Journal of Science and Transport Technology
Journal homepage: https://jstt.vn/index.php/en

UNIVERSITY OF TRANSPORT TECHNOLOGY

Integration of Computer Vision, Microscopic Traffic Simulation and Heuristics for Optimizing Motorcycle-Dominated Signalized Intersections

Tam Vu[1,*], Ngoc Viet Pham[1], Ngoc Son Nguyen[1], Quang Thai Pham[1], Thanh Hieu Pham[2], Viet Phuong Nguyen[1]
[1]Faculty of Transportation Engineering, Hanoi University of Civil Engineering, Hanoi, Vietnam
[2]Mechanics of Advanced Materials and Structures, University of Transport Technology, Hanoi, Vietnam

Article info
Type of article:
Original research paper

DOI:
https://doi.org/10.58845/jstt.utt.2025.en.5.4.141-155

***Corresponding author:**
Email address:
tamvm@huce.edu.vn

Received: 06/09/2025
Received in Revised Form: 01/12/2025
Accepted: 12/12/2025

Abstract: Traffic signal optimization in motorcycle-dominated environments remains a critical challenge in many developing cities, where the heterogeneity and high dynamics of traffic flows often limit the effectiveness of traditional control methods. This study introduces an approach that integrates computer vision, microscopic traffic simulation and heuristic optimization to design signals of motorcycle-dominated mixed traffic intersections. By leveraging visual data through modern object detection techniques, the proposed framework enables a more comprehensive and precisely of key traffic parameters including traffic volume and travel time - overcoming the limitations of conventional field surveys and loop detectors. These data are then utilized for developing and calibrating VISSIM models to accurately reflect reality. Rule-based and multi-start local search heuristics are implemented with VISSIM and Python to iteratively refine signal timing plans, aiming to minimize travel time and queue lengths at intersections. A case study conducted at a motorcycle-dominated intersection in Hanoi, Vietnam demonstrates the potential of this integration to improve both operational efficiency and adaptability of signal control systems. The chosen solution performs much better than the existing situation, with the average queue length and travel time reduced by approximately 52.5% and 16.3% correspondingly. The findings can prove the feasibility and accuracy of proposed integrated framework that traffic engineers and decision-makers might apply in motorcycle-dominated mixed traffic environments in practice.

Keywords: traffic signal optimization, computer vision, VISSIM, heuristic search, motorcycle-dominated intersections.

1. Introduction

Traffic congestion poses a significant and growing challenge to urban mobility, economic productivity, and environmental sustainability in cities worldwide. The issue is particularly acute in many Southeast Asian nations, including Vietnam, Indonesia, and Thailand, where transportation infrastructure struggles to accommodate rapidly growing numbers of vehicles, predominantly motorcycles [1, 2]. This motorcycle-dominated traffic creates a unique and complex heterogeneous flow, characterized by non-lane-based movements, high densities, and dynamic vehicle interactions, which profoundly impacts intersection performance [3, 4].

Signalized intersections, as critical nodes in the urban network, are primary sources of delay, fuel consumption, and emissions in these environments. Conventional traffic signal control strategies, often derived from principles established for homogeneous car traffic, frequently prove inadequate for managing the intricate dynamics of mixed flows with high motorcycle proportions [5]. The effectiveness of any signal optimization effort is fundamentally contingent on the accuracy and comprehensiveness of its input data, namely traffic volumes and performance measures like travel time and queue length [6]. Traditional data collection methods, including manual surveys and inductive loops, are often incapable of capturing the full complexity of motorcycle-dominated traffic streams, typically sampling only a small fraction of vehicles and struggling with vehicle classification in dense, mixed conditions [7].

Recent advancements in computer vision, particularly deep learning–based object detection models such as YOLO, have demonstrated transformative potential for traffic data extraction [8]. Nevertheless, most existing studies employ these technologies in isolation, focusing on detection accuracy rather than leveraging the extracted data for optimizing motorcycle-dominated signalized intersections. Although microscopic traffic simulation platforms such as VISSIM are widely used for performance evaluation [9], and heuristic optimization techniques are increasingly applied for signal timing optimization [10], few studies have established an integrated methodological framework that couples high-fidelity computer vision–based data extraction with simulation and iterative heuristic search, particularly under motorcycle-dominated mixed traffic conditions. As a result, this study proposes a novel integrated framework that combines computer vision, microscopic simulation (VISSIM) and heuristic search in finding near-optimal solutions for signals plan of intersections.

The remainder of this paper is structured as follows. Section 2 provides a review of related literature. Section 3 details the proposed integrated methodology. Section 4 presents a case study in Hanoi, Vietnam. Section 5 discusses the results and concludes with key findings and suggestions for future research.

2. Literature Review

The optimization of motorcycle-dominated signalized intersections can remain a significant challenge in urban areas because of complexity related to data collection and design process. The unique characteristics of motorcycle traffic - such as their small size, high maneuverability, and non-lane-based movement - render conventional traffic analysis and control strategies, often developed for homogeneous car traffic, less effective. A substantial body of research has employed microscopic traffic simulation (e.g. VISSIM) to understand these complex dynamics. Studies consistently highlight the impact of motorcycle-specific behaviors, revealing that allowing motorcycles to filter and position themselves ahead of cars can reduce start-up lost time and overall travel time [11, 12]. Further analyses of heterogeneous traffic in cities have used VISSIM to quantify key performance metrics such as delay and queue length [13].

Traditional methods relying on manual counts, loop detectors, or low-frequency GPS data have inherent limitations. Manual counts are labor-intensive and sample only a small fraction of the traffic stream, while methods using GPS trajectory data often struggle with low-frequency data, leading to estimations that may not capture full traffic variability [14]. The recent application of computer vision, specifically YOLO, for object detection and analysis points towards a paradigm shift, offering the potential to automatically and accurately track nearly all vehicles (90-95%) within view of camera on-site [8, 11]. This might provide a rich, high-resolution dataset for travel time and queue length calculation that was previously unattainable at scale. The critical advantage might be the ability to extract comprehensive travel time data directly from video feeds, which serves as the essential ground truth for building and calibrating highly accurate VISSIM models - a foundational step that is often compromised in previous studies due to data paucity.

Building upon the need for adaptive control, advanced computational intelligence approaches have been explored. Vuong et al. [15] proposed an adaptive traffic signal control method for an isolated intersection under mixed traffic conditions in Hanoi using an Adaptive Neuro-Fuzzy Inference System (ANFIS) integrated with VISSIM and MATLAB. The model employed inputs of maximum queue length and vehicle arrivals to infer phase urgency and dynamically adjust green durations within fixed cycle times derived from the Webster formula. The ANFIS controller adaptively tuned membership functions through hybrid learning to minimize delay and travel time under motorcycle-dominated traffic. While the approach demonstrates strong adaptability, it maintains a fixed cycle length and relies on a rule-based learning mechanism. This can highlight a common limitation in the field: even sophisticated adaptive controllers often operate within a constrained search space. The accuracy and performance of these systems, including the underlying simulation models used for their development, appear to be limited by accuracy of VISSIM models.

A critical aspect of modeling mixed traffic for signal optimization is the accurate conversion of heterogeneous traffic volumes into a uniform unit. A study by Roy et al. [16] optimized signal timing under heterogeneous traffic using a direction-wise dynamic Passenger Car Equivalent (PCE) model. The approach recalculated saturation flow and total intersection delay based on adaptive PCE values derived from classified vehicle counts and queuing analysis. Signal timing parameters were then adjusted to minimize total delay. Their findings indicated that static PCE assumptions significantly underestimated actual delay, highlighting the need for direction-specific calibration when optimizing signalized intersections in mixed traffic environments.

Several methods have been developed to determine the Passenger Car Unit (PCU) in heterogeneous traffic flow, such as the homogenization coefficient method, semi-empirical method, Walker's method, headway method, multiple linear regression, and simulation-based approaches [17]. However, these methods were primarily formulated for car-dominated traffic and might be therefore unsuitable for environments with dominance of motorcycles, particularly in Southeast Asian cities. Building upon and improving the aforementioned methods, Cao and Sano [17] introduced a methodology for estimating the Motorcycle Equivalent Unit (MCU) that better reflects the characteristics of urban traffic in Hanoi, as well as for evaluating the capacity of urban roads under mixed traffic conditions. Based on this conceptual background, the study develops a more context-sensitive framework to analyze motorcycle-dominated mixed traffic flow, emphasizing the dynamic relationship between vehicle speed, space occupancy, and road capacity.

Heuristic optimization methods have been applied in transportation and traffic control problems where the search space is complex and exact optimization is computationally infeasible. Unlike traditional mathematical programming techniques, heuristics aim to obtain good-though not necessarily optimal-solutions within a reasonable computational time [18]. In traffic signal optimization, heuristic approaches are particularly useful because of the nonlinear and stochastic nature of vehicle interactions in microscopic simulation models [10]. Among heuristic methods, local search is one of the most fundamental techniques. It starts from an initial solution and iteratively explores its neighborhood to find improved solutions according to a defined objective function [19]. Although local search can efficiently converge with high-quality solutions, it often gets trapped in local optima. To overcome this limitation, multi-start local search (MSLS) strategies have been developed, which repeatedly restart the search from multiple initial solutions [18]. Another commonly used approach is the greedy heuristic, which constructs a solution step-by-step by making the locally optimal choice at each iteration [20]. While greedy algorithms are computationally efficient and easy to implement, they may yield suboptimal global performance due to their myopic decision-making. Nonetheless, greedy strategies are often integrated into hybrid or simulation-based frameworks to provide quick initial estimates of signal timings or phase splits before applying more sophisticated search techniques [21]. Recent studies have increasingly revisited rule-based heuristics as an interpretable and operationally safe alternative to purely search-based methods in traffic signal control. A recent review highlights how hybrid control frameworks preserve the core “plan authority” of rule-based systems (cycle / split / offset decisions) while overlaying adaptive heuristics or data-driven adjustments, thus ensuring both real-world deployability and performance gains [22].

To address these gaps, the present study adopts a simulation-based heuristic framework using Python–VISSIM integration to iteratively search for globally improved signal timing plans by varying both green splits and cycle times. This methodology directly addresses the identified research gaps by creating a closed-loop optimization system. It first leverages the robust travel time, queue length and traffic volume data extracted by AI models that is used in the study of Vu et al. [8], to develop and calibrate a highly accurate VISSIM model. This well-tuned model then acts as a reliable evaluation function within a heuristic search process, where Python scripts systematically alter signal parameters and VISSIM simulations assess their impact on key performance indicators. This integration of high-resolution data, multi-model evaluation, and iterative simulation-based search moves beyond the isolated application of these technologies towards a holistic and data-driven optimization system, capable of conducting a global search for optimal signal timings that are specifically tailored to the complex dynamics of motorcycle-dominated intersections.

3. Methodology

The methodological framework of this research that includes sequential six-stage workflow is illustrated in the first subsection. The second subsection states application of YOLO for vehicle detection and data extraction which details the computer vision techniques used to derive accurate traffic parameters from video footage. The third subsection describes the hybrid optimization algorithm that refines the signal timings.

3.1. Methodological framework

The research methodology follows a systematic six-stage framework designed to optimize signalized intersections in motorcycle-dominated environments through the integration of computer vision, simulation modeling, and heuristic optimization in Fig. 1.

Stage 1: Collection data by camera. Field data was collected at a signalized intersection by camera. Simultaneous recording was initiated using cameras strategically positioned at each approach of the intersection, providing comprehensive, multi-directional coverage of all traffic movements, queue length and vehicle interactions. This setup enabled the detailed observation of traffic stream behavior across all approaches, forming an empirical foundation for subsequent computer vision analysis and model calibration.

Stage 2: Traffic parameters determined by computer vision. The video footage from the cameras was processed using an advanced computer vision pipeline built upon the YOLO architecture. This deep learning model was specifically fine-tuned for Vietnamese traffic conditions, enabling highly accurate detection and classification of 15 distinct vehicle types. The system employed the ByteTrack multi-object tracking algorithm to maintain consistent vehicle identities across frames, even though challenging scenarios such as occlusions at the congested intersection. This robust tracking capability was essential for generating reliable vehicle trajectories and ensuring data integrity throughout the analysis period.

The extraction of quantitative traffic parameters was achieved through strategically placed virtual detection lines at each intersection approach. Vehicle travel times were calculated with precision by recording timestamps when vehicles crossed successive detection lines. The computer vision system automatically compiled comprehensive traffic volume counts for each movement, classified by vehicle type, creating a rich dataset that captured the intricate dynamics of motorcycle-dominated traffic flow. This data extraction methodology successfully achieved approximately 90-95% coverage of the vehicle stream, far surpassing the limitations of traditional manual observation methods [8]. This will be further explained in section 3.2.

Stage 3: Develop VISSIM simulation by calibrating traffic parameters from Stage 2. A microscopic simulation model was developed in VISSIM to replicate the observed intersection geometry, signals and traffic conditions. The model was rigorously calibrated using the computer vision-derived parameters from Stage 2, including travel time. Queue length is used for the VISSIM calibration process. However, real queue length is counted by manual in this research. Multiple calibration iterations were performed until the simulated travel times and queue lengths demonstrated statistical alignment with field observations, ensuring the model accurately represented real-world intersection operations.

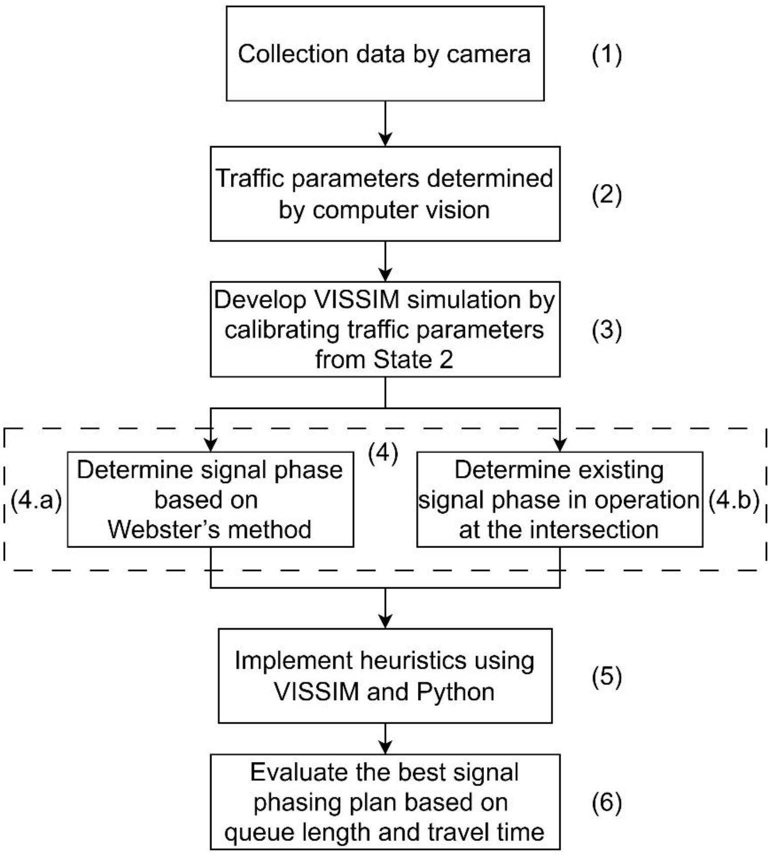

Fig. 1. Methodological framework

Stage 4a: Determine signal phase based on Webster’s method. Initial signal timing parameters were calculated using Webster's formula adapted for heterogeneous traffic conditions. Two distinct conversion approaches were implemented: PCU factors that normalize all vehicles to PCU and MCU factors that normalize all vehicles to equivalent motorcycles. This dual-conversion methodology generated alternative signal timing plans reflecting different optimization perspectives for mixed traffic environments, which will be discussed in more detail in section 3.3.

Stage 4b: Determine existing signal phase in operation at the intersection. The current signal timing plan operating at the study intersection was thoroughly documented, including phase sequences, cycle lengths, green splits, and amber intervals. This existing configuration served as the baseline scenario for comparative performance evaluation against the optimized timing plans developed through the research methodology.

Stage 5: Implement heuristics using VISSIM and Python. An integrated optimization framework was developed using Python-VISSIM COM interface to systematically refine signal timing parameters through two complementary heuristic approaches. The initial solutions for this optimization process were derived from Stage 4, comprising both the existing signal timing (Stage 4b) and the Webster-optimized plans using PCU and MCU conversions (Stage 4a). These three distinct initial solutions served as starting points for parallel optimization pathways in multi-start heuristic, ensuring comprehensive exploration of the solution space.

Based on the study of [23], the following default parameters of VISSIM in the research are used for this study: Desired speed distribution of bus and car; Motorcycle desired speed distribution; Average standstill distance; Minimum lateral distance driving at 50 km/h; motorcycle acceleration and deceleration parameters. In addition, maximum deceleration for cooperative braking is modified in this research to compare travel time between VISSIM model and reality.

The optimization employed a hybrid strategy combining Rule-Based Heuristic and Multi-Start Local Search (MSLS) methodologies. For each initial solution, the Rule-Based Heuristic first performed localized adjustments to green splits based on predefined improvement rules targeting travel time and queue length reduction. Subsequently, the MSLS algorithm initiated multiple optimization trajectories from these refined solutions, systematically exploring neighboring regions in the parameter space to escape local optima. This dual-layer approach enabled both intensive local refinement and extensive global exploration, with each candidate solution evaluated through multiple VISSIM simulation runs to ensure statistical significance and reliability of performance metrics, which will be discussed in section 3.4.

Stage 6: Evaluate the best signal phasing plan based on queue length and travel time. The performance evaluation of candidate signal timing plans was conducted based on two critical performance indicators: the average vehicle travel time through the intersection approach and the average queue length across all approaches.

3.2. Application of YOLO for vehicle detection and data extraction

Accurate vehicle detection and tracking are fundamental to extracting reliable traffic parameters. This study employs the YOLOv8 (You Only Look Once version 8) model, building upon the enhanced architecture proposed by Vu et al. [8] for robust vehicle detection in complex traffic scenarios. The model was pre-trained on a comprehensive dataset containing over one million annotated vehicle instances across 15 vehicle types, ensuring high detection accuracy in mixed traffic conditions typical of motorcycle-dominated intersections. Although newer versions of YOLO (e.g., v11 and v12) have been released, the framework of Vu et al. [8] was specifically optimized and validated for dense, heterogeneous traffic environments in Southeast Asia. Its large-scale, domain-specific dataset and proven robustness make YOLOv8 a more reliable choice for this study's application context compared to the generic configurations of later base-versions.

The detection process begins by processing input video frames through the YOLO network, which outputs bounding boxes with corresponding class labels and confidence scores for each detected vehicle. To establish temporal consistency and enable trajectory analysis across consecutive frames, ByteTrack - a multi-object tracking algorithm is integrated [24]. This tracking module associates detections over time by calculating feature similarities and motion patterns, effectively maintaining vehicle identities through occlusions and complex interactions.

The vehicle detection and tracking process was implemented using the YOLOv8 architecture. To ensure optimal performance in the specific context of Vietnamese traffic conditions, the model was pre-trained on a comprehensive local dataset containing diverse vehicle types and traffic scenarios [8]. The trained model demonstrated high detection accuracy, achieving a mean Average Precision (mAP@0.5) of 0.91 on validation set, confirming its reliability for subsequent traffic parameter extraction.

For travel time extraction, virtual detection lines are strategically placed at the entry and exit points of the intersection approach. Two distinct virtual lines are digitally established perpendicular to the traffic flow at strategic locations: one upstream (entry point) and one downstream (exit point) of the intersection approach. These lines are marked in VISSIM simulation at stage 3 in Fig. 1. The multi-object tracking algorithm maintains consistent vehicle identities (IDs) across frames, generating continuous trajectories for each detected vehicle. The travel time measurement protocol is executed as follows: when a vehicle's bounding box centroid crosses the first virtual line, a precise timestamp (T_1) is recorded. The system continues tracking the same vehicle ID until its centroid crosses the second virtual line, where a second timestamp (T_2) is registered. The travel time (TT) for that specific vehicle is then computed as:

$$TT = T_1 – T_2 \quad (1)$$

This process is automatically repeated for all successfully tracked vehicles, generating a comprehensive dataset of individual travel times across the intersection approach. This method enables the capture of nearly the entire vehicle stream, significantly surpassing the limited sampling capabilities of manual observation or fixed-loop detectors. Moreover, traffic volumes are quantified by counting vehicles crossing virtual detection lines during specific analysis intervals, employing the traffic counting model established by Vu et al. [8]. This automated vision-based approach generates a rich, high-resolution dataset of travel times and traffic composition, serving as the empirical foundation for both simulation model calibration and signal system optimization.

To assess the performance of the proposed framework, all experiments were carried out on a workstation equipped with an AMD Ryzen Threadripper 3960X 24-core processor, 128 GB of RAM, and an NVIDIA GeForce RTX 3090 GPU with 24 GB of memory. Under this hardware setup, the system consistently achieved an average inference speed of approximately 27 frames per second (FPS) on 1080p video streams, even in high-density and heterogeneous traffic conditions. These results indicate that the model is well-suited for deployment in practical urban monitoring environments.

3.3. Implement heuristics using VISSIM and Python

This section details the comprehensive heuristic optimization framework that integrates VISSIM simulations with Python control logic to systematically identify optimal signal timing parameters. The framework employs a hybrid three methodology, sequentially combining Rule-Based Heuristics, Local Search, and Multi-Start Heuristics. This structured approach, detailed in eleven sequential steps (as illustrated in Fig. 2.), ensures a thorough exploration of the solution space, beginning with rapid, rule-guided improvements and progressing to intensive local and global search to converge on a robust, high-performing signal plan.

Step 1: Initial solution selection method. The optimization process commences by selecting initial signal timing parameters derived from Stage 4 of the methodological framework. This includes choose distinct starting points: the existing signal timing configuration (State 4b), and the two Webster formula based on PCU and MCU conversions (State 4a). These diverse initial solutions ensure comprehensive exploration of the solution space from multiple perspectives.

Step 2: Create initial solution. Cycle length, phase sequence, and green splits are determined for the initial solution.

Step 3: Run initial solution in VISSIM and Python. The integrated simulation-execution framework is initiated, where Python scripts automatically configure and execute the VISSIM simulation with the initial signal timing parameters.

Step 4: Determine Outputs. Key performance indicators are extracted from the simulation output, including queue lengths for all intersection approaches and travel times for all traffic movements. These metrics are systematically recorded for subsequent comparative analysis.

Step 5: Find new solution according to rule-based heuristic. Rule-based heuristics is applied where approach-specific green times are adjusted based on queue length performance. Approaches exhibiting queue lengths longer than the average one receive a 1 second green time increase, while those with shorter-than-average queues decrease a 1 second green time reduction, maintaining constant cycle length through compensatory adjustments.

Step 6: Run new solution in VISSIM and Python. The modified signal timing parameters are implemented in the simulation environment, and the updated configuration is executed through the Python-controlled VISSIM interface to evaluate performance of the adjusted solution.

Step 7: Determine best solution from iteration 1 to existing iteration. Performance metrics from successive iterations are compared to identify progressive improvements. If convergence criteria are not met - defined as the reduction in average queue length compared to the best solution from previous iterations is less than 1 % - the process returns to Step 5 for further iterative refinement through the 5-6-7 loop.

Step 8: Retain the final best solution that is evaluated. Upon achieving convergence, the best-performing signal timing configuration from the rule-based heuristic optimization is retained as the locally optimized solution for the current initial starting point.

Step 9: Local search. To ensure robustness and escape potential local optima, a local search is conducted by exploring neighboring solutions around the identified optimum. The cycle time is increased and decreased by 4 seconds. Green splits are added for all phases in cycle time. This process re-executes Steps 2 through 8 with strategically varied initial conditions within the vicinity of the current best solution.

Step 10: Multi-start heuristics. A comprehensive search strategy is implemented by repeating the entire optimization procedure (Steps 1 through 8) for each of the three distinct initial solutions derived from different traffic conversion methodologies. They can be existing situation, Webster formula with PCU and MCU. This multi-start approach ensures global exploration of the solution space from diverse starting points.

Step 11: Chosen Solution. The final optimal signal timing plan is selected from the collection of solutions generated through the multi-start heuristic process. This chosen solution demonstrates superior performance across evaluation criteria, minimizing both average travel time and maximum queue length across the intersection approaches, representing the globally optimal configuration identified through the exhaustive search methodology.

4. Case study

Pham Ngoc Thach – Luong Dinh Cua intersection in Hanoi, Vietnam (shown in Fig. 3) is chosen as a case study. It can be named as a four-leg junction or double closely spaced T-intersections. A notable feature of the site is the presence of pedestrian overpasses, which provide an elevated and unobstructed vantage point for data acquisition. Traffic data were collected through high-definition video recordings between 13:00-14:00 on March 3, 2024. In addition to video signals data, on-site geometric surveys were conducted to record key intersection attributes, including lane widths and lengths, approach spacing, and stop-line locations. These data are inputs for developing VISSIM models.

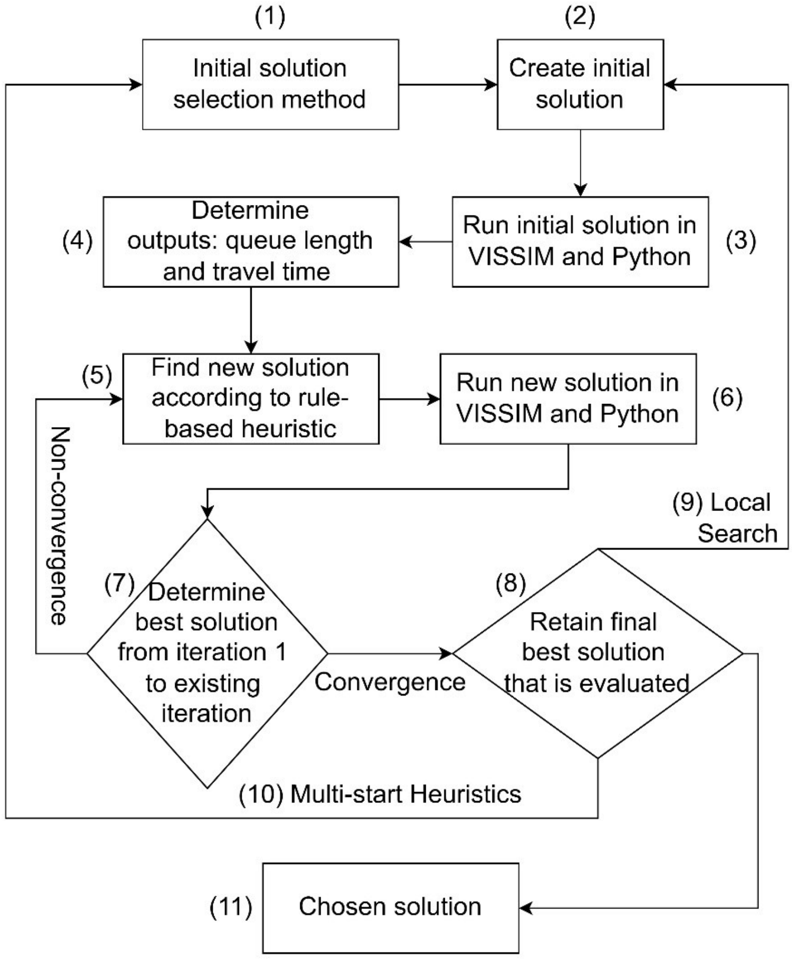

Fig. 2. Method Implement heuristics using VISSIM and Python

This case study is executed sequentially according to the stages outlined in section 3.1. The initial stage, high-resolution cameras were strategically positioned at each leg of the intersection complex to capture comprehensive traffic movements across all approaches. The video footage from all cameras was processed through an advanced computer vision pipeline utilizing the YOLOv8 architecture and ByteTrack algorithm. Virtual detection lines were strategically placed at all entry and exit points of the intersection approaches, enabling precise calculation of travel times and accurate traffic volume counts. The model initially classified vehicles into 15 detailed categories. In this case study, the results of traffic counts were grouped into four types: motorcycles, passenger cars, trucks, and coaches that are presented in Table 1, providing the essential input parameters for subsequent simulation modeling and optimization processes.

At stage 3, a detailed microscopic simulation model was subsequently developed in VISSIM to accurately replicate the intersection's geometry, lane configurations, and traffic flow characteristics.

The model incorporated the extracted traffic composition data from Stage 2, along with observed driver behavior parameters specific to motorcycle-dominated traffic conditions, including reduced vehicle following distances, and lane-changing behavior

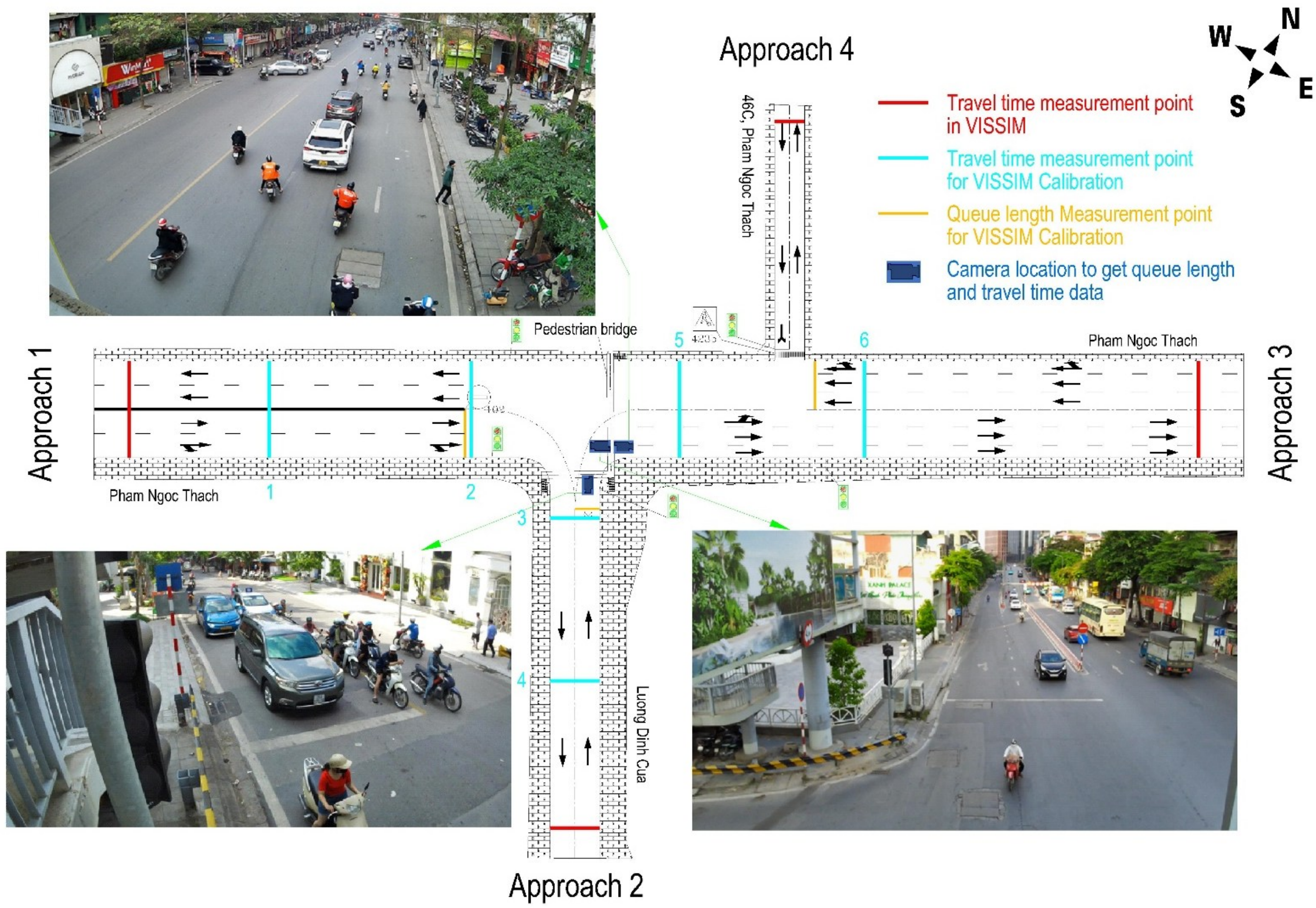

Fig. 3. Layout of Pham Ngoc Thach – Luong Dinh Cua intersection

Table 1. The traffic volume for all approaches within the intersection on March 3, 2024

No.	Movement	Motorcycle	Car	Coach	Truck	Total (veh/h)
1	Approach 1 - Go straight	3252	496	40	4	3792
2	Approach 1 - Turn right	256	36	0	0	292
3	Approach 2 - Turn left	356	68	4	0	428
4	Approach 2 - Turn right	508	84	0	0	592
5	Approach 3 - Go straight	2828	400	28	12	3268
6	Approach 3 - Turn right	180	8	0	0	188
7	Approach 4 - Turn right	288	20	0	0	308
8	Approach 4 - Turn left	192	8	0	0	200
9	Cross-section below the pedestrian bridge - Turn left	276	20	0	0	296
10	Cross-section below the pedestrian bridge - Go straight	3032	428	28	12	3500
11	Cross-section below the pedestrian bridge - Turn left	508	12	0	0	520
12	Cross-section below the pedestrian bridge - Go straight	3252	568	40	4	3864

The calibration process involved iterative adjustments to key behavioral parameters including queue length and travel time. The final calibrated model achieved strong statistical validation with minimal error margins. The validation results presented in Tables 2 and 3 demonstrate the model's accuracy in replicating real-world queue formation patterns and travel time, confirming that the parameter adjustments successfully produced a simulation environment that closely matches field conditions and provides a reliable platform for subsequent signal optimization experiments.

Table 2 presents a comparison between the observed queue lengths from the field survey and the values obtained from the VISSIM model. The differences in queue lengths between reality and VISSIM results are small, ranging from -10.5% to -6.7% (0.6m to 1.6m). As a result, the VISSIM model demonstrates relatively high accuracy in simulating queue lengths at the approaches, with only minor deviations from the field survey results.

Table 3 presents a comparison between the travel times between two cross-sections obtained from AI models and VISSIM for different directions. The results show that differences ranging from -5.7% to 7.4% are less than 4 seconds. Therefore, the VISSIM model demonstrates relatively high accuracy in travel times. Proceeding with the framework, at Stage 4a, Webster's signal timing method is applied, employing a dual-conversion methodology. The heterogeneous traffic volumes are processed using both PCU and MCU values to calculate distinct sets of optimal cycle lengths and phase splits, thereby generating two theoretically optimized scenarios for the intersection's specific traffic composition.

In addition, Stage 4b involves the detailed documentation of the intersection's existing signal control strategy. This existing operational plan serves as the crucial baseline scenario, enabling a comparative performance evaluation against the proposed models in subsequent stages.

At stage 5, the Python-VISSIM optimization framework was implemented, executing the eleven-step heuristic procedure described in Section 3.4. The algorithm systematically evaluated signal timing parameters, iterating through rule-based adjustments and multi-start local searches initiated from the three scenarios (existing situation, PCU-Webster, MCU-Webster). For existing situation, cycle time is 98 s while these values for PCU-Webster and MCU-Webster are 108 and 89 s respectively. For local search heuristic, the cycle time of each scenario is changed ±4-seconds, resulting in nine total scenarios for optimization.

At stage 6, the summarized results of this optimization process are summarized in Table 4, which presents a comparative analysis of key performance indicators across different solutions. For each scenario with different cycle time, Table 4 shows the average queue length and travel time for both the initial solution (iteration 1) and the final solution at the converged iteration (iteration n).

Table 4 indicates that the final solution at the converged iteration is better than the initial solution at iteration 1 for all scenarios. The chosen solution of all scenarios has cycle time of 89 seconds with average queue length and travel time of 5.7 meters and 42.2 second respectively. This best solution performs much better than the current signal plan on March 3, 2024, with the average queue length and travel time reduced by approximately 52.5% and 16.3%, correspondingly. This solution is achieved from the initial solution based on Webster formula with motorcycle equivalent and converged at iteration 7. This seems to be consistent with motorcycle-dominated mixed traffic environments. This is because the roadway capacity is converted into motorcycle units, and other vehicle types are also normalized into motorcycle equivalents. Based on the Webster formula, this study generated the initial solution, then applied a rule-based heuristic that produced the best solution (consistent with the proposed framework) at the 7th iteration. With the initial solution based on Webster formula with the best solution PCU, the optimal plan has a cycle time of 104s, with an average queue length and average travel time of 6.9 m and 44.7 s, respectively. This best solution also performs much better than the existing situation, reducing the average queue length and travel time by approximately 42.5% and 11.3%, respectively. These results demonstrate that the integrated framework combining computer vision-based data extraction, microscopic simulation, and heuristics to optimize signalized intersections dominated by motorcycles in this study is both feasible and effective. It can provide a practical approach for optimizing motorcycle-dominated signalized intersections, leading to measurable improvements in operational efficiency and traffic performance.

Table 2. Comparison of Queue Lengths between Field Survey and VISSIM Results

	Queue length at approach 1	Queue length at approach 2	Queue length at approach 3	Average
Reality by manual (meters)	9.0	10.0	26.0	15.0
VISSIM (meters)	8.1	9.4	24.4	14.0
Difference ratio (%)	-10.5	-6.8	-6.7	-7.5

Table 3. Comparison of Travel Times between Field Survey and VISSIM Results

Direction	Real travel time (seconds)	Travel time in VISSIM (seconds)	Difference ratio (%)
Section 2 to 1	58	61.5	5.7
Section 1 to 2	121	117.1	-3.4
Section 3 to 4	36	34.0	-5.7
Section 4 to 3	112	108.4	-3.4
Section 5 to 6	40	43.2	7.4
Section 6 to 5	153	149.3	-2.5
Average	86.7	85.6	-1.3

Table 4. Evaluation of Queue Length and Travel Time Convergence across Different Signal Cycle Times

		Average queue length (meters)	Average travel time (seconds)
Cycle time is 98 s	**Iteration 1(existing situation)**	**12.0**	**50.4**
	Converge at Iteration 10	10.4	49.7
Cycle time is 102 s	Iteration 1	14.5	53.7
	Converge at Iteration 10	11.5	51.6
Cycle time is 94 s	Iteration 1	12.0	48.9
	Converge at Iteration 6	10.4	44.8
Cycle time is 108 s	Iteration 1	7.9	48.3
	Converge at Iteration 8	7.0	44.8
Cycle time is 112 s	Iteration 1	8.2	48.5
	Converge at Iteration 10	7.5	46.4
Cycle time is 104 s	Iteration 1	7.5	48.2
	Converge at Iteration 7	**6.9**	**44.7**
Cycle time is 89 s	Iteration 1	6.8	46.1
	Converge at Iteration 7	**5.7**	**42.2**
Cycle time is 93 s	Iteration 1	6.7	44.3
	Converge at Iteration 10	6.5	43.2
Cycle time is 85 s	Iteration 1	7.2	47.5
	Converge at Iteration 6	6.1	42.6

5. Conclusion

This study proposed an integrated methodology combining computer vision, microscopic simulation, and heuristics to optimize signalized intersections dominated by motorcycles. The main contributions of this paper include: (i) Traffic volume and travel time counted by AI models, alongside manually measured queue lengths, served as inputs for developing and calibrating a VISSIM model tailored to a motorcycle-dominated mixed traffic environment. An accurate VISSIM model that appropriately reflects reality is crucial as a basis for evaluating proposed signal phasing plans; (ii) A combination of multi-start local search and rule-based heuristics was employed to identify the optimal solution for signal cycle and phasing at the intersection, based on the two key performance indicators: travel time through the intersection and queue length; (iii) A method for determining the initial solution was proposed using the traditional Webster formula with conversion to motorcycle equivalents. This approach appears suitable for mixed traffic flows with high motorcycle proportions and is demonstrated in the case study results; and (iv) The integrated model proposed in this paper runs automatically in Python via the COM interface. This significantly reduces processing time for finding a near-optimal solution.

The research model was applied to an intersection in Hanoi. This intersection is more complex than a typical four-legged junction, as it can be considered as two closely spaced three-legged intersections. The results indicate that the chosen solution was found when the initial solution was determined using the Webster formula with MCU. This chosen solution showed considerable improvement over the existing situation in terms of both average queue length and average travel time through the intersection. This demonstrates the feasibility and effectiveness of the proposed research. Given the widespread installation of low-cost cameras at intersections and the popularity of integration with programming environments (e.g. Python) of microscopic traffic simulation (e.g. VISSIM and SUMO), traffic engineers and decision-makers can apply the integrated model from this paper in practice.

However, this study has following limitations. From the camera data, the vehicle queue length before the stop line at the signalized intersection has not been determined by AI models. Future research will address this issue by developing AI models to automatically extract this parameter, thereby providing optimal input for developing and calibrating microscopic traffic simulation models best suited for simulating motorcycle-dominated mixed traffic signalized intersections.

While the integration of multi-start local search and rule-based heuristics in this study proved effective and reduced computation time for finding a near-optimal solution, future work will explore the combination with other Heuristics (e.g., Simulated Annealing) and Reinforcement Learning to identify even better solutions.

Furthermore, this study optimized the signal timing based on the morning peak hour demand pattern. While this demonstrates the framework's effectiveness under high-demand conditions, the performance of the optimized plan during off-peak hours or other distinct periods, such as the evening peak with potentially different directional splits. Future research should therefore focus on evaluating the robustness of the optimized solutions across multiple time periods and developing adaptive signal control strategies that can dynamically respond to temporal variations in traffic demand.

Acknowledgments

This work was supported by the Vietnamese Ministry of Education and Training under Grant B2025-XDA-03.

References

[1] K. Wilinski, S. Pathak. (2022). Mobility in The Developing Country. The Case Study of Bangkok Metropolitan Region. *Komunikácie*. 24(3), A112-A122. https://doi.org/10.26552/com.C.2022.3.A112-A122

[2] D.N. Huu, V.N. Ngoc. (2021). Analysis Study of Current Transportation Status in Vietnam’s Urban Traffic and The Transition to Electric Two-wheelers Mobility. *Sustainability*, 13(10), 5577. https://doi.org/10.3390/su13105577

[3] T.V. Mathew, P. Radhakrishnan. (2010). Calibration of Microsimulation Models for Nonlane-based Heterogeneous Traffic at Signalized Intersections. *Journal of Urban*

Planning and Development, 136(1), 59-66. https://doi.org/10.1061/(ASCE)0733-9488(2010)136:1(59)

[4] L. Ambarwati., A.J. Pel, R. Verhaeghe, B. Arem. (2014). Empirical Analysis of Heterogeneous Traffic Flow and Calibration of Porous Flow Model. *Transportation Research Part C: Emerging Technologies*, 48, 418-436. https://doi.org/10.1016/j.trc.2014.09.017

[5] Ž. Majstorović, L. Tišljarić, E. Ivanjko, T. Carić. (2023). Urban Traffic Signal Control Under Mixed Traffic Flows: Literature review. *Applied Sciences*, 13(7), 4484. https://doi.org/10.3390/app13074484

[6] D. Leitner, P. Meleby, L. Miao. (2022). Recent Advances in Traffic Signal Performance Evaluation. *Journal of Traffic and Transportation Engineering (English Edition)*, 9(4), 507-531. https://doi.org/10.1016/j.jtte.2022.06.002

[7] Z. Marszalek, K. Duda. (2024). Validation of Multi-Frequency Inductive-Loop Measurement System for Parameters of Moving Vehicle Based on Laboratory Model. *Sensors*, 24(22), 7244. https://doi.org/10.3390/s24227244

[8] T. Vu, H.N. Thai, V.N. Pham, H.T. Vu, A.T. Luong, T.V. Luong. (2025). Counting Mixed Traffic Volumes at Motorcycle-Dominated Intersections by Using Computer Vision. *International Journal of Intelligent Transportation Systems Research*, 23, 146-164. https://doi.org/10.1007/s13177-024-00442-z

[9] S. Hadi, Khairurrasyid. (2024). Performance Analysis of Unsignalized Intersection Using PTV VISSIM Software Modeling (Case Study of Sakra 4-way Intersection, East Lombok). *IOP Conference Series: Earth and Environmental Science, 5th International Conference on Coastal and Delta Areas, 1321, 012027. IOP Publishing*. doi:10.1088/1755-1315/1321/1/012027

[10] B. Park, J. Schneeberger. (2003). Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System. *Transportation Research Record*, 1856(1), 185-192. https://doi.org/10.3141/1856-20

[11] A. Charef, Z. Jarir, M. Quafafou. (2024). The Impact of Motorcycle Positioning on Start-up Lost Time: The Empirical Case Study of Signalized Intersections in Marrakech Using VISSIM. *Engineering, Technology & Applied Science Research*, 14(3), 14313-14318. https://doi.org/10.48084/etasr.7141

[12] N.F. Paiman, A. Hamzah, M.S. Solah, A.H. Ariffin, M.S.A. Khalid, K.A.A. Kassim, S.Z. Ishak, H. Imanaga, H. Ishida. (2020). Motorcycle Positioning in Queues at Signalized Intersections in City of Klang Valley. *Jurnal Kejuruteraan SI*, 3(1), 89-93. DOI:10.17576/jkukm-2020-si3(1)-14

[13] X.-C. Vuong, R.-F. Mou, H.-S Nguyen, T.-T. Vu. (2019). Signal Timing Optimization of Isolated Intersection for Mixed Traffic Flow in Hanoi City of Vietnam using VISSIM. *International Conference on Smart Vehicular Technology, Transportation, Communication and Applications. Springer, pp. 133-139.* https://doi.org/10.1007/978-3-030-04582-1_15

[14] L. Tang, Z. Kan, X. Zhang, X. Yang, F. Huang, Q. Li. (2016). Travel Time Estimation at Intersections Based on Low-frequency Spatial-temporal GPS Trajectory Big Data. *Cartography and Geographic Information Science*, 43(5), 417-426. https://doi.org/10.1080/15230406.2015.1130649

[15] X.C. Vuong, R.-F. Mou, T.T. Vu, H.V. Nguyen. (2021). An Adaptive Method for An Isolated Intersection Under Mixed Traffic Conditions in Hanoi Based on ANFIS Using VISSIM-MATLAB. *IEEE Access*, 9, 166328-166338. doi: 10.1109/ACCESS.2021.3135418

[16] B. Roy, S.A. Suma, MD. Hadiuzzaman, S. Barua, SK. Mashrur. (2021). Optimization of Delay Time at Signalized Intersections Using

Direction-Wise Dynamic PCE Value. *International Journal of Transportation Engineering*, 8(3), 279-298. 10.22119/ijte.2020.225672.1514

[17] N.Y. Cao, K. Sano. (2012). Estimating Capacity and Motorcycle Equivalent Units on Urban Roads in Hanoi, Vietnam. *Journal of Transportation Engineering*, 138(6), 776-785. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000382

[18] E.G. Talbi. (2009). Metaheuristics: From Design to Implementation. *John Wiley & Sons*.

[19] J.K. Lenstra, E. Aarts. (2018). Local Search in Combinatorial Optimization. *Princeton University Press.*

[20] R. Mart, P.M. Pardalos, M.G. Resende. (2018). Handbook of Heuristics. *Springer Cham.*

[21] J. Lee, B. Park, I. Yun. (2013). Cumulative Travel-time Responsive Real-time Intersection Control Algorithm in The Connected Vehicle Environment. *Journal of Transportation Engineering*, 139(10), 1020-1029. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000587

[22] F. Kurniawan, H. Agustian, D. Dermawan, R. Nurdin, N. Ahmadi, O. Dinaryanto. (2025). Hybrid Rule-Based and Reinforcement Learning for Urban Signal Control in Developing Cities: A Systematic Literature Review and Practice Recommendations for Indonesia. *Applied Sciences*, 15(19), 10761. https://doi.org/10.3390/app151910761

[23] T. Vu, J. Preston. (2023). Microscopic Simulation Model for Motorcycle Dominated Networks: A Case Study of a VISSIM Simulation Model for a Mixed Traffic Corridor. *IOP Conference Series: Materials Science and Engineering. IOP Publishing*, 1289, 012045. doi:10.1088/1757-899X/1289/1/012045

[24] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, X. Wang. (2022). ByteTrack: Multi-object Tracking by Associating Every Detection Box. *Computer Vision – ECCV 2022. Springer Cham, pp. 1-21.*