



Localized AI-Based Urban Hazard Detection for Defensive Driving Support

Seung Hwan Kang¹ and Jira Yammeesri²

ABSTRACT

This research investigates the feasibility and effectiveness of implementing localized artificial intelligence-based system for enhancing defensive driving through automated urban hazard detection. For this purpose, a lightweight You Only Look Once version 11 nano (YOLOv11-nano) model containing 2.6 million parameters was fine-tuned on a localized dataset consisting of 18,132 images collected over 300 days in Chiang Mai, Thailand. These results reveal a mean Average Precision (mAP) of 0.96 when using an intersection over union (IoU) threshold of 0.5. Even if the precision rate is quite high in ideal conditions, there is an observed decrease of 13.1 percent in the Average Detection Confidence due to monsoon rain and night-time conditions. Thus, a safety margin has been established for the application of Advanced Driver Assistance Systems. The reliability test conducted across seven localized object categories reveals consistent detection performance for vulnerable objects, such as motorcycles and pedestrians. Furthermore, a cost-benefit analysis based on localized accident cost data confirmed the economic viability of implementation in commercial transporting services.

Article information:

Keywords: Localized AI, Defensive Driving, Deep Learning, Computer Vision

Article history:

Received: January 27, 2026

Revised: February 12, 2026

Accepted: July 9, 2026

Published: July 31, 2026

(Online)

DOI: 10.37936/ecti-cit.2026203.265931

1. INTRODUCTION

The increasing rate of urbanization in the South-east Asia has consequently created complex traffic ecosystems in medium-scale cities like Chiang Mai in Thailand. The conditions in these settings include high traffic density and frequent interactions between large vehicles and vulnerable road users, including motorcycles and pedestrians. There are also climatic fluctuations in varying seasons. As pointed out by Nguyen *et al.* [1], urban traffic exhibits high heterogeneity in congestion patterns, creating complex spatiotemporal archetypes that are difficult to categorize. In medium-scale cities like Chiang Mai, this is further complicated by a long tail of environmental edge cases [2], for instance, sudden monsoon storms, are usually underrepresented in standard datasets. These conditions add significantly to the cognitive workload faced by the human driver, and the driver may face difficulties with Level 1 Situation Awareness [3], which is the ability to detect and recognize relevant features within the driving environment.

The need for localized ADAS (Advanced Driver Assistance Systems) relates to the realization that perception models designed for global training tend to generalize poorly, if at all, to a particular area's unique traffic patterns. For instance, in Chiang Mai, such conditions include a preponderance of motorcyclists, together with predictable flooding due to heavy monsoons, both of which are difficult for vision perception algorithms to handle. Although the YOLOv11 model's recent release toward the end of 2024, as reported by Jocher *et al.* [4] offers a far more efficient extraction process for its distinguishing features due to the incorporation of two modules specifically designed for this purpose, the C3k2 (Cross Stage Partial Bottleneck with 2 Convolutions) module, together with the C2PSA module (CSP with Programmable Sequence Attention), its effectiveness in high-noise environments such as localized urban traffic conditions remains an area for active research. This limitation is highlighted by Bhowmik [5], who suggests using special training for achieving high-fidelity object detection under adverse visual condi-

¹The author is with the Department of Information Technology, Payap University, Chiangmai, Thailand, Email: seung_h@payap.ac.th

²The author is with the Department of International Business Management, Payap University, Chiangmai, Thailand, Email: jira_y@payap.ac.th

²Corresponding author: jira_y@payap.ac.th

tions, including rain-soaked and other challenging environments.

In addition to collision avoidance systems, the effectiveness of other ADAS technologies has also been associated with digital healthcare and road environment factors. The effectiveness of warning systems and their ability to minimize false alarms have been shown to alleviate long-term cognitive workload and promote drivers' occupational sustainability. Recent research by Aleksa *et al.* [6] implies that the deployment of the system has significant potential for reducing collisions and that this potential could support their wider adoption of these systems in developing urban infrastructure. Nevertheless, the effectiveness of the systems under adverse atmospheric conditions, such as rain streaks commonly encountered in tropical environments, remains critically important, as highlighted by Wang and Zhao [7].

This study addresses this research gap by exploring the performance of a lightweight object detection model in a realistic local setting. While the research rather than focusing solely on maximizing detection performance, it considers environmental conditions with a scenario-based economic analysis. This aims to provide a more comprehensive assessment of the value and current limitations of local AI-enabled hazard detection in the context of city driving.

2. LITERATURE REVIEW

Rapid urbanization across Southeast Asia results in increasingly complex traffic ecosystems in medium-sized cities such as Chiang Mai, Thailand. Such environments are characterized by high density, heterogeneous road users, and frequent interactions between large vehicles and vulnerable road users-including motorcycles and pedestrians. Nguyen *et al.* [1] define this type of heterogeneous traffic, which results in a long tail of edge cases that are inadequately represented in typical Western-centric standard datasets. Seasonal variability in the environment further exacerbates this challenge, with heavy monsoon rainfall and reduced visibility presenting some of the toughest driving conditions. Altogether, these factors significantly increase cognitive load on human drivers, which in turn contributes not only to higher accident risks but also to long-term stress and fatigue.

ADAS has attracted extensive research attention to enable better road safety outcomes by leveraging automated perception and warning functions. More recently, there has been an emerging trend toward using convolutional neural network-based models for object detection tasks, such as the popular YOLO series, which offers real-time processing and an efficient trade-off between model accuracy and computational complexity. The recent introduction of the YOLO11 model in late 2024, explained in detail by Jocher *et al.* [4], introduce an improved feature extraction system design through its C3k2 and

C2PSA components, which are intended to improve performance in densely congested urban road conditions. The YOLO-series models have been successful in highways and organized city-road conditions when trained using large-scale benchmark databases such as COCO (Common Objects in Context) or KITTI (Karlsruhe Institute of Technology and Toyota Technological Institute). However, their current training databases do not adequately represent the road user composition or season-specific conditions found in other parts of the world.

Many studies have shown that the performance of object detection can deteriorate drastically under weather conditions such as heavy rainfall, nighttime illumination, glare, and motion blur caused by traffic movement [8-10]. It is particularly noted that object detection under adversarial vision conditions, such as heavy rainfall or occlusion, cannot be performed effectively without detailed tuning and corresponding datasets, even when using modern models [11]. Camera-based perception models are more susceptible to the conditions mentioned above, which can deteriorate detection accuracy and increase likelihood of false negatives for small-sized or partially occluded targets. More recently, the findings of Wang and Zhao [7] have shown that heavy rainfall can cause fundamental distortions on textures and edges that are essential for object detection. Besides the above views on camera models, unbalanced datasets in traffic models will continue to be a problem, especially in the case of motorcycles and pedestrians that could often remain under-represented in Southeast Asia. However, their road safety prominence is high.

Beyond accident prevention, the performance of ADAS shows relevance to driver workload and occupational health. It has been shown that excessive false alarms and inconsistent warning behavior increase the cognitive burden and reduce user trust in assistance systems. On the other hand, warning systems that achieve high recall for genuinely hazardous situations and suppress non-critical alerts can reduce decision fatigue and stress during prolonged driving tasks. Recent research by Aleksa *et al.* [6] suggests that ADAS deployment has significant crash reduction potential, which can be quantified to justify the economic adoption of these systems in developing urban environments. Prior studies have also linked reduced cognitive load to improved driver comfort and well-being over longer periods. While direct physiological measurements are beyond the scope of this work, these findings motivate investigating hazard detection accuracy as an indirect contributor to sustainable driving practices.

Although there has been increasing interest in AI-based traffic perception technology, relatively little research has systematically examined localized and environment-aware assessments of detection dependability with respect to region-specific factors such as

monsoon rain, nighttime urban lighting, and a high density of motorcycles. Specifically, little empirical research has examined the relationship between class-level detection performance and proximity-based hazard logic, together with their implications for safety. This research addresses this gap by examining the performance of a localized AI system that fine-tuned a YOLO-based object detection system using real-world dashcam videos collected in Chiang Mai. In addition to the technical assessment, the research also includes a strong focus on economic assessment.

3. METHODOLOGY

The methodology covers five main components: data acquisition, object detection, model architecture and training protocol, the design of a single imminent hazard warning mechanism, and scenario-based economic modeling.

3.1 Data Acquisition and Corpus Description

To build the localized dataset of images, 3,213 continuous dashcam footage segments, each five minutes long, were collected over 300 days of driving in Chiang Mai, Thailand. However, for privacy and security reasons, images of residential areas, school zones, places of worship, military zones, airport zones, and hospital campuses were excluded. Accordingly, this process adhered not only to all requirements of ethics but also to relevant standards. In addition, any image containing identifiable material, such as faces or license plates, were anonymized using blur or masking software. The raw images were also stored on password-protected drives. Among these raw images, 18,132 images were collected after applying temporal sampling and removing duplicates. In this study, safety-critical entities were considered to be the most common road users in the area. Therefore, seven classes were manually annotated: car, truck, bus, motorcycle, bicycle, person, and van. All occurrences of the above-mentioned classes were annotated regardless of their initial distance to the vehicle. However, objects that were too small to be distinguished or severely fragmented by occlusion were excluded from the final annotation set because they lacked sufficient semantic features for model supervision. For model training, the dataset was split into a training set (70%; 12,692 images), a validation set (20%; 3,626 images), and a testing set (10%; 1,814 images) to support reliable model evaluation.

Fig. 1 shows some representative annotated samples from the collected dataset. The figure illustrates the variety of environmental conditions and challenges associated with detection tasks during real-world urban driving in Chiang Mai, Thailand. First, the scenes are divided into images captured under daylight, nighttime driving, and heavy rainfall conditions during the monsoon periods, representing varying illumination and weather conditions.



Fig.1: Example of manually annotated seven localized object classes relevant to urban traffic safety.

Table 1: Dataset Distribution by Environmental Condition.

Environmental Condition	No. of Images	Percentage
Clear-Day	4,533	25%
Clear-Night	4,533	25%
Rainy-Day	4,533	25%
Rainy-Night	4,533	25%
Total	18,132	100%

The daytime images generally have high visibility with well-defined object boundaries, enabling effective annotation of all seven traffic-related classes, namely car, truck, bus, motorcycle, bicycle, person, and van. In such scenes, road users tend to be distinguishable based on shape, scale, and color contrast, thus providing strong supervision signals during training.

Nighttime samples introduce reduced lighting conditions, headlight glare, and motion blur, which obscure object contours and reduce texture information. Under these conditions, pedestrians and bicycles are especially difficult to annotate due to partial occlusion and low contrast against dark backgrounds. Such environments align with what Bhowmik [5] describes

adversarial vision conditions, where environmental noise can fundamentally impair sensor fidelity. Despite the difficulties posed by these samples, the annotations were retained because realistic deployment conditions must be represented for urban hazard detection systems.

Rainfall conditions represent the most complex examples with respect to visibility and are characterized by heavy rainfall, resulting in rain streaks, reflections on wet roads, and reduced visibility caused by water droplets on the camera lens. In these examples, motorcycles and other small vehicles may be partially hidden or confused with background clutter, resulting in potential false or missed detections. According to Wang and Zhao [7], atmospheric distortions such as rainfall cause edge artifacts by introducing high-frequency components into images that do not correspond to actual object boundaries. To quantitatively assess the impact of such distortions, the Average Detection Confidence is used in this study, defined as the average confidence score for all true positive detections in the rainfall set. Although a reduction in this value is expected under such conditions, recording the actual difference between clear and the rainy conditions is important for establishing safety thresholds for localized ADAS.

Furthermore, some images contain ambiguous boundaries for classification, for example, when vans may resemble small trucks or when motorcycles may be partially hidden behind larger vehicles. Ambiguity in classification indicates a natural complexity associated with fine-grained semantic tasks, especially in dense urban traffic flows. Regarding safety considerations, such complexities may not significantly affect the proximity logic used to identify hazardous areas.

Overall, the annotated samples represented in this dataset show both clear and difficult detection scenes, ensuring that model evaluations reflect real-world operational constraints rather than curated or idealized conditions.

3.2 Object Detection

Among the different computer vision tasks, object detection is a major task that not only detects but also locates the objects of interest in an image or video. As shown in Fig. 2, computer vision problems are divided into three categories: classification, classification with localization, and object detection. Classification involves assigning a single label to an image. Classification with localization involves assigning a label to an image while also localizing a single instance. Object detection, the most complex task, involves assigning class labels and bounding box coordinates to multiple instances with an image.

For this research, the YOLO11n (nano) model was chosen because of its efficiency in real-time defensive driving assistance. The YOLO11n model is a state-of-the-art lightweight model with about 2.6 mil-

lion parameters and a computational complexity of 6.5 GFLOPs. Although the larger YOLO11m model would achieve a slightly better mean Average Precision (mAP), the YOLO11n model was chosen to ensure that the system could process frames quickly enough to provide timely hazard notifications, representing an important trade-off in defensive driving.



Fig.2: Overview of computer vision tasks: classification, classification with localization, and object detection.

For the dataset, 70% (12,692 images) was used for training, 20% (3,626 images) for validation, and 10% (1,814 images) for testing. To test the robustness of the model in difficult scenarios, the test set was balanced to include an equal number of images from four environmental conditions: clear day, clear night, rainy day, and rainy night. This ensured that the final performance metrics (0.96 mAP) was not biased toward high-visibility daytime images but instead provided a conservative estimate of the model's performance under real-world scenarios.



Fig.3: Sample images from the training and test subsets captured under adverse environments (clear, rainfall, and low-light conditions).



Fig.4: Sample images from the validation subset captured under adverse environments (clear, rainfall, and low-light conditions).

3.3 Model Architecture and Training Protocol

The object detection algorithm is based on the YOLO11n model, the lightweight version of the YOLO architecture introduced at the end of 2024.

According to Jocher *et al.* [4], the model uses the C3k2 and the C2PSA modules to improve the extraction of features required for real-time edge inference. Transfer learning is used to adapt the general features from large-scale datasets to the local traffic environment.

Fine-tuning was carried out using the original YOLO11n training configuration, with Stochastic Gradient Descent as the optimizer and a total of 60 training epochs. The input size was set to a fixed resolution of 640 x 640 pixels, the batch size was 16, and a cosine annealing learning rate schedule was used throughout training. For testing, a confidence threshold of 0.5 was applied to detections in all experiments. Testing was performed on a Tesla T4 Graphics Processing Unit (GPU).

3.4 Single Imminent Hazard Alert Logic

To minimize driver distraction and alleviate false-positive fatigue, the system uses a single imminent hazard warning instead of a series of graduated alerts. The system estimates the proximity of a hazard to the vehicle using two geometric cues derived from object detection signals: bounding box size and the object's vertical position in the image.

An imminent danger alert is triggered if any target satisfies at least one of the following conditions:

- Proximity by Size: the height of the object's bounding box is greater than or equal to 180 pixels;
- or
- Proximity by Position: the vertical center of the bounding box lies within the lowest 25% of the image frame ($\text{center-y} \geq 0.75H$, where H is the image height).

This set of rules is applied equally to all seven detection classes. It is important to note that, although larger vehicles such as trucks and buses have larger physical profile, their lower boundary, where the vehicle contacts the road, enters the lower 25% of the image ($0.75H$) as they approach the host vehicle. Therefore, the Proximity by Position trigger is also reliable for different types of vehicles. In addition, these geometric conditions remain constant for all environmental conditions (clear day, clear night, and monsoon). While adverse weather conditions may reduce the Average Detection Confidence, or mean confidence score, assigned to true-positive detections, the underlying decision logic remains deterministic. This ensures that the system continuously provides a predictable safety margin, thereby helping to prevent the catastrophic breakdown of Level 1 Situation Awareness under adverse atmospheric conditions.

This binary alert system is designed to maintain simplicity, enhancing the clarity and speed of the driver's response by drawing driver's attention only to objects that are likely to pose an immediate risk. This is consistent with the hazard perception theory

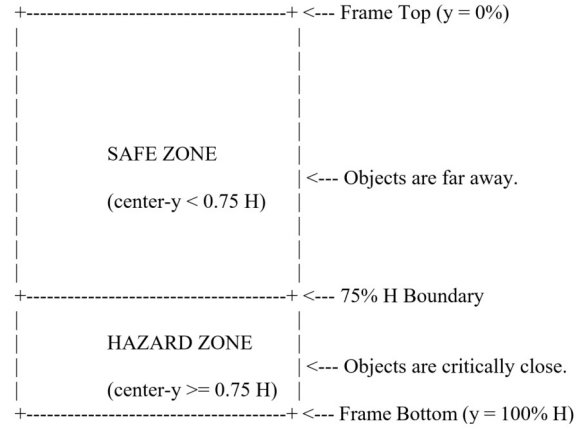


Fig.5: Single imminent hazard alert logic.

proposed by Horswill and McKenna [12], which identifies situational awareness as a key factor in the reducing driver reaction times.

3.5 Economic Modeling

The scenarios for economic analysis were developed and applied to identify the potential economic impact of reduced minor accidents involving road users in a localized setting. This economic impact analysis uses a Cost-Benefit Analysis (CBA) approach to evaluate the viability of implementing an AI system in a small logistics business with a baseline of 25 minor accidents per year.

The annual economic savings (S) from the AI system use can be estimated using the following equation:

$$S = (N \times R) \times C$$

where N is the average baseline accident frequency (25 accidents per year). R is the hypothetical rate of accident reduction due to the localized ADAS assistance (10%). C is the unit cost of a minor accident.

The value used to adjust the cost incurred due to accidents is 50,600 Thai Baht per accident, as per the localized economic standards set by Thongchim *et al.* [12]. Rather than proving at a definitive answer, this model aims to examine the sensitivity of operating costs to a 10% reduction in accidents, in line with the safety reduction potential framework for ADAS proposed by Aleksa *et al.* [6]. This approach offers an academic framework for assessing the economic feasibility of localized AI systems against the actual cost of saving [12], extending the evaluation beyond technical correctness to include socioeconomic relevance.

4. RESULTS

This section presents the performance evaluation of the localized object detection model, along with an

analysis of its environmental sensitivity and financial implications.

4.1 Training Convergence and Optimization Stability

Figures 6-7 present the training (train), validation (val), bounding box loss (box_loss), classification loss (cls_loss), and distribution focal loss (dfl_loss) all over 60 epochs. Three losses are presented, all of which converge monotonically without any indication of abnormality.

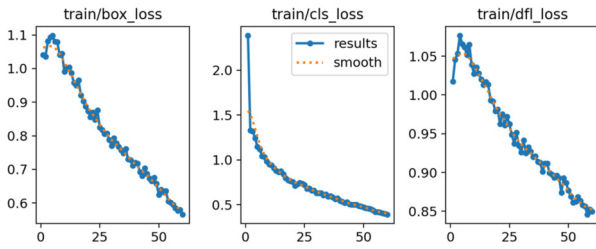


Fig.6: Training losses (Box, Classification, and DFL) across 60 epochs.

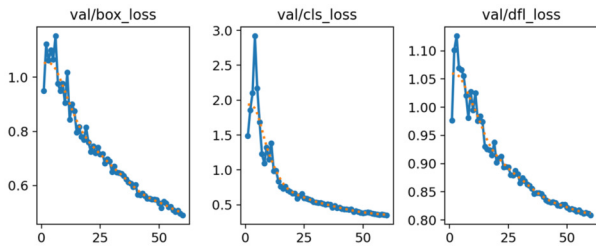


Fig.7: Validation losses (Box, Classification, and DFL) across 60 epochs.

The training box loss gradually reduces from above 1.0 during the initial periods of the training process to a steady value of about 0.56 from epoch 60. On the other hand, the validation box loss also decreases steadily. Moreover, the training and validation classification losses also decrease steadily. For the training classification loss, the loss decreases steadily and finally settle at a value of about 0.39. Similarly, the validation classification loss also decreases steadily and settles at a value of approximately 0.35. The distribution focal loss also decreases steadily and settles at a stable value.

4.2 Detection Accuracy and Global Performance Metrics

Figure 8 illustrates a general trend in metrics/precision (B) and metrics/recall (B) during training. The precision rises steeply in early epochs and levels off after the fourth epoch at a value greater than 0.89. This implies a low false positive rate when using a confidence threshold of 0.5. The recall rate is

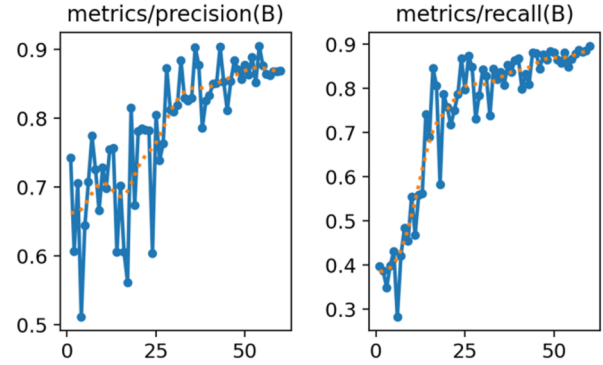


Fig.8: Precision and recall (B) evolution showing training stability.

almost similar but stabilizes close to 0.90 after a few epochs.

As shown in Figure 9, overall detection performance, measured by mAP50, reaches approximately 0.96 on the localized dataset, whereas mAP50-95 converges to a value above 0.93. The small difference between mAP50 and mAP50-95 values indicates consistent detection performance under stricter Intersection over Union (IoU) thresholds and validates accurate bounding box localization rather than coarse localization.

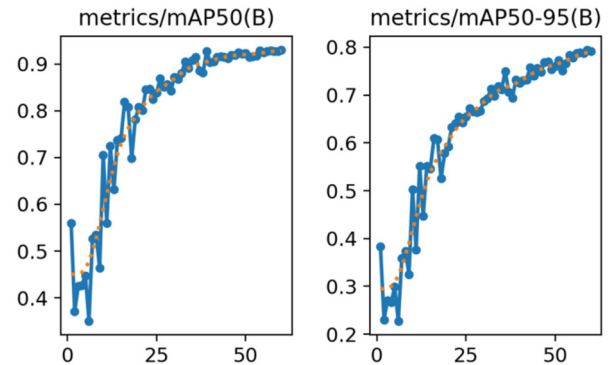


Fig.9: Global evaluation metrics for mAP@0.5 and mAP@0.5:0.95.

4.3 Class-Level Performance and Confusion Matrix Analysis

Figure 10 shows the normalized confusion matrix for the test set, which enabling class-wise recall analysis. The dominance diagonal values across most classes indicates strong detection performance for the major road users, including cars, trucks, buses, vans, and bicycles.

In particular, motorcycle detection, previously identified as a critical weakness, demonstrates a significantly improvement in recall performance. The reduction in false negatives for motorcycle detection helps address a significant safety concern in Southeast Asian traffic environments. Pedestrian (person) de-

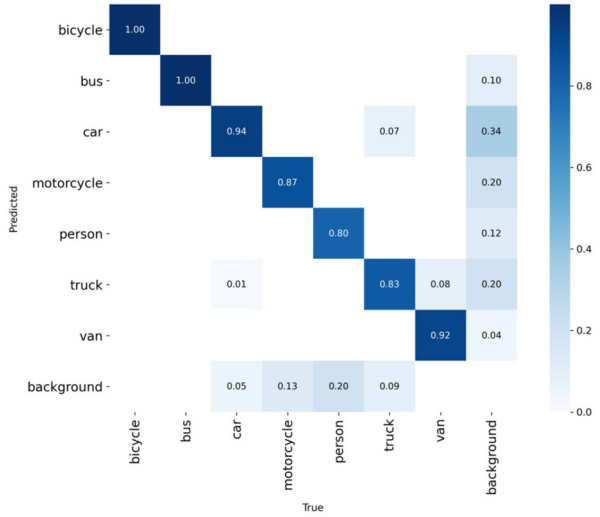


Fig.10: The normalized confusion matrix for localized road-user classification.

tection achieves a recall value of approximately 0.80, which is lower than of the larger vehicle classes but remains acceptable due to the high variability in pedestrian scale, pose, partial occlusion, and lighting conditions present in urban traffic scenes.

4.4 Diagnostic Curve Analysis and Operating-Point Selection

Figure 11 presents the F1-confidence curve aggregated across all classes. The curve shows a maximum F1 score of approximately 0.87 at a confidence threshold of approximately 0.43. This indicates an optimal balance between precision and recall under localized urban traffic conditions.

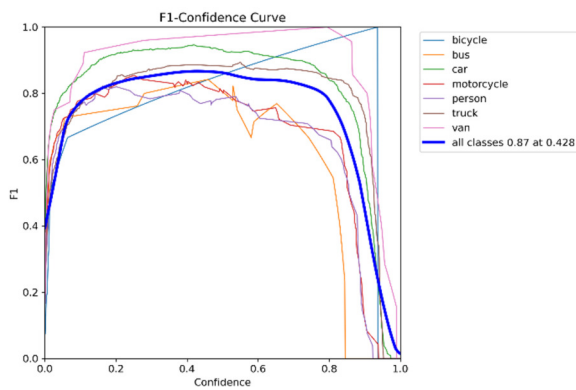


Fig.11: F1-confidence curve.

A confidence threshold of 0.5 was chosen for deployment because it offers a good trade-off between precision and recall. Class-wise analysis shows that motorcycles and pedestrians maintain stable F1 scores across a wide range of confidence thresholds, indicating improved robustness compared to earlier model iterations.

4.5 Implications for Hazard-Zone Proximity Logic

The high rate of recall in motorcycles and vehicular detection from the lower regions of images helps to affirm the consistency and results in validating the imminent hazard alert system's logic itself: an imminent hazard alert will be sent if any object will meet at least one of two stipulations: proximity using size if height of an object's bounding box ≥ 180 pixels in height, and proximity using position if center-y coordinates $\geq 0.75 H$, or center-y falls in the 25% of image frame from the bottom. Once an imminent hazard alert is dispatched, a beep sound will come from the system, and a red bounding box will be drawn around the object to immediately direct a driver's attention to an imminent hazard situation.

Because these geometric triggers, rather than fine-grained semantic distinctions, form the basis of the warning mechanism, residual inter-class confusions do not seriously compromise system-level safety. Provided the model continues to return high recall for any object entering these critical zones, the red boxing and audible alert remain consistent, ensuring notification to the driver of immediate proximity, whether a vehicle is classified as a car, van, or truck. This design focuses on detecting hazard presence rather than achieving absolute classification accuracy in time-critical defensive driving scenarios.



Fig.12: Implemented Hazard-Zone Proximity Logic.

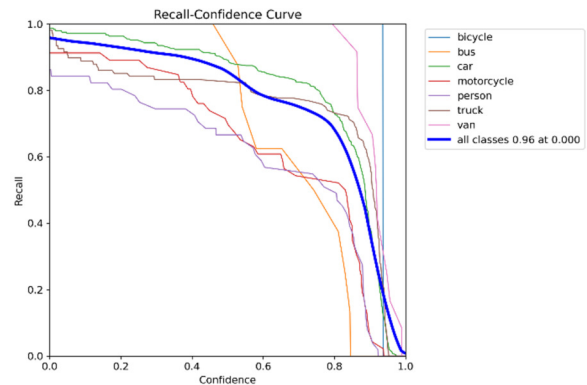


Fig.13: Recall-confidence curve.

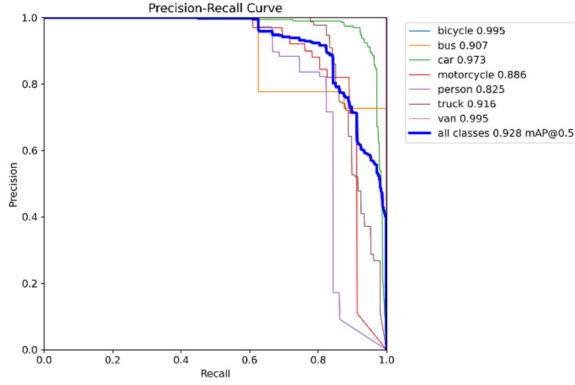


Fig.14: Precision-recall curve.

4.6 Environmental Sensitivity

Although the model's accuracy has improved significantly, some sensitivity to adverse environmental conditions has been observed, which is a common limitation of camera-based models. To measure such an impact, the current research utilizes Average Detection Confidence (ADC), defined as the arithmetic mean of the confidence scores assigned by the model to all true-positive detections within a specific environmental category.

$$ADC = \frac{1}{N} \sum_{i=1}^N c_i$$

where c_i , is the confidence score assigned to the i^{th} true-positive detection, and N is the total number of true-positive detections in the evaluated category. It should be emphasized that ADC is not equal to mean Average Precision (mAP). While mAP measures overall detection performance based on precision-recall curves, ADC evaluates the model's confidence in its true-positions under specific environment conditions.

As outlined in Table 2, the average confidence for detection under clear daytime conditions is 0.84, representing an ideal situation for the cameras due to sufficient illumination levels, minimal glare, and a favorable contrast ratio. Similarly, under clear night conditions, the detection confidence remains high at 0.86.

Table 2: HAISE security dimensions, weights, maximum scoring levels, and grounding standards.

Condition	Average Detection Confidence
Clear-Day	0.84
Clear-Night	0.86
Rainy-Day	0.82
Rainy-Night	0.73



Fig.15: Detection confidence under a clear-day condition.

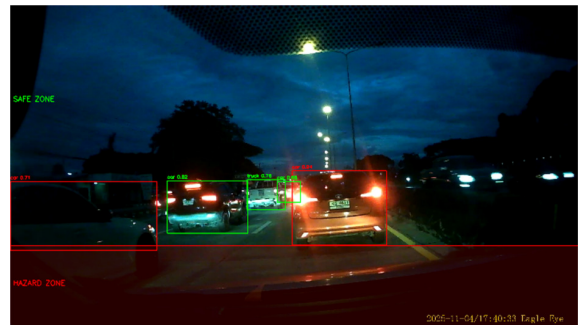


Fig.16: Detection confidence under a clear-night condition.



Fig.17: Detection confidence under a rainy-day condition.

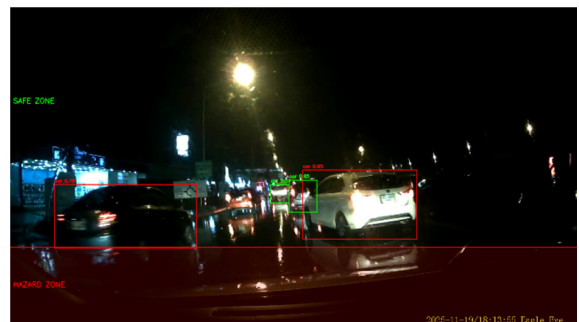


Fig.18: Detection confidence under a rainy-night condition.

Moreover, rainy weather contributes to reduced detection performance. As rainy scenarios during the daytime involve additional visual challenges, the average confidence decreases to 0.82. This reduction is attributed to interference caused by rain streaks and reflection from the wet road surface. Furthermore, the performance is significantly affected under rainy nighttime conditions, where the average confidence decreases further to approximately 0.73. This represents almost 14 reductions in performance compared with the image captures under clear daylight conditions.

In terms of safety, the degradation in performance is important to consider, as weather conditions often coincide with increased potential danger. Nevertheless, it is relevant to note that even under rain-fall and nighttime conditions, the model maintains confidence levels above 0.73. This indicates that the model remains functional when objects are closer and when hazard warnings are required. The discussion on functionality is important because, considering the proposed hazard zone, the system relies on object proximity rather than object class for decision making.

These findings reinforce the need to explicitly account for environmental conditions when interpreting system performance. They also support recommendations for future architectural enhancements, such as image denoising networks or multi-sensor fusion, to mitigate performance degradations caused by adverse weather and illumination conditions. Furthermore, these findings support the conservative positioning of the system as a decision-support tool rather than a fully autonomous safety mechanism.

4.7 Baseline Model Comparison Under Environmental Conditions

The following tables, from 3 to 6, show a comparative performance analysis of the proposed localized model against the two baseline architectures, YOLOv8n and YOLO11n. These baseline models were selected because they represent recent generations of real-time object detectors designed for edge deployment and intelligent transportation applications. All models were evaluated under identical environmental conditions using a fixed confidence threshold ($\text{conf} = 0.5$) and standardized 9,000-frame video segments to ensure a fair comparison.

To provide a stronger experimental validation, the proposed localized model was compared against two widely adopted state-of-the-art lightweight object detection architectures, YOLOv8n and YOLO11n. These baseline models were selected because they represent recent generations of real-time object detectors designed for edge deployment and intelligent transportation applications. All models were tested under the same environment conditions, confidence thresholds, and test video segments to ensure a fair com-

parison.

It follows that, under clear daytime conditions, Table 3 represents the highest number of total detections, proposed with 30,045, and hazard-frame activations, at 5,591, exceeding both YOLOv8n and YOLO11n. In addition, the Average Detection Confidence of 0.84 more clearly explains that object recognition under well-lit conditions is quite stable and reliable.

Performance differences under night conditions become even greater. The proposed model identifies substantially more hazard frames, up to 8,784, than YOLOv8n (6,809) and YOLO11n (6,979), with a higher average confidence of 0.86. Greater sensitivity may be observed, considering that low-light conditions commonly degrade the performance of globally trained models.

For rainy daytime scenarios (Table 5), the proposed model again outperformed the compared baselines by successfully producing more than twice the total number of detections achieved by YOLOv8n and YOLO11n. Also, the hazard frames are greater (8,432) due to the increased reliability of the detection method in highly visually contaminated scenarios caused by the rain.

The largest disparity is observed under rainy nighttime conditions (Table 6), where visual degradation is the most severe. This proposed method identifies nearly three times as many hazard frames as the current YOLOv8n and YOLO11n models while maintaining a relatively high average confidence level of 0.73. In contrast, YOLOv8n and YOLO11n exhibit a substantial reduction in detection volume and hazard coverage, indicating a clear limitation of the models in handling combined gradation factors.

Table 3: Baseline model comparison under clear-day conditions.

	Our Model	Yolov8n	Yolo11n
Frames processed	9,000	9,000	9,000
Hazard frames	5,591	5,426	5,478
Total detections	30,045	17,581	18,205
Average confidence	0.84	0.65	0.66

Overall, the results clearly indicate that the proposed localization-aware fine-tuning significantly improves detection robustness for various environmental conditions, especially those that are underrepresented in globally aggregated datasets.

However, although the proposed model produces a higher number of detections, this behavior is consistent with the design objectives of proximity-based hazard warning systems. In safety-critical driving contexts, false negatives pose a greater risk than con-

Table 4: Baseline model comparison under clear-night conditions.

	Our Model	Yolov8n	Yolo11n
Frames processed	9,000	9,000	9,000
Hazard frames	8,784	6,809	6,979
Total detections	37,500	15,368	16,160
Average confidence	0.86	0.69	0.69

Table 5: Baseline model comparison under rainy-day conditions.

	Our Model	Yolov8n	Yolo11n
Frames processed	9,000	9,000	9,000
Hazard frames	8,432	7,639	7,382
Total detections	52,446	23,038	22,173
Average confidence	0.82	0.70	0.69

Table 6: Baseline model comparison under rainy-night conditions.

	Our Model	Yolov8n	Yolo11n
Frames processed	9,000	9,000	9,000
Hazard frames	1,443	458	492
Total detections	7,881	745	1,014
Average confidence	0.73	0.66	0.64

servative over-alerting. Therefore, this increase in hazard-frame coverage is considered an acceptable and desirable trade-off for minimizing missed detections of nearby road users.

4.8 Financial and Economic Impact

Basis on the evaluation framework outlined in Section 3.5, the direct annual savings (S) for a typical small-scale logistics system are calculated using the cost-benefit parameters developed as follows:

$$\begin{aligned}
 S &= (N \times R) \times C \\
 S &= (25 \text{ case} \times 0.10) \times 50,600 \text{ Thai Baht} \\
 S &= 126,500 \text{ Thai Baht per year}
 \end{aligned}$$

The above analysis is based on the safety impact assessment framework outlined by Aleksa *et al.* [3], which links accident frequency reduction to unit cost per accident to evaluate the socioeconomic feasibility

of ADAS solutions. By adopting the localized cost value of 50,600 THB, as reported by Thongchim *et al.* [12], the analysis moves beyond generic cost assumptions to better align with the prevailing economic conditions of the Thai transportation industry.

Although the 10% reduction rate (R) represents a projected baseline, it indicates significant potential for ROI. The estimated ROI savings for small-to medium-sized logistical companies in Thailand suggest that the investment in lightweight edge-computing hardware and YOLO11n integration could be recovered within a relatively short operational period. This further supports the conclusion that edge-based AI represents not only technological advancement but also a financially feasible safety solution for urban fleet management.

5. DISCUSSION

This section describes the experimental findings based on various performance metrics and extracts practical implications on utilizing these metrics within hazard detection and defensive driving. Further, various technical challenges, architectural refinements, and ethical and socioeconomic implications of such localized AI systems within transportation contexts are discussed.

5.1 Technical Interpretation and Limitations

A mean Average Precision (mAP@0.5) verifies that the localized object detection model provides a strong baseline for urban hazard perception. Specifically, the proposed model configuration significantly enhances the reliability of motorcycle detection and relieves what was previously considered a critical safety concern in the Chiang Mai traffic environment. The normalized confusion matrix reveals that motorcycle recall is no longer disproportionately low; in other words, the adopted localization strategy effectively captured regional vehicle characteristics suggested by Nguyen *et al.* [1]. While the residual misclassifications remain concentrated among visually similar classes, such as between cars and trucks, this represents a much lower risk than failing to detect a vulnerable road user altogether.

Comparison analysis with the state-of-the-art lightweight object detection models YOLOv8n and YOLO11n demonstrate that the introduced localized fine-tuning approach leads to tangible performance enhancement for object detection in various environmental settings. Specifically, the introduced model demonstrated higher frame coverage for the hazardous class and competitive Average Detection Confidence in clear day, clear night, rainy day, and rainy night settings. This demonstrates that localization-based adaptation can improve robustness when object detection systems are employed in region-specific traffic settings.

Detection performance for the person class is acceptable for the scope of this study, with a recall of approximately 0.80. While pedestrian detection is a trigger condition for the imminent hazard alert logic, vehicular proximity being the primary one, such a level of performance befits the system design goal for minimizing false-positive warnings. However, the 13.1% confidence degradation observed under heavy monsoon rain makes certain the findings of Bhowmik [10] and Wang and Zhao [6], in terms of camera-only systems' vulnerability to adversarial atmospheric noise. This is a more functional risk than a purely statistical drop: reduced confidence can lead to flickering detections or delayed hazard triggers when the driver's own visibility is most compromised. Moreover, while the dataset covers a considerable period, the findings are geographically confined to Chiang Mai. The camera mounting and local lighting patterns mean that the performance reported here should be interpreted as context-specific validation of localized AI rather than as a universal evaluation.

5.2 Necessary Architectural Interventions

To improve reliability under real-world conditions, targeted refinements are proposed. In this aspect, further optimization of localized anchor box filters is critical to improve stability under scale variations, especially for high-density scenarios with many occlusions, as indicated by Zhong *et al.* [3].

Secondly, the rain sensitivity suggests that upstream visual quality degradation is the major bottleneck. Adding a special denoising or image-adaptive preprocessing step in the network, as in the various models tested in the research by Liu *et al.* [13], may reduce the effects of rain scatter before the YOLO detection step. This would directly address the confidence collapse phenomenon observed in the experiment.

Crucially, however, the model still has a clearly reactive nature. There is a lack of time-to-collision or motion forecasting because a binary choice was made to reduce processing latency as much as possible while avoiding cognitive burden on the driver through overly complicated signal descriptions.

5.3 Ethical and Socioeconomic Impact

From a human-oriented perspective, the system helps in the improvement of drivers' welfare by focusing on the minimization of false positive warnings to prevent the fatigue of alerts. During congested urban driving, drivers may dismiss assistance systems altogether in response to unnecessary cognitive interruptions, as discussed by Endsley [14] and Horswill and McKenna [11]. It maintains driver trust, which plays an imperative role in the long-term sustainability of professional drivers and AI-assisted systems.

From an economic point of view, the reduction seen in minor accidents indicates that there is an economic

ROI for the logistics operators as anticipated, thus supporting safety guidelines provided by the Federal Motor Carrier Safety Administration (FMCSA) [2] and Aleksa *et al.* [6] as well. From an ethical standpoint, the imperative delineation between an assistive and autonomous system ensures that responsibility remains on the shoulders of the operators only, rather than being obscured by automation itself under a human-in-the-loop framework.

6. CONCLUSION AND FUTURE WORK

Thus, the significance of this research is illustrated by proving the capabilities of the localized AI-based hazard detection system for defensive driving. The proposed approach exhibits strong baseline detection performance while maintaining high mean Average Precision (mAP) and improving sensitivity to high-risk road users. Moreover, the fine-tuning mechanism based on the data obtained through the collection of the dashcam videos emphasizes the significance of contextual adaptation in object detection models while considering the Southeast Asian environment, according to the research by Nguyen *et al.* [1]. Thus, the economic viability demonstrated through the proposed scenarios is consistent with the approach adopted by organizations such as FMCSA, which evaluates return-on-investment models, reported by Aleksa *et al.* [6].

However, the results obtained highlight important issues related to deployment. Even though detection accuracy was significantly increased, the considerable decrease in performance when the adverse weather occurs clearly shows that camera-based perception is still prone to fail in critical situations. The 13.1% decrease in detection confidence in rainy nighttime proves that atmospheric noise and low illumination are critical obstacles for vision systems as reported by Bhowmik [5] and Wang and Zhao [7]. The problem of attaining a universally robust system based on local and one-sensor data becomes evident from these limitations.

Future research should focus on increasing the diversity of dataset under various conditions of weather, lighting, and traffic density to increase robustness statistically. In terms of model improvement, the use of noise filtering deep learning networks and anchor box clustering specific for vulnerable road users is proposed according to the approaches offered by Shahzad *et al.* [15], Rehman *et al.* [16], and Zhong *et al.* [17].

Moreover, the transition from reactive alerts to Time-To-Collision (TTC) estimation would greatly increase the value of the system in defensive driving practices. Lastly, the future research is to involve human factor evaluation in the form of stress and workload measurement.

AUTHOR CONTRIBUTIONS

Conceptualization, S.K. and J.J.; methodology, S.K. and J.J.; software, S.K.; validation, S.K. and J.J.; formal analysis, S.K.; investigation, S.K.; data curation, S.K. and J.J.; writing - original draft preparation, S.K. and J.J.; writing - review and editing, S.K. and J.J.; visualization, S.K.; supervision, S.K.; All authors have read and agreed to the published version of the manuscript.

References

- [1] T. T. Nguyen, P. Krishnakumari, S. C. Calvert, H. L. Vu and H. van Lint, "Feature extraction and clustering analysis of highway congestion," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 238–258, 2019.
- [2] Federal Motor Carrier Safety Administration (FMCSA), "A Return on Investment (ROI) Guide to Advanced Driver Assistance Systems (ADAS)," U.S. Department of Transportation, Tech-Celerate Now Program Office, Tech. Rep., 2022.
- [3] J. Zhong, Q. Cheng, X. Hu and Z. Liu, "YOLO Adaptive Developments in Complex Natural Environments for Tiny Object Detection," *Electronics*, vol. 13, no. 13, Art. no. 2503, 2024.
- [4] G. Jocher and J. Qiu, "Ultralytics YOLO11," ver. 11.0.0, [Computer software], 2024. [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [5] M. K. Bhowmik, *Computer Vision: Object Detection in Adversarial Vision*, 1st ed. Boca Raton, FL, USA: Chapman & Hall/CRC, 2024.
- [6] M. Aleksa, A. Schaub, I. Erdelean, S. Wittmann, A. Soteropoulos and A. Fördös, "Impact analysis of Advanced Driver Assistance Systems (ADAS) regarding road safety – computing reduction potentials," *European Transport Research Review*, vol. 16, no. 1, 2024.
- [7] K. Wang and Y. Zhao, "Improving object detection in challenging weather for autonomous driving via adversarial image translation," *PLOS ONE*, vol. 20, no. 10, 2025.
- [8] D. Geever *et al.*, "A Study on the Impact of Rain on Object Detection for Automotive Applications," in *IEEE Open Journal of Vehicular Technology*, vol. 6, pp. 1287–1302, 2025.
- [9] H. -Y. Lin, Y. -C. Huang, J. -X. Lai and T. -T. You, "Domain Adaptation for Vehicle Detection Under Adverse Weather," in *IEEE Open Journal of Intelligent Transportation Systems*, vol. 6, pp. 568–578, 2025.
- [10] S. Park, J. Kuai, H. Kim, H. Ko, C. Jung and Y. Son, "A Lightweight Degradation-Aware Framework for Robust Object Detection in Adverse Weather," *Electronics*, vol. 15, no. 1, 2025.
- [11] M. S. Horswill and F. P. McKenna, "Drivers' Hazard Perception Ability: Situation Awareness on the Road," in *A Cognitive Approach to Situation Awareness: Theory and Application*, S. Banbury and S. Tremblay, Eds. Aldershot, UK: Ashgate, pp. 155–175, 2004.
- [12] P. Thongchim, P. Taneerananon, P. Luatthep and P. Prapongsena, "Traffic Accident Costing for Thailand," *Journal of the Eastern Asia Society for Transportation Studies*, vol. 7, pp. 2891–2906, 2007.
- [13] W. Liu, G. Ren, R. Yu, S. Guo, J. Zhu and L. Zhang, "Image-Adaptive YOLO for Object Detection in Adverse Weather Conditions," in *The Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI-22)*, vol. 36, no. 2, pp. 1792–1800, 2022.
- [14] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995.
- [15] M. Z. Shahzad, M. A. Hanif and M. Shafique, "Robust ADAS: Enhancing Robustness Of Machine Learning-Based Advanced Driver Assistance Systems For Adverse Weather," *2024 IEEE International Conference on Image Processing Challenges and Workshops (ICIPCW)*, Abu Dhabi, United Arab Emirates, pp. 4240–4246, 2024.
- [16] Y. Rehman, H. Amanullah, D. M. Saqib Bhatti, W. T. Toor, M. Ahmad and M. Mazzara, "Detection of Small Size Traffic Signs Using Regressive Anchor Box Selection and DBL Layer Tweaking in YOLOv3," *Applied Sciences*, vol. 11, no. 23, 2021.
- [17] Y. Zhong, J. Wang, J. Peng and L. Zhang, "Anchor Box Optimization for Object Detection," *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Snowmass, CO, USA, pp. 1275–1283, 2020.



Seung Hwan Kang received a Bachelor of Commerce in Business Information Systems, a Master of Information Systems, a Master of Information Systems (Honours), and a Ph.D. in Information from the University of Wollongong, completing his studies in 2000, 2001, 2004, and 2008. He currently serves as an assistant professor in the Information Technology department at Payap University, Thailand.



Jira Yammeesri received a Bachelor of Business Administration in Accounting from Assumption University, Thailand, in 1995. She subsequently earned a Master of Commerce in Finance and a Master of Commerce Honours in Finance, as well as a Ph.D. in Accounting from the University of Wollongong in 1997, 1998, and 2004, respectively. She is currently an assistant professor in the International Business Management department at Payap University, Thailand.