hoang and Trung-Nghia Le and Khac-Tuan Nguyen and Dat-Thanh Dinh and Thanh-An Nguyen and Hai-Dang Nguyen and Trong-Tung Nguyen and Xuan-Nhat Hoang and Viet-Khoa Vo-Ho and Trong-Le Do and Lam Nguyen and Minh-Quan Le and Hoang-Phuc Nguyen-Dinh and Trong-Thang Pham and Xuan-Vy Nguyen and E-Ro Nguyen and Quoc-Cuong Tran and Hung Tran and Hieu Dao and Mai-Khiem Tran and Quang-Thuc Nguyen and The-Anh Vu-Le and Tien-Phat Nguyen and Gia-Han Diep and Minh N. Do, iTASK - Intelligent Traffic Analysis Software Kit, CVPR AI City Challenge 2020 (HCMUS Team)

Available: https://github.com/selab-hcmus/AI_City_2020

[6] Yingying Li, Jie Wu, Xue Bai, Xipeng Yang, Xiao Tan, Guanbin Li, Shilei Wen, Hongwu Zhang, Errui Ding. "Multi-Granularity Tracking with Modularized Components for Unsupervised Vehicles Anomaly Detection." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020.

Available: https://openaccess.thecvf.com/content_CVPRW_2020/papers/w35/Li_Multi-Granularity_Tracking_With_Modularlized_Components_for_Unsupervised_Vehicles_Anomaly_Detection_CVPRW_2020_paper.pdf

[7] Github**,** Dung Vo, Tung Minh Tran, Nguyen D. Vo and Khang Nguyen, UIT-Anomaly: A Modern Vietnamese Video Dataset for Anomaly Detection, NICS2021, pp.352 - 357,2021

Available: https://github.com/nguyenvd-uit/uit-together-dataset/blob/main/UIT-Anomaly.md.

[8] Github, Vu, Tu N and Dinh, Toan T and Vo, Nguyen D and Tran, Tung Minh and Nguyen, Khang, 8th NAFOSTED Conference on Information and Computer Science (NICS), pp. 266 - 271, 2021

Available: https://github.com/nguyenvd-uit/uit-together-dataset/blob/main/VNAnomaly.md

[9] Vietnam News, "Traffic congestion still a big problem for Ha Noi," Vietnam News, May 2023.

Available: https://vietnamnews.vn/society/738234/traffic-congestion-still-a-big-problem-for-ha-noi.html.

[10] Mordor Intelligence, "ANOMALY DETECTION MARKET SIZE & SHARE ANALYSIS - GROWTH TRENDS & FORECASTS (2023 - 2028)", Mordor Intelligence

Available: Anomaly Detection Market Size & Share Analysis - Industry Research Report - Growth Trends (mordorintelligence.com)

[11] ReportLinker, "Global Anomaly Detection Market to Reach US$8.6 Billion by the Year 2026", GlobeNewswire

Available: Global Anomaly Detection Market to Reach US$8.6 Billion by (globenewswire.com)

[12] Du, X., Wang, X., Gozum, G., & Li, Y. (2022). Unknown-Aware Object Detection: Learning What You Don't Know from Videos in the Wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 13678-13688).

Available: https://arxiv.org/abs/2203.03800

[13] Fontanel, D., Tarantino, M., Cermelli, F., & Caputo, B. (2022). Detecting the unknown in Object Detection. arXiv preprint arXiv:2208.11641.

Available: https://arxiv.org/abs/2208.11641

[14] Aboah, A. (2021). A vision-based system for traffic anomaly detection using deep learning and decision trees. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 4207-4212).

Available: https://arxiv.org/abs/2104.06856

[15] AI City Challenge. (n.d.). Home. Retrieved May 8, 2023

Available: https://www.aicitychallenge.org

[16] BAI, Shuai, et al. Traffic anomaly detection via perspective map based on spatial-temporal information matrix. In: CVPR Workshops. 2019. p. 117-124.

Available: https://openaccess.thecvf.com/content_CVPRW_2019/papers/AI%20City/Bai_Traffic_Anomaly_Detection_via_Perspective_Map_based_on_Spatial-temporal_Information_CVPRW_2019_paper.pdf Appendices

[17] K. Doshi and Y. Yilmaz, "Fast Unsupervised Anomaly Detection in Traffic Videos," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 2658-2664, doi: 10.1109/CVPRW50498.2020.00320.

Available: https://ieeexplore.ieee.org/abstract/document/9150642

[18] P. Aboah, R. K. Ward and K. N. Plataniotis, "A Vision-Based System for Traffic Anomaly Detection Using Deep Learning," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2021, pp. 3192-3199. doi: 10.1109/CVPRW50498.2021.00409. [Online].

Available:
https://openaccess.thecvf.com/content/CVPR2021W/AICity/papers/Aboah_A_Vision-Based_System_for_Traffic_Anomaly_Detection_Using_Deep_Learning_CVPRW_2021_paper.pdf

[19] M. F. M. Ariffin and N. A. M. Isa, "An Approach to Detecting and Classifying Type of Dementia Using Electroencephalogram Signals," in IEEE Access, vol. 7, pp. 188691-188703, 2019, doi: 10.1109/ACCESS.2019.2954592.

Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7169508

[20] S. J. Koopman and J. W. Weibull, "Notes and communication: On the possibility of predicting ethical behaviour in bargaining games," in Information Sciences, vol. 178, no. 20, pp. 3995-4002, Oct. 2008.

Available: https://www.sciencedirect.com/science/article/abs/pii/S0020025507001648#preview-section-abstract.

[21] M. Li, Y. Li, Z. Yang and J. Li, "A hybrid feature selection method for intrusion detection based on improved binary cuckoo search," in IEEE Access, vol. 7, pp. 47900-47913, 2019, doi: 10.1109/ACCESS.2019.2901906. [Online].

Available: https://ieeexplore-ieee-org.ezproxy.lib.rmit.edu.au/stamp/stamp.jsp?tp=&arnumber=8649753.

[22] A. R. Hasan, M. N. Uddin, M. M. Uddin and T. Hasan, "A survey of anomaly detection techniques in internet of things," World Electric Vehicle Journal, vol. 13, no. 4, pp. 69, Dec. 2022. doi: 10.3390/wevj13040069. [Online].

Available: https://www.mdpi.com/2032-6653/13/4/69.

[23] W. Li, S. Li, Q. Liu and Z. Li, "Anomaly Detection and Identification for IoT Security: A Survey," IEEE Access, vol. 8, pp. 177399-177417, Aug. 2020, doi: 10.1109/ACCESS.2020.3014013. [Online].

Available: https://ieeexplore-ieee-org.ezproxy.lib.rmit.edu.au/stamp/stamp.jsp?tp=&arnumber=9040873.

[24] Kontopoulos, A. Makris, K. Tserpes and V. Bogorny, “TraClets: Harnessing the power of computer vision for trajectory classification,” *arXiv preprint arXiv:2205.13880*, 2022. [Online]

Available: https://arxiv.org/pdf/2205.13880.pdf?fbclid=IwAR06GdguehSAmu-ZOUWMkwrQzvSvx2OSMe_9kkdI941QXeK2CTSehxdujvQ

[25] C. Zhang and K. Ren, "LRATD: a lightweight real-time abnormal trajectory detection approach for road traffic surveillance," *Neural Comput & Applic*, vol. 34, pp. 22417–22434, Sep. 2022.

Available: https://link.springer.com/article/10.1007/s00521-022-07626-2#citeas

[26] WorldAviation 4K. " Crazy yet Organized Traffic in Vietnam Ho Chi Minh." YouTube, 15 May 2022.

Available: https://www.youtube.com/watch?app=desktop&v=fm0cX7P6PSw

[27]Y. Wang, Y. Zhang, Y. Liu and J. Zhang, "Self-Supervised Learning for Anomaly Detection With Dynamic Local Augmentation," IEEE Access, vol. 9, pp. 108758-108768, 2021, doi: 10.1109/ACCESS.2021.3119354.

Available: Discovering Anomalous Data with Self-Supervised Learning – Google Research Blog

[28] J. -G. Lee, J. Han and X. Li, "Trajectory Outlier Detection: A Partition-and-Detect Framework," 2008 IEEE 24th International Conference on Data Engineering, Cancun, Mexico, 2008, pp. 140-149, doi: 10.1109/ICDE.2008.4497422.

Available: https://ieeexplore.ieee.org/document/4497422

[29] Sohn, Kihyuk, et al. "Learning and evaluating representations for deep one-class classification." arXiv preprint arXiv:2011.02578 (2020).

Available: https://arxiv.org/pdf/2011.02578.pdf

List of Figure

Appendix

Figure 1 – Vietnam Traffic Scene (source: [26]):

Figure 2 – Detail view of Box Level Tracking Branch (source: [6]):

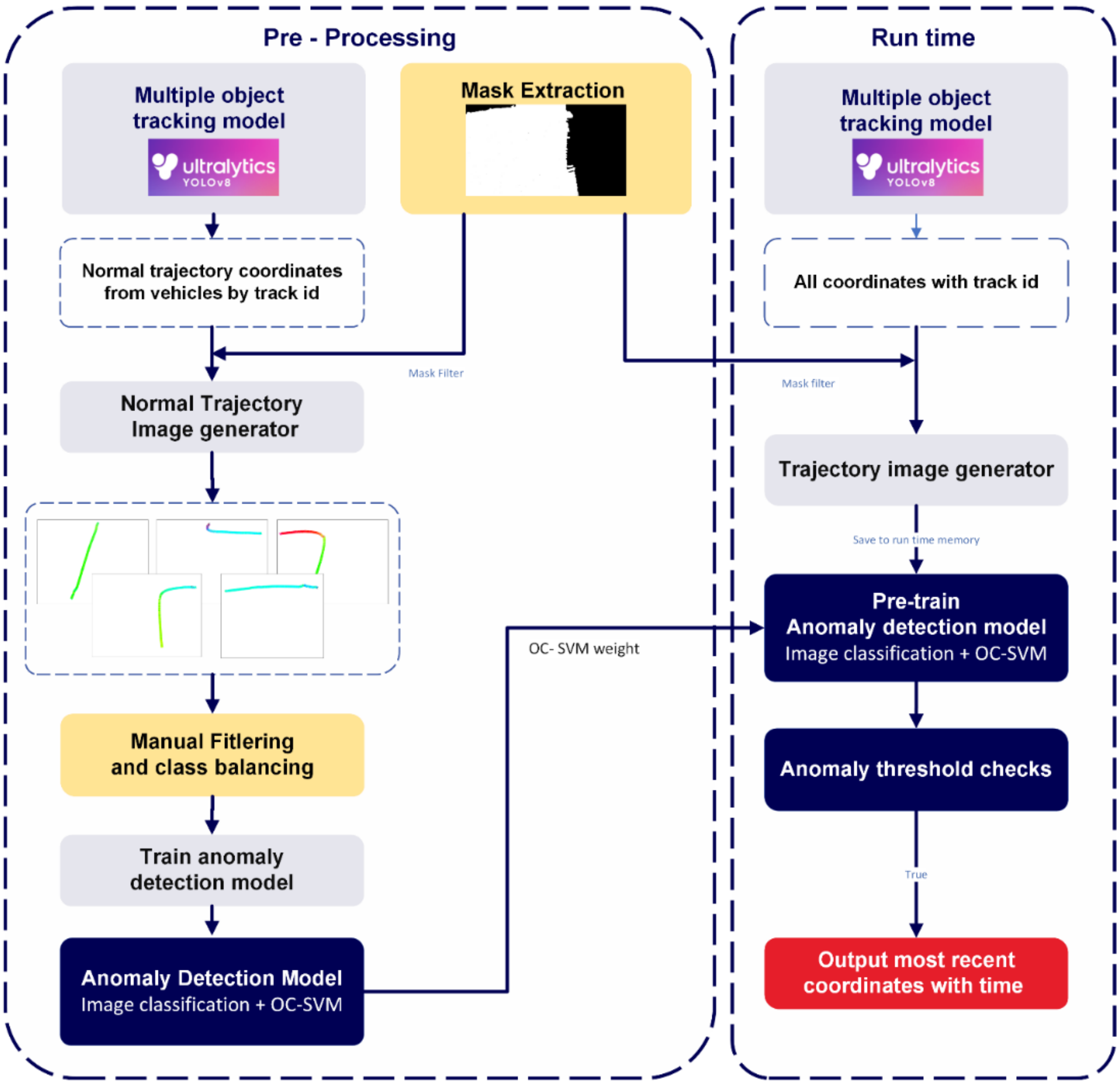
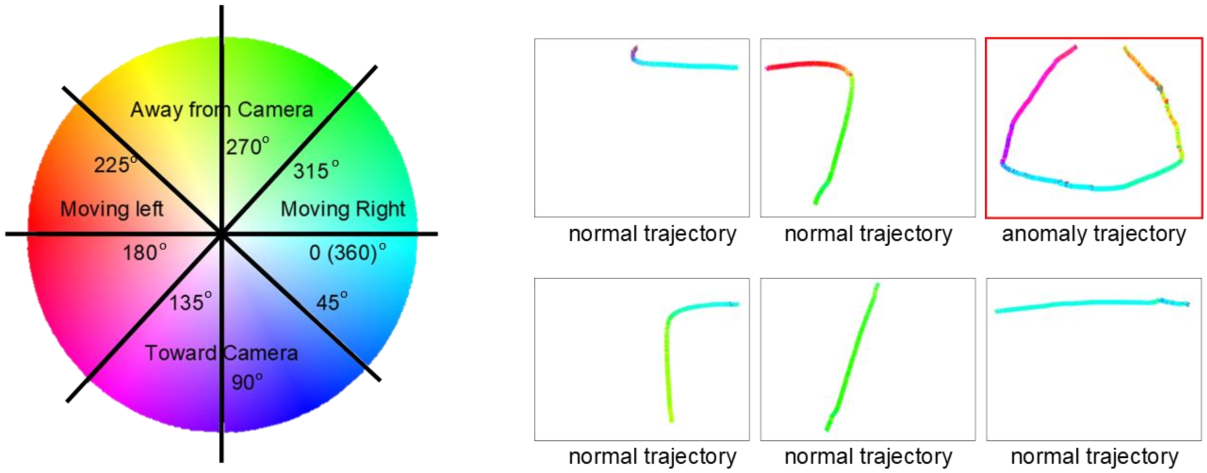


Figure 3 – Color meaning and sample trajectory images generated by this step:



Meeting minutes

Ice Tea Team

Kick-off meeting – week 2

Meeting No:01

Meeting Details

Date:	July 13, 2023
Time:	02:00 pm
Attendees:	Ngô Quang Khải - s3836387
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Update progression that we have completed detection and tracking.
2	Report: algorithm for detection is slow.
3	Request support of other methods to improve detection, reduce memory usage by file writing and reading.

Action Items

No	Item	Who	By
1	Request permission to use google collab.	Duc	
2	Implement the suggested method	Khang, Khai	

Ice Tea Team

Meeting – week 3

Meeting No:02

Meeting Details

Date:	July 20, 2023
Time:	02:00 pm
Attendees:	Ngô Quang Khải - s3836387
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Update progression: Improve the detection and tracking algorithm
2	Report error occurred in runtime and configuration.
3	Request support to reduce output times, video for tracking and detection.

Action Items

No	Item	Who	By
1	Continue the development.	Khang, Khai, Duc	
2	Preparing document	Khoa	

Ice Tea Team

Kick-off meeting – week 4

Meeting No:03

Meeting Details

Date:	July 27, 2023
Time:	02:00 pm
Attendees:	Ngô Quang Khải - s3836387
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	

Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Update progression that we have completed the new modification on detection and tracking system.
2	Report limitation: the system is not running in real-time.

Action Items

No	Item	Who	By
1	Continue improving the performance	All members	
2	Modify system for real-time usage	All members	

Ice Tea Team

Kick-off meeting – week 5

Meeting No:04

Meeting Details

Date:	August 03, 2023
Time:	02:00 pm
Attendees:	Ngô Quang Khải - s3836387
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823

Apologies:	
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Update progression that we are working on real-time tracking and detection.

Action Items

No	Item	Who	By
1	Continue the development	All members	

Ice Tea Team

Meeting – week 6

Meeting No:05

Meeting Details

Date:	August 10, 2023
Time:	02:00 pm
Attendees:	Ngô Quang Khải - s3836387
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Update progression that we have completed real-time running and received output.
2	Report error in logistic issue.

Action Items

No	Item	Who	By
1	Continue to improve performance.	All members	
2	Start training AI to improve confidence score	Duc	

Ice Tea Team

Meeting – week 7

Meeting No:06

Meeting Details

Date:	August 17, 2023
Time:	02:00 pm
Attendees:	Ngô Quang Khải - s3836387
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	No meeting from Academic advisor’s announcement.

Action Items

No	Item	Who	By
1			

Ice Tea Team

Meeting – week 8

Meeting No:07

Meeting Details

Date:	August 24, 2023
Time:	02:00 pm
Attendees:	Ngô Quang Khải - s3836387
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Update progression on trajectory clustering
2	Question on optimal flow.

Action Items

No	Item	Who	By

1	Progress the development	All members	

Ice Tea Team

Meeting – week 9

Meeting No:08

Meeting Details

Date:	August 31, 2023
Time:	02:00 pm
Attendees:	Ngô Quang Khải - s3836387
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Report to the advisors that we were updating our detection model to improve efficiency.

Action Items

No	Item	Who	By

1	Progress the development	All members	

Ice Tea Team

Meeting – week 10

Meeting No:09

Meeting Details

Date:	September 07, 2023
Time:	02:00 pm
Attendees:	
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	Ngô Quang Khải - s3836387
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Report the current progress, new model, image classification.
2	Request for videos from Vietnam Traffic.

Action Items

No	Item	Who	By
1	Progress the development	All members	
2	Prepare document, cover art, poster	Khoa	

Ice Tea Team

Meeting – week 11

Meeting No:10

Meeting Details

Date:	September 14, 2023
Time:	02:00 pm
Attendees:	
	Trần Nguyễn Anh Khoa - s3863956
	Nguyễn Hữu Đức - s3669698
	Trần Vĩnh Khang - s3855823
Apologies:	Ngô Quang Khải - s3836387
Copy To:	

Information /Decision

Item No.	Discussion Summary
1	Report on progress. Mr. Tien has approved our methodology.
2	Mr.Tien ‘s request to send a compiled document at the start of each meeting, so he could comprehend the problem.

Action Items

No	Item	Who	By
1	Complete the project	All members	
2	Create posters, cover art, project description for showcase.	Khoa, Khai	