

# Fourier Domain Adaptation for Traffic Light Detection in Adverse Weather\*

Ishaan Gakhar<sup>1</sup>, Aryaman Gupta<sup>1†</sup>, Aryesh Guha<sup>2†</sup>, Amit Agarwal<sup>3</sup>, Ujjwal Verma<sup>4</sup>

<sup>1</sup> Manipal Institute of Technology, Manipal Academy of Higher Education, India

<sup>2</sup> Dept. of Electrical and Electronics Engineering, MIT, MAHE, India

<sup>3</sup> Enterprise Analytics & Data Science AI, Center of Excellence, Wells Fargo ISPL

<sup>4</sup> Dept. of Electronics and Communication Engineering, MIT, MAHE, India

{ishaangakhar04@gmail.com, aryamangupta004@gmail.com, aryeshguha1510@gmail.com,  
 aagarwal3@cs.iitr.ac.in, ujjwal.verma@manipal.edu}

## Abstract

*Traffic light detection under adverse weather conditions remains largely unexplored in ADAS systems, with existing approaches relying on complex deep learning methods that introduce significant computational overheads during training and deployment. This paper proposes Fourier Domain Adaptation (FDA), which requires only training data modifications without architectural changes, enabling effective adaptation to rainy and foggy conditions. FDA minimizes the domain gap between source and target domain by manipulation of frequency components which allows to reduce the high frequency adverse weather effects of fog and rain, creating a dataset for reliable performance under adverse weather.*

*Since low-frequency components encode global structural information while high-frequency components capture finer details, this transformation helps the model retain structural consistency while adapting to new environmental conditions. By training models with FDA-augmented data and labels of the source domain (target domain labels are absent), the model becomes more robust to domain shifts caused by adverse weather conditions. This method is especially helpful for mitigating the effects of rain or fog, as it can be challenging to find data with these conditions along with proper annotations. The source domain merged LISA and S<sup>2</sup>TLTD datasets, processed to address class imbalance. Established methods simulated rainy and foggy scenarios to form the target domain. Semi-Supervised Learning (SSL) techniques were explored to leverage data more effectively, addressing the shortage of comprehensive datasets and poor performance of state-of-the-art models under hostile weather.*

*Experimental results show FDA-augmented models outperform baseline models across mAP50, mAP50-95, Precision, and Recall metrics. YOLOv8 achieved a 12.25% average increase across all metrics. Average improvements of*

*7.69% in Precision, 19.91% in Recall, 15.85% in mAP50, and 23.81% in mAP50-95 were observed across all models, demonstrating FDA's effectiveness in mitigating adverse weather impact. These improvements enable real-world applications requiring reliable performance in challenging environmental conditions.*

## 1. Introduction

1

Environmental perception, along with behaviour decision and motion control, is an essential task in the domain of autonomous driving. With artificial intelligence and computer vision techniques playing integral roles in autonomous driving, object detection plays a crucial role in understanding surrounding environmental scenarios [26, 30]. Adverse weather conditions compromise sensor accuracy and real-time performance, resulting in image degradation issues. These weather anomalies affect the object detection performance of autonomous vehicles [16, 24].

Detecting and recognizing traffic lights is important for ensuring safety in the field of autonomous driving. Vision-based traffic light detection and recognition in traffic scenes is still a challenge [1, 29] to be successfully overcome by the autonomous driving industry, especially when there are disturbances due to weather conditions that contribute to the degradation in the performance of object detection models. State-of-the-art (SOTA) object detection algorithms, like YOLOv5 [9], YOLOv6 [13], YOLOv8 [10] and YOLOv10 [23], perform well in clear weather conditions. However, when used in detrimental weather conditions, such as en-

<sup>†</sup> indicates equal contribution. \*The views expressed in this article are of the authors and do not represent the views of Wells Fargo. This article is for informational purposes only. Nothing contained in the article should be constructed as investment advice. Wells Fargo makes no express or implied warranties and expressly disclaims all legal, tax and accounting implications related to this article.

vironments with rain or fog, a significant deterioration in their performance is observed, as mentioned in [12]. Consequently, there is a pressing need for methods that enable these models to adapt to the variability introduced by challenging weather conditions, ideally without relying on annotated labels as a prerequisite.

Unsupervised Domain Adaptation (UDA) is a form of transfer learning that involves adapting a model trained on one domain that has an abundance of annotations (source domain) to demonstrate results in a related yet distinct domain (target domain) characterized by a different data distribution that lacks annotations. Training the model solely within the source domain fails to produce satisfactory outcomes owing to the occurrence of covariate shift, as mentioned in [28]. The primary objective of domain adaptation techniques is to facilitate knowledge transfer from a well-labelled source domain to a target domain that lacks labels. This process is essential as minor alterations in low-level statistics notably impact the model's effectiveness. This is particularly relevant for leveraging existing datasets with annotations to enhance performance on related tasks that lack such annotations, like Advanced Driver Assistance Systems (ADAS), security and surveillance systems, etc. Domain adaptation thus plays a pivotal role in addressing the challenge of data scarcity in supervised learning scenarios by bridging the discrepancy between different but related domains.

This work proposes Fourier Domain Adaptation (FDA) [28], a non-parametric domain adaptation algorithm. Here, the Fast Fourier Transform (FFT) of each input image (clear weather) in the source domain is computed along with the FFT of a randomly chosen target image (rainy or foggy). Then, the low-level frequencies of the target image replaces the low-level frequencies of the source image, before the images are reconstructed using Inverse FFT (iFFT). The source domain is formed by merging two widely used datasets, LISA [7] and SJTU: Small Traffic Light Dataset (S<sup>2</sup>TLD) [27], and employing various augmentations to address the class imbalance against images containing yellow traffic lights. By Leveraging pre-existing techniques for fog addition [21] and rain addition (<https://github.com/ShenZheng2000/Rain-Generation-Python>), the target domain for traffic light detection in rainy and foggy scenarios was created. This enables models to learn domain-invariant features, enhancing their ability to detect traffic lights compared to models trained on clear weather conditions as observed. State-of-the-art UDA algorithms often entail training a deep neural network, which undergoes difficult adversarial training to make the model invariant in the binary selection of source/target domain [28] whereas FDA requires no trainable parameters or complex architectures. It facilitates the learning of domain-invariant features by introducing variability, which helps

the model learn high-level characteristics without relying on low-level characteristics from the sensor, the illuminant, or other sources of low-level variability. Therefore, this non-parametric, traditional signal processing algorithm can be used as an alternative to dense deep-learning based, computationally heavy, parametric state-of-the-art unsupervised domain adaptation (UDA) algorithms [6]. This serves as motivation to apply FDA to weather-affected images. The source dataset was benchmarked using various state-of-the-art methods, and their performance was qualitatively and quantitatively evaluated across different metrics.

In order to tackle the challenge posed by the limited availability of annotated datasets for traffic light detection in hostile weather, semi-supervised learning (SSL) techniques were employed. SSL, which leverages both labelled and unlabeled data, is particularly advantageous in situations where obtaining a comprehensive labelled dataset is costly or logistically difficult. By integrating a smaller set of accurately labelled traffic light images with a larger pool of unlabeled images, SSL allows leveraging the already limited data and demonstrates that even 50% of the available data can be used to produce comparable results. This approach aims to mitigate data scarcity inherent in the domain. Through the application of semi-supervised learning, it becomes feasible to develop more effective algorithms that can accurately identify and classify traffic lights in diverse driving conditions at much less cost, ultimately contributing to advancements in intelligent transportation systems. Hence, the contributions in this work can be summarized as:

- Proposing and demonstrating the effectiveness of a non-parametric method of Fourier Domain Adaptation for Traffic Light detection in adverse weather conditions requiring no pretraining or dense architectures.
- Leveraging Semi-Supervised Learning in the realm of Traffic Light Detection as a potential solution to the lack of datasets in the domain and demonstrating comparable performance with only 50% labelled data.
- Combining widely used datasets and mitigating class imbalance to synthetically generate realistic rainy and foggy images employing effective pre-existing methods.

## 2. Related Works

### 2.1. Object Detection

In recent years, significant advancements have been made in object detection, driven by the development of deep learning-based techniques. One of the most influential works is the YOLO (You Only Look Once) series, introduced by [18], which revolutionized object detection by achieving real-time performance while maintaining high accuracy. YOLOv3 [17] and its successors, YOLOv5 [9] and YOLOv8 [10], have further improved detection capabilities

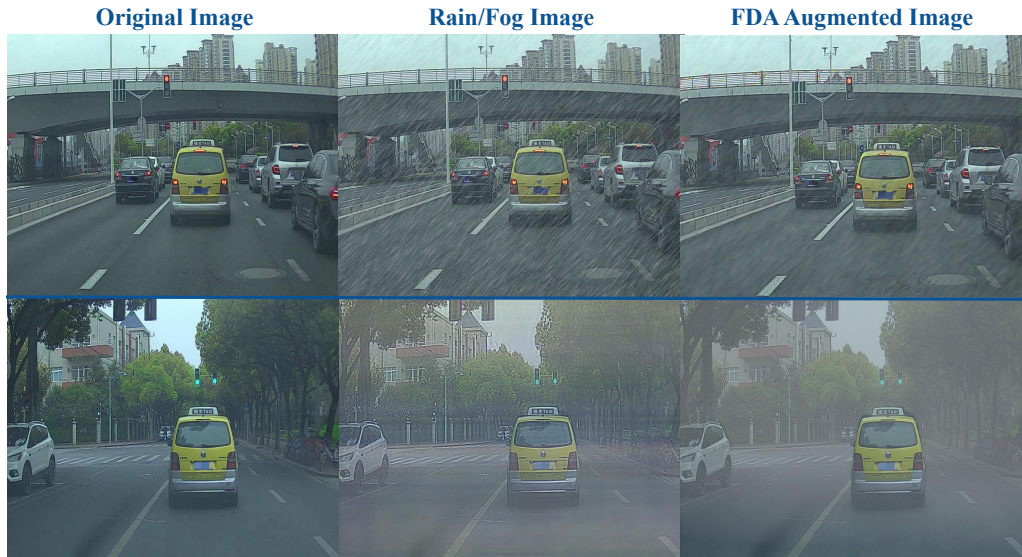


Figure 1. Impact of FDA on Rainy and Noisy images. The first column contains original images, the 2nd column adds rain/fog to the corresponding image and the 3rd column presents the effects of FDA on the images of the 2nd column. Each row represents a different instance from the dataset. As evident, applying FDA to the images increases visibility and the ability to detect the traffic light.

by incorporating advanced features and optimization techniques. Additionally, Mask R-CNN by He et al. [5] extended Faster R-CNN by adding a branch for predicting segmentation masks, thereby providing a unified approach for object detection and instance segmentation. A novel object detection model that leverages transformers instead of traditional CNNs was introduced in [2]. It simplifies the object detection pipeline by directly predicting bounding boxes and class labels from image pixels, eliminating the need for anchor boxes and non-maximum suppression.

## 2.2. Object Detection in Unfavourable Weather

Object detection in hostile weather conditions has been a topic of recent interest to improve the effectiveness of autonomous vehicles and surveillance systems. The use of YOLOv5 for real-time identification of cars, traffic lights, and pedestrians in various weather conditions, including challenging scenarios like rain and fog, is explored in [19]. A novel approach has been used in [4] combining visible and infrared image features using a CNN. By fusing these complementary features, the method enhances the accuracy of detection in low-visibility conditions. A network based on the joint optimal learning of an image-defogging module (IDOD) and YOLOv7 detection modules is proposed in [16]. This approach is designed for low-light foggy images, where the SAIP module's parameters are predicted by a separate CNN network, and the AOD module performs defogging by optimizing the atmospheric scattering model. The model proposed in [8] significantly improves the metrics but focuses solely on detecting traffic signs rather than

lights. Given that lights have higher intensity, their varying brightness levels can introduce difficulty in distinguishing shapes and colors, which may affect detection accuracy.

## 2.3. Traffic Light Detection

The detection and recognition of traffic lights are crucial components for the safety and efficiency of autonomous vehicles and intelligent transportation systems. In [11], the authors first passed the images through a deep neural object detection architecture to get traffic light candidates and then used a reward-based system where each detected traffic light earns rewards concerning features extracted from both spatial and temporal domains to reduce false positives. One drawback noted was that the classifier performed poorly for yellow lights, likely due to fewer training examples for them. To detect traffic lights at night and at a distance, the paper by Phuc et al. [15] used an approach combining the benefits of hand-crafted features and deep learning techniques. However, this method fails to detect traffic lights in complex regions with many traffic light-like objects. An improved YOLOv4 with a lightweight ShuffleNetv2 backbone, along with an enhanced K-Means clustering algorithm for bounding boxes and a novel CS2A mechanism, was introduced by [31] for traffic light recognition on mobile devices. These updates, combined with data augmentation techniques, improve recognition accuracy, reduce model size, and enhance detection speed and small target recognition. At the same time, though, the overall mAP is limited due to a small dataset, and the model cannot generalize very well due to the poor fitting ability of YOLOv4.



## 2.4. Fourier Domain Adaptation (FDA) for Classification

Domain adaptation is a crucial technique in computer vision, designed to bridge the gap between source and target domains with different distributions. Traditional methods, such as those developed by Tzeng et al. [22] and Long et al. [14], have significantly advanced the field by focusing on feature alignment and adversarial training. An extension of domain adaptation is FDA, which aligns the spectral properties of images through the Fourier transform. Yang and Soatto [28] pioneered this approach for cross-domain semantic segmentation. This method has been further explored in various applications; for example, in [25], the authors have used the concept of swapping the low-frequency part of the source and target domains used in FDA using the concept of Farthest Point Sampling (FPS).

## 2.5. Semi-Supervised Learning for Classification

Previous work has addressed object detection under adverse weather conditions [12] [20]. However, to the best of our knowledge, no datasets specifically for traffic light detection under adverse weather conditions exist at the time of writing. In this context, semi-supervised learning becomes increasingly important as it leverages both labelled and unlabelled data, addressing the challenge of manually annotating large datasets. To address this issue for the more general domain of traffic signs, the authors have used a combination of weakly supervised and semi-supervised learning to address the problems of imbalanced data and optimizing pseudo-labels generated on unlabelled data. A ranking-based pseudo-label generation system has been employed in [26], where certain objects are given more importance than others in the object detection system leading to an improvement in performance over supervised learning methods.

## 3. Methodology

This work aims to overcome the effects of adverse weather conditions and the lack of datasets for such applications. To this end, two widely used datasets were combined and modified, followed by Fourier Domain Adaptation to modify the spectral signals obtained by using the Fast-Fourier Transform (FFT). The spectral signals are then converted back to the image space using Inverse Fast-Fourier Transform (iFFT). This produces images that consist of high-level characteristics of the source domain (clear weather) while incorporating the low-level characteristics of the target domain (rainy or foggy conditions).

Fourier Domain Adaptation (FDA) [28] is a non-parametric technique that leverages Fast Fourier transforms to improve domain adaptation by modifying the frequency components of images. The key idea is to replace the low-frequency components of a source domain image (e.g., clear

weather) with those from a target domain image (e.g., rainy or foggy conditions). Since low-frequency components encode global structural information while high-frequency components capture finer details, this transformation helps the model retain structural consistency while adapting to new environmental conditions. By training models with FDA-augmented data and labels of the source domain (target domain labels are absent), the model becomes more robust to domain shifts caused by adverse weather conditions. This method is especially helpful for mitigating the effects of rain or fog, as it can be challenging to find data with these conditions along with proper annotations. Rain and fog primarily cause blurring, which reduces visibility. FDA curbs this issue by exposing the models to frequency-adjusted samples during training. This approach helps the models generalize more effectively to real-world conditions.

$$D^s = \{(x_i^s, y_i^s) \sim P(x^s, y^s)\}_{i=1}^{N_s}, \quad x^s \in R^{H \times W \times 3} \quad (1)$$

As defined in (1), the source dataset with clear weather conditions,  $D^s$ , is utilized, where  $x^s$  represents colour images, and  $y^s$  denotes the corresponding bounding box labels. The dataset  $D^s$  consists of  $N_s$  such image-label pairs, each sampled from the joint probability distribution  $P(x^s, y^s)$ . Similarly, the target dataset is defined as  $D^t = \{x_i^t\}_{i=1}^{N_t}$ , which contains target images with adverse weather conditions and no ground truth bounding box labels. Here,  $H$  and  $W$  represent the height and width of the image, respectively.

This transformation is implemented by (Fast) Fourier Transform (FFT) as in [3].  $\mathcal{F}$  represents the Fourier transform, which converts an image into spectral components (phase and amplitude), whereas  $\mathcal{F}^{-1}$  represents the inverse Fourier transform, converting spectral signals back into the image space.

The Fourier Transform  $\mathcal{F}$  for an image  $x$  with a single channel, gives:

$$\mathcal{F}(x)(m, n) = \sum_{h, w} x(h, w) e^{-j2\pi(\frac{h}{H}m + \frac{w}{W}n)}, \quad j^2 = -1 \quad (2)$$

Here,  $h$  and  $w$  denote pixel indices in the vertical and horizontal directions, respectively, while  $m$  and  $n$  represent the corresponding frequency coordinates in the Fourier domain. Since low-frequency components correspond to the central region of the frequency spectrum, we define a binary mask  $M_\beta$  that selects this central low-frequency area as:

$$M_\beta(h, w) = \mathbf{1}_{(h, w) \in [-\beta H : \beta H, -\beta W : \beta W]} \quad (3)$$

The origin (0,0) is taken to be at the center of the image in the frequency domain. The parameter  $\beta \in (0, 1)$  controls the size of this low-frequency region as a proportion of the image dimensions and is thus resolution-independent. This makes Fourier Domain Adaptation (FDA) applicable

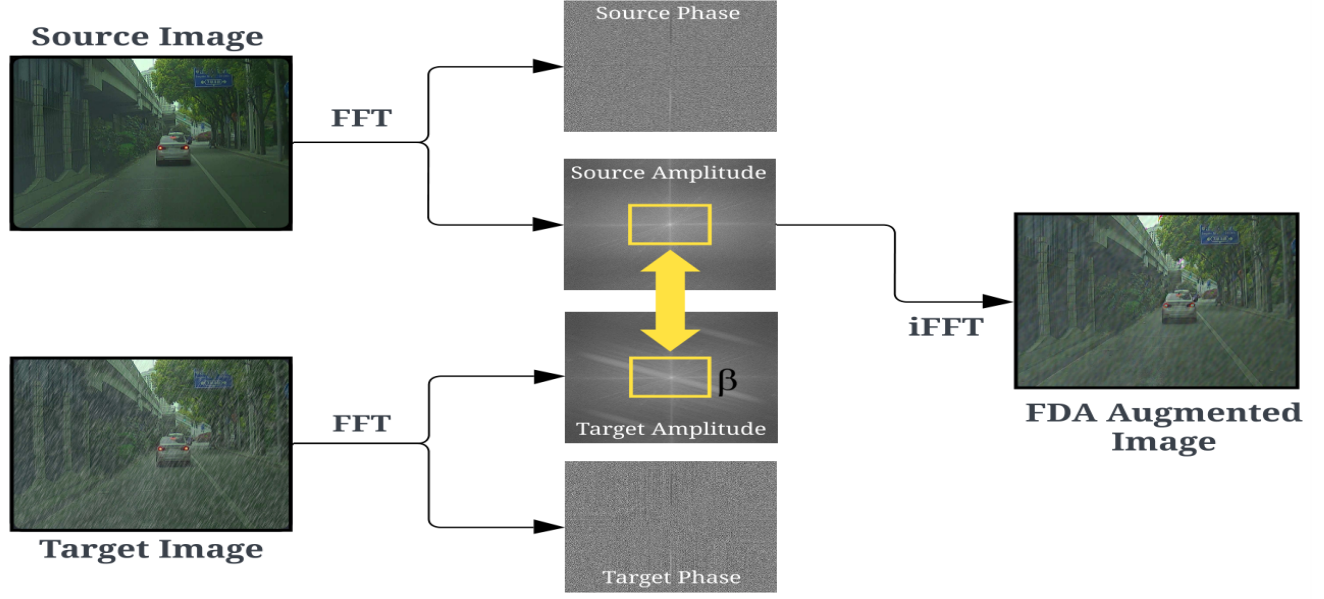


Figure 2. Shown above is the utilization of a target image to provide the desired style to a source image. The technique achieves this by swapping the low-frequency component of the source image (upper) with that of the target image (lower). The resulting FDA-augmented image exhibits a reduced perceptual domain gap, which helps object detection models adapt better to adverse weather conditions, as mentioned in the benchmarking section. It is important to note that the phase (high-frequency component) and amplitude (low-frequency component) images retain the same resolution as the original images; they are depicted smaller in the diagram solely for visual representation.

even when the source and target images differ in resolution. Given two randomly sampled images from the source and target pair,  $x^s \sim D^s$  (clear weather) and  $x^t \sim D^t$  (adverse weather), Fourier Domain Adaptation can be formalized as:

$$x^{s \rightarrow t} = \mathcal{F}^{-1} \left[ M_\beta \circ \mathcal{F}^A(x^t) + (1 - M_\beta) \circ \mathcal{F}^A(x^s), \mathcal{F}^P(x^s) \right] \quad (4)$$

Where  $\mathcal{F}^A, \mathcal{F}^P$  : represent the amplitude and phase components of the Fourier transform  $\mathcal{F}$  of the corresponding image.

The low-frequency amplitude components of the source image  $x^s$  are replaced with those from the target image  $x^t$ , showing adverse weather. The modified spectral representation of  $x^s$ , with its phase component retained, is then transformed back into the image  $x^{s \rightarrow t}$  using iFFT. This resultant image  $x^{s \rightarrow t}$  retains the content of  $x^s$  but adopts the appearance of a sample from  $D^t$ .

This process is illustrated in Fig. 2, where the mask  $M_\beta$  has been depicted. The optimal  $\beta$  value chosen for the FDA was 0.15. This value was used to construct the adapted dataset  $D^{s \rightarrow t}$ . The final merged dataset containing the FDA-augmented images and the ground truth bounding box labels of source dataset  $D^s$  is used to train various state-of-the-art object detection models, such as YOLOv5,

YOLOv6, and YOLOv8. The effect of FDA is illustrated in Fig. 1 and Fig. 4.

As discussed in Section 1, there exists a lack of datasets for traffic light detection in hostile weather. To validate the proposed method for learning domain-invariant features, we merged two widely used datasets: LISA and S<sup>2</sup>TLD. Here, a severe class imbalance against yellow lights was noted, and augmentations were applied to increase the proportion of yellow lights in the dataset. Experiments were conducted with varying proportions and several models to choose an optimal percentage, to which rain and fog were added artificially. Furthermore, semi-supervised learning was leveraged to demonstrate comparable results with only 50% labelled data. These details are provided in the following section.

The details regarding the choice of hyperparameters for the addition of fog and rain, their effects and the details regarding convergence are mentioned in the supplementary material.

## 4. Experiments

To validate our method, two widely used datasets - S<sup>2</sup>TLD and LISA were merged for experiments.

**The Shanghai Jiao Tong University (SJTU): Small Traffic Light Dataset (S<sup>2</sup>TLD)** [27] is a dataset designed

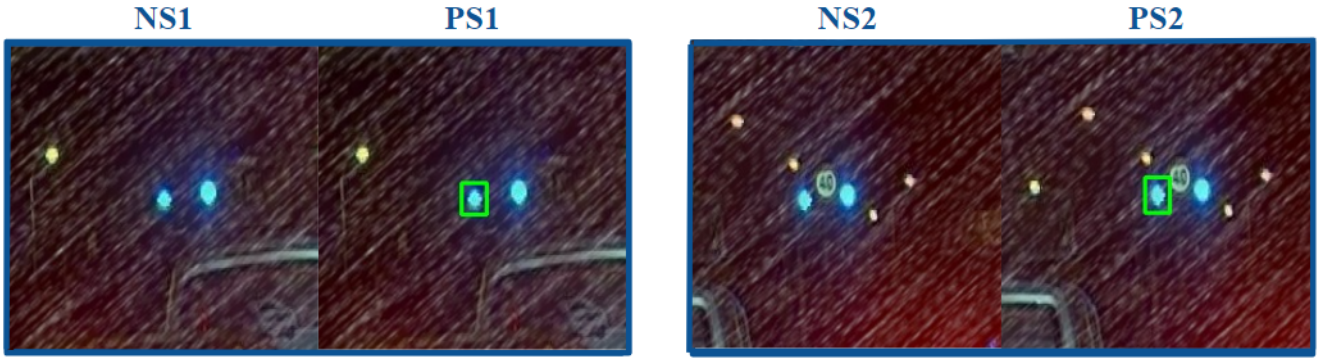


Figure 3. Qualitative comparison on the  $D^l$  dataset for detections from YOLOv8 under adverse weather conditions. The images NS1 and NS2 represent negative samples where the model trained only on the  $D^s$  dataset failed to detect the traffic lights. In contrast, PS1 and PS2 are the same positive samples where the model trained on the FDA-augmented dataset  $D^{s \rightarrow t}$  successfully detected the traffic lights. This highlights how FDA improves the model’s robustness by making it better suited to handle challenging visual conditions such as rain or fog.

for traffic light detection and automotive navigation tasks. It contains images with resolutions of approximately 1920 x 1080 pixels and 1280 x 720 pixels. The dataset consists of 14,130 instances categorized into five classes: red, yellow, green, off, and wait on of which only the red, green, and yellow classes were used.

Table 1. Brief summarization of the individual datasets chosen. The  $S^2TLD$  is a multiclass dataset, which is why its sum exceeds 100%. As evident, there innately exists a severe class imbalance against the ‘Yellow’ class.

Dataset	#Images	#Red	#Green	#Yellow
LISA	38323	48.51%	48.60%	2.88%
$S^2TLD$	5153	73.29%	56.88%	3.41%

**LISA Traffic Light Dataset** [7] is a comprehensive resource for traffic light recognition (TLR) tasks used for Autonomous Driving Systems. It consists of images of 14 classes of the resolution 1280 x 960. It contains annotated traffic lights (of which only the ones containing red, yellow, and green were used), captured using a stereo camera mounted on a vehicle.

FDA employs a hyperparameter ( $\beta$ ), which determines the size of the spectral neighbourhood swapped between the source and target images.  $\beta = \{0, 0.05, 0.10, 0.15\}$  were utilized for experiments. Higher values of  $\beta$  tend to create artefacts, as mentioned in [28]. As explained in Section 3,  $\beta$  is independent of the resolution of images, which allows this method to be adapted to images of various sizes and crops. FDA was performed on the source domain images using the target domain images to generate a modified dataset,  $D^{s \rightarrow t}$ .

For experiments, the final dataset was split into 70% for training, 20% for validation, and 10% for testing. All the images are resized to the resolution of 1280 x 1080 using

bicubic interpolation, and the labels are adjusted accordingly. The training set consisted of images from  $D^{s \rightarrow t}$ , the FDA-augmented dataset, while the validation and test sets used images from  $D^t$ , the target domain with rain and fog effects.

Furthermore, we explore Semi-Supervised Learning (SSL) to mitigate the lack of datasets available by initially training YOLOv6m on 50% of the annotated training data to generate pseudo-labels for the remaining unlabeled half. This new compiled dataset of 50% original labels, and 50% pseudo-labels is used to benchmark models. A confidence threshold of 50% was applied to filter the predicted labels, ensuring that only high-confidence detections were retained.

To address the class imbalance evident in Table 1, various proportions of yellow traffic light images were evaluated to identify the optimal percentage for benchmarking. The models were trained on datasets with various percentages of yellow images, including 5%, 7%, 9%, 11%, and 13%. These experiments indicate that the dataset comprising 13% yellow images led to convergence, suggesting that generalizability across all models is achieved largely at this percentage, making it suitable for the source domain. Detailed results of these experiments, along with the augmentations performed in order to mitigate class imbalance, are provided in the supplementary material.

## 5. Results

The performance of the proposed approach is evaluated against confidence-dependent and confidence-independent metrics like Precision, Recall, mAP50, and mAP50-95. Here, the qualitative and quantitative results are presented.

As demonstrated in Table 2, the class imbalance has been mitigated by incorporating augmented images. The models



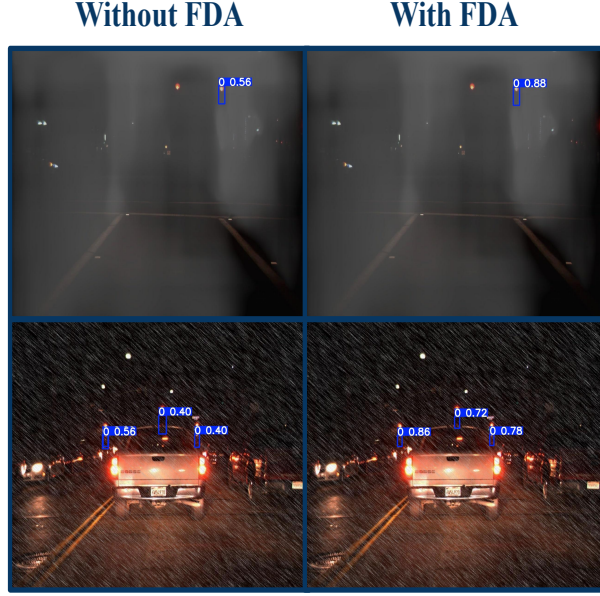


Figure 4. Comparisons of outputs from inferring on YOLOv8m. The first image in each row is the output of the model trained on  $D^s$ , and the second is the output when trained on  $D^{s \rightarrow t}$ . Each bounding box has the label (0 implies Red) and the associated confidence score.

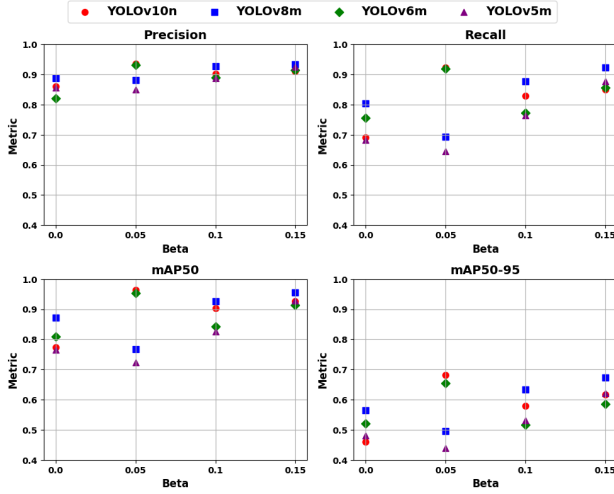


Figure 5. Comparisons of results of various models with different  $\beta$  values. It is observed that these models perform best and yield the most consistent results at  $\beta = 0.15$ .

exhibit excellent performance, especially in detecting yellow lights across various degrees of FDA, as indicated by the parameter  $\beta$ , demonstrating the effectiveness of the proposed methodology. For  $\beta = 0.15$ , the average percentage increase in metrics from training on  $D^s$  to  $D^{s \rightarrow t}$  for mAP50 and mAP50-95 for the yellow class is 16.12% in YOLOv8, 22.63% in YOLOv6, and 31.80% in YOLOv5, respectively.

As shown in Fig. 5, when  $\beta = 0$  (unchanged source do-

main), models trained on  $D^s$  and inferred on  $D^t$  exhibit sub-par performances. This degradation in metrics highlights how exhaustive experimentation ensures the addition of not only effective but also realistic rain and fog. It simulates real-world weather conditions in the label-less target domain dataset, thereby validating the observed performance drop and pointing to its real-world applications. Performance across all metrics improves with the increase in  $\beta$ , where the best results are observed at  $\beta = 0.15$ . According to Fig. 5, results at  $\beta = 0.15$  are higher and most tightly clustered, showing better performance and consistency across metrics, with YOLOv8 displaying the highest performance metrics - Precision of 0.93, Recall of 0.92, mAP50 of 0.96, and mAP50-95 of 0.67. An average increase of **12.25%** across all metrics for YOLOv8 was observed when training on  $D^{s \rightarrow t}$ . The authors believe YOLOv8m outperforms YOLOv10n because the nano version is parametrically smaller, leading to reduced performance on complex features. However, the nano version was used as it is less computationally intensive compared to the bulkier medium version. The average increase across all metrics, across all models is observed to be **16.81%**. On average, percentage increases of 7.69% in Precision, 19.91% in Recall, 15.85% in mAP50, and 23.81% in mAP50-95 is observed across all models when utilizing  $D^{s \rightarrow t}$  instead of  $D^s$ . This increase in metrics validates the proposed methodology of introducing FDA to curb the effect of adverse weather, pointing to a solution applicable to real-world scenarios.

The effect of the proposed methodology is also evident through the qualitative results, as observed in Fig. 4. Looking closely, it is noted that models tend to detect traffic lights more accurately and with a much higher confidence score. Furthermore, on observing Fig. 3, it is noted through the Negative Samples (NS1, NS2) and Positive Samples (PS1, PS2) that the effect of FDA-augmented images extends to real-world scenarios where training on clear weather ( $D^s$ ) may be inadequate and lead to missing detections. As evident in the figure, training on FDA-augmented images results in such traffic lights being detected with accurate bounding boxes. This exemplifies the applicability of the proposed methodology in real-world scenarios. In cases with insufficient data on traffic scenes in challenging weather conditions, FDA can be applied to existing data in clear environments. This can help the models generalize and perform better in unforeseen weather conditions.

Referring to the SSL results in Table 3, it is observed that training with only 50% of the actual ground truth labels, while generating pseudo labels for the remaining data using YOLOv6, yields comparable performance. Evaluation was conducted using the same models as ones for fully supervised learning, i.e YOLOv5, YOLOv6, YOLOv8, and YOLOv10 on the NVIDIA A6000 for 20 epochs. Although the results may appear marginally suboptimal compared to

Table 2. Class-wise results when trained on  $D^{s \rightarrow t}$  and inferred on the  $D^t$  (foggy and rainy) dataset. As evident across all metrics, the quantitative results of different models are impressive, especially for the Yellow images, showcasing the effectiveness of countering class imbalance.

Model	$\beta$	Class	Precision	Recall	mAP50	mAP50-95
YOLOv10n	0.15	Red	0.938	0.893	0.952	0.698
		Green	0.912	0.809	0.91	0.623
		Yellow	0.882	0.847	0.918	0.529
	0.1	Red	0.942	0.871	0.94	0.669
		Green	0.895	0.802	0.887	0.586
		Yellow	0.874	0.817	0.883	0.485
	0.05	Red	0.949	0.941	0.973	0.749
		Green	0.922	0.898	0.952	0.687
		Yellow	0.934	0.931	0.966	0.609
YOLOv8m	0.15	Red	0.945	0.942	0.97	0.741
		Green	0.922	0.888	0.942	0.678
		Yellow	0.934	0.939	0.957	0.603
	0.1	Red	0.949	0.899	0.952	0.704
		Green	0.915	0.835	0.9	0.629
		Yellow	0.919	0.898	0.929	0.565
	0.05	Red	0.908	0.71	0.811	0.563
		Green	0.856	0.664	0.738	0.491
		Yellow	0.883	0.703	0.751	0.431
YOLOv6m	0.15	Red	0.934	0.89	0.934	0.662
		Green	0.899	0.813	0.895	0.588
		Yellow	0.908	0.865	0.914	0.506
	0.1	Red	0.932	0.817	0.89	0.599
		Green	0.878	0.732	0.82	0.519
		Yellow	0.859	0.769	0.818	0.432
	0.05	Red	0.952	0.936	0.968	0.730
		Green	0.913	0.895	0.94	0.658
		Yellow	0.931	0.926	0.953	0.574
YOLOv5m	0.15	Red	0.940	0.898	0.950	0.697
		Green	0.909	0.839	0.902	0.615
		Yellow	0.930	0.897	0.933	0.551
	0.1	Red	0.912	0.785	0.867	0.614
		Green	0.849	0.725	0.779	0.516
		Yellow	0.901	0.704	0.832	0.466
	0.05	Red	0.889	0.669	0.78	0.52
		Green	0.804	0.62	0.689	0.437
		Yellow	0.857	0.648	0.703	0.361

fully supervised baselines, the implications are significant: this semi-supervised learning (SSL) approach demonstrates strong potential in scenarios where annotated data is inherently limited. Given the challenges in acquiring high-quality, annotated data in this domain—owing to both logistical and safety-related constraints—methods that effectively exploit limited supervision become essential. The results indicate that, even in the absence of 50% of the original ground truth annotations, the proposed SSL-based pipeline achieves performance levels comparable to fully supervised models. This underscores the effectiveness of pseudo-labelling strategies in practical settings where annotated data is scarce and expensive to obtain. By effectively utilizing limited supervision, the method demonstrates significant potential for deployment in data-constrained environments.

Table 3. Comparisons of results of various models trained using SSL. As mentioned in the Benchmarking and Results section, pseudo-labels are generated for 50% of  $D^{s \rightarrow t}$  using YOLOv6m, while the rest are original labels.

Model	Class	Precision	Recall	mAP50	mAP50-95
YOLOv10n	All	0.82	0.772	0.838	0.47
	Red	0.816	0.832	0.876	0.54
	Green	0.786	0.718	0.805	0.455
	Yellow	0.86	0.766	0.835	0.414
YOLOv8m	All	0.849	0.826	0.878	0.501
	Red	0.842	0.876	0.908	0.565
	Green	0.808	0.771	0.833	0.478
	Yellow	0.898	0.832	0.892	0.46
YOLOv6m	All	0.82	0.807	0.843	0.472
	Red	0.861	0.862	0.89	0.543
	Green	0.79	0.751	0.81	0.455
	Yellow	0.808	0.808	0.828	0.417
YOLOv5m	All	0.838	0.796	0.85	0.47
	Red	0.828	0.854	0.887	0.532
	Green	0.795	0.745	0.8	0.441
	Yellow	0.892	0.791	0.862	0.438

## 6. Conclusion

In this paper, we present a novel approach for Traffic Light Detection under Adverse Weather conditions, tackling two key challenges: domain shift and limited labelled data. To address domain shift between clear and low-visibility scenarios (rain and fog), we employ Fourier Domain Adaptation (FDA). As the low-frequency area ( $\beta$ ) increases from 0.05 to 0.15, performance improves, with  $\beta = 0.15$  yielding the best results. At this setting, YOLOv8 achieves Precision 0.93, Recall 0.92, mAP50 0.96, and mAP50-95 0.67—an average gain of **12.25%** over baseline. Overall, FDA provides an average improvement of **16.81%** across models, with increases of 7.69% in Precision, 19.91% in Recall, 15.85% in mAP50, and 23.81% in mAP50-95. Semi-supervised learning (SSL) further shows competitive results using only 50% labelled data, with pseudo-labels generated by YOLOv6m. For rigorous evaluation, we curate a dataset of traffic light images in rainy and foggy conditions. These results highlight the effectiveness of our approach in mitigating domain shift and label scarcity, underscoring its applicability in real-world adverse weather.

## 7. Acknowledgment

We would like to thank Mars Rover Manipal, an interdisciplinary student project of MAHE, for providing the essential resources and infrastructure that supported our research.

## References

- [1] Mario Bijelic, Tobias Gruber, Fahim Mannan, Florian Kraus, Werner Ritter, Klaus Dietmayer, and Felix Heide. See-



- ing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. 1
- [2] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. End-to-end object detection with transformers, 2020. 3
- [3] Matteo Frigo and Steven G Johnson. Fftw: An adaptive software architecture for the fft. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*, pages 1381–1384. IEEE, 1998. 4
- [4] Yajing Han and Dean Hu. Multispectral fusion approach for traffic target detection in bad weather. *Algorithms*, 13(11), 2020. 3
- [5] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017. 3
- [6] Judy Hoffman, Eric Tzeng, Taesung Park, Jun-Yan Zhu, Phillip Isola, Kate Saenko, Alexei Efros, and Trevor Darrell. CyCADA: Cycle-consistent adversarial domain adaptation. In *Proceedings of the 35th International Conference on Machine Learning*, pages 1989–1998. PMLR, 2018. 2
- [7] Morten Bornø Jensen, Mark Philip Philipsen, Andreas Møgelmoose, Thomas Baltzer Moeslund, and Mohan Manubhai Trivedi. Vision for looking at traffic lights: Issues, survey, and perspectives. *IEEE Transactions on Intelligent Transportation Systems*, 17(7):1800–1815, 2016. 2, 6
- [8] Bin Ji, Jiafeng Xu, Yang Liu, Pengxiang Fan, and Mengli Wang. Improved yolov8 for small traffic sign detection under complex environmental conditions. *Franklin Open*, 8: 100167, 2024. 3
- [9] Glenn Jocher. Ultralytics yolov5. <https://github.com/ultralytics/yolov5>, 2020. Version 7.0, DOI: 10.5281/zenodo.3908559. 1, 2
- [10] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. Ultralytics yolov8, 2023. 1, 2
- [11] Jinkyu Kim, Hyunggi Cho, Myung Hwangbo, Jaehyung Choi, John Canny, and Youngwook Paul Kwon. Deep traffic light detection for self-driving cars from a large-scale dataset. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 280–285, 2018. 3
- [12] Debasis Kumar and Naveed Muhammad. Object detection in adverse weather for autonomous driving through data merging and yolov8. *Sensors*, 23(20), 2023. 2, 4
- [13] Chuyi Li, Lulu Li, Yifei Geng, Hongliang Jiang, Meng Cheng, Bo Zhang, Zaidan Ke, Xiaoming Xu, and Xiangxiang Chu. Yolov6 v3.0: A full-scale reloading, 2023. 1
- [14] Mingsheng Long, Yue Cao, Zhangjie Cao, Jianmin Wang, and Michael I. Jordan. Transferable representation learning with deep adaptation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12):3071–3085, 2019. 4
- [15] Phuc Manh Nguyen, Vu Cong Nguyen, Son Ngoc Nguyen, Linh My Thi Dang, Ha Xuan Nguyen, and Vinh Dinh Nguyen. Robust traffic light detection and classification under day and night conditions. In *2020 20th International Conference on Control, Automation and Systems (ICCAS)*, pages 565–570, 2020. 3
- [16] Yongsheng Qiu, Yuanyao Lu, Yuantao Wang, and Haiyang Jiang. Idod-yolov7: Image-dehazing yolov7 for object detection in low-light foggy traffic environments. *Sensors*, 23(3):1347, 2023. 1, 3
- [17] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018. 2
- [18] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection, 2016. 2
- [19] Teena Sharma, Benoit Debaque, Nicolas Duclos, Abdellah Chehri, Bruno Kinder, and Paul Fortier. Deep learning-based object detection and scene perception under bad weather conditions. *Electronics*, 11(4), 2022. 3
- [20] Huiping Su, Hongbo Gao, Xinmiao Wang, Xiaozhao Fang, Qingchao Liu, Guang-Bin Huang, Xin Li, and Qi Cao. Object detection in adverse weather for autonomous vehicles based on sensor fusion and incremental learning. *IEEE Transactions on Instrumentation and Measurement*, 73:1–10, 2024. 4
- [21] Le-Anh Tran, Chung Nguyen Tran, Dong-Chul Park, Jordi Carrabina, and David Castells-Rufas. Toward improving robustness of object detectors against domain shift. In *2024 International Conference on Green Energy, Computing and Sustainable Technology (GECOST)*, pages 01–05. IEEE, 2024. 2
- [22] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017. 4
- [23] Ao Wang, Hui Chen, Lihao Liu, Kai Chen, Zijia Lin, Jungong Han, and Guiguang Ding. Yolov10: Real-time end-to-end object detection, 2024. 1
- [24] Gang Wang, Jingming Guo, Yupeng Chen, Ying Li, and Qian Xu. A pso and bfo-based learning strategy applied to faster r-cnn for object detection in autonomous driving. *IEEE Access*, 7:18840–18859, 2019. 1
- [25] Jing Wang, Qiufeng Wu, Tianci Liu, Yuqi Wang, Pengxian Li, Tianhao Yuan, and Ziyang Ji. Fourier domain adaptation for the identification of grape leaf diseases. *Applied Sciences*, 14(9), 2024. 4
- [26] Runsheng Xu, Hao Xiang, Xin Xia, Xu Han, Jinlong Li, and Jiaqi Ma. Opv2v: An open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication, 2022. 1, 4
- [27] Xue Yang, Junchi Yan, Wenlong Liao, Xiaokang Yang, Jin Tang, and Tao He. Srdet++: Detecting small, cluttered and rotated objects via instance-level feature denoising and rotation loss smoothing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 2, 5
- [28] Yanchao Yang and Stefano Soatto. Fda: Fourier domain adaptation for semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4085–4095, 2020. 2, 4, 6
- [29] Shizhe Zang, Ming Ding, David Smith, Paul Tyler, Thierry Rakotoarivelo, and Mohamed Ali Kaafar. The impact of adverse weather conditions on autonomous vehicles: How rain,

snow, fog, and hail affect the performance of a self-driving car. *IEEE vehicular technology magazine*, 14(2):103–111, 2019. [1](#)

- [30] Xiangmo Zhao, Pengpeng Sun, Zhigang Xu, Haigen Min, and Hongkai Yu. Fusion of 3d lidar and camera data for object detection in autonomous vehicle applications. *IEEE Sensors Journal*, 20(9):4901–4913, 2020. [1](#)
- [31] Ying Zhao, Yiyuan Feng, Yueqiang Wang, Zhihan Zhang, and Zhihao Zhang. Study on detection and recognition of traffic lights based on improved yolov4. *Sensors*, 22(20), 2022. [3](#)