A new multimodal RGB and polarimetric image dataset for road scenes analysis

Rachel Blin
Normandie Univ
INSA Rouen, UNIROUEN,
UNIHAVRE, LITIS, 76000 Rouen, France
rachel.blin@insa-rouen.fr

Samia Ainouz
Normandie Univ
INSA Rouen, UNIROUEN,
UNIHAVRE, LITIS, 76000 Rouen, France
samia.ainouz@insa-rouen.fr

Stéphane Canu
Normandie Univ
INSA Rouen, UNIROUEN,
UNIHAVRE, LITIS, 76000 Rouen, France
stephane.canu@insa-rouen.fr

Fabrice Meriaudeau
University of Burgundy
UBFC, ImViA,
71200 Le Creusot, France
fabrice.meriaudeau@u-bourgogne.fr

Abstract

Road scene analysis is a fundamental task for both autonomous vehicles and ADAS systems. Nowadays, one can find autonomous vehicles that are able to properly detect objects present in the scene in good weather conditions but some improvements are left to be done when the visibility is altered. People claim that using some non conventional sensors (infra-red, Lidar, etc.) along with classical vision enhances road scene analysis but still when conditions are optimal. In this work, we present the improvements achieved using polarimetric imaging in the complex situation of adverse weather conditions. This rich modality is known for its ability to describe an object not only by its intensity but also by its physical information, even under poor illumination and strong reflection. The experimental results have shown that, using our new multimodal dataset, polarimetric imaging was able to provide generic features for both good weather conditions and adverse weather ones. By combining polarimetric images with an adapted learning model, the different detection tasks in adverse weather conditions were improved by about 27%.

1. Introduction

The development of autonomous driving and ADAS systems those past few years has been possible thanks to the great improvements in understanding road scenes. Object detection is a fundamental step to achieve a reliable road scene analysis. Deep neural networks have shown great performances to accomplish this task with a high accuracy and by processing several frames per second. One can cite the example of Faster R-CNN [24], which achieved a high precision and more recently SSD [19], RetinaNet [17] and YOLOv3 [23] which were able to be as precise as Faster R-CNN with a higher frame rate. Nowadays, autonomous vehicles can be found in some parts of the world like the Waymo car in Arizona [13] and the Rouen Normandy autonomous lab[1] in France. Those systems achieved a high autonomy in driving especially in good weather conditions and urban areas at a cost of low speed limit. However, when it comes to adverse weather conditions, object detection is still an issue that autonomous vehicles are facing. Indeed, in such weather conditions, the visibility of an object is altered and the conventional imaging, which relies mostly on an object's color and shape to achieve a good detection, fails to characterize them properly. This is the reason why non conventional sensors could be the best alternative to conventional ones for improving road scenes analysis [6]. Aldibaja *et al*. used Lidar to improve localization accuracy under snowy and rainy weather conditions [3]. Pinchon *et al*. showed that by using an infra-red camera, vehicles could be detected at a larger distance range under foggy weather conditions [22]. However, their results were limited because of the noise in adverse weather that had to be removed prior to process the Lidar data and because of the impossibility to detect vehicle lights, road markings and traffic signs

[1] https://www.rouennormandyautonomouslab.com/

using infra-red camera in foggy weather.

Polarimetric imaging is a rich modality that enables to characterize objects not only by their color or their shape but also regarding their physical information. Each pixel of a polarimetric image encodes information about the nature of the object's roughness, its orientation and its reflection [28]. Polarimetric imaging has many applications in a wide range of fields. In the 3D vision field, the use of polarisation by Berger *et al*. [5] helped to enhance indoor autonomous navigation. Polarization was also used by Zhu *et al*. [29] to estimate the depth map of an object and by Morel *et al*. [21] to reconstruct 3D objects using the angle and degree of polarization (respectively the orientation and quantity of polarized light in the wave) and the normal surface. The link between polarization parameters and the fog granularity was exploited by Schechner *et al*. [25] to achieve scene restoration and in the same time depth estimation with high accuracy. Finally, in the field of autonomous driving and ADAS systems, recent works have came up with new polarization-based features showing a significant improvement in road scene analysis. Fan *et al*. [11] showed that fusing polarimetric features with RGB ones enhanced car detection. They first performed a selection among five relevant polarimetric features to get the most informative one. After experiments, it turned out that the Angle Of Polarization (AOP) was the best feature to achieve this task. They trained an AOP and a color-based Deformable Part Models (DPM) [12] independently producing two different score maps. The polar-based model was taken as a confirmation to the color-model using an AND-fusion scheme in order to get the final detection bounding boxes. By taking these two complementary information, it was proven that the false alarm rate was largely reduced and the detection accuracy improved. Kamman *et al*. [15] used a radar and a reflective system to capture the polarization of the electromagnetic wave to reduce the detection of ghost objects by radar. Their reflective system enabled to characterize the nature of the object detected and to avoid the confusion between an object and its ghost equivalent. Recently, Blin *et al*. [7] used polarimetric images in order to improve object detection in road scenes in adverse weather conditions. New data formats with polarimetric features were constitued to best characterize objects. By using polarimetric imaging and an adapted learning model, both vehicles and pedestrian detection were improved under adverse weather conditions while other methods failed. In this work, the experiments were extended by proposing a large variety of labelled polarimetric road scenes but also their equivalent in RGB format in order to reiterate the experiment on both modalities for fair comparison. A new dataset containing multimodal RGB and polarimetric images for road scene analysis in adverse weather conditions is then proposed. The whole acquired dataset for our experiments is available at: http://pagesperso.litislab.fr/rblin/databases/

The contributions of this paper are threefold:

- as far as our knowledge can go, we propose the first multimodal dataset, containing pairwise RGB and polarimetric version of the same road scene,
- we propose new polarimetric image formats optimizing object detection in adverse weather,
- we enhance object detection in road scenes in adverse weather conditions.

2. Polarization formalism

In its propagation plan, the electrical field of a progressive transverse wave, with a pulsation ω and a phase ϕ, in the orthonormal basis $B = \{u_x, u_y\}$, is given by:

$$\begin{aligned}\vec{E}(t) = E_x(t)\cos(-k\vec{z} + \omega t)\vec{u_x} \\ + E_y(t)\sin(-k\vec{z} + \omega t + \phi(t))\vec{u_y}\ ,\end{aligned} \tag{1}$$

where k is the wave number, $\vec{z}$ the direction of propagation and E_x, E_y are respectively the amplitudes of $\vec{E}(t)$ according to u_x and u_y.

Polarization is the property of light waves able to describe the direction in which the wave is travelling. It is better understood by introducing the three states of polarization of the light [4]:

- the wave is totally polarized when the direction of its electrical field is well determined, *i.e*. elliptic, linear or circular,
- the wave is unpolarized in the case the light waves oscillate in totally random directions,
- it is partially polarized when it is a combination of a polarized part and an unpolarized part.

Polarimetric imaging consists in the giving of the polarization state of the reflected light (Stokes vector) for each pixel of a scene. It is used to dissociate metallic object from dielectric surface [8]. In the mechanism of the polarization, when an unpolarized light is being reflected by an object, it becomes partially linearly polarized. The reflected wave light can be described by a measurable vector, called the linear Stokes vector, $S = \begin{bmatrix} S_0 & S_1 & S_2 \end{bmatrix}^\top$. It is defined as the co-variance parameters of the Electrical field components of equation 1:

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \end{bmatrix} = \begin{bmatrix} \langle E_x^2 \rangle + \langle E_y^2 \rangle \\ \langle E_x^2 \rangle - \langle E_y^2 \rangle \\ 2\langle E_x E_y \cos(\phi) \rangle \end{bmatrix}\ ,$$

where σ_x and σ_y are respectively the components according to u_x and u_y of the phase of the electromagnetic wave. $\langle . \rangle$ is referred to the temporal mean of the signal.

By its construction, the Stokes parameters satisfy the physical admissibility constraints defined by:

$$S_0 > 0 \quad \text{and} \quad S_0^2 \geqslant S_1^2 + S_2^2 \ .$$

It is important to notice that the first constraint means that any object reflects an energy. The second constraint means that the total energy is always greater than the sum of the partial ones. The wave is thus totally polarized if the equality holds meaning that $S_0^2 = S_1^2 + S_2^2$. It is partially polarized if we have strict inequality and unpolarized if $S_0 > 0$, $S_1 = 0$, $S_2 = 0$.

In order to obtain polarimetric images, a polarizer oriented at a specific angle α_i is placed between the scene and the sensor. The sensor measures an intensity according to each polarizer rotation angle α_i. To get the three Stokes parameters, at least three different orientations of the polarizer are needed. For our acquisition, a Polarcam 4D Technology polarimetric camera is used. The technology of such a sensor uses 4 linear polarizers oriented at four different angles ($0°$, $45°$, $90°$ and $135°$) enabling to get simultaneously four different intensities $I(\alpha_i)_{i=1:4}$ of the same scene. The relationship between each intensity $I(\alpha_i)$ and the Stokes parameters is :

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

In a more compact way, equation 2 can be written as:

$$I = AS \ , \tag{3}$$

where $I = \begin{bmatrix} I_0 & I_{45} & I_{90} & I_{135} \end{bmatrix}^\top$ refers to the four intensities according to each angle of the polarizer. $S = \begin{bmatrix} S_0 & S_1 & S_2 \end{bmatrix}^\top$ is the Stokes vector and $A \in \mathbb{R}^{4\times 3}$ is the calibration matrix of the linear polarizer, defined as:

$$\begin{aligned} A &= \frac{1}{2}\begin{bmatrix} 1 & \cos(2\alpha_1) & \sin(2\alpha_1) \\ 1 & \cos(2\alpha_2) & \sin(2\alpha_2) \\ 1 & \cos(2\alpha_3) & \sin(2\alpha_3) \\ 1 & \cos(2\alpha_4) & \sin(2\alpha_4) \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \ . \end{aligned} \tag{4}$$

Knowing the intensities $I(\alpha_i)_{i=1:4}$ reaching the camera and the calibration matrix A, the only unknowns in equation 3 are the Stokes parameters. As the matrix A is not square, the most used solution in the literature to get the Stokes parameters for each pixel is the least mean square solution. The Stokes vector is then calculated by:

$$S = \tilde{A}I \ , \tag{5}$$

where $\tilde{A} = (A^tA)^{-1}A^t$ the pseudo-inverse matrix of A. The proposed mean square solution is submitted to some additional constraints on the acquired intensities. Indeed, if we combine equations 3 and 5, we get:

$$I = A\tilde{A}I \ . \tag{6}$$

This equality holds if and only if:

$$I_0 + I_{90} = I_{45} + I_{135} \ . \tag{7}$$

Note that from equations 3 and 5 and 7, the Stokes vector can be given by:

$$S = \begin{bmatrix} I_0 + I_{90} \\ I_0 - I_{90} \\ I_{45} - I_{135} \end{bmatrix} \ . \tag{8}$$

Other physical parameters can be obtained from the Stokes parameters, the Angle Of Polarization (AOP) and the Degree Of Polarization (DOP) [2]. They can be determined with the following formulas:

$$AOP = \frac{1}{2}\arctan\left(\frac{S_2}{S_1}\right) \ , \tag{9}$$

$$DOP = \frac{\sqrt{S_1^2 + S_2^2}}{S_0} \ . \tag{10}$$

The $DOP \in [0, 1]$ is one of the most important physical properties. It refers to the quantity of polarized light in a wave. It is equal to one for a totally polarized light, between 0 and 1 for the partially polarized light and up to zero for an unpolarized light. The $AOP \in \left[-\frac{\pi}{2}; \frac{\pi}{2}\right]$ is the orientation of the polarized part of the wave with regards to the incident plan. Figure 1 illustrates an example of a road scene and its representation with different polarimetric parameters. From now on AOP and DOP are respectively referred as ϕ and ρ.

3. The proposed method

In this section, we are going to introduce the new acquired dataset named PolarLITIS. The different formats that constitute the dataset are presented followed by the whole protocol used for the detection purpose.

3.1. The PolarLITIS dataset

In a previous work [7], a first version of a polarimetric dataset containing diverse polarimetric encoded road

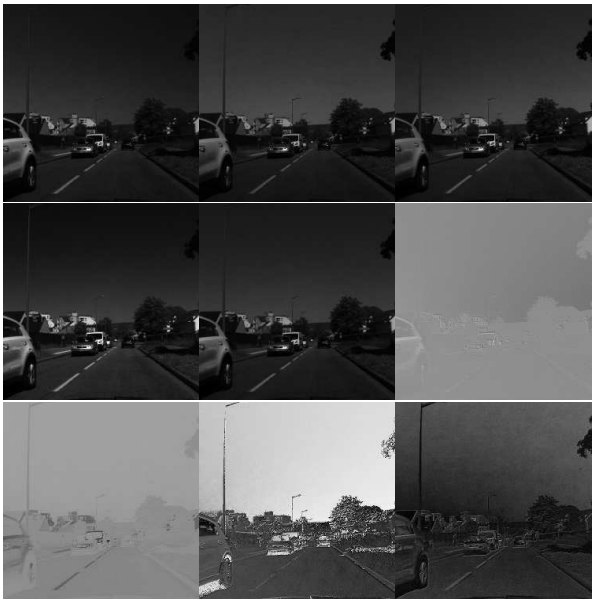

Figure 1: From left to right and from top to bottom: I_0, I_{45}, I_{90}, I_{135}, S_0, S_1, S_2, ϕ and ρ.

scenes in sunny and foggy weather conditions was presented. This dataset was limited as the RGB counterpart of the polarimetric image was missing . In order to carry out a fair comparison between the detection in both modalities, the new dataset acquired in this work is more diverse and contains a couple of polarimetric and RGB images for each scene. The training, validation and testing sets come from this new dataset.

For the acquisition setup, a GoPro and a polarimetric camera were embedded behind the windshield of a car. The different acquisitions were done in various places, including highways, cities, small villages, parkings and academic areas. The images contain scenes under three different weather conditions ; foggy, sunny and cloudy. The training set is composed of sunny and cloudy scenes, the validation set of cloudy scenes and the testing set is exclusively composed of foggy scenes. We chose to process by this way because the foggy scenes were taken in the same place which limits the variability of this set. By including foggy images into the training set, there is a high risk of overfitting in the training process. Also, knowing that the polarimetric features are robust to lighting variations, the detection on adverse weather conditions scenes should be possible even by using a network trained only on good weather conditions. Figure 2 illustrates this variety.

To sort the dataset, the number of frames per second (fps) each camera could capture was taken into account. The polarimetric camera is able to capture 25 fps and the GoPro was set at 30 fps. In order to get diverse enough frames in the final dataset, one out of 50 frames for the

Figure 2: Examples of the scenes contained in our dataset. First row is the the intensity I_0 of the polarimetric scene and second row is the RGB equivalent.

Class	Train	Val	Test
Images	1640	420	509
car	6061	2102	9265
person	527	134	442
bike	39	7	7
motorbike	14	5	0

Table 1: Dataset features. Here the number of instances of each object is given. Those properties are available for both RGB and polarimetric datasets.

polarimetric camera and one out of 60 frames for the GoPro were kept.

It is important to note that the output of the polarimetric camera has a resolution of 500×500 pixels and the GoPro images are 3648×2736. Moreover, unlike the polarimetric camera, the GoPro is equipped with a fisheye lens. To get the closest content possible, the RGB images were cropped into 906×945 pixels so they could have the same content. The deformation caused by the fisheye lens is also reduced because the interesting content of the images is mostly located at their center.

Four different classes were labeled in this dataset. The first class, named 'car', contains all kinds of vehicles, including trucks, buses and vans. The second class, named 'person', contains all kind of road users, except car drivers. The third and fourth classes are respectively 'bike' and 'motorbike'. Table 1 sums up all the properties of each class over the dataset.

3.2. Encoding polarimetric images for machine learning

In order to compare polarimetric images to RGB ones, it is important to give images constituted of three channels, each one corresponding to a polarimetric information. Because pre-trained networks are used for the experiments, it is important to keep three channels for polarimetric data to achieve efficient training. To this end, five different polarimetric data formats are con-

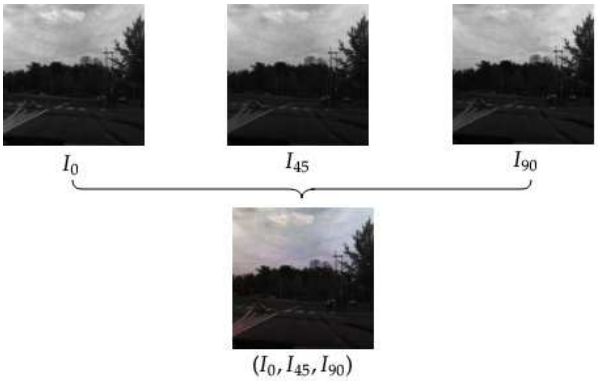

Figure 3: Intensity image. I_0, I_{45} and I_{90} are placed respectively as the RGB configuration.

stituted. In order to get more homogeneous images, each channel of the polarimetric images is normalized between 0 and 255. This way, the polarimetric data is coded in the same range as the RGB ones to be processed in the same way by neural networks.

3.2.1 Intensities images

This data format gathers three intensities $I(\alpha_i)_{i=1:3}$ associated to three angles of the linear polarizer α_i. The choice of three angles instead of four comes from equation 7 meaning that the fourth channel could be deduced from the three others. This is the reason why the intensities I_0, I_{45} and I_{90} are considered to contain all the necessary information for the learning process and that I_{135} would be redundant. The data format referred to as intensities images is I_0, I_{45} and I_{90}. An example of such coding is illustrated in Figure 3.

3.2.2 Stokes images

The linear Stokes vector is a rich polarimetric feature that directly describes information on the reflected light wave. The three Stokes parameters are thus chosen to constitute another data format. We will refer to them as the Stokes images and an illustration of this data format can be found in Figure 4.

3.2.3 Pauli inspired images

This data format is a mix between the intensities $I(\alpha_i)_{i=1:4}$ and the linear Stokes vector. It is inspired by the Pauli decomposition of the polarimetric information contained in polarization-encoded SAR images. The Pauli decomposition has shown high performances in image classification [20], [26]. Unlike, the four intensity images I_0, I_{45}, I_{90} and I_{135}, the polarimetric SAR imaging are rather encoded as I_{HH}, I_{HV}, I_{VH} and I_{VV}

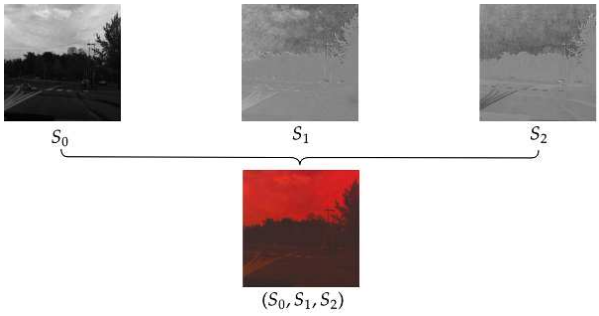

Figure 4: Stokes image. S_0, S_1 and S_2 are placed respectively as the RGB configuration.

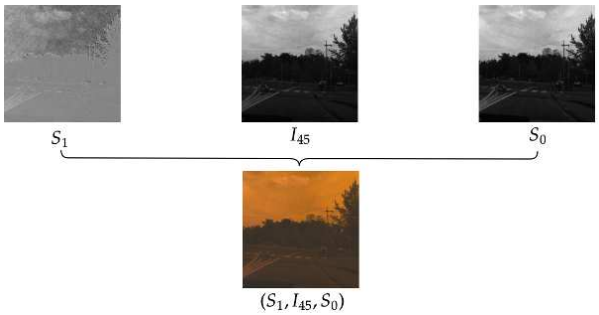

Figure 5: Pauli inspired image. S_1, I_{45} and S_0 are placed respectively as the RGB configuration.

which respectively refer to orientations H (horizontal) and V (vertical) of the received and the transmitted light wave. For instance, the crossed polarization I_{HV} represents the intensity of an horizontal transmitted light by a source and a vertical received light wave by the sensor. The original Pauli decomposition is carried out for $I_{HH} - I_{VV}$, I_{HV} and $I_{HH} + I_{VV}$ components placed respectively as the R, G and B coding of the image. According to the rotation of the polarizer, similarities between I_{HH} and I_0 could be noticed meaning that the polarizer oriented at an angle of $0°$ can be assimilated to an horizontal polarization filter. In the same way, I_{VV} can be assimilated to I_{90} for the vertical filter. We chose to assimilate I_{VH} to I_{45} as it corresponds to the mean orientation between the horizontal and the vertical filters. From equation 8, the Stokes parameters S_0 and S_1 are calculated as $S_0 = I_0 + I_{90}$ and $S_1 = I_0 - I_{90}$. Following the Pauli inspired format, the data were encoded as $(I_0 - I_{90}, I_{45}, I_0 + I_{90})$ which means (S_1, I_{45}, S_0) respectively coded in the same (R, G, B) configuration. An illustration of this format can be found in Figure 5.

3.2.4 HSV images

From the polarimetric features, it is possible to obtain the HSV (Hue, Saturation, Value) format of the scene [27]. This equivalence between the HSV encoding and

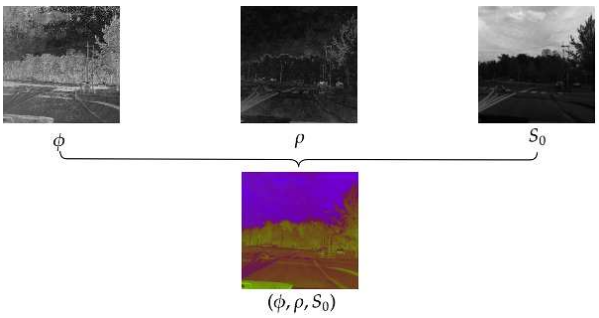

Figure 6: HSV image. ϕ, ρ and S_0 are placed respectively as the RGB configuration.

the polarization parameters is based on the intuition that the angle of polarisation ϕ corresponds to the Hue channel, the degree of polarization ρ to the Saturation one and the total intensity S_0 to the Value of each pixel. An illustration of this data format can be found in Figure 6. It is important to note that to respect the HSV format, the Hue channel is normalized between 0 and 180.

3.2.5 Poincaré inspired images

This data format is inspired by the representation of the Stokes vector, normalized by its first component S_0, in the Poincaré sphere [1]. In the case of linear polarization, the normalized Stokes vector is of dimension three instead of dimension four in a general Stokes formalism. The Stokes vector normalized by its first element is no longer represented in a Poincaré sphere, but in a unitary circle. An illustration of this representation can be found in Figure 7. This representation illustrates equations 9 and 10 as well. The projection of ρ on the x and y axes of the unit circle results in two components, $\rho\cos(2\phi)$ and $\rho\sin(2\phi)$. These two components are the ones used to constitute this new data format defined S_0, $\rho\cos(2\phi)$ and $\rho\sin(2\phi)$ respectively encoded in the same R, G and B configuration. An illustration of this data format can be found in Figure 8.

To sum up, we came up with five polarimetric data formats which are adapted to be processed by deep neural networks. Tabular 2 recaps the content of each channel of the five polarimetric data formats.

4. Experimental setup

In order to compare the detection in road scenes in adverse weather conditions using RGB and polarimetric encoded images, we carried the following experiment. Before going into further details regarding the description of this experiment, it is important to remind the composition of our dataset. Because our training set is composed of 1640 images, it is impossible to train the model from scratch without over-fitting a deep neural network [9]. This is the reason why we chose to fine-tune a network, pre-trained on another dataset. By using a network that has already converged into a larger dataset as a basis for our experiments, we make sure the network learns the general features of our dataset while saving some computation time.

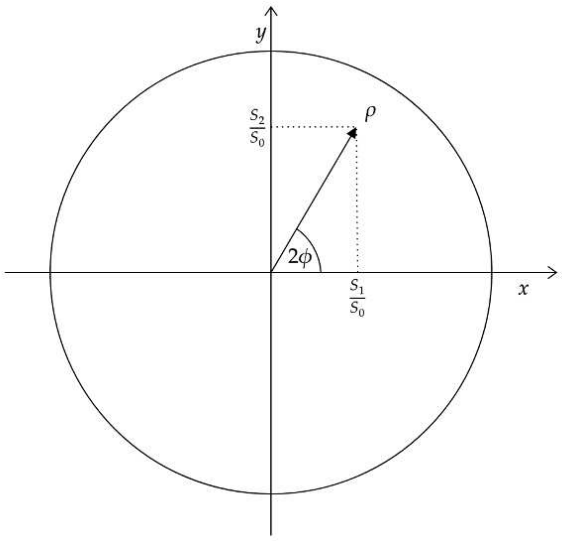

Figure 7: Representation of the polarized light wave in the unitary circle.

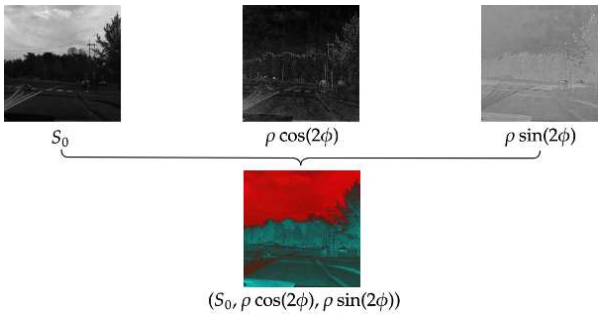

Figure 8: Poincaré image. S_0, $\rho\cos(2\phi)$ and $\rho\sin(2\phi)$ are placed respectively as the RGB configuration.

Data format	Channel 1	Channel 2	Channel 3
I	I_0	I_{45}	I_{90}
S	S_0	S_1	S_2
Pauli	S_1	I_{45}	S_0
HSV	ϕ	ρ	S_0
P	S_0	$\rho\cos(2\phi)$	$\rho\sin(2\phi)$

Table 2: Summary of the different polarimetric data formats. Here I, S, Pauli, HSV and P stand respectively for intensities images, Stokes images, Pauli inspired images, HSV images and Poincaré inspired images.

The goal of the experiment is to evaluate the best modality for describing a road scene in adverse weather conditions. We chose to use RetinaNet network using a ResNet50 [14] as a backbone for this task. From now on, this architecture will be referred as RetinaNet-50. The choice of this architecture was motivated by the fact it has a frame rate of 14 fps and because it outperforms Faster R-CNN on PASCAL VOC 2007 [10], which was known for its high accuracy. As mentionned previously, a network pre-trained on a larger dataset is needed for this task. We selected a publicly available model of RetinaNet-50 pre-trained on the MS COCO dataset [18] as a basis for the experiments. In order to find the best data format, the pre-trained network is fine-tuned on each one of them separately and evaluated on their respective test set. In order to be able to have a strong basis for the comparison, the same operation was also carried out on the RGB dataset. Figure 9 sums up the different experiments.

Regarding the training hyperparameters, the ones suggested in RetinaNet's article [17] were selected, *i.e.* a learning rate of 10^{-5} and the Adam optimizer [16]. Because the fine tuning process doesn't need a lot of epochs to converge, all the networks were trained on 20 epochs each. The optimal weights were found according to the lowest value of the validation loss.

5. Discussion and results

Due to the fact the classes 'bike' and 'motorbike' don't have a lot of instances in the dataset, the experiments are evaluated only on classes 'car' and 'person'. For this section, the following formulas for the mean average precision (mAP) and the error rate evolution (ER) [10] were used:

$$mAP^d = \frac{AP_p^d + AP_c^d}{2} \quad ,$$

where mAP^d is the mean average precision associated to the data format d, AP_p^d and AP_c^d are the average precision respectively for the classes 'person' and 'car' for the related data format d,

$$ER_o^d = \frac{1 - AP_o^d - (1 - AP_o^{RGB})}{1 - AP_o^{RGB}} \times 100 \quad ,$$

where ER_o^d is the error rate evolution associated to the polarimetric data format d for object $o \in \{'person', 'car'\}$, AP_o^{RGB} is the average precision for object o with the RGB data format while AP_o^d denotes the average precision on the object o and the related polarimetric data format d.

After fine tuning RetinaNet-50 pre-trained on MS COCO on all the data formats, the obtained models were evaluated on the testing set. As a reminder, the training process focused on sunny and cloudy scenes and the testing process on foggy scenes. The idea behind this configuration is to prove that the polarimetric features learnt in good weather conditions were invariant in adverse weather conditions. The obtained results of this training are summed up in Table 3.

Format	RGB	I	S	Pauli	HSV	P
car	73.28	**77.18**	72.43	**75.09**	30.86	39.10
person	80.97	**89.27**	**86.01**	**88.45**	57.52	69.14
mAP	77.13	**83.23**	**79.22**	**81.77**	44.19	54.12

Table 3: Comparison of the detection using RetinaNet-50 on the different data formats described in Table 2. In blue we have the RGB detection scores in percentage and in bold the detection scores that overcome it.

The obtained results showed that three out of five polarimetric data formats overcame RGB in terms of detection in adverse weather conditions. Regarding the Intensities images, we can notice a 15% augmentation regarding vehicles detection, a 44% augmentation regarding pedestrian detection which leads to a 27% augmentation in the mAP. For the Stokes images, there's an augmentation of 26% regarding pedestrian detection that leads to a 9% augmentation of the mAP. As for the Pauli inspired images, there's a 7% augmentation regarding car detection and a 39% augmentation regarding pedestrian detection that leads to a 20% augmentation of the mAP. Figure 10 illustrates the detection results on the same scene, showing the differences between RGB detection and its polarimetric equivalent in each data format.

The added value of polarimetric-encoded imaging for detecting objects in road scenes in adverse weather conditions is proven by this experiment. Because it characterizes an object not only regarding its shape but by its reflection, polarimetry provides strong features to well describe an object. Because these features are invariant to strong illumination and altered visibility, they are a strong asset to analyze road scenes in adverse weather conditions, where we can no longer rely on basic features such as color or shape to describe an object.

6. Conclusion and perspectives

In this work polarimetric features associated to deep neural networks proved to be a real added value to detect objects in a road scene, especially in adverse weather conditions where the visibility is altered. Experiments showed that when using polarimetric encoded data, a 44% augmentation regarding pedestrian detection and a 15% augmentation regarding car detection are noticed compared to RGB detection. These results are encouraging and promising to achieve better results regarding

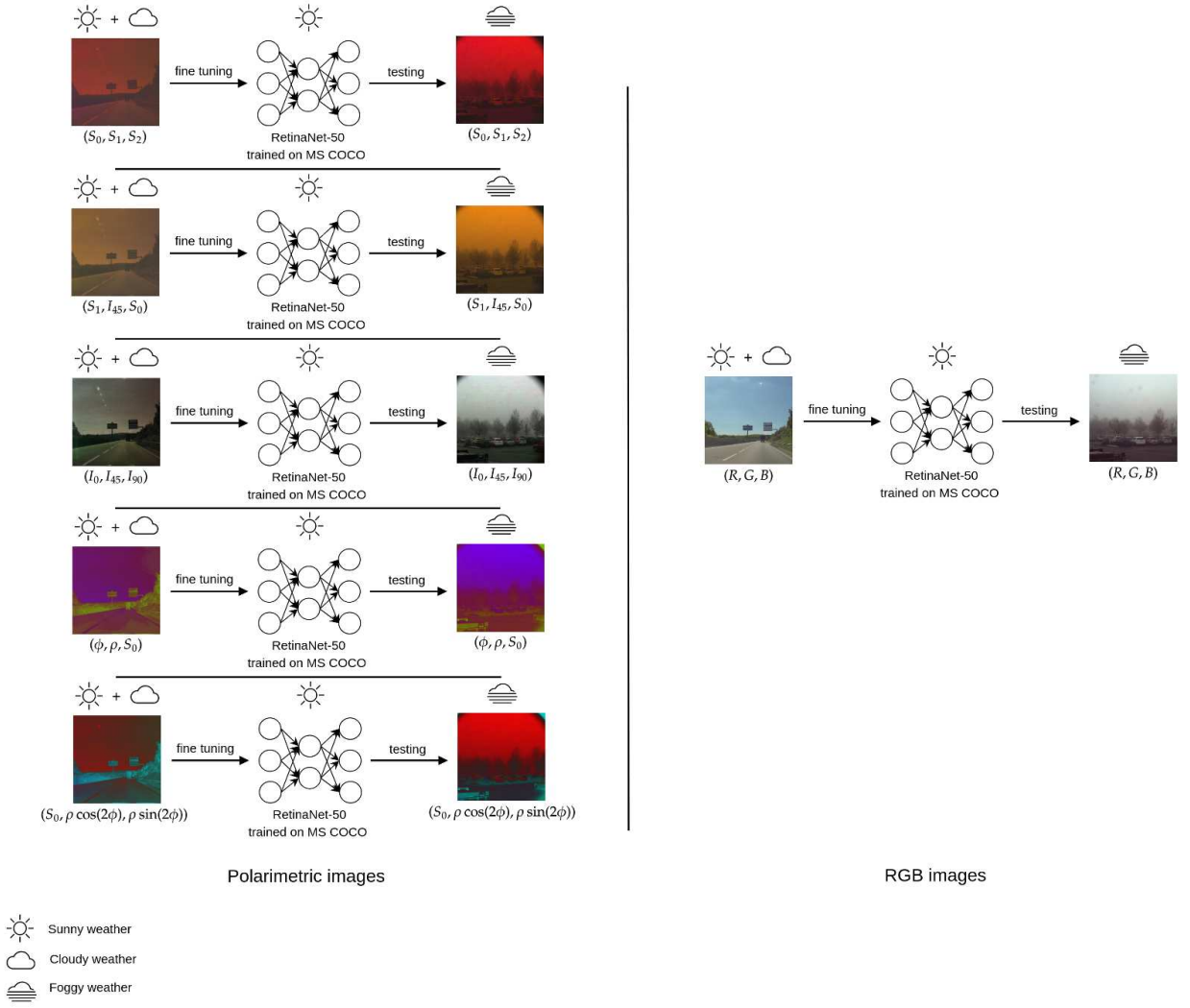

Figure 9: Experimental setup to evaluate each data format and compare their results to the RGB dataset. Here RetinaNet-50 pre-trained on MS COCO is fine tuned on each data format separately.

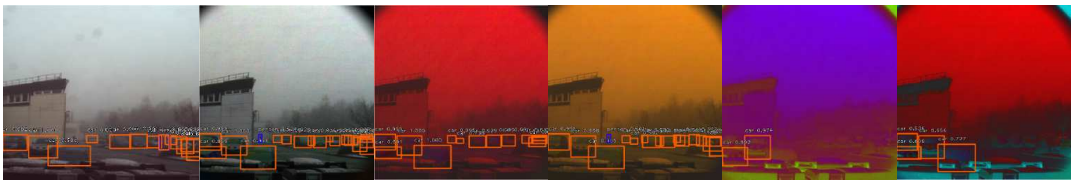

Figure 10: Detection of RetinaNet-50 on the same scene. From left to right, RGB, I, S, Pauli, HSV and P.

road scene analysis in any visibility condition.

We aim in the close future to study the best fusion scheme between polarisation and RGB images, enabling to process all the polarimetric features to provide a better detection of road scenes. More adverse weather conditions should be incorporated into the training process in order to improve the detection results. The polarimetric dataset should then be extended in order to characterize a road scene in other weather conditions as in presence of a heavy rain or in snowy days where the reflection of the scene could be strong.

Acknoledgements

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

References

[1] Samia Ainouz. *Analyse et traitement d'images multidimensionnelles de polarisation*. PhD thesis, Strasbourg 1, 2006.

[2] Samia Ainouz, Olivier Morel, David Fofi, Saleh Mosaddegh, and Abdelaziz Bensrhair. Adaptive processing of catadioptric images using polarization imaging: towards a pola-catadioptric model. *Optical engineering*, 52(3):037001, 2013.

[3] Mohammad Aldibaja, Noaki Suganuma, and Keisuke Yoneda. Improving localization accuracy for autonomous driving in snow-rain environments. In *2016 IEEE/SICE International Symposium on System Integration (SII)*, pages 212–217. IEEE, 2016.

[4] Michael Bass, Eric W Van Stryland, David R Williams, and William L Wolfe. *Handbook of optics*, volume 2. McGraw-Hill New York, 1995.

[5] Kai Berger, Randolph Voorhies, and Larry H Matthies. Depth from stereo polarization in specular scenes for urban robotics. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1966–1973. IEEE, 2017.

[6] Mario Bijelic, Tobias Gruber, and Werner Ritter. Benchmarking image sensors under adverse weather conditions for autonomous driving. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1773–1779. IEEE, 2018.

[7] Rachel Blin, Samia Ainouz, Stéphane Canu, and Fabrice Meriaudeau. 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, 2019.

[8] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[9] Gavin C Cawley and Nicola LC Talbot. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11(Jul):2079–2107, 2010.

[10] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2):303–338, 2010.

[11] Wang Fan, Samia Ainouz, Fabrice Meriaudeau, and Abdelaziz Bensrhair. Polarization-based car detection. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3069–3073. IEEE, 2018.

[12] Pedro Felzenszwalb, David McAllester, and Deva Ramanan. A discriminatively trained, multiscale, deformable part model. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008.

[13] Samuel Gibbs. Google sibling waymo launches fully autonomous ride-hailing service. *The Guardian*, 7, 2017.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[15] Alexander Kamann, Patrick Held, Florian Perras, Patrick Zaumseil, Thomas Brandmeier, and Ulrich T Schwarz. Automotive radar multipath propagation in uncertain environments. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 859–864. IEEE, 2018.

[16] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[17] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988, 2017.

[18] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

[19] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European conference on computer vision*, pages 21–37. Springer, 2016.

[20] Gerard Margarit, Jordi J Mallorqui, and Xavier Fabregas. Single-pass polarimetric sar interferometry for vessel classification. *IEEE transactions on geoscience and remote sensing*, 45(11):3494–3502, 2007.

[21] Olivier Morel, Christophe Stolz, Fabrice Meriaudeau, and Patrick Gorria. Active lighting applied to three-dimensional reconstruction of specular metallic surfaces by polarization imaging. *Applied optics*, 45(17):4062–4068, 2006.

[22] Nicolas Pinchon, Olivier Cassignol, Adrien Nicolas, Frédéric Bernardin, Patrick Leduc, Jean-Philippe Tarel, Roland Brémond, Emmanuel Bercier, and Johann Brunet. All-weather vision for automotive safety: which spectral band? In *International Forum on Advanced Microsystems for Automotive Applications*, pages 3–15. Springer, 2018.

[23] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[24] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.

[25] Yoav Y Schechner, Srinivasa G Narasimhan, and Shree K Nayar. Polarization-based vision through haze. *Applied optics*, 42(3):511–525, 2003.

[26] Michal Shimoni, Dirk Borghys, Roel Heremans, Christiaan Perneel, and Marc Acheroy. Fusion of polsar and polinsar data for land cover classification. *International Journal of Applied Earth Observation and Geoinformation*, 11(3):169–180, 2009.

[27] J Scott Tyo, Bradley M Ratliff, and Andrey S Alenin. Adapting the hsv polarization-color mapping for regions with low irradiance and high polarization. *Optics letters*, 41(20):4759–4762, 2016.

[28] Lawrence B Wolff and Andreas G Andreou. Polarization camera sensors. *Image and Vision Computing*, 13(6):497–510, 1995.

[29] Dizhong Zhu and William A. P. Smith. Depth from a

polarisation + rgb stereo pair. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.