

Road Scene Analysis: A Study of Polarimetric and Color-based Features under Various Adverse Weather Conditions

Rachel Blin¹^a, Samia Ainouz¹^b, Stéphane Canu¹^c and Fabrice Meriaudeau²^d

¹Normandie Univ., INSA Rouen, LITIS, 76000 Rouen, France

²University of Burgundy, UBFC, ImViA, 71200 Le Creusot, France

Keywords: Road Scene, Object Detection, Adverse Weather Conditions, Polarimetric Imaging, Data Fusion, Deep Learning.

Abstract: Autonomous vehicles and ADAS systems require a reliable road scene analysis to guarantee road users' safety. While most of autonomous systems provide an accurate road objects detection in good weather conditions, there are still some improvements to be made when visibility is altered. Polarimetric features combined with color-based ones have shown great performances in enhancing road scenes under fog. The question remains to generalize these results to other adverse weather situations. To this end, this work experimentally compares the behaviour of the polarimetric intensities, the polarimetric Stokes parameters and the RGB images as well as their combination in different fog densities and under tropical rain. The different detection tasks show a significant improvement when using a relevant fusion scheme and features combination in all the studied adverse weather situations. The obtained results are encouraging regarding the use of polarimetric features to enhance road scene analysis under a wide range of adverse weather conditions.


1 INTRODUCTION


Autonomous vehicles and ADAS systems have shown outstanding improvements these past few years thanks to a more accurate and reliable road scene analysis. Object detection, which is essential to understand road scenes, has widely contributed to these improvements. While autonomous vehicles can be found in several places, e.g. the Waymo car¹ in Arizona and the Rouen Normandy autonomous lab² in France, they are restricted to a small driving area and can not go above a fixed speed limit. Moreover, all these systems are limited when the visibility is altered, in case of adverse weather conditions, since they struggle reliably to detect road objects. Because conventional imaging is very sensitive to lighting variations, it fails to properly characterize objects in these complex weather situations. Using non-conventional sensors is nowadays the best alter-


native solution to improve road scenes analysis when the conditions are not optimal (Bijelic et al., 2018). Lidar is used to improve localization accuracy under rain and snow (Aldibaja et al., 2016). Lidar combined with color-based images also helps improving road object detection in low-light conditions (Rashed et al., 2019). An infra-red camera enables to detect vehicles at a larger distance range than a color-based sensor under fog (Pinchon et al., 2018). In (Major et al., 2019), a radar is used to enhance road object detection for fast moving vehicles, when the line-of-sight is reduced, like in highway environments. More recently, the use of polarimetric imaging helped enhancing road scenes understanding under fog, since it is able to generalize the road objects' features learnt in good weather conditions to foggy scenes, to the contrary of conventional color-based sensors (Blin et al., 2020).

These preliminary results are encouraging but are limited since they explore only one fog density. A wider study on various adverse weather conditions needs to be carried out to extend these results to all kinds of low visibility conditions. In this work, the behaviour of polarimetric and color-based features is studied in several adverse weather conditions, including several densities of fog and tropical rain. To

^a <https://orcid.org/0000-0003-1036-154X>

^b <https://orcid.org/0000-0002-2699-4002>

^c <https://orcid.org/0000-0002-7602-4557>

^d <https://orcid.org/0000-0002-8656-9913>

¹<https://waymo.com/>

²<https://www.rouennormandyautonomouslab.com/>

this end, a dataset containing road scenes under 11 different weather conditions is constituted. Two fusion schemes on several polarimetric features as well as polarimetric and color-based features are also explored. As a matter of fact, fusing multimodal features is paramount to enable better scene perception for autonomous driving. Indeed, if the road scene analysis relies on only one modality, any perturbation in the sensor could alter the scene analysis and directly impact the vehicle decision and then the road users' safety (Feng et al., 2020). Moreover, combining information from multimodal sensors also enables to describe a road scene under different angles (Nie et al., 2020; Gu et al., 2018) and provides a more relevant scene description when the visibility is altered (Bijelic et al., 2020).

The contributions of this paper are fourfold:

- Constitution of a multimodal polarimetric and color-based dataset for road scene analysis in various adverse weather conditions,
- Study of different polarimetric and color-based features in such complex situations,
- Exploration of two fusion schemes of polarimetric and color-based features,
- Enhancement of object detection in road scenes in adverse weather situations.

The code used in this work can be downloaded at: <https://github.com/RachelBlin/keras-retinanet> and the WPol dataset can be downloaded at: <https://zenodo.org/record/5547801#.Ycintrso9uQ>

2 POLARIZATION FORMALISM AND RELATED WORK

When an electromagnetic light wave is being reflected by an object, it is partially linearly polarized. Polarimetric imaging captures the polarization state of the reflected light, i.e. the direction in which the wave is travelling. It enables to characterize an object by its physical properties, giving information on the nature of the object's surface the light wave impinges on. The polarization state of the reflected light can be described by the linear Stokes vector $S = [S_0 \ S_1 \ S_2]^T$, which is a set of measurable parameters (Bass et al., 1995). A polarizer oriented at a specific angle α_i is required to capture polarimetric images, measuring an intensity I_{α_i} . To reconstruct the Stokes vector characterizing the reflected wave, at least three intensities $I_{\alpha_i, i=1:3}$, corresponding to the polarizer oriented at three different angles $\alpha_{i,i=1:3}$, are needed. A Polar-

cam 4D Technology polarimetric camera³ is used in this work. This camera contains four different linear polarizers, oriented at four different angles, including 0° , 45° , 90° and 135° . Four intensities of the same scene, $I = [I_0 \ I_{45} \ I_{90} \ I_{135}]^T$, are simultaneously captured with this polarimetric camera. The relationship between an intensity $I_{\alpha_i, i=1:4}$ captured by the camera and the Stokes parameters S is given by:

$$I_{\alpha_i} = \frac{1}{2} \begin{bmatrix} 1 & \cos(2\alpha_i) & \sin(2\alpha_i) \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ S_2 \end{bmatrix}, \forall i = 1, 4. \quad (1)$$

This equation can be reformulated by:

$$I = AS, \quad (2)$$

where $A \in \mathbb{R}^{4 \times 3}$ is the calibration matrix of the linear polarizer (Blin et al., 2020).

To reconstruct the Stokes vector from the intensities, the following equation is used:

$$S = \begin{bmatrix} I_0 + I_{90} \\ I_0 - I_{90} \\ I_{45} - I_{135} \end{bmatrix}. \quad (3)$$

Other physical features can be obtained from the Stokes parameters: the Angle Of Polarization (ϕ) and the Degree Of linear Polarization (ρ) (Ainouz et al., 2013). They are expressed in the following way:

$$\phi = \frac{1}{2} \arctan2 \left(\frac{S_2}{S_1} \right), \quad (4)$$

$$\rho = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}, \quad (5)$$

where $\rho \in [0, 1]$ refers to the quantity of the linear polarized light in a wave and $\phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ is the orientation of the polarized part of the wave with regards to the incident plan.

Recent works have demonstrated the impact of polarimetric features to enhance road scene analysis. The fusion of polarimetric and color-based features enable to increase car detection (Fan et al., 2018). The Angle Of Polarisation (ϕ) is selected to perform this task as it provides the best results. Two different score maps are obtained by training independently a polarimetric-based and a color-based Deformable Part Models (DPM) (Felzenszwalb et al., 2008). These two score maps, are used to perform an AND-fusion between polarimetric and color features, resulting in the final detection bounding boxes. This pipeline, more specifically, the fusion of these

³<https://www.4dtechnology.com/>

two complementary modalities largely, improves the detection accuracy by reducing the false alarm rate.

Other experiments focus on road objects detection under fog, using polarimetric images (Blin et al., 2019). New data formats using polarimetric features are created to perform this task. These data formats provide the necessary features to characterize road objects when the visibility is altered. When using an adapted deep learning framework combined with these data formats, both vehicles and pedestrian detection are enhanced compared to color-based models. The PolarLITIS dataset is created (Blin et al., 2020) in order to expand and confirm these results. It is the first public multimodal dataset, containing polarimetric road scenes and their color-based equivalent in various weather conditions (sunny, cloudy and foggy), labeled with bounding boxes to perform object detection.

Our interest in this work is to focus on the exploration of various adverse weather conditions, using polarimetric and color-based features. Road scenes under 10 different fog densities, from 15m to 70m visibility, as well as tropical rain are collected. The relevance of the polarimetric intensities I and the Stokes vector S to enhance road scene analysis in various weather situations is studied as well as their limits.

3 THE PROPOSED METHOD

In this section, information regarding the collected data constituting the WPolar dataset (Weather Polarimetric) are first given. The experimental setup, used to evaluate the performances of polarimetric and color-based features for road objects detection in various adverse weather conditions, is then sketched.

3.1 The WPolar Dataset

In the PolarLITIS dataset (Blin et al., 2020), the testing set is exclusively constituted of road scenes under fog. However, since it contains outdoor scenes, it is difficult to deduce the density of the fog. Knowing the properties of adverse weather situations enable to provide direct information on the visibility of the scene, as well as, a more precise idea on the ability of polarimetric features to see beyond human perception. However, in outdoor scenes, it is difficult to get such information. Moreover, studying the behaviour of polarimetric features in diverse weather conditions is paramount to define the extent of their added value as well as their limits to describe road scenes in complex situations.

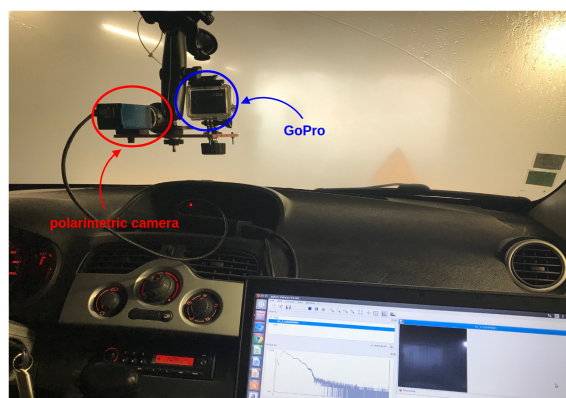


Figure 1: Acquisition setup.

Disposing of road scenes in various adverse weather conditions is difficult in a reasonable amount of time. This is due to the fact that weather conditions causing low visibility are not the most common. To palliate these limitations, an acquisition campaign was made in the CEREMA tunnel⁴, simulating road scenes under adverse weather conditions. This facility is 30m long and simulates fog from 15m visibility and different densities of rain. It enables to collect road scenes under various weather conditions while having exact information on the scene visibility.

During this acquisition campaign, a Polarcam4D Technology polarimetric camera is placed next to a GoPro (RGB camera), behind the windshield, at the height of the driver's eye. This setup collects paired multimodal scenes. An illustration of the acquisition setup can be found in Figure 1.

The polarimetric camera disposes of a standard lens and has a resolution of 500×500 while the RGB camera has a fisheye lens and a resolution of 3648×2736 . To get the closest content possible in the multimodal images constituting each pair, the RGB images are cropped to 1115×1165 . By keeping the center of the RGB images, the deformation induced by the fisheye lens is attenuated, which implies a closest scene geometry between the two modalities.

Regarding the content of each road scene, acquisitions were made under 11 different adverse weather conditions, including 10 different fog densities and tropical rain. The 10 different fog densities are characterized by their visibility distance, based on human perception, include 15m, 20m, 25m, 30m, 35m, 40m, 45m, 50m, 60m and 70m. The collected scenes visibility is altered due to the weather conditions but also because of the presence of water on the windshield. As a matter of fact, under tropical rain, drops of water

⁴<https://www.cerema.fr/fr/innovation-recherche/innovation/offres-technologie/plateforme-simulation-conditions-climatiques-degradees>

modify the view and a veil of water covers the windshield while the wiper clears the view (see fifth and sixth columns of Figure 2 for an illustration of these phenomena). Fog also regularly leads to an accumulation of droplets on the windshield, which gives a blurring effect to the scene.

The Cerema tunnel is composed of a tar ground, on which different kind of landmarks can be placed. Several road signs are also available to simulate real road scenes. The roof is transparent, which enables natural sunlight to lighten the tunnel. Once the desired road scene constituted, road users, including motorized vehicles, pedestrian and cyclists can circulate in the tunnel. All these elements enable to simulate road scenes with the same content and road users as a real one. Since the acquisitions were made within a 30 meters tunnel, the images are selected to maximize the variability of the scenes. The selected road scenes ensure the presence of the different road users at several distances and various combinations of road users in each image. To provide strictly paired images, each pair of images was selected manually. Examples of multimodal images in the database are shown in Figure 2.

The collected images were labelled for road objects detection. To this end, three classes are used, which represent the most common road users: 'car', 'person' and 'bike'. The 'car' class includes all four wheels vehicles, the class 'person' includes the pedestrian and cyclists and the class 'bike' designates the bikes without their cyclists. The number of instances of each class are summarized up in Table 1. Note that the RGB images under fog with respectively 25m and 30m visibility are not available due to a technical issue during the acquisition campaign.

Since the images were collected in a tunnel, this limits the variability of the dataset. This is the reason why the images are only used for evaluation purposes. Including these scenes in the training process is very likely to bias the evaluation due to over-fitting. On top of that, not including road scenes in adverse weather in the training and validation sets enables to evaluate the ability of polarimetric features learnt in good weather conditions to describe reliably road objects when the visibility is altered. Consequently, the final training and validation sets are the ones of the PolarLITIS dataset (sunny and cloudy weather) and the collected images constitute the testing set. The properties of the final dataset are given in Table 2.

The dataset can be downloaded at: <https://doi.org/10.5281/zenodo.5547801>.

4 EXPERIMENTS

In order to study the behaviour of the polarimetric features in several weather conditions, some experiments are carried out. The first goal of this experiment is to evaluate how invariant are the polarimetric features characterizing road objects to the visibility conditions. The second goal of the experiment is to evaluate the relevance of multimodal fusion to enhance road object detection in several adverse weather situations. The polarimetric features selected for this experiment are the intensity images $I = (I_0, I_{45}, I_{90})$ and the Stokes images $S = (S_0, S_1, S_2)$ described in (Blin et al., 2020). The color-based features are RGB information.

As mentioned in Table 2, the dataset used for the experiments is constituted of sunny and cloudy road scenes. This design enables to evaluate if the road objects features learnt in good weather conditions are still valid to detect objects when the visibility is altered. The RetinaNet network (Lin et al., 2017), using a ResNet152 backbone (He et al., 2016) is used for this task. This object detector, thanks to the focal loss, focuses on hard misclassified examples during its training process. This property is useful to process datasets with unbalanced classes, as it is the case herein. It is also able to process images in real time with a high accuracy, which is paramount to perform object detection in road scenes. All the experimental setups are sketched in Figure 3.

A Late fusion scheme is used to fuse the different modalities of this experiment.

To this end, the Double soft-NMS filter and the OR filter are constituted. The Double soft-NMS filter consists in using the soft-NMS algorithm (Bodla et al., 2017) on each modality to filter the raw predicted bounding boxes on each modality a first time and filter a second time the concatenation of the obtained bounding boxes using the soft-NMS algorithm. The OR filter consists in filtering the raw bounding boxes of each modality separately, using the soft-NMS bounding boxes and apply an OR operation on the two obtained sets of bounding boxes. To fuse the polarimetric and color-based images, the offset between these two modalities is computed.

Since the training set is composed of 1640 images and the validation set of 420 images, it is paramount to pre-train the network on a larger dataset. The BDD100K dataset (Yu et al., 2020) is selected for this task since it is rather large and aims to detect objects in road scenes in good weather conditions. On top of that, it contains all the classes of the dataset described in Table 2, making fine-tuning towards this dataset easier. Once the architecture pre-trained on



Figure 2: Example of acquisitions at the CEREMA tunnel. The first row are the polarimetric intensity images $I = (I_0, I_{45}, I_{90})$ and the second row is their equivalent in RGB.

Table 1: Properties of the WPolar dataset, for each weather condition. The first 10 columns refer to the different visibility distances under fog. It is important to note that the number of instances are valid for the polarimetric and the RGB modalities.

Weather conditions	15m	20m	25m	30m	35m	40m	45m	50m	60m	70m	rain
images	44	44	44	44	44	44	44	44	44	44	44
car	25	43	44	42	44	44	44	44	43	44	44
person	54	71	83	90	65	49	70	58	74	74	30
bike	9	13	6	14	9	8	25	14	35	19	0

Table 2: Dataset properties. The training and validation sets are the ones of PolarLITIS(Blin et al., 2020) dataset. The testing set contains all the acquisitions of the WPolar dataset.

Properties	Train	Validation	Test
Weather conditions	sunny/cloudy	cloudy	fog/rain
images	1640	420	484
car	6061	2102	461
person	527	134	718
bike	39	7	152

BDD100K, it is fine-tuned on the PolarLITIS dataset, on each modality separately (I , S and RGB).

Regarding the training hyperparameters, the ones provided by the RetinaNet’s article are kept, i.e. a learning rate of 10^{-5} and the Adam optimizer (Kingma and Ba, 2014). Each training process is repeated five times to provide reliable results. Note that the different architectures are trained for 50 epochs on the BDD100K dataset and for 20 epochs on the PolarLITIS dataset. The optimal weights are selected according to the lowest value of the validation loss.

5 DISCUSSION AND RESULTS

To evaluate the increase in detection scores, the error rate evolution is computed as follows:

$$ER_o^M = \frac{1 - AP_o^M - (1 - AP_o^I)}{1 - AP_o^I} \times 100, \quad (6)$$

where ER_o^M is the error rate evolution between the intensities polarimetric data format I and the modality $M \in \{S, RGB, I+S (2 \times \text{soft-NMS}), I+S (\text{OR filter}), I+RGB (2 \times \text{soft-NMS}), I+RGB (\text{OR filter})\}$ for object $o \in \{‘person’, ‘car’, mAP\}$, AP_o^I is the average precision for object o with data format I , while AP_o^M denotes the average precision on the object o with modality M . A negative error rate is associated to an increase of AP_o^M with regards to AP_o^I and a positive error rate is associated to a decrease of AP_o^M with regards to AP_o^I .

The intensity images I are used as a reference to compute the error rate evolution since they provide the best results in the literature (Blin et al., 2020).

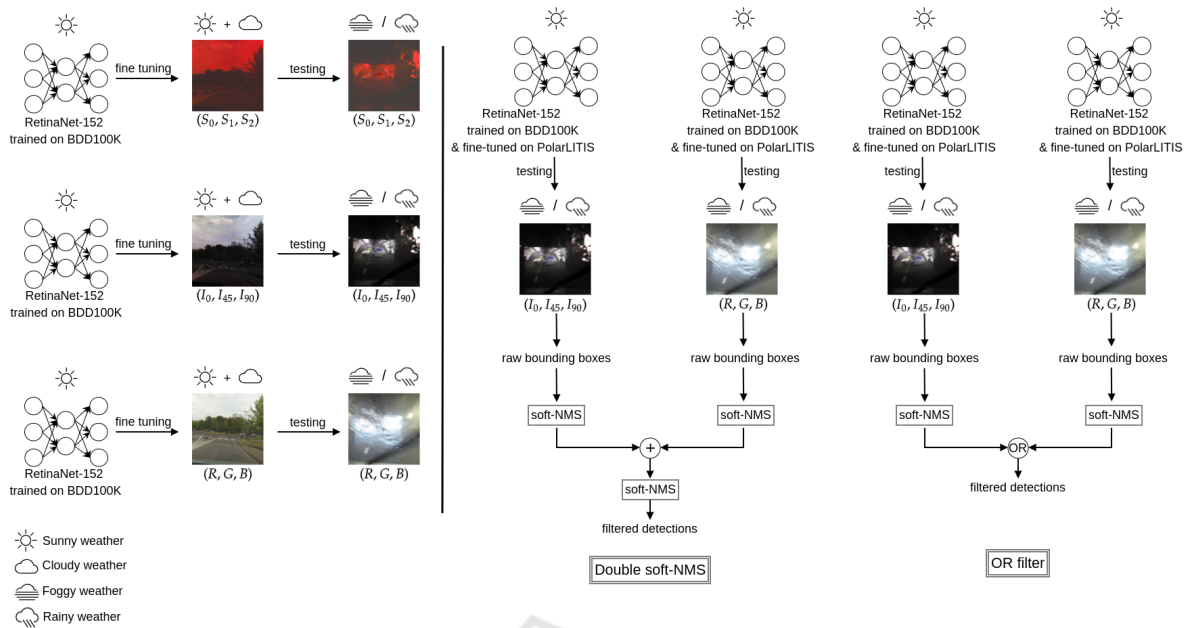


Figure 3: Experimental setup. On the left, the training process on each modality, respectively I , S and RGB is illustrated. On the right, the two fusion schemes (Double soft-NMS filter and OR filter) are illustrated with I and RGB fusion and can be extended to other modalities combinations.

As the PolarLITIS does not contain enough instances of the class bike, they are not taken into account during the evaluation process. It is important to remind that the architectures used for this experiment are exclusively trained on good weather conditions (sunny and cloudy) and tested exclusively on adverse scenes (foggy scenes with different visibility and rain). This pipeline enables to evaluate how polarimetric and color-based features vary with the visibility conditions. On top of that, since the acquisitions are made into a tunnel, the glare is an additional visibility alteration. The results of the experiments can be found in Table 3.

As can be seen in this table, regarding the three data formats, the polarimetric detection scores overcome the RGB detection scores in every adverse situation. We can also notice that the Stokes images S are more adapted to detect road objects in foggy scenes when the fog visibility is lower than 30m with up to 15% amelioration in the mAP . Stokes images are also more adapted to detect objects under tropical rain with a 30% increase in the mAP . When processing scenes under fog with other visibility distances, the intensity images are more adapted. These results are also summarized in Figure 5 regarding fog detection. Note that even if there is a gap between the detection scores corresponding to 35m and 40m visibility, it can be due to a higher number of non-ideal images (i.e. with harder objects to detect) contained in the class 40m visibility or ideal images contained in

the 35m visibility class. Nevertheless, the increasing curve tendency shows enhanced detection scores with a greater visibility distance. These results give a first intuition on the use of fusing intensities and Stokes images in order to improve road object detection in every situation.

Regarding the fusion schemes, when fusing intensities and Stokes images using a Late fusion scheme with the Double soft-NMS filter, it leads up to a 27% increase of the mAP for road object detection under fog and to a 42% increase of the mAP for road object detection under tropical rain. The same fusion scheme with an OR filter is less adapted since it takes into account the false positives, which are more numerous in adverse weather conditions as seen in Figure 4. It enables a slight amelioration for intensities and Stokes images fusion with up to a 5% amelioration of the mAP for road object detection under fog and a 2% increase of the mAP under tropical rain.

As for the polarimetric and color-based fusion, as mentioned previously, the Late fusion scheme using an OR filter is not adapted to fuse RGB and intensity images since it takes the false positives into account. The same pipeline using soft-NMS filter applied to these two modalities, however, is not adapted to every situation. They overcome the intensities and Stokes images fusion in foggy scenes when the visibility is the greatest, i.e. of 70m. The mAP in this situation is increased by 10%.

From all these results, we can conclude that when

Table 3: Comparison of the detection scores. The best detection scores for each adverse weather condition are in blue. The crosses (✖) remind that the RGB images of foggy scenes with 25m and 30m visibility are not available for this experiment.

Modality	Class	15m	20m	25m	30m	35m	40m	45m	50m	60m	70m	rain
<i>I</i>	person	49.46 ± 2.6	33.96 ± 5.1	35.63 ± 2.8	48.73 ± 4.6	68.76 ± 3.2	60.61 ± 3.1	71.89 ± 2.3	74.72 ± 1.3	76.49 ± 1.5	71.97 ± 2.1	73.68 ± 2.8
	car	0 ± 0	6.99 ± 1.2	41.65 ± 3.7	69.21 ± 4.3	85.21 ± 5.4	58.85 ± 7.4	75.40 ± 2.7	79.83 ± 8.5	90.97 ± 3.2	86.56 ± 4.6	85.69 ± 7.3
	<i>mAP</i>	24.73 ± 1.3	20.48 ± 2.4	38.64 ± 3.2	58.97 ± 4.1	76.98 ± 2.1	59.73 ± 4.2	73.65 ± 2.2	77.28 ± 4.7	83.73 ± 1.7	79.27 ± 2.6	79.69 ± 4.1
<i>S</i>	person	50.64 ± 5.1	51.58 ± 8.1	44.47 ± 3.4	66.69 ± 2.5	68.88 ± 4.1	68.61 ± 4.2	70.24 ± 2.2	77.29 ± 3.3	68.40 ± 2.2	68.43 ± 2.9	78.26 ± 6.1
	car	0.48 ± 0.6	13.73 ± 1.9	37.59 ± 8.5	47.16 ± 10.2	77.06 ± 7.2	53.51 ± 17.7	63.76 ± 14.4	64.61 ± 15.6	70.35 ± 7.4	70.78 ± 9.7	93.10 ± 4.4
	<i>mAP</i>	25.56 ± 2.7	32.66 ± 3.6	41.03 ± 4.6	56.92 ± 5.6	72.97 ± 4.9	61.06 ± 10.2	67.00 ± 8.2	70.95 ± 9.4	69.37 ± 4.1	69.61 ± 6.0	85.68 ± 4.7
ER_o^S	person	-2.3	-26.7	-13.7	-35.0	-0.4	-20.3	5.9	-10.2	34.4	12.6	-17.4
	car	-0.5	-7.3	7.0	71.6	55.1	13.0	47.3	75.5	228.3	117.4	-51.8
	<i>mAP</i>	-1.1	-15.3	-3.9	5.0	17.4	-3.3	25.2	27.9	88.3	46.6	-29.5
RGB	person	14.56 ± 4.9	16.27 ± 3.2	✖	✖	18.19 ± 2.5	16.68 ± 1.8	18.80 ± 3.1	12.00 ± 4.3	24.26 ± 4.3	18.22 ± 1.9	21.08 ± 4.3
	car	0.00 ± 0.0	0.00 ± 0.0	✖	✖	0.00 ± 0.0	6.82 ± 0.0	8.89 ± 0.4	1.84 ± 4.1	28.48 ± 0.7	20.90 ± 0.9	14.52 ± 3.5
	<i>mAP</i>	7.28 ± 2.5	8.14 ± 1.6	✖	✖	9.10 ± 1.2	11.75 ± 0.9	13.85 ± 1.7	6.90 ± 3.4	26.37 ± 2.1	19.56 ± 1.2	17.80 ± 4.3
ER_o^{RGB}	person	69.1	26.8	✖	✖	161.9	11.5	188.9	248.1	222.2	191.8	199.8
	car	0.0	7.5	✖	✖	576.1	126.4	270.4	386.7	692.0	488.5	497.3
	<i>mAP</i>	23.2	15.5	✖	✖	294.9	191.5	226.9	309.8	352.6	288.0	304.7
<i>I+S</i> Double soft-NMS	person	54.14 ± 3.1	53.03 ± 7.4	45.07 ± 3.2	69.23 ± 3.1	72.62 ± 2.7	69.84 ± 4.0	76.93 ± 2.1	78.99 ± 1.9	77.00 ± 1.4	74.51 ± 2.5	83.24 ± 2.8
	car	0.45 ± 0.6	13.60 ± 2.1	46.85 ± 5.3	71.19 ± 4.4	88.85 ± 2.4	61.70 ± 10.5	74.58 ± 4.3	75.40 ± 11.5	90.28 ± 3.8	84.59 ± 5.3	93.35 ± 3.7
	<i>mAP</i>	27.30 ± 1.7	33.31 ± 3.4	45.96 ± 3.3	70.21 ± 3.4	80.74 ± 2.5	65.77 ± 5.5	75.76 ± 3.1	77.19 ± 6.3	83.64 ± 2.3	79.55 ± 3.6	88.29 ± 2.3
ER_o^{I+S} (2x soft-NMS)	person	-10.3	-28.9	-14.7	-40.0	-12.4	-23.4	-17.9	-16.9	-2.2	-9.1	-36.3
	car	-0.5	-7.1	-8.9	-6.4	-24.6	-6.9	3.33	22.0	7.6	14.6	-53.5
	<i>mAP</i>	-3.4	-16.1	-11.9	-27.4	-16.3	-15.0	-8.0	0.4	0.6	-1.4	-42.3
<i>I+S</i> OR filter	person	51.05 ± 2.6	41.25 ± 5.5	37.49 ± 3.3	50.25 ± 4.9	69.10 ± 3.0	63.97 ± 1.7	72.30 ± 2.4	75.48 ± 1.8	76.81 ± 1.4	72.72 ± 2.2	73.61 ± 2.9
	car	0.08 ± 0.2	8.34 ± 2.7	40.30 ± 5.6	70.08 ± 3.3	83.42 ± 4.4	59.04 ± 7.1	74.66 ± 3.7	78.84 ± 7.9	90.82 ± 3.3	86.40 ± 4.0	86.64 ± 6.6
	<i>mAP</i>	25.56 ± 1.3	24.79 ± 2.4	38.90 ± 3.5	60.17 ± 3.8	76.26 ± 1.1	61.51 ± 4.0	73.48 ± 2.5	77.16 ± 4.3	83.82 ± 1.6	79.56 ± 2.4	80.13 ± 3.7
ER_o^{I+S} (OR filter)	person	-3.1	-11.0	-2.9	-3.0	-1.1	-8.5	-1.5	-3.0	-1.4	-2.7	0.3
	car	-0.1	-8.2	2.3	-2.8	12.1	-0.5	3.0	4.9	1.7	1.2	-6.6
	<i>mAP</i>	-1.1	-5.4	-0.4	-2.9	3.1	-4.4	0.6	0.5	-0.6	-1.4	-2.2
<i>I+RGB</i> Double soft-NMS	person	49.65 ± 2.7	38.06 ± 3.8	✖	✖	69.07 ± 4.5	61.08 ± 3.1	72.94 ± 2.6	75.10 ± 1.5	76.08 ± 3.1	72.83 ± 2.5	72.95 ± 2.7
	car	0.00 ± 0.0	5.49 ± 2.0	✖	✖	83.64 ± 5.2	59.26 ± 6.26	76.80 ± 2.4	78.39 ± 9.7	87.59 ± 3.6	89.95 ± 3.2	77.23 ± 7.5
	<i>mAP</i>	24.83 ± 1.4	21.77 ± 1.5	✖	✖	76.35 ± 2.5	60.17 ± 3.5	74.87 ± 2.2	76.74 ± 5.4	81.84 ± 3.2	81.39 ± 2.6	75.09 ± 3.6
ER_o^{I+RGB} (2x soft-NMS)	person	-0.4	-6.2	✖	✖	-1.0	-1.2	-3.7	-1.5	1.7	-3.1	2.8
	car	0.0	1.6	✖	✖	10.6	-1.0	-5.7	7.1	37.4	-25.2	59.1
	<i>mAP</i>	-0.1	-1.6	✖	✖	2.7	-1.1	-4.6	2.4	11.6	-10.2	22.6
<i>I+RGB</i> OR filter	person	49.58 ± 2.6	31.34 ± 5.8	✖	✖	68.74 ± 3.3	60.26 ± 3.1	71.35 ± 1.9	74.19 ± 2.0	73.89 ± 1.5	72.43 ± 2.2	73.72 ± 2.8
	car	0.00 ± 0.0	6.29 ± 1.9	✖	✖	84.94 ± 5.5	59.10 ± 7.4	74.88 ± 2.8	79.12 ± 9.1	89.89 ± 3.0	86.54 ± 4.6	81.72 ± 7.2
	<i>mAP</i>	24.79 ± 1.3	18.82 ± 3.0	✖	✖	76.84 ± 2.0	59.68 ± 4.2	73.12 ± 1.8	76.65 ± 5.5	81.89 ± 1.8	79.49 ± 2.6	77.72 ± 3.9
ER_o^{I+RGB} (OR filter)	person	-0.2	-4.0	✖	✖	0.1	0.9	1.9	2.1	11.1	-1.6	-0.2
	car	0.0	0.8	✖	✖	18.3	-0.6	2.1	3.5	12.0	0.1	27.7
	<i>mAP</i>	-0.1	2.1	✖	✖	0.6	0.1	2.0	2.8	11.3	-1.1	9.7



Figure 4: Examples of false positives detection in adverse weather conditions. Blue red and orange bounding boxes respectively denotes car, person and bike detection.

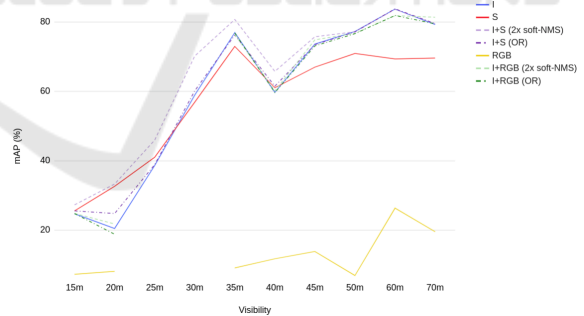


Figure 5: Evolution of the *mAP* in foggy scenes while varying the visibility distance. *I*, *S* and the RGB scores are respectively in blue, red and yellow (full lines). The fusion scores of *I* and *S* are respectively in pale and dark purple for the Double soft-NMS and the OR filters (dashed lines). The fusion scores of *I* and RGB are respectively in pale and dark green for the Double soft-NMS and the OR filters (dashed lines).

the visibility is very low, as it is the case in very dense fog and tropical rain, the Stokes and intensity images fusion provide the best results. The color-based and polarimetric fusion is beneficial in adverse road scenes with a better visibility, such as light fog. How-

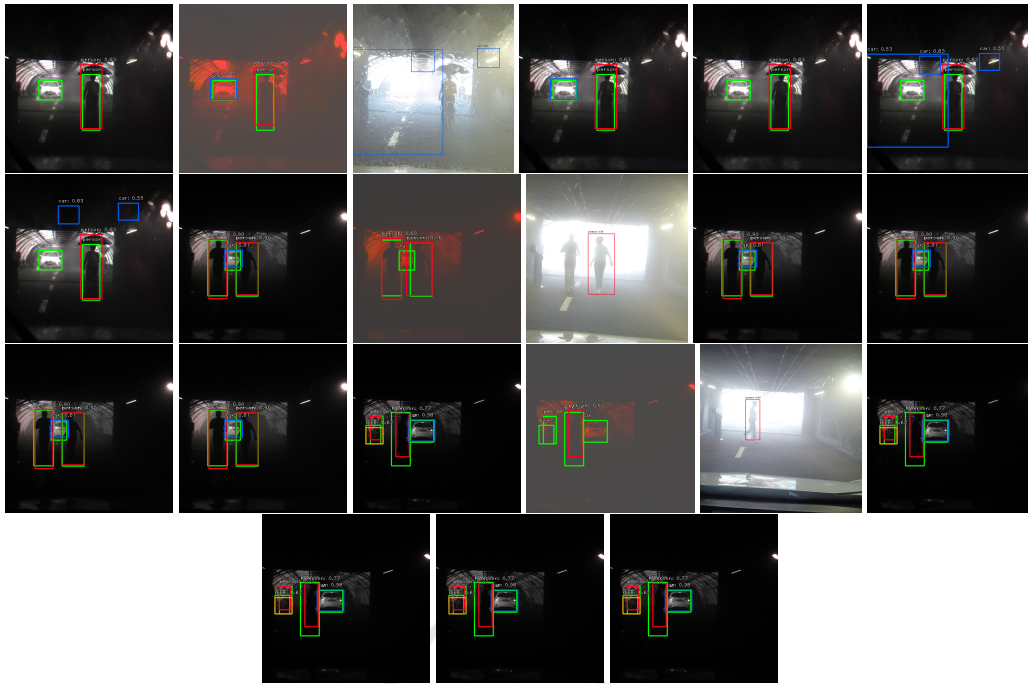


Figure 6: Detection results in several adverse weather conditions. From top to bottom: tropical rain and fog with respectively 35m and 60m visibility. From left to right: I , S , RGB, $I+S$ (Double soft-NMS), $I+S$ (OR filter), $I+RGB$ (Double soft-NMS) and $I+RGB$ (OR filter). Bounding boxes in green, blue, red and orange denote respectively the ground truth, car, person and bike detection.

ever, as can be seen in Figure 5, when the visibility is lower than 30m, the polarimetric features learnt in good weather conditions are not able to detect efficiently all road objects in adverse scenes. This limitation could be overcome by including adverse situations in the training process. Despite this limitation, the experimental results show that under tropical rain and under fog from 30m visibility, polarimetric features are a real added value to enhance road objects detection. Moreover, as it can be seen in Figure 2, polarimetric features are more robust to the glare and to drops or veils of water on the windshield, causing deformation and loss of information in color-based images. Overall, polarimetric features are more adapted than color-based features to characterize objects in unexpected visibility alterations, as it is illustrated in Figure 6.

6 CONCLUSION AND PERSPECTIVES

In this work, polarimetric features, including polarimetric intensities and Stokes parameters, prove to be a real added value to enhance object detection in a wide range of adverse weather conditions. The experimental results demonstrate that, unlike color-based

features, polarimetric features are invariant to the visibility variations induced by fog and rain or the glare. This property implies that the features learnt in good weather conditions are still valid to detect road objects in adverse weather. Using a well chosen fusion scheme, polarimetric intensity images combined with Stokes images lead up to a 27% increase of road object detection under fog and to a 42% increase of road object detection under tropical rain. The combination of polarimetric and color-based features, however, finds its utility to analyze road scenes when the visibility gets better. Overall, polarimetric features are more robust than color-based ones to unexpected visibility changes.

It is important to note that, the polarimetric features learnt in good weather conditions show limits to efficiently describe a road scene when the visibility is very low. This limitation should be palliated by including polarimetric road scenes in adverse weather conditions in the training process to increase road object detection in low visibility. This work also aims in a close future to repeat the experiments on real road scenes in several adverse weather conditions such as hail or snow.

ACKNOWLEDGEMENTS

This work is supported by the ICUB project 2017 ANR program : ANR-17-CE22-0011.

REFERENCES

- Ainouz, S., Morel, O., Fofi, D., Mosaddegh, S., and Benschair, A. (2013). Adaptive processing of catadioptric images using polarization imaging: towards a pola-catadioptric model. *Optical engineering*, 52(3):037001.
- Aldibaja, M., Sukanuma, N., and Yoneda, K. (2016). Improving localization accuracy for autonomous driving in snow-rain environments. In *2016 IEEE/SICE International Symposium on System Integration (SII)*, pages 212–217. IEEE.
- Bass, M., Van Stryland, E. W., Williams, D. R., and Wolfe, W. L. (1995). *Handbook of optics*, volume 2. McGraw-Hill New York.
- Bijelic, M., Gruber, T., Mannan, F., Kraus, F., Ritter, W., Dietmayer, K., and Heide, F. (2020). Seeing 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*, pages 11682–11692.
- Bijelic, M., Gruber, T., and Ritter, W. (2018). Benchmarking image sensors under adverse weather conditions for autonomous driving. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1773–1779. IEEE.
- Blin, R., Ainouz, S., Canu, S., and Meriaudeau, F. (2019). Road scenes analysis in adverse weather conditions by polarization-encoded images and adapted deep learning. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 27–32. IEEE.
- Blin, R., Ainouz, S., Canu, S., and Meriaudeau, F. (2020). A new multimodal rgb and polarimetric image dataset for road scenes analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 216–217.
- Bodla, N., Singh, B., Chellappa, R., and Davis, L. S. (2017). Soft-nms—improving object detection with one line of code. In *Proceedings of the IEEE international conference on computer vision*, pages 5561–5569.
- Fan, W., Ainouz, S., Meriaudeau, F., and Benschair, A. (2018). Polarization-based car detection. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3069–3073. IEEE.
- Felzenszwalb, P., McAllester, D., and Ramanan, D. (2008). A discriminatively trained, multiscale, deformable part model. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE.
- Feng, D., Haase-Schütz, C., Rosenbaum, L., Hertlein, H., Glaeser, C., Timm, F., Wiesbeck, W., and Dietmayer, K. (2020). Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges. *IEEE Transactions on Intelligent Transportation Systems*.
- Gu, S., Lu, T., Zhang, Y., Alvarez, J. M., Yang, J., and Kong, H. (2018). 3-d lidar+ monocular camera: An inverse-depth-induced fusion framework for urban road detection. *IEEE Transactions on Intelligent Vehicles*, 3(3):351–360.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988.
- Major, B., Fontijne, D., Ansari, A., Teja Sukhavasi, R., Gowalkar, R., Hamilton, M., Lee, S., Grzechnik, S., and Subramanian, S. (2019). Vehicle detection with automotive radar using deep learning on range-azimuth-doppler tensors. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*.
- Nie, J., Yan, J., Yin, H., Ren, L., and Meng, Q. (2020). A multimodality fusion deep neural network and safety test strategy for intelligent vehicles. *IEEE Transactions on Intelligent Vehicles*, pages 1–1.
- Pinchon, N., Cassignol, O., Nicolas, A., Bernardin, F., Leduc, P., Tarel, J.-P., Brémond, R., Bercier, E., and Brunet, J. (2018). All-weather vision for automotive safety: which spectral band? In *International Forum on Advanced Microsystems for Automotive Applications*, pages 3–15. Springer.
- Rashed, H., Ramzy, M., Vaquero, V., El Sallab, A., Sistu, G., and Yogamani, S. (2019). Fusemodnet: Real-time camera and lidar based moving object detection for robust low-light autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*.
- Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., Madhavan, V., and Darrell, T. (2020). Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2636–2645.