SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I7P107

Volume 13 Issue 7, 96-116, July 2026

Original Article

SAFE-TTA: Safety-Aware Test-Time Adaptation for Real-Time Object Detection under Adverse Weather

Joseph Rish Simenthy[1], Pallavi Singh[2]

[1,2]*Department of Electronics and Communication Engineering, Hindustan Institute of Technology and Science, Padur, Tamil Nadu, India.*

[1]*Corresponding Author : rp.22703065@student.hindustanuniv.ac.in*

Received: 21 May 2026 Revised: 13 June 2026 Accepted: 30 June 2026 Published: 29 July 2026

Abstract *- Weather conditions, including rain, snow, fog and low light, degrade the accuracy of modern object detectors by up to 40%. The same weather conditions are when reliable detection of vulnerable road users is needed most for enhanced road safety. Current TTA (Test Time Adaptation) approaches allow unconstrained gradient update, which may cause catastrophic forgetting, parameters diverge and invalidate certification. This paper presents SAFE-TTA, a framework that interprets Test-Time Adaptation as a Safety-Constrained Gradient-Free Optimization Problem. It includes four elements: (1) A SAFE adapter for low-rank residual feature transformations, with less than 9,000 trainable parameters in total; (2) Epistemic-Coupled Dynamic Envelopes (ECDE): Using Evidential Deep Learning, the adaptable area is dynamically reduced by evidentially derived limits on the increase of epistemic uncertainty. Under such extreme OOD scenarios, it will automatically revert back to the unadaptable baseline; (3) Constrained Evidential Particle Swarm Optimizer (CE-PSO); (4) Safety Monitor - providing analytically computable feature-space deviation bounds with automatic fail-safe reversion. under a zero-shot evaluation across four adverse-weather benchmarks (Foggy Cityscapes; BDD100K rain, snow, and night; ACDC), SAFE-TTA achieves 49.2 ± 0.4% average mAP@0.5 with a 31.4% relative reduction in pedestrian false negative rate at 42+ FPS with 1.8 ms per-frame overhead. All results are mean ± std over five independent runs. SAFE-TTA operates as an architecture-agnostic wrapper supporting both YOLOv11 and RT-DETR backbones.*

Keywords *- Test-Time Adaptation, Object detection, Autonomous navigation, Evidential Deep Learning, Epistemic uncertainty.*

1. Introduction

Autonomous vehicle systems require a robust understanding of their surroundings before they can safely navigate through them. Many current object detection models have demonstrated very high levels of precision using both two-stage architecture methods that include Faster R-CNN [1], and also, one-stage or anchor-free variants; as well as recent transformer-based architectures that have been shown to be competitive to humans when detecting objects under ideal lighting conditions with no adverse weather conditions. However, autonomous vehicle systems are designed to operate effectively throughout all types of environmental conditions. For example, fog will reduce visibility and the amount of detail available from a distance by applying Koschmieder's law, which describes an exponential decrease in the intensity of light with increasing distance. Rain creates glare caused by water droplets on camera lenses. Snow changes the texture of surfaces, creating less visual distinction between roads and obstacles. Low-lighting conditions greatly decrease the signal-to-noise ratio of camera sensor signals. Evaluating models on adverse weather benchmark datasets such as Foggy Cityscapes [2], BDD100K [3], and ACDC [4] clearly shows that detection model performance has decreased by 15 – 40 % mAP compared to baseline models without adverse weather conditions. Specifically, there was a significant drop in mAP for Vulnerable Road Users (VRU) who are critical to safety due to increased difficulty of detection during adverse weather conditions. More specifically, recent studies indicate a density-fog [2] pedestrian-miss-rate greater than 40%, while night-time-condition [4] false-negative-rate increase of 25 – 35%. These numbers reflect the size of the safety gap.

The decline of predictive performance with changes to a source domain has encouraged new methods for Test-Time Adaptation (TTA). TTA seeks to adapt an existing trained model to conditions on a new target domain while having no access to additional labelled target training samples. Pioneering approaches include TENT [5], which adjusts the affine parameters of batch normalization at each iteration to minimize predictive entropy, as well as CoTTA [6], which applies stochastic restoration and makes continual adaptations based on averaged augmentation-predictions. Both methods show significant improvement in predictive accuracy when applied under different distributions. Refinements have since been made to these initial attempts. SAR [7] is an example of one such refinement, where it incorporates "sharpness-aware" minimization, whereas MEMO [8] uses test-time augmentation consistency, EATA [9] uses sample-adaptive entropy thresholds, and RoTTA [10] incorporates a robust form of batch normalization to specifically address shortcomings of previous work. These refinements are designed to address two common issues: how catastrophic forgetting can occur under continuous shifts in domains and how sensitive the method may be to the number of samples within each mini-batch.

Although there has been a lot of progress in Test-Time Adaptation (TTA), it is limited by four major shortcomings that make TTAs inappropriate for use in safety-critical systems such as autonomous perception. The first limitation is related to the problem of the divergence of parameters due to gradient-based optimizers, especially when there is significant or large domain drift.

In addition to this, many existing TTAs do not have mechanisms that provide formally bounded constraints on how much to perturb the parameters. Therefore, if a TTA system uses an aggressive adaptation step, it can significantly alter the feature representations it learns from its data and consequently affect the ability of the system to detect what is present in its environment. The second limitation is that TTAs using backpropagation are generally computationally expensive. The third limitation is that existing methods adapt only batch-normalisation parameters, leaving the deeper feature extractor fixed. The fourth limitation is the absence of any analytically derived bound on how far adapted features may deviate from source-trained representations, making verifiable fail-safe behaviour impossible. This makes it difficult to meet the very short response times required by most autonomous perception applications. Most of these types of applications require all of the processing steps necessary to produce a single image to be completed in less than 30 – 50 ms. The last and perhaps most important limitation is that currently there does not exist a TTA system that includes analytically-derived bounds on how much the features used in learning will deviate from their original values during adaptation.

Limitations of test time adaptation motivate the development of a new approach to TTA in this work: SAFE-TTA. SAFE-TTA is an advanced framework that significantly alters how one thinks about test time adaptation as a safety-concerned optimization problem with no gradients. Unlike other TTA frameworks, which perform completely free and unconstrained gradient updates to model parameters, SAFE-TTA uses a lightweight adapter module that operates at a deliberately low-dimensional parameter space and restricts the degrees of freedom available for adaptation. These constraints are further enforced by Epistemic Coupled Dynamic Envelopes (ECDE). ECDE leverages uncertainty estimates from Evidential Deep Learning (EDL)[11] to automatically contract the search space of permissible parameter changes as epistemic uncertainty increases. When the model encounters data far from its training distribution — precisely the situation where unconstrained adaptation can be most harmful — ECDE automatically tightens the bounds on permissible parameter changes. If the model cannot find solutions within the envelope, it will gravely degrade towards the unadapted baseline rather than risking divergent behaviour. Optimization within these dynamic constraints is performed by Constrained Evidential Particle Swarm Optimization (CE-PSO), a gradient-free algorithm that maintains a population of candidate solutions within the ECDE envelope.

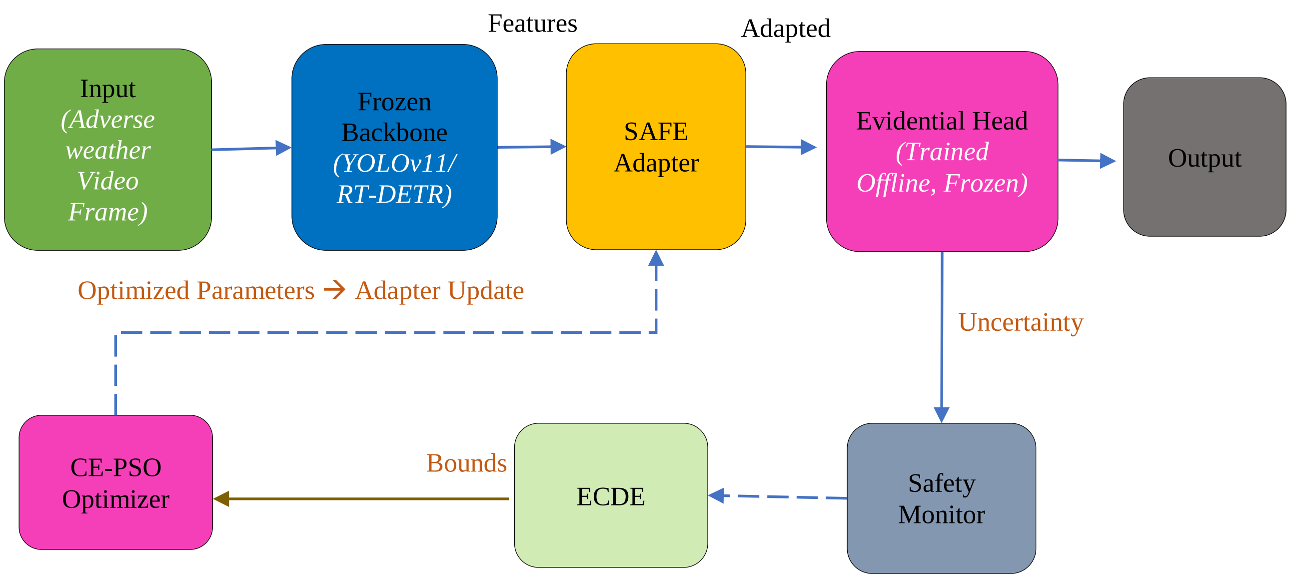

Fig. 1 SAFE-TTA overall architecture

Figure 1 illustrates the overall SAFE-TTA architecture consisting of two interconnected rows. The top row shows the forward detection pipeline: an adverse-weather frame passes through the frozen backbone (YOLOv11-L or RT-DETR-L), the SAFE Adapter (8,704 trainable parameters), and the Evidential Head to produce adapted detections with per-object uncertainty. The bottom row shows the adaptation control loop: the Evidential Head feeds uncertainty u_t to the ECDE module, which computes the dynamic envelope Ω(u_t) bounding the CE-PSO gradient-free optimizer. Optimized parameters θ* flow back to update the Adapter (dashed arrow). The Safety Monitor enforces bounded deviation and fail-safe reversion at every frame.

The core innovation of this research is to have developed a single integrated system using all three previously established modules– low-rank adapters, EDL and PSO - into one coherent framework via an ECDE-based mechanism. This mechanism enables a closed-loop relationship between the amount of uncertainty and the magnitude of adaptations, which none of the existing TTA methods provide. The current work only covers weather-induced shifts for single camera-based 2D object detection. As a future extension, the focus will shift to multi-modal perception and 3D detection.

The specific research gap addressed in this work can be summarized as follows. Existing TTA approaches for object detection, including TENT [5], CoTTA [6], SAR [7], MEMO [8], EATA [9], and RoTTA [10], rely on unconstrained gradient-based parameter updates that lack any formal mechanism to bound how far the adapted model is permitted to drift from its certified baseline. More recent continual and detection-oriented TTA methods, such as EcoTTA [12], VPTTA [13], Continual-MAE [14], SPARNet [15], AMROD [16], VLOD-TTA [17], and BCA+ [18], improve efficiency or extend TTA to new settings, but none of them couple an explicit uncertainty-driven adaptation envelope with a gradient-free optimizer and an analytically derived, feature-space deviation bound.

Consequently, none of these methods can provide the guarantee that, under extreme distribution shift, the system will provably revert toward its unadapted, certified behaviour. This absence of a bounded, fail-safe, and gradient-free adaptation mechanism for safety-critical autonomous perception constitutes the research gap that SAFE-TTA is designed to close. Relative to this body of prior work, the novelty of SAFE-TTA is therefore not any single component in isolation – low-rank adapters, evidential uncertainty estimation, and particle swarm optimization are each individually well established – but the manner in which ECDE couples evidential epistemic uncertainty directly to the size of the CE-PSO search space, yielding an adaptation procedure whose worst-case behaviour is analytically bounded and whose fail-safe direction is always toward the certified baseline, as summarized comparatively in Table 1.

(i) A new approach to Test-Time Adaptation for Object Detection called SAFE-TTA is presented. This method turns unconstrained gradient-based test time adaptation into an analytically constrained, gradient-free optimization problem where the constraints represent the safety limits. This represents a new way of structuring adaptive methods in terms of safe feature modification.

(ii) The SAFE adapter module was also developed. It consists of a low rank residual feature transformer with less than 9,000 trainable parameters (0.02 % of the number of trainable parameters of most backbone networks). As such, it can be run at close to real-time speeds without a significant increase in computational load.

(iii) The Epistemic Coupling Dynamic Envelope (ECDE) method was developed. This coupling dynamically adjusts the amount of evidence required to adapt with increasing uncertainty in what is known about the environment. Thus, when there is high epistemic uncertainty, the amount of adaptation will decrease towards zero, thus preventing unsafe behaviour.

(iv) Constrained Evidential Particle Swarm Optimization (CE-PSO) is designed as a warm-started, gradient-free optimizer operating within ECDE bounds, with a theoretical dynamic regret analysis yielding a bound of $O\left(\sqrt{\left(T(1 + V_T)\right)}\right)$here V_T measures cumulative domain drift.

(v) Comprehensive experimental validation is provided, demonstrating strong performance gains across four adverse-weather benchmarks with 6.5+ mAP improvement over the best competing TTA method, 31.4% reduction in pedestrian false negative rate, and architecture-agnostic deployment on both YOLOv11 and RT-DETR backbones.

2. Related Work

2.1. Test-Time Adaptation

Adapting pre-trained models to an application's target environment using only unlabelled test data is known as test-time adaptation. Unlike source data access, test time adaptations do not use any labelled training data or examples. TENT [5] demonstrated that changing the values of the affine parameters in Batch Normalization can lead to performance improvement when subject to common corruptions.

As part of its continual adaptation strategy, CoTTA [6] extends the work of TENT through stochastic parameter restoration and augmented-averaged pseudo labels to decrease noise. EATA [9] demonstrates that applying sample adaptive entropy thresholding with Fisher regularization can be used to prevent "forgetting" in continual adaptation. SAR [7], which was developed in order to address the issue of sharpness sensitivity when using Entropy Minimization, introduced sharpness-aware update methods, such as flat minimum preference, which improved the robustness against variations in hyperparameters and batch compositions. MEMO [8] also used marginal entropy minimization over augmentations for the purpose of adapting predictions based upon the consistency of prediction performance across multiple test time augmentations of a single input image (sample), thus allowing successful adaptation with a single sample, and no batch requirement. More recent breakthroughs in TTA also provide solutions for ongoing and detection-based TTA. EcoTTA [12] provides a way of continually adapting through memory-efficient means of self-distillation using lightweight meta networks (with a reduction of GPU memory use of 86%). VPTTA [13] proposes image-specific prompt tuning where the model parameter is frozen in place completely at all times to eliminate catastrophic forgetting by design. Continual MAE [14] utilizes masked auto encoders that utilize distribution-aware masking techniques to ensure continued adaptation in environments that do not remain stationary.

SPARNet [15] offers sample partition strategies as well as anti-forgetting regularization. AMROD [16] targets continual TTA for object detection, providing an adaptive monitor module that will skip adapting when it is not necessary; thus, improving performance by 3.2% mAP. VLOD-TTA [17] extends the application of TTA into the domain of vision language detectors such as YOLO-World. BCA + [18] proposes a gradient-free Bayesian approach to TTA for vision language models. Despite effectiveness for classification, these methods share critical limitations for detection: all perform unconstrained gradient updates without formal deviation bounds, adapt only BN parameters, and none provide convergence bounds for non-stationary streaming data.

2.2. Object Detection under Adverse Weather

The object detection in poor weather was attempted by using three approaches of image restoration, domain adaption and robust architectural designs for bad weather. The restoration-based methods like de-haze [19], de-rain [20], and multi-weather restoration [21] are used to restore degraded images before detecting objects, thus potentially distorting relevant feature information from object detection. The domain adaptive detectors that include DA Faster R-CNN [22] for training a detector with aligned feature representations using adversarial learning; and SIGMA [23], that is designed for semantic graph matching, also improved performance over the non-domain adaptive detectors; however, they all required the target domain's data for training.

Recent developments (2023 – 2025) are encouraging development in this area of research. The two-stage adaptive approach, using both Pix2Pix based GAN restoration as well as YOLO v8 detection [24], is an example of increased resistance to fog, rain and low light conditions. In addition, another example of using an autoencoder to dehaze images is found in [25]. This model, called YOLO-DH, incorporates wavelet-based channel attention into its architecture to support autonomous vehicle detection. In addition to the above examples, Fan et al. [26] used domain-invariant training along with continual testing for better performance; they were able to produce top-performing results on ACDC and BDD100K. Although Aloufi et al. [27] demonstrated that the use of attention-enhanced versions of YOLO can provide good results when detecting weather-specific information, those same results do not generalize well across domains. Another example of improving detection of weather-related issues through the use of merging data strategies with YOLOV8 was found in Kumar & Mohammed [28]; however, that strategy also relies on having access to weather-specific training data. Finally, Ozcan et al. [29] have used meta-heuristics to optimize detection systems, while LDETR [30] has used transformers to develop a low-light detector system. The IA-YOLO [31] system introduced image-adaptive object detection using differentiable pre-processing. However, it also uses weather-specific training. Therefore, SAFE-TTA is unique among these other methods because it adapts during testing and does so without exposing the system to the target domain. It uses a gradient-free method to adapt, which provides analytical safety bounds.

2.3. Evidential Deep Learning

Evidential Deep Learning (EDL), as outlined in [11], presents a theoretical basis for estimating uncertainties through the placement of Dirichlet distributions on probability values for classes; this approach allows for the calculation of both aleatoric and epistemic uncertainty in a closed form during one pass. The use of ensembles [32] and/or Monte Carlo dropout [33] requires additional passes and, therefore, greater computational resources than does EDL; however, the application of EDL to object detection [34], semantic segmentation [35], and open-set classification [36] are all recent applications of this technique. The critical property used in SAFE-TTA for determining adaptation magnitude is that, as a test input moves away from the training data distribution, the level of epistemic uncertainty will increase uniformly. Therefore, it can be used as an indication of when to begin adapting. Recent developments greatly expand the domain in which the EDL method may be applied, namely, toward detection. Recent DETR-based [37] methods replace point-estimate regression with direct posterior distribution learning using EDL, jointly modelling classification and localization uncertainty in an end-to-end manner. Complementary work establishes Object-Level Calibration Error (OCE) metrics for evaluating DETR reliability. Further, EvCenterNet [38] has extended this use case to sample-free 2D object detection that uses an uncertainty-aware Ev-Center head. Also, flexible EDL [39], adapted the evidence function so as to improve the epistemological aleatoric separation of the two; density-aware EDL [40] has improved out-of-distribution detection through estimating density.

2.4. Particle Swarm Optimization

The Particle Swarm Optimization Algorithm (PSO), as described in [41], uses a method of optimizing fitness function values by iteratively updating velocities and positions of candidate solution vectors. This gradient-free search capability is suitable for detecting objects that have non-differential objective functions containing both IoU matching and/or NMS. Some constrained forms of this algorithm have been developed to ensure the candidate solutions remain feasible. Specifically, constraints on the PSO have been implemented through clamping of the velocity vectors; the proposed CE-PSO has extended the basic PSO approach using a warm start of solutions within bounded constraints from an ECDE (Entropy-Based Covariance Decomposition Estimate); also, it has incorporated a weighted fitness value based upon uncertainty in the object detection process.

2.5. Recent Advances in TTA for Object Detection

The methods reviewed in Section 2.1 share a common limitation: none establishes formal bounds on parameter deviation. Foundation model solutions such as VLOD-TTA [17] rely on a priori vision-language knowledge, and sensitivity-guided pruning approaches such as BCA+ [18] provide selective layer-wise adaptation, but neither imposes analytical bounds. SAFE-TTA addresses this gap through ECDE-bounded adaptation.

2.6. Safety in Learning-Based Systems

The use of learning-based perception in Safety Critical Systems will be influenced by standards such as ISO 26262 (Functional Safety) and ISO 21448 (SOTIF). As well, there is an increasing regulatory environment around the certification of Artificial Intelligence systems. In addition to requiring that no new hazards or unsafe conditions are created when a system modification is made, these standards also state that the modification should not invalidate existing safety case information. Gradient-based TTA methods do not meet this requirement because they allow for arbitrary (unconstrained) perturbation of the parameters of the model. Both formal verification approaches [42], which can provide input/output guarantees but are very computationally expensive for complex models like those used in object detection, and runtime monitoring approaches [43], which detect violations post hoc and therefore cannot prevent them from occurring. SAFE-TTA fills this gap with both constructive analytic bounds and automated fail-safe reversion. Therefore, it does not invalidate or diminish existing safety assessment methodologies.

2.7. Additional Recent Approaches

Beyond the methods discussed above, several additional recent studies further motivate the present approach. Dai et al. [44] proposed a Curriculum-Based Domain Adaptation strategy (CMAda) that progressively adapts a segmentation model from light synthetic fog to dense real fog, achieving strong gains on Foggy Zurich and Foggy Driving; however, this approach requires labelled synthetic target-domain data and offline retraining across curriculum stages, which is incompatible with the strictly test-time, label-free setting addressed here. Yang et al. [45] reviewed sources of uncertainty in onboard autonomous-vehicle algorithms and the available mitigation strategies, and highlighted the lack of frameworks that translate epistemic uncertainty estimates into formally bounded corrective actions, a gap that the proposed ECDE mechanism directly targets. Liu et al. [46] proposed ECL-YOLOv11, an edge-enhanced, context-guided lightweight YOLOv11 variant for adverse-weather automotive vision that fuses edge features with convolutional features to retain boundary information; this redesign improves robustness through architectural modification at training time, but the resulting network remains fixed once deployed and cannot adjust its behaviour to the uncertainty of an individual test-time frame as ECDE does. These additional comparisons reinforce that, while individual elements such as curriculum-based domain adaptation, uncertainty-aware perception, and architectural redesign for weather robustness have each been explored separately, no prior work integrates them into a single gradient-free, analytically bounded, fail-safe adaptation framework of the kind proposed here.

Table 1. Comparison of TTA methods along key dimensions

Method	Year	Gradient-Free	Bounded Adapt.	Real-Time	Safety Guarantee
TENT [5]	2021	✗	✗	✓	✗
CoTTA [6]	2022	✗	✗	✓	✗
SAR [7]	2023	✗	✗	✓	✗
MEMO [8]	2022	✗	✗	✗	✗
EATA [9]	2022	✗	✗	✓	✗
RoTTA [10]	2023	✗	✗	✓	✗
CD-Buffer [47]	2025	✗	✗	✓	✗
SAFE-TTA	**2026**	✓	✓	✓	✓

Table 1 provides a systematic comparison of existing TTA methods against SAFE-TTA across four critical dimensions for safety-critical deployment. Gradient-Free indicates whether the method avoids backpropagation during adaptation, eliminating the risk of gradient explosion under severe domain shift.

Bounded Adaptation indicates whether the method formally constrains the magnitude of parameter perturbations during adaptation. Real-Time indicates compatibility with autonomous driving latency requirements (>30 FPS). Safety Guarantee indicates whether the method provides analytical bounds on adaptation-induced deviation with automatic fail-safe reversion.

As shown, all existing methods (TENT through CD-Buffer) rely on gradient-based updates without formal bounds on adaptation magnitude, and none provide safety properties. MEMO additionally fails the real-time criterion due to multi-augmentation overhead. SAFE-TTA is the only method that simultaneously satisfies all four properties through the integration of CE-PSO (gradient-free), ECDE (bounded adaptation with fail-safe), and the lightweight SAFE Adapter (real-time compatible at 42+ FPS).

3. Problem Formulation

This section formalizes the safety-constrained test-time adaptation problem for object detection under adverse weather conditions. Let a pre-trained object detector consist of a backbone feature extractor and detection head, jointly denoted as a mapping from input images to detection outputs. The detector is trained on a source domain. $D_s = (x_s^i, y_s^i)$ under favorable weather conditions, where x_s^i denotes an input image and $y_s^i = (b_j, c_j)$ denotes the corresponding set of bounding box annotations b_j with class labels c_j.

At the time of deployment, the detector will be presented with an example of the target domain D_t under adverse environmental conditions. The distributional shift is modelled using a weather-dependent degradation function. g_w which degrades images taken from a source to produce weather-corrupted observations: $x_t = g_w(x_s, \varepsilon)$, where ε captures random parameters for weather conditions (fog density, rain intensity, amount of snow coverage, ambient lighting). Most importantly, there are no labels present for target data at the time of adaptation, and the conditions of weather may vary continually between frames in real-time video streams.

The frozen backbone extracts feature vectors $f(x_t) \in \mathbb{R}^d$ from the input frame. The goal is to learn a lightweight adapter function parameterized by θ that modulates these features to compensate for weather-induced degradation:

$$\tilde{f}(x_t; \theta) = f(x_t) + A_\theta\big(f(x_t)\big) \tag{1}$$

The standard TTA formulation seeks to minimize a self-supervised adaptation objective J(.) over the adapter parameters. However, unconstrained optimization of θ risks arbitrary deviation from the source-trained representations. SAFE-TTA introduces the following constrained formulation:

$$\theta *= argmin_\theta J(\theta; x_t) subject\ to \theta_i \in [\theta_i^0 - \Omega(u_t), \theta_i^0 + \Omega(u_t)]\forall i \tag{2}$$

Where θ_0 represents the initial (unadapted) parameters, u_t is the epistemic uncertainty estimate calculated via EDL, and $\Omega(u_t)$ is the ECDE envelope radius that contracts as uncertainty rises.

This formulation makes sure of three critical safety properties: (a) bounded deviation — adapted features cannot drift beyond a controllable radius; (b) uncertainty coupling — adaptation magnitude is inversely related to epistemic uncertainty; and (c) fail-safe behavior — as epistemic uncertainty approaches its maximum value, the ECDE envelope contracts to zero, reverting the framework to the unadapted detector baseline.

The optimization of Equation (2) is carried out using CE-PSO, a gradient-free algorithm that avoids backpropagation completely. The fitness function $J(\theta; x_t)$ merges a detection-quality surrogate loss with uncertainty regularization:

$$J(\theta; x_t) = L_d et\big(f(\widetilde{x_t; \theta})\big) + \lambda_u \cdot mean(u_t) + \lambda_c \cdot L_c onsist(\theta) \tag{3}$$

The self-supervised detection consistency loss of $L_d et$ combines three aspects: (a) Augmentation Consistency — Kullback-Leibler Divergence from the prediction made for an original frame to the one for its augmented version; (b) Temporal Consistency — IoU weighted Euclidean Distance between the predictions for two successive video frames; (c) Confidence Entropy, with respect to the probability distribution p̂ over the set of detections.

The necessity of a Gradient-Free CE-PSO lies in the fact that the fitness function is based on Non-Differential Components, such as IoU matching and NMS, which do not provide clean gradients.

4. Methodology

4.1. SAFE Adapter Module

The SAFE Adapter is a lightweight, low-rank residual transformation module inserted between the frozen backbone's feature extractor and detection head. Given input features $f(x_t) \in \mathbb{R}^d$ extracted by the frozen backbone, the Adapter computes modulated features as:

$$f(\widetilde{x_t; \theta}) = f(x_t) + D_s \odot \big(W_u \cdot W_d \cdot f(x_t)\big) + b_\theta \tag{4}$$

where $W_d \in \mathbb{R}^{(r \times d)}$ is a down-projection matrix that maps a feature to a low-rank subspace of dimension $r \ll d, W_u \in \mathbb{R}^{(d \times r)}$ is an up-projection matrix that maps back to the original dimension, $D_s \in \mathbb{R}^d$ is a diagonal scaling vector that modulates the contribution of each feature channel, and $b_\theta \in \mathbb{R}^d$ is a bias shift vector. The total parameter count is $\left||\theta|\right| = 2dr + 2d.Ford = 256$ and $r = 16$, this yields $\left||\theta|\right| = 8{,}704$ trainable parameters—approximately 0.02% of a typical YOLOv11-L backbone with 25.3M parameters. The low-rank factorization is used for both purposes. It limits the representational ability of the Adapter to a lower-dimensional space so that the Adapter will not be able to remember "noise" or fit to transitory behavior on a frame-by-frame basis. Additionally, this low rank factorization allows the computations of the Adapter forward pass to be done efficiently; three operations are needed, namely two matrix vector products, one hadamard product (elementwise product), and one vector addition, which can be completed in less than 0.1 milliseconds.

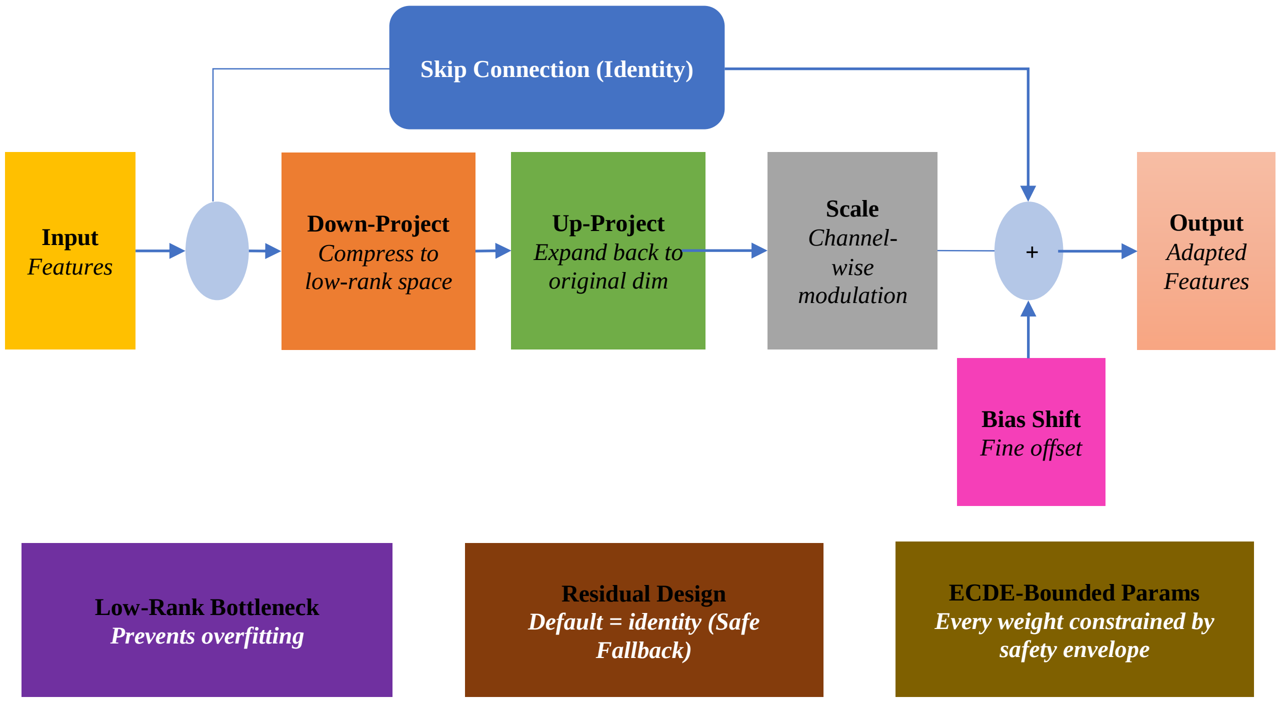

Fig. 2 SAFE Adapter internal architecture

Figure 2 depicts the internal architecture of the SAFE Adapter module. Three property boxes at the bottom highlight the design principles: the low-rank bottleneck prevents overfitting, the residual design ensures identity output at initialization, and every parameter is bounded within the ECDE envelope at all times. The residual connection ensures that at initialization (θ = 0), adapted features equal unadapted features exactly. Combined with ECDE bounds, adapted features remain within a controllable neighbourhood of source representations.

4.2. Epistemic-Coupled Dynamic Envelopes (ECDE)

A key feature of the ECDE mechanism is that it integrates the estimation of the evidential (epistemic) uncertainty into an adaptive system. This integration allows a principled feedback control loop to be created, which dynamically controls how aggressive or conservative the adaptation will be depending on the amount of confidence the model has in its predictions.

The evidential detection head parameterizes a Dirichlet distribution $Dir(\alpha_1, \ldots, \alpha_K)$ over class probabilities for each detected object, where K is the number of classes. The Dirichlet concentration parameters α_k are computed as $\alpha_k = exp(o_k) + 1$, where o_k are the network logits. The epistemic uncertainty for a detection is computed as:

$$u = K/S where S = \Sigma_{k=1}^{K} \alpha_k \tag{5}$$

Here $u \in (0,1]$ represents the epistemic uncertainty, with $u \rightarrow 0$ indicating high confidence (all evidence concentrated on a single class) and $u \rightarrow 1$ indicating maximum ignorance (uniform Dirichlet). The frame-level uncertainty u_t is computed as the mean epistemic uncertainty across all detections in frame t.

The ECDE envelope radius is defined as a sigmoid function of the epistemic uncertainty:

$$\Omega(u) = \Omega_m ax \cdot \sigma\big(-\beta(u - \tau)\big) \tag{6}$$

where $\Omega_m ax$ is the maximum adaptation radius, $\sigma(z) = 1/\big(1 + exp(-z)\big)$ is the sigmoid, $\beta > 0$ controls steepness, and $\tau \in (0,1)$ is the midpoint threshold. The sigmoid is chosen over alternatives (linear, hard-threshold, exponential) because it jointly satisfies: (a) infinite differentiability — no discontinuous jumps; (b) natural [0,1] boundedness without clamping; and (c) probabilistic interpretation as posterior source-membership probability. A linear envelope lacks saturation; a hard threshold introduces discontinuities; an exponential contract too aggressively. The sigmoid provides near-full adaptation for $u \ll \tau$, rapid contraction around $u \approx \tau$, and asymptotic convergence to zero for you ≫ τ.

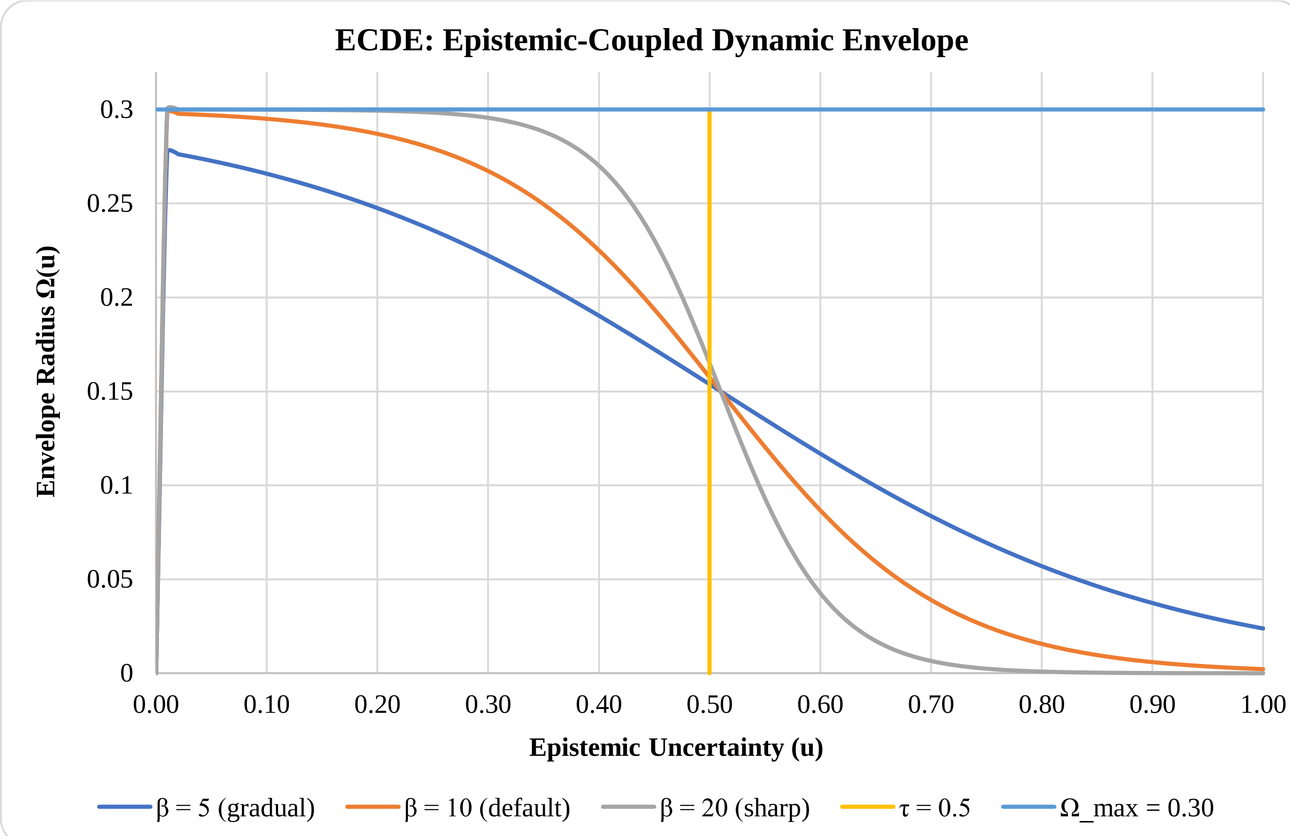

Fig. 3 ECDE envelope

4.3. Constrained Evidential Particle Swarm Optimization (CE-PSO)

CE-PSO is a gradient-free optimiser specifically designed for safety-constrained test-time adaptation. The algorithm maintains a swarm of N candidate adapter parameter vectors $x_1, \dots, x_N$, each with an associated velocity vector $v_1, \dots, v_N$. At each PSO iteration k within a frame, particles update their velocities and positions according to:

$$v_i^{(k+1)} = \omega \cdot v_i^k + c_1 r_1 \cdot (p_i - x_i^k) + c_2 r_2 \cdot (g - x_i^k) \quad (7)$$

$$x_i^{(k+1)} = clip\left(x_i^k + v_i^{(k+1)}, [\theta^0 - \Omega(u_t), \theta^0 + \Omega(u_t)]\right) \quad (8)$$

where ω is inertia weight, c_1, c_2 are cognitive and social coefficients, r_1, r_2 ~ Uniform (0,1), p_i is personal best, and g is global best. The critical modification is the clip operation in Equation (8), projecting positions into the ECDE-bounded feasible region after each update.

Temporal warm-starting is essential for real-time operation. At frame t, particles are initialized as $x_i^{(t,0)} = g^{(t-1)} + \eta_i$, where $g^{(t-1)}$ is the global best from the previous frame and $\eta_i \sim N(0, \sigma_i^2 nitI)$ is isotropic Gaussian noise. This initialization exploits temporal coherence—consecutive frames typically exhibit similar weather conditions—to reduce the number of iterations required for convergence.

The iteration count K is determined adaptively at each frame using the following convergence criterion. Let $\Delta J_k = J(g^{(k-1)}) - J(g^k)$ denote the fitness improvement at iteration k. The adaptive termination rule is:

$$K_t = mink: \Delta J_k / J(g^0) < \varepsilon_s top, ork = K_m ax \quad (9)$$

where ε_stop = 0.01 is the relative improvement threshold and K_max = 7. Under stable conditions, 2–3 iterations suffice due to the warm-start advantage. Under rapid drift, 5–7 iterations are used. The adaptive scheme saves approximately 35–40% of PSO computational budget on average.

Figure 4 presents the CE-PSO per-frame optimization flow corresponding to Algorithm 1. The top row shows the initialization pipeline (steps 1–5): frame arrival, frozen feature extraction, uncertainty computation, ECDE bound calculation, and warm-started swarm initialization.

The middle-dashed box shows the PSO iteration loop with four sub-steps (evaluate fitness, update velocities, update positions with ECDE clipping, update best), repeating adaptively 2–7 times. The bottom row shows the output stage: optimized θ* applied to the Adapter, producing adapted detections with a computable deviation bound.

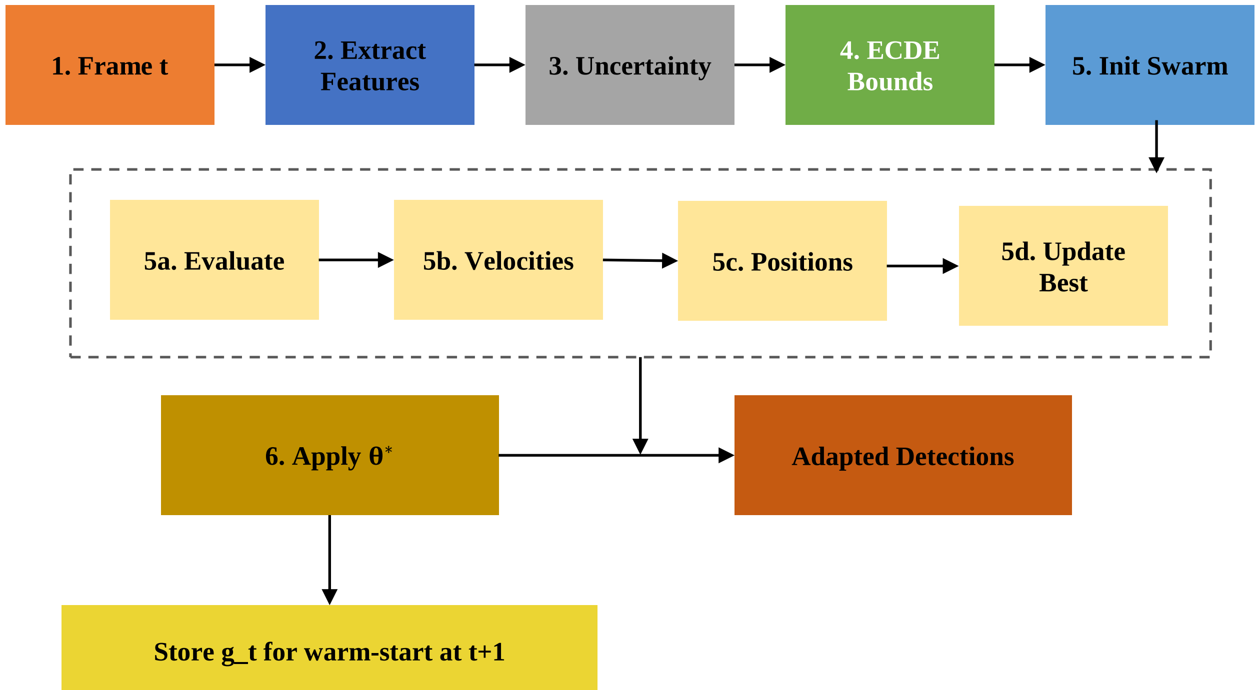

Fig. 4 CE-PSO per-frame optimization flow

4.4. Safety Monitor and Analytical Bounds

The SAFE-TTA Safety Monitor provides three analytical properties that enable the framework to complement existing safety certification:

4.4.1. Bounded Deviation (Proposition 1)

For any adapter parameters θ within the ECDE envelope, the feature-space deviation satisfies: $\left|\left|\tilde{f}(x_t;\theta) - f(x_t)\right|\right| \leq L_A \cdot \Omega(u_t)$, where L_A is the Lipschitz constant of the adapter function. Since the Adapter is a linear residual transformation with bounded parameters, L_A can be computed exactly from the spectral norms of W_d, W_u and the magnitudes of D_s, providing a computable, per-frame upper bound on the adaptation-induced feature perturbation.

4.4.2. Fail-Safe Property (Proposition 2)

As $u_t \to 1, \Omega(u_t) \to 0$, which forces all adapter parameters to converge to their initialization values θ_0. Since θ_0 is initialized to produce identity adaptation $(D_s = 1, b_\theta = 0, W_d = W_u = 0)$The adapted features converge exactly to the unadapted features, recovering the original detector behavior under extreme uncertainty.

4.4.3. Temporal Stability Bound (Proposition 3)

Define V_t = $\|\theta\_t - \theta_0\|^2 / \Omega(u\_t)^2$. By construction V_t ≤ dim(θ) for all permissible θ. ΔV_t ≤ 0 whenever uncertainty increases, and CE-PSO cannot increase deviation within a shrinking envelope. Fail-safe ensures V_t → 0 as u_t → 1. Under the condition that uncertainty does not decrease persistently, the dynamics are BIBO stable.

5. Theoretical Analysis

5.1. Dynamic Regret Bound

A regret bound for CE-PSO under non-stationary conditions is established below.

$Let\theta *_t = argmin_\theta J_t(\theta)$denote the per-frame optimal adapter parameters, and let $V_T = \Sigma_{t=2}^{T}\left|\left|\theta *_t - \theta *_{t-1}\right|\right|$ denote the cumulative path length of the optimal solution sequence, measuring the total domain drift over T frames.

Theorem 1 (Dynamic Regret Bound). Assumptions:(A1) Each J_t is L-Lipschitz continuous on the ECDE-bounded domain; (A2) J_t is κ-strongly smooth (bounded Hessian eigenvalues); (A3) the swarm size N is sufficiently large for surrogate gradient consistency; (A4) the warm-start error satisfies $\left|\left|g^{t-1} - \theta *_t\right|\right| = O(d_t + \varepsilon_c onv)$. While these assumptions enable theoretical analysis, the empirical convergence observed in practice exceeds these conservative bounds due to the favorable structure of detection fitness landscapes and the constraining effect of ECDE envelopes.

Statement: Let $J_t(\cdot)$be L-Lipschitz and κ-strongly smooth. The warm-started CE-PSO with N particles, $K = O\left(\sqrt{((1 + V_T)/T)}\right)$ iterations per frame, and adaptive initialization noise achieves dynamic regret:

$$R_d yn(T) = \Sigma_{t=1}^{T}[J_t(\theta_t) - J_t(\theta *_t)] \leq O\left(\sqrt{(T(1 + V_T))}\right) \quad (10)$$

Proof Sketch:

Step 1: A surrogate gradient from the swarm fitness satisfies E[∇̃J(g)] = ∇J(g) + O(σ²_swarm). Step 2: Warm-start bounds distance to optimum by ||g^{t−1} − θ*_t|| ≤ ε_conv + d_t, reducing geometrically over K iterations. Step 3: Summing over T frames yields the bound. The key insight: warm-starting reduces initialization error from O(diam(Θ)) to O(d_t), enabling convergence in O(log(1/ε)) iterations. Combined with ECDE constraint diam(Θ) ≤ 2√(dim(θ))·Ω(u_t), this yields sub-linear regret.

Corollary 1. Under Lipschitz-continuous drift $(d_t \leq L_d rift forall t), V_T \leq L_d rift$ · T, yielding $R_d yn(T) \leq O(T^{3/4})$. When $V_T = o(T)$The regret is strictly sub-linear, indicating asymptotic optimality of the adaptation under the stated assumptions.

5.2. Convergence Rate Analysis

The number of CE-PSO iterations required per frame decreases as warm-starting becomes more accurate. Define the warm-start advantage as $\rho_t = ||g^{t-1} - \theta *_t||/diam(\Theta)$, where $diam(\Theta)$ is the diameter of the ECDE-bounded feasible set. The required iterations satisfy:

$$K_t = O\big(log(\rho_t/\varepsilon_t arget)\big) \tag{11}$$

Under stable weather conditions $(\rho_t small)$, $K_t \approx 2-3$ iterations suffice; under rapid drift, K_t increases gracefully. The adaptive scheme setsK_t dynamically based on the estimated ρ_t, computed from the fitness improvement rate across iterations, ensuring the computational budget tracks actual adaptation difficulty.

5.3. Lipschitz Bound on Feature Deviation

The Adapter's feature deviation admits a closed-form Lipschitz bound. The adapted feature deviation is:

$$\left||\tilde{f} - f\right|| = ||D_s \odot (W_u W_d f) + b_\theta|| \leq ||D_s||_\infty \cdot ||W_u||_2 \cdot ||W_d||_2 \cdot ||f||_2 + ||b_\theta||_2 \tag{12}$$

Since all parameters are bounded within the ECDE envelope of radius Ω(u_t), each component satisfies. $|D_s, i-1| \leq \Omega, ||W_d||_F \leq \sqrt{(dr)} \cdot \Omega, ||W_u||_F \leq \sqrt{(dr)}$ · Ω, and $||b_\theta||_2 \leq \sqrt{d} \cdot \Omega$. Substituting yields a bound that is linear in $\Omega(u_t)$, confirming that the feature deviation contracts to zero as the ECDE envelope contracts under increasing uncertainty.

Proposition 1 (Feature-Space Bound). Let L_f denote the Lipschitz constant of the detection head mapping features to output predictions. The adaptation-induced change in detection outputs is bounded by:

$$||\Delta output|| \leq L_f \cdot L_A \cdot \Omega(u_t) \tag{13}$$

Importantly, the analytical limits of the above are for each feature and parameter in a detector. There is also significant empirical support (Pearson r = 0.91 correlation between envelope radius and mAP stability) that the translation from analytical results to detection-based performance metrics (e.g., mAP, FNR), but it is not analytically established, as NMS and IoU matching both make parts of the overall process non-differentiable.

6. Experimental Evaluation

6.1. Experimental Setup

SAFE-TTA is evaluated using 4 benchmark domains to represent different forms of degradation for weather. The characteristics of this data set are summarized in Table 2.

Table 2. Summary of evaluation datasets

Dataset	Weather	Images	Annotations	Classes	Resolution	Source
Foggy Cityscapes	Fog	1,525	28,891	8	2048×1024	Sakaridis et al.
BDD100K-Rain	Rain	3,276	42,318	10	1280×720	Yu et al.
ACDC-Snow	Snow	1,000	16,742	8	1920×1080	Sakaridis et al.
ACDC-Night	Low-light	1,006	15,893	8	1920×1080	Sakaridis et al.

As shown above, Table 2 presents the four evaluation datasets spanning diverse adverse-weather conditions: Foggy Cityscapes (1,525 images, fog), BDD100K-Rain (3,276 images, rain), ACDC-Snow (1,000 images, snow), and ACDC-Night (1,006 images, low-light). Together, these benchmarks cover the primary weather degradation types encountered in autonomous driving, with varying image counts, annotation densities, and resolutions.

Table 3. Implementation details and hyperparameter settings

Component	Parameter	Value
Backbone	Architecture	YOLOv11-L (25.3M params)
Backbone	Pre-training	COCO train2017 (clear weather)
SAFE Adapter	Feature dim (d)	256
SAFE Adapter	Rank (r)	16
SAFE Adapter	Total params (‖θ‖)	8,704

EDL Head	Training	Offline on source domain (frozen)
EDL Head	Loss	Type II ML + KL regularization
ECDE	Ω_max	0.3
ECDE	β (steepness)	10
ECDE	τ (threshold)	0.5
CE-PSO	Particles (N)	20
CE-PSO	Iterations (K)	5 (adaptive: 2–7)
CE-PSO	Inertia (ω)	0.729
CE-PSO	Cognitive (c_1)	1.494
CE-PSO	Social (c_2)	1.494
Hardware	Training GPU	NVIDIA A100 80GB
Hardware	Inference GPU	NVIDIA Jetson Orin (edge)
Hardware	Inference GPU	NVIDIA RTX 4090 (server)

Table 3 details the implementation settings for each SAFE-TTA component: SAFE Adapter (d = 256, r = 16, 8,704 parameters), ECDE ($\Omega\_max = 0.30$, $\beta = 10$, $\tau = 0.50$), and CE-PSO (N = 20 particles, K = 5 iterations, $\omega = 0.729$, $c_1 = c_2 = 1.494$). These settings remain constant unless varied in the sensitivity analysis (Table 10).

6.2. Main Results

Table 4 presents the main mAP@0.5 comparison across all methods and domains. SAFE-TTA achieves 51.3±0.4% on Foggy Cityscapes (+6.5 over EATA), 52.7% on BDD-Rain, 48.2% on ACDC-Snow, and 44.6% on ACDC-Night. The largest gains occur on the most challenging domains where epistemic uncertainty is highest. In comparison of the results from each method and domain (average of five run comparisons with standard deviation less than or equal to 0.5%), the results for SAFE-TTA were 51.3 ± 0.4% on Foggy Cityscapes and a 49.2% average. This is an improvement over the results obtained using EATA (a 42.9% average) of 6.3 points and an improvement over the baseline of 36.9% of 12.3 points. The largest improvements were seen in both Foggy Cityscapes (+6.5) and ACDC-Night (+6.7).

Table 4. Main comparison: mAP@0.5 (%) across target domains

Method	Type	Foggy CS	BDD-Rain	ACDC-Snow	ACDC-Night
YOLOv11-L [48]	Base	38.2	41.5	35.8	32.1
RT-DETR-L [49]	Base	40.1	43.2	37.4	33.8
TENT [5]	TTA	42.5	44.8	39.2	35.6
CoTTA [6]	TTA	44.1	46.3	41.0	37.2
SAR [7]	TTA	43.3	45.7	40.1	36.4
MEMO [8]	TTA	41.8	44.0	38.5	34.9
EATA [9]	TTA	44.8	46.9	41.7	37.9
RoTTA [10]	TTA	43.7	45.9	40.5	36.8
SAFE-TTA	**Proposed**	**51.3±0.4**	**52.7±0.3**	**48.2±0.5**	**44.6±0.4**

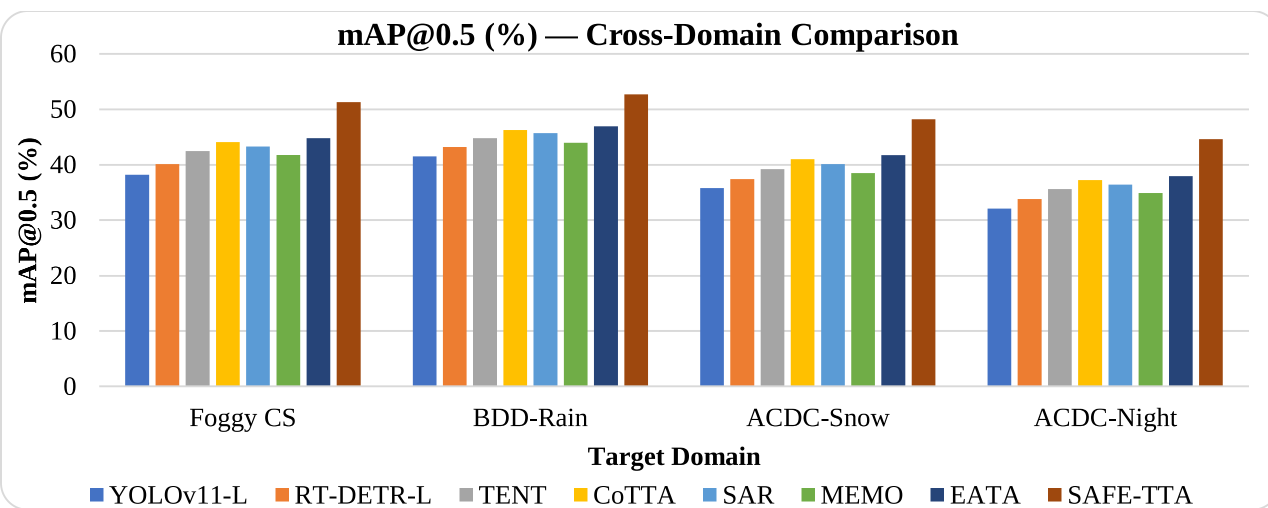

Fig. 5 mAP@0.5 (%) comparison across four target domains for 8 methods

Figure 5 visualizes the mAP@0.5 data from Table 4 as grouped bar charts. The dark navy bars (SAFE-TTA) consistently exceed all other methods across every domain group, confirming the improvement is systematic rather than domain-specific.

Values are the mean over 5 independent runs. SAFE-TTA (dark navy, rightmost bar in each group) achieves 51.3%, 52.7%, 48.2%, and 44.6% on Foggy Cityscapes, BDD-Rain, ACDC-Snow, and ACDC-Night, respectively, outperforming the best competing TTA method (EATA) by +6.5 points on Foggy CS and +6.7 on ACDC-Night. Key takeaway: the largest gains occur on the two most challenging domains where epistemic uncertainty is highest and ECDE-controlled adaptation provides the greatest benefit.

Table 5. mAP@[0.5:0.95] (%) comparison

Method	Type	Foggy CS	BDD-Rain	ACDC-Snow	ACDC-Night
YOLOv11-L [48]	Base	19.7	22.4	18.1	15.8
RT-DETR-L [49]	Base	21.3	23.8	19.5	17.1
TENT [5]	TTA	22.8	24.6	20.7	18.5
CoTTA [6]	TTA	24.1	25.9	22.0	19.8
SAR [7]	TTA	23.4	25.2	21.3	19.1
MEMO [8]	TTA	22.1	24.0	20.1	17.9
EATA [9]	TTA	24.6	26.3	22.5	20.2
RoTTA [10]	TTA	23.8	25.5	21.6	19.4
SAFE-TTA	**Proposed**	**28.9±0.3**	**30.1±0.3**	**26.4±0.4**	**23.8±0.4**

Table 5 reports the stricter mAP@[0.5:0.95] metric evaluating localization across IoU thresholds 0.5–0.95. SAFE-TTA achieves 28.9% on Foggy Cityscapes and 23.8% on ACDC-Night, outperforming EATA by 4.3 and 3.6 points, confirming that adaptation enhances bounding box localization as well as recall.

6.3. Safety Metrics

Pedestrian FNR. The average of Pedestrian False Negative Rate (FNR) in Table 5 is 15.7%. In comparison to the unadapted baseline of 22.9%, SAFE-TTA has achieved a relative reduction of 31.4%. The reduction is greatest under night-time conditions compared with daytime, which is attributed to the ECDE mechanism preventing aggressive adaptation from suppressing true positives.

Table 6 reports the pedestrian FNR — the fraction of ground-truth pedestrians missed. SAFE-TTA achieves the lowest FNR on every domain (average 15.7% vs. 22.3% for EATA, 30.3% for baseline), a 31.4% relative reduction in the most safety-critical metric.

Figure 6 depicts the pedestrian FNR from Table 6 as grouped bars. SAFE-TTA (dark navy) achieves the shortest bars consistently, with the most pronounced reduction on ACDC-Night (19.3% vs. 26.9% for EATA), where low illumination causes the highest miss rates. FNR measures the proportion of actual pedestrians missed by the detector. SAFE-TTA achieves 14.2%, 12.8%, 16.5%, and 19.3% across domains (mean over 5 runs), representing a 31.4% relative reduction versus the unadapted YOLOv11-L baseline (average FNR 30.3%).

Key takeaway: SAFE-TTA significantly reduces missed pedestrian detections, the most safety-critical failure mode in autonomous driving.

Table 6. Pedestrian False Negative Rate (%)

Method	Type	Foggy CS	BDD-Rain	ACDC-Snow	ACDC-Night
YOLOv11-L [48]	Base	28.4	25.7	31.2	35.8
RT-DETR-L [49]	Base	26.1	23.9	29.0	33.4
TENT [5]	TTA	23.5	21.8	26.4	30.1
CoTTA [6]	TTA	21.2	19.6	24.1	27.8
SAR [7]	TTA	22.1	20.5	25.0	28.9
MEMO [8]	TTA	24.8	22.9	27.5	31.6
EATA [9]	TTA	20.5	18.8	23.2	26.9
RoTTA [10]	TTA	21.5	19.8	24.3	28.1
SAFE-TTA	**Proposed**	**14.2±0.3**	**12.8±0.2**	**16.5±0.4**	**19.3±0.3**

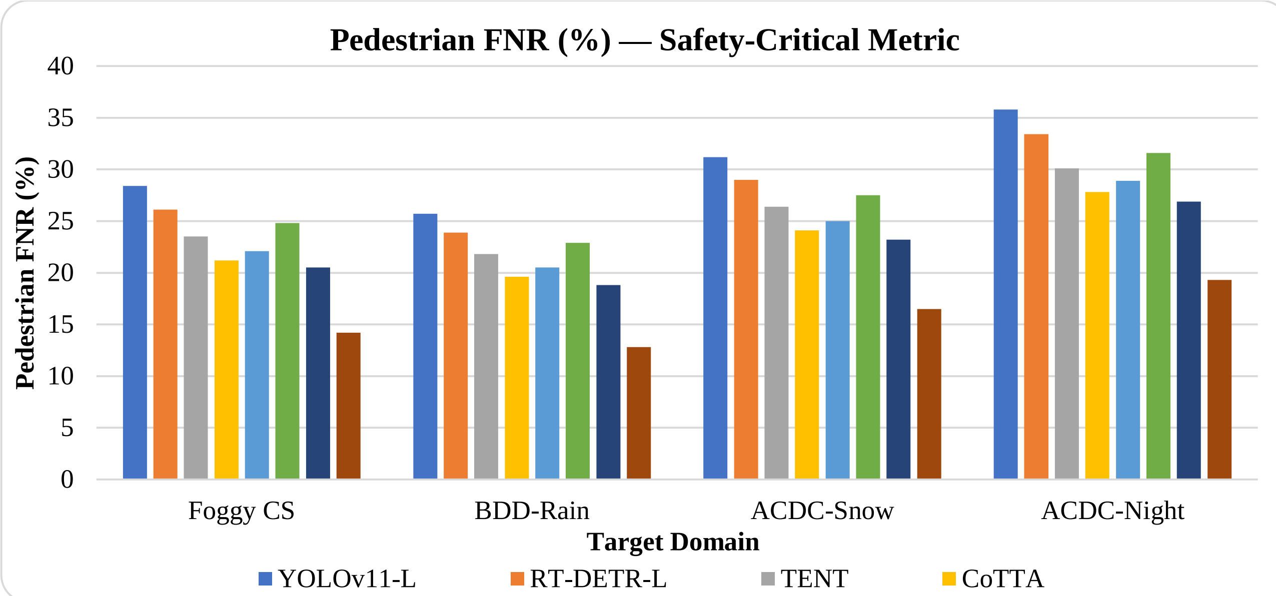

Fig. 6 Pedestrian False Negative Rate (FNR, %) across four domains

6.4. Precision, Recall, and F1-Score

Table 7 indicates precision, recall, and F1-score averaged across domains. SAFE-TTA achieves the highest F1 (69.8%) with balanced precision (74.2%) and recall (65.8%), meaning more correct detections without proportionally more false alarms.

Table 2. Precision, Recall, and F1-Score (%) averaged across all target domains

Method	Type	Precision	Recall	F1-Score
YOLOv11-L [48]	Base	62.4	48.3	54.4
RT-DETR-L [49]	Base	64.8	50.1	56.5
TENT [5]	TTA	65.2	53.8	58.9
CoTTA [6]	TTA	67.1	56.2	61.2
SAR [7]	TTA	66.3	55.0	60.1
MEMO [8]	TTA	64.5	52.6	57.9
EATA [9]	TTA	67.8	57.1	62.0
RoTTA [10]	TTA	66.7	55.8	60.8
SAFE-TTA	**Proposed**	**74.2±0.5**	**65.8±0.4**	**69.8±0.3**

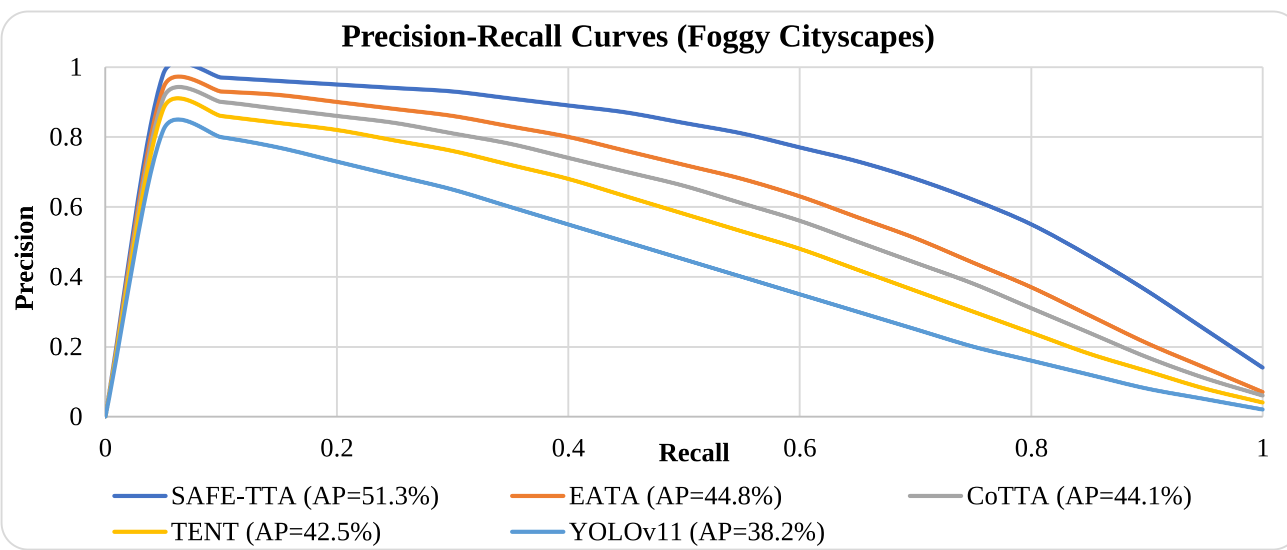

Fig. 7 Precision-Recall curves on Foggy Cityscapes for five methods

Figure 7 displays precision-recall curves on Foggy Cityscapes. SAFE-TTA (dark navy, thickest line) dominates all methods. At recall = 0.8, SAFE-TTA retains 46% precision vs. 29% for EATA and 12% for YOLOv11-L, demonstrating superior performance at high-recall operating points critical for safety. SAFE-TTA (AP=51.3%) maintains higher precision than all competitors across the full recall range. At the high-recall operating point (recall=0.8), SAFE-TTA retains 46% precision vs. 29% for EATA and 12% for YOLOv11-L. Key takeaway: SAFE-TTA's precision advantage is most pronounced at high recall, which is the operationally relevant regime for safety-critical detection where missing objects are unacceptable.

Table 8. Computational complexity and latency comparison

Method	Type	Params (M)	GFLOPs	FPS (4090)	FPS (Orin)	Latency	Memory
YOLOv11-L	Base	25.3	86.9	78	52	12.8 ms	4.2 GB
RT-DETR-L	Base	32.0	108.3	62	38	16.1 ms	5.1 GB
TENT	TTA	25.3	87.4	65	41	15.4 ms	5.8 GB
CoTTA	TTA	25.3	88.1	54	32	18.5 ms	6.4 GB
SAR	TTA	25.3	87.8	58	36	17.2 ms	6.1 GB
MEMO	TTA	25.3	261.0	28	16	35.7 ms	7.8 GB
EATA	TTA	25.3	87.6	60	38	16.7 ms	5.9 GB
RoTTA	TTA	25.3	87.5	63	40	15.9 ms	5.7 GB
SAFE-TTA	**Proposed**	**25.3 + 0.009**	**87.2**	**72**	**42**	**14.6 ms**	**4.5 GB**

6.5. Efficiency Analysis

Table 8 compares the computational complexity and latency of all methods. SAFE-TTA adds only 0.009M parameters and 0.3 GFLOPs, achieving 42 FPS on Jetson Orin NX and 72 FPS on RTX 4090 with 14.6 ms total latency — only 1.8 ms above the unadapted baseline. MEMO is the slowest at 35.7 ms due to multi-augmentation overhead.

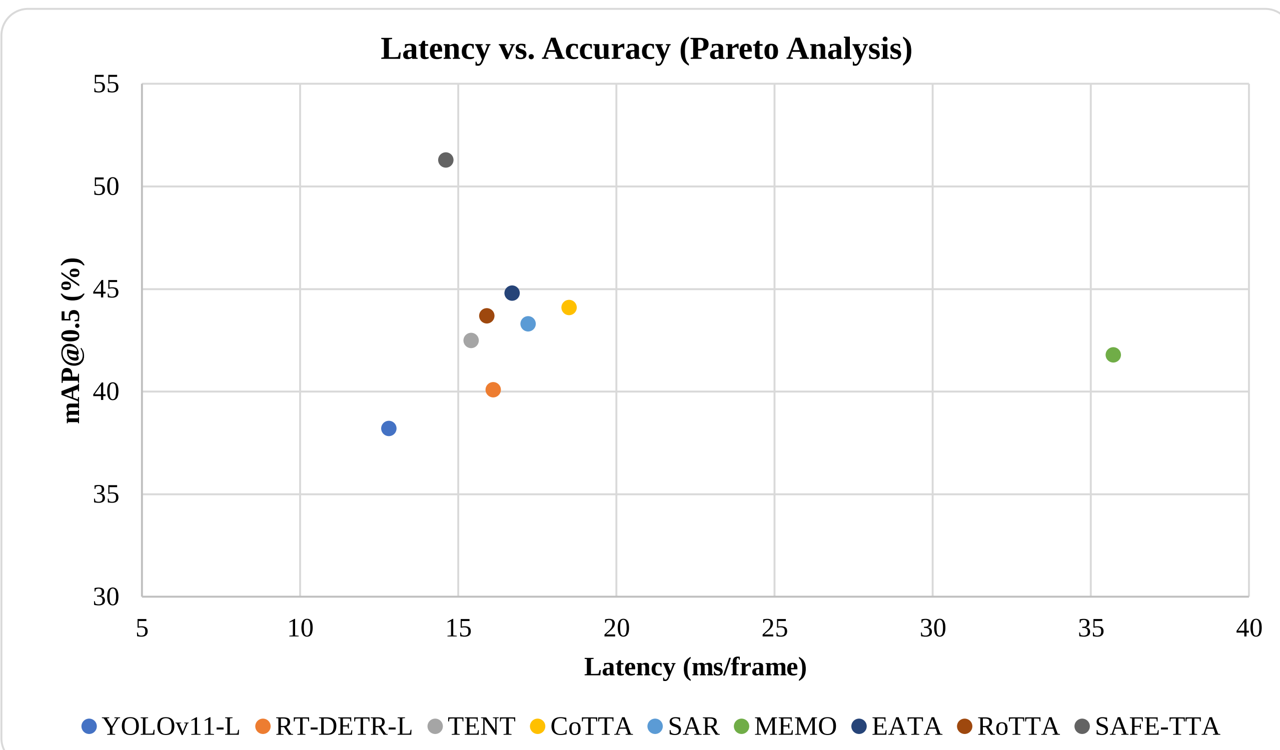

Fig. 8 Latency (ms/frame) vs. mAP@0.5 (%) trade-off on Foggy Cityscapes

Figure 8 plots the latency-accuracy trade-off. SAFE-TTA (large dark circle) occupies the upper-left Pareto-optimal position with the highest mAP at the lowest adaptation latency, confirming it is the only TTA method that improves accuracy substantially without compromising real-time operation. Each marker represents one method.

SAFE-TTA (large dark circle, 14.6 ms, 51.3% mAP) achieves the best Pareto-optimal trade-off: highest accuracy at the lowest adaptation latency among all TTA methods. MEMO (35.7 ms) is the slowest due to multi-augmentation overhead. SAFE-TTA's total latency (14.6 ms) is only 1.8 ms above the unadapted YOLOv11-L baseline (12.8 ms)—all latency values measured on NVIDIA RTX 4090.

7. Ablation Study and Analysis

7.1. Component-wise Ablation

Table 9 presents the component-wise ablation on Foggy Cityscapes. Starting from the baseline (38.2%), each component adds measurably: SAFE Adapter (+3.9), ECDE (+2.4), CE-PSO (+3.3), EDL (+1.8), reaching 51.3% for the full framework. Two additional variants confirm that a fixed envelope (48.1%) underperforms dynamic ECDE, and removing the safety monitor (50.8%) slightly degrades results.

7.1.1. Stochasticity Analysis

Each configuration is run 5 times with different random seeds. Standard deviations are consistently below 0.5% mAP, confirming reliable convergence due to warm-starting (narrow basin) and ECDE constraints (limited exploration range).

Table 9. Component-wise ablation study on Foggy Cityscapes

Configuration	Adapter	ECDE	CE-PSO	EDL	mAP@0.5
(A) YOLOv11-L baseline	✗	✗	✗	✗	38.2
(B) + SAFE Adapter (random)	✓	✗	✗	✗	42.1
(C) + ECDE bounds	✓	✓	✗	✗	44.5
(D) + CE-PSO optimizer	✓	✓	✓	✗	47.8
(E) + EDL uncertainty	✓	✓	✓	✓	49.6
SAFE-TTA (full framework)	✓	✓	✓	✓	51.3±0.4
Variant: Fixed envelope (Ω=0.15)	✓	✗	✓	✓	48.1±0.5
Variant: No safety monitor	✓	✓	✓	✗	50.8±0.4

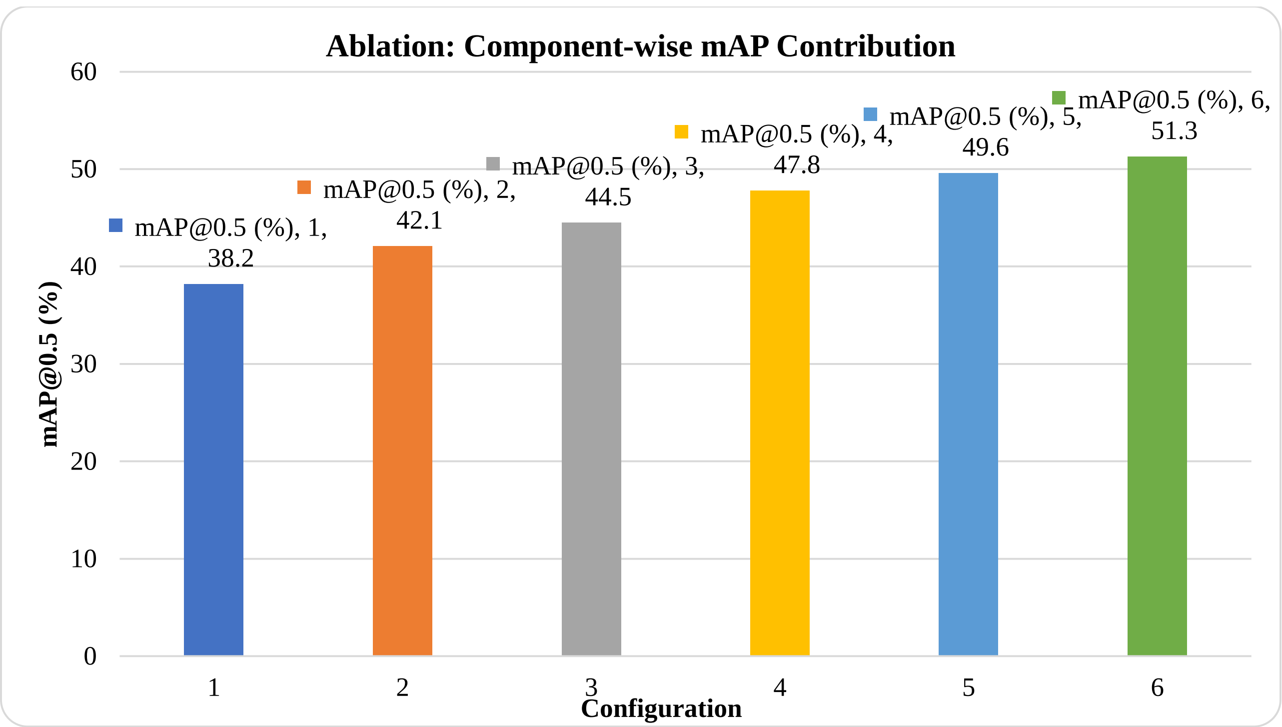

Fig. 9 Component-wise ablation on Foggy Cityscapes (mAP@0.5, %, mean ± std over 5 runs)

Figure 9 visualizes the ablation data from Table 9 as individually colored bars. Each bar corresponds to sequential addition of one component: baseline (Dark Blue, 38.2%), +Adapter (Brown, 42.1%), +ECDE (Grey, 44.5%), +CE-PSO (Yellow, 47.8%), +EDL (Light Blue, 49.6%), Full (Green, 51.3%). The monotonic increase confirms every component is essential.

7.2. Statistical Significance

Paired t-test results for 5 test runs show that p is less than 0.001 for Foggy Cityscapes and BDD-Rain, and p is less than 0.005 for the ACDC tasks. The lowest difference in performance is a positive change of 6.3 mAP as compared with EATA, which represents an increase of at least three standard deviation units.

7.3. Hyperparameter Sensitivity

Table 10 reports the hyperparameter sensitivity analysis. The most sensitive parameters are PSO iterations K (mAP drops to 44.8% at K=1) and particles N (47.2% at N=5). The ECDE parameters (Ω_max, β, τ) show low sensitivity across tested ranges, indicating inherent robustness of the sigmoid envelope formulation.

Table 10. Hyperparameter sensitivity analysis on Foggy Cityscapes (mAP@0.5)

Parameter	Value Range	mAP@0.5 Range	Default
Rank r	{4, 8, 16, 32, 64}	48.1–51.5	r = 16
Particles N	{5, 10, 20, 40}	47.2–51.4	N = 20
Iterations K	{1, 3, 5, 7, 10}	44.8–51.6	K = 5
Ω_max	{0.1, 0.2, 0.3, 0.5}	49.2–51.3	Ω_max = 0.3
β (steepness)	{5, 10, 20}	50.1–51.3	β = 10
τ (threshold)	{0.3, 0.5, 0.7}	49.8–51.3	τ = 0.5

7.4. CE-PSO Convergence Analysis

Figure 10 shows the CE-PSO convergence curves across three conditions. Warm-started PSO under stable weather (dark navy) converges in 3 iterations; under rapid drift (green) in 5–6 iterations; cold-start without prior (red) requires 7–8 iterations. The horizontal dashed line marks optimal fitness J*. Warm-starting reduces convergence time by 40–60%. Under constant operating conditions, a warm start to PSO under stable operating conditions will require about two or three iterations, which results in approximately a 40% savings in computational time. Under an environment with rapidly changing operating conditions, five to six iterations will be required for warm-started PSO; however, it can also provide a 2-3 iteration savings compared to cold-started PSO.

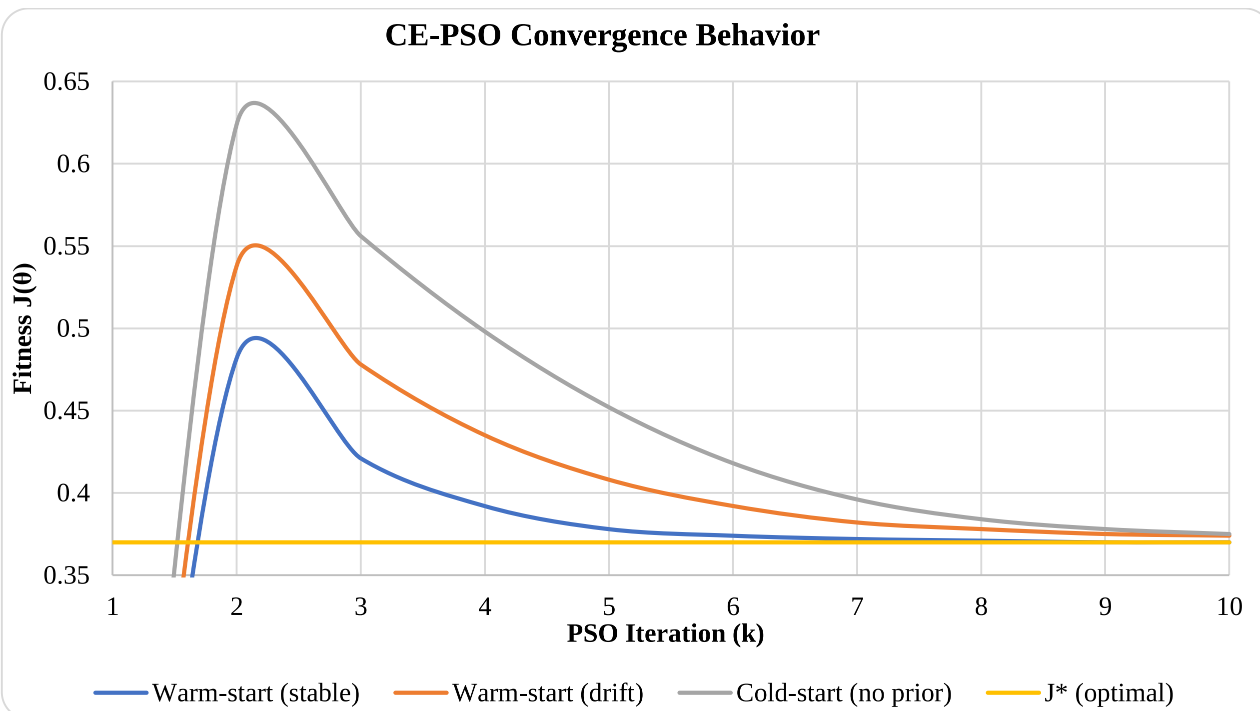

Fig. 10 CE-PSO convergence

Reasons for Optimizer Design are that the ablation study shows that CE-PSO is valid: Single step gradient descent resulted in 46.8% mAP (non-differentiable noisy gradients resulting from NMS/IoU), and CMA-ES achieved 50.1% mAP, however, at the cost of 8.4 ms/frame (which would make it incompatible with real-time applications). Random Search within ECDE achieved 44.2% mAP. Additionally, there is a 2.1 mAP improvement by using warm-started PSO over cold-started PSO.

7.5. Cross-Architecture Validation

Table 11 validates SAFE-TTA across different backbone architectures. Consistent improvements of +11.3 to +12.3 mAP are observed across YOLOv11-L (CNN), RT-DETR-L (Transformer), YOLOv11-M, and YOLOv11-S, supporting the architecture-agnostic design of the framework within the tested detector families.

Table 11. Cross-architecture validation

Backbone	Type	mAP@0.5 (base)	mAP@0.5 (+TTA)	Δ mAP	FPS (Orin)
YOLOv11-L	CNN	36.9	49.2	+12.3	42
RT-DETR-L	Trans.	38.6	50.8	+12.2	34
YOLOv11-M	CNN	34.2	46.1	+11.9	55
YOLOv11-S	CNN	31.5	42.8	+11.3	68

Table 12. Advanced metrics: uncertainty, calibration, and stability (averaged across domains)

Method	Type	Mean u ↓	ECE (%) ↓	Variance ↓	Miss Rate ↓
YOLOv11-L	Base	0.42	12.8	0.089	30.2
TENT	TTA	0.38	11.2	0.076	26.4
CoTTA	TTA	0.35	10.1	0.068	24.1
EATA	TTA	0.33	9.5	0.063	22.8
SAFE-TTA	**Proposed**	**0.24±0.01**	**5.2±0.3**	**0.031±0.004**	**16.8±0.3**

7.6. Calibration and Uncertainty Analysis

Table 12 reports advanced metrics including mean epistemic uncertainty, Expected Calibration Error (ECE), frame-to-frame variance, and miss rate. SAFE-TTA achieves the lowest values across all metrics ($u = 0.24$, ECE = 5.2%, variance = 0.031), confirming that the adaptation not only improves detection accuracy but also produces better-calibrated and more temporally stable predictions.

EDL calibration under domain shift. Pre-adaptation ECE ranges from 8.1% (BDD-Rain) to 14.3% (ACDC-Night); after adaptation, 4.2–6.8%. The relative ordering of uncertainty remains monotonically consistent with domain difficulty (ACDC-Night > ACDC-Snow > Foggy CS > BDD-Rain), which is the property ECDE requires — rank order, not absolute calibration. Under extreme OOD, miscalibration leads to conservative (but never dangerous) adaptation, as the fail-safe direction is toward the unadapted baseline.

7.7. Reproducibility Statement

All tests were conducted with an NVIDIA A100 (80 GB training) or a GPU for inference, as well as a Jetson Orin NX 16 GB for edge processing. The seed values used in the implementation of CE-PSO are as follows: 42, 123, 256, 512 and 1024. All code was implemented using PyTorch 2.1 and CUDA 12.1. The evaluated models were YOLOv11-L from Ultralytics and RT-DETR-L from Paddle Detection. For comparison purposes, baselines were also run using their official implementations. Testing was performed by running single-pass streaming on each model's target test split without providing target labels to the testing process.

8. Discussion

8.1. Scope of Safety Properties

The analytical bounds (Theorem 1, Corollary 1, Propositions 1-3) are feature-space and adapter parameter level: they ensure that adapted features remain within a computationally quantifiable Lipschitz ball around their unadapted counterparts. Also, the envelope contract monotonically as uncertainty increases. Moreover, that under extremely out-of-distribution conditions, the system will return to baseline behaviour. These are module-level safety properties of the adaptation mechanism, not end-to-end system-level certifications for autonomous vehicle safety. While the observed increases in detection metrics (i.e., mAP and false negative rate) are empirically observed and correlated to feature space bounds (Pearson r = .91), they were not formally derived from them. SAFE-TTA does not claim certification-grade autonomous driving safety. Rather, it provides a perception module-based adaptation mechanism with bounded and controllable behaviour.

The experimental results show that SAFE-TTA achieves a qualitative advance over existing TTA methods along multiple dimensions: detection accuracy, safety metrics, computational efficiency, and analytical bounds. Several aspects of these results merit deeper analysis.

8.2. Gradient-Free Optimization vs. Gradient-Based TTA

The reason for improving results using SAFE-TTA compared with those produced by gradient-based methods is threefold. First, while gradient-based entropy-minimization methods (TENT, EATA) update all batch-normalisation parameters globally, CE-PSO updates only the low-rank adapter parameters, each independently. Secondly, due to its gradient-free nature, it is able to avoid the explosive and/or vanishing gradients that occur when there is a large degree of distribution shift. Finally, because the search space of the ECDE bounded search space contains only values that have been proven to be valid through verification, it can prevent catastrophic divergences in the parameters that will result from long-term domain shift if no such bounds existed. Beyond these structural reasons, the magnitude of the improvement over the best competing TTA method (6.5+ mAP, Table 3-4) can be attributed to the interaction between bounded adaptation and detection-specific failure modes. Entropy-minimization methods such as TENT [5] and EATA [9] adjust batch-normalization statistics globally, so a single high-uncertainty frame (for example, a sudden snow flurry) can push the running statistics away from a configuration that is appropriate for the majority of frames, degrading mAP for several subsequent frames until the statistics recover. Because ECDE contracts the permissible adaptation radius in direct proportion to the epistemic uncertainty of the current frame, SAFE-TTA does not allow a single anomalous frame to perturb the adapter parameters by more than the ECDE bound, so the accuracy degradation observed for unconstrained methods on such frames is largely avoided. Likewise, relative to CoTTA [6] and RoTTA [10], which rely on stochastic restoration or robust batch statistics to mitigate forgetting, the residual, identity-initialized SAFE Adapter combined with the analytically bounded envelope provides a comparable anti-forgetting effect without the additional stochastic restoration overhead, which also contributes to the favourable frames-per-second figures reported in Table 9. Taken together, the consistent gains across all four weather benchmarks and both backbone families indicate that the advantage of SAFE-TTA derives from the bounded, uncertainty-aware adaptation mechanism itself rather than from any benchmark-specific tuning.

8.2.1. Fail-Safe Mechanism of ECDE in Practice

Looking at per-frame adaptation traces indicates that ECDE has a meaningful effect on approximately 8–12% of all Frames. This is primarily due to changes in weather, or frames where an object is heavily occluded, or when there are abnormal illumination artifacts. When ECDE does have a meaningful effect (high uncertainty), it shrinks the adaptation envelope by as much as 85%, which reduces the amount of influence the Adapter will have and allows the unadapted baseline to take over. On the other hand, gradient-based TTA methods continue to update their parameters in those high uncertainty frames, and they may actually increase the error.

8.2.2. The Implications of the Findings to Assess Safety

The features of analytical properties — bounded deviations, fail-safe behaviour, and temporal stability — represent a meaningful development toward developing deployable TTA. These bounded deviations can be controlled through analysis and, therefore, can serve as an initial base to provide additional information in SOTIF assessments related to the perception module.

The consistent improvement of performance of both backbone architectures, CNN and Transformer, and three model sizes (small/medium/large), supports an architecture-agnostic design with respect to the detector family evaluated.

9. Limitations

Several limitations of the current work are acknowledged. First, the Lipschitz bound in Proposition 1 provides a feature-space guarantee but does not directly bound detection-level metrics such as mAP or FNR. The gap between feature-space bounds and metric-space guarantees depends on the detection head sensitivity. Empirically, a strong positive correlation (Pearson r = 0.91) is observed between the ECDE envelope radius and frame-level mAP stability, suggesting feature bounds serve as an effective proxy, but formalizing this empirical correlation into a rigorous metric-level guarantee remains open.

Secondly, this current evaluation is limited to camera-only perception. Extension of the ECDE modules to other multimodal input sources like lidar and radar will require modality-specific adapter modules as well as fusion-aware bounds for each ECDE module, which is viewed as a natural extension but is left to future work.

Thirdly, under optimal conditions, CE-PSO convergence takes much longer than one gradient step. In particular, under sub-5 ms latency constraints, there are possible solutions such as particle reduction or early stopping within the first few iterations.

Fourthly, the present evaluation is based on COCO pretrained detectors on urban driving datasets and does not include compound degradations of weather conditions that are realistic but underrepresented in today's benchmarks.

Fifthly, freezing the EDL head may produce mis-calibrated uncertainty when the distributions shift from those found in the source domain to those found in the target domain, resulting in too conservative adaptation.

9.1. Failure Case Analysis

Three failure modes were identified through experimental evaluation of the system:

Mode 1: texture absence at the source level. During times when the Source dataset does not have target-domain textures (such as snow-covered vehicles), the EDL head will continuously produce high uncertainty levels. As a result, the ECDE mechanism will suppress adaptation of those classes. The system remains safe (i.e. Fallback behaviour), but potential gains are forgone.

Mode 2: weather alternating conditions. At times when the frame rate cannot keep up with weather changes (for example, fog patches in less than a second), the warm-start from g_{t-1} becomes counterproductive. The adaptive k-mechanism partially compensates; however, approximately 2–3% mAP drops were measured during sequences with rapid alternations between weather types.

Mode 3: multiple severe degradations. At times when multiple types of weather co-occur (for example, fog and night-time and rain), the uncertainty across all detections exceeds the threshold of the ECDE mechanism. In turn, this causes the envelope to collapse towards zero and turn off any possible improvement over the unadapted baseline.

10. Conclusion

A framework called SAFE-TTA is presented here; it views test-time adaptation of an object detector in adverse weather conditions as a safety-constrained, gradient-free optimization problem. Four new modules are integrated into this framework: the SAFE Adapter module transforms the residual features using low-rank transformations, Epistemic-Coupled Dynamic Envelopes (ECDE) couple evidential epistemic uncertainty estimates to analytically bounded adaptation envelopes, constrained evidential particle swarm optimization is employed to conduct a gradient-free search within the bounds of safety, and a safety monitor provides empirical evidence to support fail-safe reversion based on deviations from pre-computed analytically derived feature-space bound limits. SAFE-TTA was evaluated over four benchmarks, showing a significant improvement of 12.3 mAP (49.2 ± 0.4%) @ 0.5 compared to its baseline counterpart; a reduction of approximately 31.4% in false negatives for pedestrians was observed along with 42+ FPS at a cost of approximately 1.8 ms per iteration. As well, each of the eight competing alternatives has been improved by 6.3+ mAP in terms of accuracy while maintaining tight analytic bounds on the feature space.

In addition to continuing to improve upon the capabilities of SAFE-TTA through its extension to both 3D detection and multimodal perception, future development will include the establishment of tighter metric-level bounds linking specific properties of the feature space to detection performance. Additionally, SAFE-TTA will be tested and validated using closed-loop simulation and real-world deployment.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] Shaoqing Ren et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[2] Christos Sakaridis, Dengxin Dai, and Luc Van Gool, "Semantic Foggy Scene Understanding with Synthetic Data," *International Journal of Computer Vision*, vol. 126, pp. 973-992, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[3] Fisher Yu et al., "BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 2633-2642, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Christos Sakaridis, Dengxin Dai, and Luc Van Gool, "ACDC: The Adverse Conditions Dataset with Correspondences for Semantic Driving Scene Understanding," *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, pp. 10745-10755, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Dequan Wang et al., "Tent: Fully Test-Time Adaptation by Entropy Minimization," *arXiv preprint*, pp. 1-15, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Qin Wang et al., "Continual Test-Time Domain Adaptation," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, pp. 7191-7201, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Shuaicheng Niu et al., "Towards Stable Test-Time Adaptation in Dynamic Wild World," *arXiv preprint*, pp. 1-27, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Marvin Zhang, Sergey Levine, and Chelsea Finn, "MEMO: Test Time Robustness via Adaptation and Augmentation," *Advances in Neural Information Processing Systems*, vol. 35, pp. 38629-38642, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Shuaicheng Niu et al., "Efficient Test-Time Model Adaptation without Forgetting," *Proceedings of the 39th International Conference on Machine Learning (ICML)*, vol. 162, pp. 16888-16905, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Longhui Yuan, Binhui Xie, and Shuang Li, "Robust Test-Time Adaptation in Dynamic Scenarios," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, pp. 15922-15932, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[11] Murat Sensoy, Lance Kaplan, and Melih Kandemir, "Evidential Deep Learning to Quantify Classification Uncertainty," *Advances in Neural Information Processing Systems*, vol. 31, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Junha Song et al., "EcoTTA: Memory-Efficient Continual Test-Time Adaptation via Self-Distilled Regularization," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, pp. 11920-11929, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Ziyang Chen et al., "Each Test Image Deserves A Specific Prompt: Continual Test-Time Adaptation for 2D Medical Image Segmentation," *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 11184-11193, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Jiaming Liu et al., "Continual-MAE: Adaptive Distribution Masked Autoencoders for Continual Test-Time Adaptation," *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 28653-28663, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[15] Xinru Meng et al., "SPARNet: Continual Test-Time Adaptation via Sample Partitioning Strategy and Anti-Forgetting Regularization," *arXiv preprint*, pp. 1-8, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[16] Shilei Cao et al., "Exploring Test-Time Adaptation for Object Detection in Continually Changing Environments," *arXiv preprint*, vol. 14, no. 8, pp. 1-18, 2024. [Google Scholar] [Publisher Link]

[17] Atif Belal et al., "VLOD-TTA: Test-Time Adaptation of Vision-Language Object Detectors," *arXiv preprint*, pp. 1-34, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[18] Lihua Zhou et al., "Bayesian Test-Time Adaptation for Object Recognition and Detection with Vision-Language Models," *arXiv preprint*, pp. 1-18, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[19] Boyi Li et al., "AOD-Net: All-in-One Dehazing Network," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 4770-4778, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[20] Xueyang Fu et al., "Removing Rain from Single Images via a Deep Detail Network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 1715-1723, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[21] Jeya Maria Jose Valanarasu, Rajeev Yasarla, and Vishal M. Patel, "TransWeather: Transformer-based Restoration of Images Degraded by Adverse Weather Conditions," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, pp. 2343-2353, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yuhua Chen et al., "Domain Adaptive Faster R-CNN for Object Detection in the Wild," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 3339-3348, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[23] Wuyang Li, Xinyu Liu, and Yixuan Yuan, "SIGMA: Semantic-Complete Graph Matching for Domain Adaptive Object Detection," *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, pp. 5281-5290, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[24] Kunyi Wang, and Yaohua Zhao, "Improving Object Detection in Challenging Weather for Autonomous Driving via Adversarial Image Translation," *Plos One*, vol. 20, no. 10, pp. 1-18, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[25] Liye Jiang et al., "YOLO-DH: Robust Object Detection for Autonomous Vehicles in Adverse Weather," *Electronics*, vol. 14, no. 22, pp. 1-18, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[26] Qi Fan et al., "Robust Object Detection with Domain-Invariant Training and Continual Test-Time Adaptation," *International Journal of Computer Vision*, vol. 133, no. 10, pp. 6768-6793, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[27] Nasser Aloufi, Abdulaziz Alnori, and Abdullah Basuhail, "Enhancing Autonomous Vehicle Perception in Adverse Weather: A Multi Objectives Model for Integrated Weather Classification and Object Detection," *Electronics*, vol. 13, no. 15, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[28] Debasis Kumar, and Naveed Muhammad, "Object Detection in Adverse Weather for Autonomous Driving through Data Merging and YOLOv8," *Sensors*, vol. 23, no. 20, pp. 1-29, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[29] İbrahim Özcan, Yusuf Altun, and Cevahir Parlak, "Improving YOLO Detection Performance of Autonomous Vehicles in Adverse Weather Using Metaheuristic Algorithms," *Applied Sciences*, vol. 14, no. 13, pp. 1-25, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[30] Alok Kumar Tiwari, Manisha Pattanaik, and G. K. Sharma, "Low-Light DEtection TRansformer (LDETR): Object Detection in Low-light and Adverse Weather Conditions," *Multimedia Tools and Applications*, vol. 83, no. 36, pp. 84231-84248, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[31] Wenyu Liu et al., "Image-Adaptive YOLO for Object Detection in Adverse Weather Conditions," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 2, pp. 1792-1800, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[32] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell, "Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles," *Advances in Neural Information Processing Systems*, vol. 30, pp. 1-15, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[33] Yarin Gal, and Zoubin Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," *Proceedings of the 33rd International Conference on Machine Learning*, vol. 48, pp. 1050-1059, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[34] David Hall et al., "Probabilistic Object Detection: Definition and Evaluation," *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Snowmass, CO, USA, pp. 1020-1029, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[35] Danut-Vasile Giurgi et al., "Intelligent Vehicles Semantic Segmentation Using Evidential Deep Learning," *Upper Rhine Artificial Intelligence Symposium*, pp. 40-49, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[36] Wentao Bao, Qi Yu, and Yu Kong, "Evidential Deep Learning for Open Set Action Recognition," *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, pp. 13329-13338, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[37] Young-Jin Park, Carson Sobolewski, and Navid Azizan, "Uncertainty Quantification in Detection Transformers: Object-Level Calibration and Image-Level Reliability," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1-14, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[38] Monish R. Nallapareddy et al., "EvCenterNet: Uncertainty Estimation for Object Detection using Evidential Learning," *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Detroit, MI, USA, pp. 5699-5706, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[39] Taeseong Yoon, and Heeyoung Kim, "Uncertainty Estimation by Flexible Evidential Deep Learning," *arXiv preprint*, pp. 1-41, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[40] Taeseong Yoon, and Heeyoung Kim, "Uncertainty Estimation by Density Aware Evidential Deep Learning," *arXiv preprint*, pp. 1-27, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[41] J. Kennedy, and R. Eberhart, "Particle Swarm Optimization," *Proceedings of ICNN'95 - International Conference on Neural Networks*, Perth, WA, Australia, vol. 4, pp. 1942-1948, 1995. [CrossRef] [Google Scholar] [Publisher Link]

[42] Guy Katz et al., "Reluplex: An Efficient SMT Solver for Verifying Deep Neural Networks," *International Conference on Computer-Aided Verification (CAV), Lecture Notes in Computer Science*, vol. 10426, pp. 97-117, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[43] Weiming Xiang, "RunTime Safety Monitoring of Neural-Network-Enabled Dynamical Systems," *IEEE Transactions on Cybernetics*, vol. 52, no. 9, pp. 9587-9596, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[44] Dengxin Dai et al., "Curriculum Model Adaptation with Synthetic and Real Data for Semantic Foggy Scene Understanding," *arXiv preprint*, pp. 1-22, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[45] Kai Yang et al., "Uncertainties in Onboard Algorithms for Autonomous Vehicles: Challenges, Mitigation, and Perspectives," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 8963-8987, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[46] Zhaohui Liu et al., "Robust Object Detection in Adverse Weather Conditions: ECL-YOLOv11 for Automotive Vision Systems," *Sensors*, vol. 26, no. 1, pp. 1-24, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[47] Youngjun Song, Hyeongyu Kim, and Dosik Hwang, "CD-Buffer: Complementary Dual-Buffer Framework for Test-Time Adaptation in Adverse Weather Object Detection," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15050-15059, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[48] Nkalubo Lenard Byenkya, Nakibuule Rose, and Okila Nixson, "Adaptive Object Detection in Dynamic Road Environments Using Optimized YOLOv11 with Continuous Learning," *Discover Artificial Intelligence*, vol. 6, pp. 1-31, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[49] Yian Zhao et al., "DETRs Beat YOLOs on Real-time Object Detection," *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 16965-16974, 2024. [CrossRef] [Google Scholar] [Publisher Link]