

Utilisation de la polarimétrie pour la segmentation de scènes extérieures avec un réseau convolutif

Marc Blanchon¹

Olivier Morel¹

Yifei Zhang¹

Ralph Seulin²

Nathan Crombez²

Désiré Sidibé¹

¹ ERL VIBOT CNRS 6000, ImViA, Université de Bourgogne Franche-Comté (UBFC)

² EPAN Research Group, Université de Technologie de Belfort-Montbéliard (UTBM)

marc.blanchon@u-bourgogne.fr

Résumé

Dans cet article, nous proposons une nouvelle méthode pour une application de segmentation de scène au niveau pixelique en utilisant la polarimétrie. Pour répondre à la difficulté de détection des zones très réfléchissantes comme l'eau et les fenêtres, nous utilisons l'information de l'angle et du degré de polarisation de la lumière réfléchi par ces objets. Nous utilisons un réseau d'apprentissage profond, basé sur une architecture encodeur-décodeur, pour la segmentation des régions d'intérêt. Différentes méthodes d'augmentation ont été développées pour obtenir une quantité suffisante de données, tout en préservant les propriétés physiques des images polarimétriques. De plus, nous introduisons un nouvel ensemble de données comprenant à la fois des images RGB et polarimétriques avec des annotations manuelles de vérité terrain pour sept classes différentes. Les résultats expérimentaux de cet ensemble de données montrent que l'apprentissage peut bénéficier de la polarimétrie et obtenir de meilleurs résultats de segmentation que lors de l'utilisation de la modalité RGB. En particulier, nous obtenons une amélioration de 38,35% et 22,92% de la précision pour la segmentation des fenêtres et des voitures respectivement.

Mots Clef

Polarimétrie, deep learning, segmentation, augmentation, zones reflectives.

Abstract

In this paper, we propose a novel method for pixel-wise scene segmentation application using polarimetry. To address the difficulty of detecting highly reflective areas such as water and windows, we use the angle and degree of polarization of these areas, obtained by processing images from a polarimetric camera. A deep learning framework, based on encoder-decoder architecture, is used for the segmentation of regions of interest. Different methods of augmentation have been developed to obtain a sufficient

amount of data, while preserving the physical properties of the polarimetric images. Moreover, we introduce a new dataset comprising both RGB and polarimetric images with manual ground truth annotations for seven different classes. Experimental results on this dataset, show that deep learning can benefit from polarimetry and obtain better segmentation results compared to RGB modality. In particular, we obtain an improvement of 38.35% and 22.92% in the accuracy for segmenting windows and cars respectively.

Keywords

Polarimetry, deep learning, segmentation, augmentation, reflective areas.

1 Introduction

La segmentation est une étape cruciale pour la compréhension de scènes complexes. Ces dernières années, les réseaux convolutifs (deep learning) se sont montrés très performants pour la segmentation d'images [7, 6, 1].

La présence possible de zones réfléchissantes est un point clé difficile à résoudre pour les algorithmes de segmentation. La méthode de segmentation doit pouvoir différencier un objet physique et sa projection sur une surface spéculaire.

Le domaine de la segmentation de scènes complexes est ouvert car de nombreuses applications pourraient en bénéficier. En effet, la robotique et les voitures autonomes pourraient profiter de ces capacités. Par exemple, si un système est capable de comprendre une scène avec des zones complexes (réfléchissantes), alors il est possible de les éviter. Des recherches ont été menées sur la détection de la boue [20], ainsi que sur la détection de l'eau [27, 13].

Pour détecter/segmenter les zones réfléchissantes, nous proposons d'utiliser la polarimétrie qui permet de mesurer et de récupérer les changements dans les ondes lumineuses. Les techniques de SFP (shape-from-polarization) utilisent la polarimétrie pour extraire des informations géo-

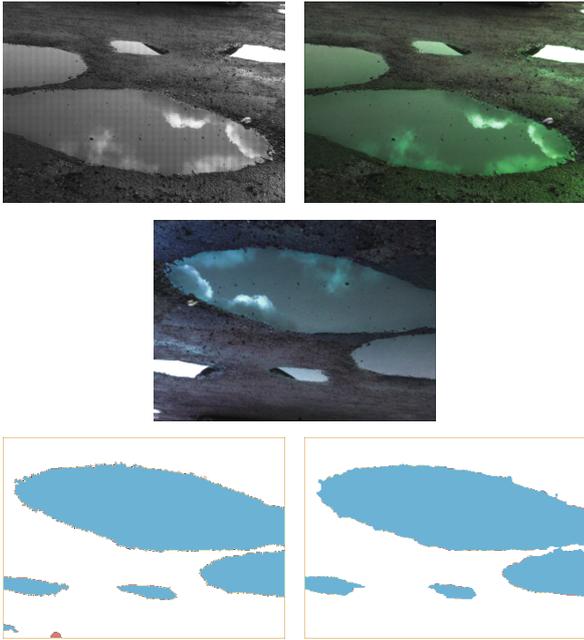


FIGURE 1 – De l’image polarimétrique brute à la segmentation. En haut : à gauche l’image polarimétrique brute, à droite l’image transformée en HSL (Hue Saturation Luminance). L’image du milieu est l’image augmentée avec une signification physique appropriée. En bas : en bas à gauche se trouve la vérité terrain faite à la main et à droite la prédiction du réseau d’apprentissage profond pour l’image du milieu.

métriques d’objets hautement réfléchissants [19, 12]. Par la suite, les caméras polarimétriques ont connu un grand développement qui a conduit à une plus grande facilité d’utilisation et une plus grande praticité. Les caméras basées sur la Division de Plan Focal (DoFP) permettant la capture d’une image en utilisant quatre polariseurs différents ce qui est similaire à l’acquisition de quatre images avec quatre polariseurs orientés chacun différemment. En combinant les avantages des différents types de données, une caméra polarimétrique traitera les données non réfléchissantes comme une caméra standard en niveaux de gris, tandis que les zones réfléchissantes observeront les changements dans les informations de l’image.

Étant donné que l’objectif de ce travail est de mesurer et de qualifier l’utilité d’une modalité complexe appliquée à une tâche spécifique, il n’est pas nécessaire de complexifier la tâche au stade initial du traitement. Par conséquent, nous avons choisi d’utiliser un réseau largement mis à l’épreuve au cœur de cette étude : SegNet [1], compte tenu de sa robustesse et sa modularité. Comme le montre la figure 1, cet article permet de comprendre et d’exploiter ce nouveau type d’information dans le contexte de l’apprentissage profond.

Le présent papier propose les principales contributions suivantes :

- Introduction de la polarimétrie dans le domaine de

l’apprentissage des caractéristiques pour discuter des avantages et des inconvénients de ces données. De plus, une base de données a été créée pour les besoins expérimentaux.

- Création de nouvelles techniques permettant d’augmenter les données polarimétriques en préservant les propriétés physiques de cette modalité.
- Détection et segmentation des zones réfléchissantes grâce à des techniques d’apprentissage profond.

Les différents travaux passés sur lesquels se fonde le présent document sont présentés à la section 2. Ensuite, les différents processus de notre mise en œuvre sont présentés dans la section 3. La quatrième section résume toutes les étapes nécessaires à l’expérimentation. Aussi, cette section présente les résultats des deux modalités utilisées (polarimétrie et RGB) et les discussions qui permettront de comparer les résultats obtenus ainsi que leurs interprétations. La dernière section conclut ce travail et offre une ouverture sur les travaux futurs.

2 Travaux Connexes

2.1 Segmentation de scènes

La segmentation sémantique au niveau pixélique est la capacité de donner une étiquette (label, signification) à chaque pixel d’une image. Cette tâche exige un apprentissage précis des caractéristiques d’un ensemble d’images. Plus précisément, de nombreuses applications de segmentation sémantique ont été développées ; parmi les plus représentées : segmentation de scènes de route [16], compréhension de scènes intérieures [7, 18].

La première segmentation remarquable basée sur l’apprentissage profond est la FCN de Long et al. [10] qui permet la segmentation d’images de toutes tailles sans couches entièrement connectées. Sur la base de ce travail, de multiples réseaux, chacun avec de meilleures performances, ont été publiés : SegNet [1], DeepLab [3, 4, 5], Image-to-Image [8], Conditional Generative Adversarial Networks [25].

2.2 Polarimétrie

La polarimétrie est la science de la mesure de l’état de polarisation de la lumière. En conséquence, une caméra polarimétrique [26] permet de récupérer les changements de lumière dans l’environnement capturé. En raison de ce comportement, l’information provenant de cette caméra peut être utilisée pour distinguer les surfaces d’une scène.

Comme le montre la figure 2, les images polarimétriques peuvent être utilisées avantageusement, car la réflexion a un impact direct sur l’image.

Par exemple, Kai Berger et al. ont proposé une méthode d’estimation de profondeur à partir d’images polarimétriques en milieu urbain [2], traitant la modalité comme une caméra RGB commune. D’autres systèmes basés sur la polarisation ont été proposés pour la détection d’étendues d’eau à l’aide d’informations polarisées. Par exemple, Nguyen et al. ont proposé une méthode de détection de l’eau avec un système stéréo polarisé [13] atteignant une

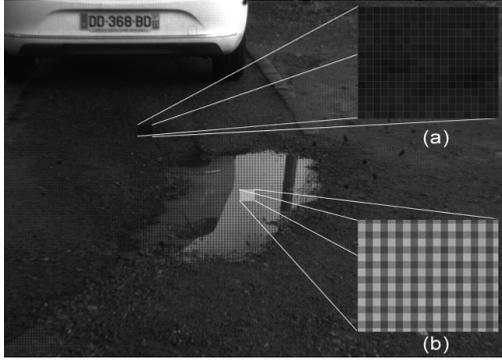


FIGURE 2 – Influence des réflexions sur la polarimétrie. (a) est un zoom sur la zone non polarisée et (b) sur une zone polarisée. De toute évidence, sur une surface polarisée, la micro grille apparaît et révèle un changement d'intensité en fonction du polariseur affecté.

précision approximative de 65%, dépassant la précision de la méthode de référence qui donnait une précision de environ 45% [27]. Rankin et Matthies ont proposé une application en reconnaissance de la boue pour la robotique autonome et ont offert un benchmark complet pour les processus de segmentation [21]. L'un des inconvénients de ces méthodes antérieures est le manque d'automatisation des tâches ou la difficulté de déploiement. En revanche, une approche d'apprentissage profond (Deep Learning) permet de créer un modèle qui peut être réutilisé et ré-ajusté au fur et à mesure de son évolution.

Malgré les aspects utiles et informatifs du système polarimétrique, l'utilisation de ces caméras a été très limitée, en raison de la limitation du matériel et de l'intégration automatique. La technologie DoFP [15, 14, 11] a été introduite, ce qui permet une intégration plus facile. La DoFP permet d'avoir les filtres polarisés en réseau directement sur le capteur. Dans cette conception, quatre filtres polarisés, avec des angles uniques, sont utilisés pour capturer quatre mesures différentes instantanément en une seule prise de vue. De nombreuses applications de traitement d'images et de vision par ordinateur peuvent bénéficier de la récente caméra polarimétrique DoFP.

Dans cet article, nous introduisons la polarimétrie dans le domaine de la segmentation de scènes extérieures.

3 Méthode proposée

3.1 Pré-traitement de la donnée polarimétrique

Contrairement aux autres types d'images standard (RGB, niveaux de gris, etc.), l'image fournie par une caméra DoFP est composée de 2x2 super pixels. Par conséquent, nous utilisons une méthode d'interpolation [22] afin de récupérer les images polarimétriques. L'idée clé derrière cette transformation est d'extraire trois images simple canal pour représenter trois notions physiques : l'angle de polarisa-

tion (AoP), le degré de polarisation (DoP) et l'intensité (I). L'AoP représente la valeur de l'angle de polarisation à chaque pixel tandis que la DoP est la force de l'état de polarisation de la lumière entrante pour chaque pixel.

Dans la nature, la lumière est principalement partiellement polarisée linéairement ce qui réduit les paramètres de Stokes à trois paramètres comme ci-dessous :

$$S = \begin{pmatrix} s_0 \\ s_1 \\ s_2 \\ 0 \end{pmatrix} = \begin{pmatrix} P_0 + P_{90} \\ P_0 - P_{90} \\ P_{45} - P_{135} \\ 0 \end{pmatrix}, \quad (1)$$

où $s_{\{0,1,2\}}$ sont les trois premiers paramètres de Stokes, et $P_{\{0,45,90,135\}}$ les images de sortie d'intensité correspondant à l'orientation du polariseur. L'AoP et la DoP peuvent être déduits en fonction de :

$$\text{DoP} = \sqrt{\frac{s_1^2 + s_2^2}{s_0}}, \quad (2)$$

$$\text{AoP} = \frac{1}{2} \tan^{-1}\left(\frac{s_1}{s_2}\right). \quad (3)$$

Le dernier paramètre I correspond directement à la première valeur du vecteur de Stokes s_0 .

Après ce calcul, trois images de description en niveaux de gris des données polarimétriques brutes sont obtenues. Nous avons choisi de construire une image HSL (Hue Saturation Luminance) cartographiant les trois sources d'information précédentes. Cet espace colorimétrique permet un comportement spécifique par canal qui correspond aux données fournies par l'AoP, la DoP et I. La teinte est généralement une valeur 360° périodique, la saturation est une valeur entre zéro et un comme pour la valeur de la luminance. Pour répondre aux pré-requis de cet espace couleur, nous avons adapté et/ou normalisé nos images en fonction de chaque canal et les avons ensuite fusionnées ensemble [26].

$$H \rightarrow 2 * \text{AoP}, \quad S \rightarrow \text{DoP}, \quad L \rightarrow I/255. \quad (4)$$

L'image HSL peut être vue comme une seule image à 3 canaux. Ceci permet à n'importe quel réseau DL pré-initialisé RGB de traiter ces images. Il est alors possible d'augmenter les données en profitant de la représentation HSL.

3.2 Augmentation de la donnée polarimétrique

Comme expliqué précédemment, l'information polarimétrique caractérise la représentation vectorielle de la lumière. Par conséquent, toute image n'a une signification unique que pour ces paramètres et orientations précis de la caméra. La procédure d'augmentation consiste à créer de nouvelles images avec l'application d'une transformation et/ou d'une interpolation. Les contraintes induites par le type de données sont exportées vers toute transformation appliquée. Les canaux de luminance et de saturation ne sont pas soumis à ces contraintes car leurs valeurs attribuées sont invariantes autour de l'axe optique. Au

contraire, la teinte est affectée par cette transformation. Il est nécessaire de recalculer la teinte en cohérence avec les propriétés physiques de la caméra et de la scène. Dans ce cas unique, l’angle de polarisation aura une signification physique cohérente.

En tournant la caméra dans le sens inverse des aiguilles d’une montre, l’angle de polarisation est tourné dans le sens des aiguilles d’une montre. Soit θ l’angle de rotation appliqué à la caméra, R_θ l’opération de rotation et H le canal de teinte de l’image :

$$H_{\text{rotated}} = R_\theta(H_{\text{prev}} - \theta). \quad (5)$$

A la fin de ce calcul, l’image gardera ses propriétés physiques et sera tournée de θ . Comme le montre le tableau 1, un ensemble de transformations a été développé pour donner la possibilité d’augmenter n’importe quel ensemble de données d’images polarimétriques et il est remarquable que seul le canal de teinte nécessite quelques modifications pour respecter les propriétés physiques. La translation n’est qu’un décalage dans les images, ce qui signifie qu’il n’y a pas de modification du point de vue de la caméra. Comme une caméra polarimétrique dépend de la position et du point de vue réel, le canal de teinte reste invariant à la translation. Par contre, si l’objectif de la caméra a un grand angle, alors dans ce cas une transformation supplémentaire sera nécessaire (Tableau 1 -*).

3.3 Segmentation en utilisant SegNet

Nous utilisons SegNet [1] en raison de sa robustesse et de sa courte durée d’entraînement. Le SegNet a une conception encodeur-décodeur et une architecture composée de 36 couches. Dans notre application, le point clé de cette conception réside dans la partie encodeur. Il est composé de 13 couches et s’adapte parfaitement à la configuration ConvNet B du VGG-16 [23]. Par conséquent, une méthode de transfert par apprentissage [17, 24](transfer learning) permet la pré-initialisation du réseau. Compte tenu de cette approche, un entraînement efficace peut être utilisé en évitant un apprentissage intégral coûteux et fastidieux.

3.4 Un nouveau dataset : PolaBot

Une série d’acquisition a été réalisée afin de fournir un nouvel ensemble de données multimodales PolaBot avec des images polarimétriques. À notre connaissance, aucun ensemble de données de ce type n’a encore été publié. De plus, afin de rendre cet ensemble de données fiable pour différents domaines (robotique, navigation autonome, etc.), les acquisitions ont été réalisées avec un système multimodal de quatre caméras calibrées. Trois modalités synchronisées sont représentées, deux RGB avec un point de vue différent, une caméra NIR (proche infrarouge) et une caméra polarimétrique. En outre, cette collecte d’informations permettra de disposer d’un repère solide et efficace, tout en permettant de comparer la modalité standard à la polarimétrie pour des scènes et une applica-

tion identiques. Cet ensemble de données est disponible à l’adresse suivante : <http://vibot.cnrs.fr/polabot.html>.

4 Experimentations

Pour confirmer notre hypothèse selon laquelle les données polarimétriques seraient plus efficaces que les données dites standard pour notre application, des expériences ont été menées, permettant ainsi une comparaison.

Toutes les expériences ont été réalisées sur le même serveur dédié composé d’un GPU Nvidia Titan Xp (12 Go de mémoire), de 128 Go de RAM et de deux processeurs accumulant un total de 24 cœurs physiques (48 threads). Pour le réseau SegNet, les paramètres internes de l’entraînement doivent être définis ainsi que la fonction de perte et l’optimiseur. Nous avons décidé d’utiliser Adam [9] comme optimiseur et comme fonction de perte l’entropie croisée, définie comme suit :

$$CEL(p, q) = - \sum_{\forall x} p(x) \log(q(x)), \quad (6)$$

où x représente la classe, $p(x)$ est la prédiction pour la classe x et $q(x)$ la vérité terrain. De plus, pour tout l’entraînement, le taux d’apprentissage a été initialisé à $\eta = 10^{-4}$ et à un maximum de 500 epochs.

4.1 Métriques

Pour mesurer l’efficacité de l’entraînement, une série de métriques communes a été utilisée au cours du processus : MIoU (Intersection Moyenne sur l’Union), Score F1, précision moyenne et précision globale. L’IoU est défini comme suit :

$$IoU = \frac{\text{Area of Intersection}}{\text{Area of Union}}. \quad (7)$$

TABLE 1 – Procédure d’augmentation par canal. Ici "-" représente invariant, "*" représente que sous certaines conditions, ce paramètre peut-être modifié.

	AoP (H)	DoP (S)	Intensity (L)
Découpage	-	-	-
Rotation	$R_{\theta}(H - \theta) \pmod{360}$	-	-
Symétrie (Flip)	$-H \pmod{360}$	-	-
Translation	-*	-*	-

Une autre mesure largement utilisée est le score F1. Cette métrique a un comportement similaire au MIoU puisque le score parfait est 1. Cette métrique est une combinaison du rappel et de la précision, qui correspondent respectivement à la pertinence et à la robustesse des résultats :

$$\text{F1 Score} = 2 * \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (8)$$

Enfin, la précision par classe est la mesure de l’ajustement du modèle entraîné pour chaque classe :

$$\text{Accuracy}_C = \frac{\sum_i [p(i) = C \cap GT(i) = C]}{\sum_i [GT(i) = C]} \quad (9)$$

où C est la classe, $p(i)$ est la classe prédite de pixel i et $GT(i)$ la vérité terrain.

4.2 Résultats

Un nuancier est utilisé. Par conséquent, pour les images suivantes, chaque zone de couleur de l’image aura une signification indiquée dans le tableau 2.

Chaque classe a une signification claire sauf les labels "Non Catégorisé" et "Aucune". "Aucune" correspond à des zones segmentées à la main mais considérées comme non révélatrices par rapport à notre application. La classe "Non Catégorisé", par contre, provient d’erreurs de segmentation manuelle. Cette classe est la huitième classe mais n’est pas nécessairement cohérente. Par conséquent, les résultats pour cette classe seront négligés dans le calcul des précisions mais pris en compte dans les conclusions tirées.

Résultats de l’apprentissage. Des métriques pour chaque epochs ont été calculées. Cette procédure permet de voir l’évolution de la condition du modèle.

Comme le montre la figure 3, les deux courbes sont différentes en fonction des données fournies au réseau. Premièrement, il est possible de voir que les deux processus ne se sont pas arrêtés en même temps. Alors que le réseau avec des données polarimétriques atteignait 500 epochs, le réseau traitant des données RGB s’est interrompu à 432 epochs. En effet, nous avons mis en place un critère d’arrêt pour empêcher le réseau de trop se dégrader. Ce résultat signifie que le système SegNet RGB a enregistré une diminution de ses paramètres de validation pour plus de 10 epochs consécutives. Cependant, notre processus permet une récupération d’état afin d’évaluer les résultats dits

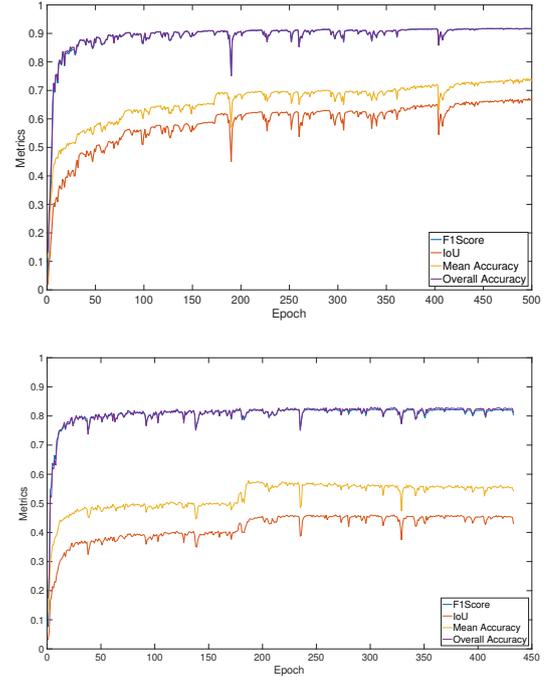


FIGURE 3 – Résultats de l’entraînement - En haut, le graphique correspond à l’estimation des métriques pour les données polarimétriques pendant l’entraînement. Le graphique du bas correspond à l’entraînement avec des données RGB.

"optimaux".

Dans un deuxième temps, il est possible de remarquer les différences entre les valeurs métriques. Le SegNet polarimétrie atteint une valeur MIoU de 0,66, un score F1 de 0,91 et une précision moyenne de 0,73. D’autre part, le SegNet RGB semble être moins efficace avec des scores plus bas : MIoU de 0,42, score F1 de 0,8 et une précision moyenne de 0,54.

Il est possible de conclure cette estimation de l’apprentissage en indiquant que le réseau SegNet polarimétrie semble avoir été plus optimisé pendant la phase d’apprentissage.

Résultats de test. Les résultats des tests correspondent aux résultats obtenus à la sortie du réseau. Comme indiqué dans le tableau 3, afin de comparer les impacts de chaque type de données, leur précision respective par classe a été

TABLE 2 – Charte de couleurs. Ce tableau de couleurs permet l’uniformité dans la visualisation des résultats (chaque classe a une couleur associée).

Signification	Non Catégorisé	Ciel	Eau	Fenêtre	Route	Voiture	Bâtiments	Aucune
Couleur	Noir	Vert	Bleu	Jaune	Orange	Rouge	Gris	Blanc

calculée pour les images RGB et la polarimétrie ensuite suivie d’une comparaison par différence :

$$\text{Accuracy}_{\text{Diff}} = \text{Accuracy}_{\text{Pol}} - \text{Accuracy}_{\text{RGB}}. \quad (10)$$

La Figure 4 montre les résultats obtenus à la sortie du SegNet polarimétrie et la Figure 5 ceux du SegNet RGB. La segmentation est correcte dans les deux cas et offre visuellement de bons résultats.

4.3 Discussion

Comme le montre le tableau 3, une très grande précision peut être observée dans toutes les classes segmentées en utilisant des données polarimétriques. Le jeu de données n’étant pas générique, le ciel reste sur le même ton (bleu), ce qui confère un avantage significatif en faveur de la modalité RGB. Les autres classes où le modèle RGB est meilleur sont : la route, l’eau et "aucune". Ces différences sont minimales et peuvent être expliquées de plusieurs manières. Une de nos hypothèses concerne la différence de segmentation manuelle pour la vérité du terrain. RGB et polarimétrie ont été segmentés indépendamment, augmentant les incertitudes. La difficulté de segmentation de certaines classes doit être prise en compte. Une autre façon d’examiner ces résultats consiste à examiner les avantages et les inconvénients des caméras par rapport à l’ensemble de données. Par exemple, la route peut être polarisée si la température est élevée ; par conséquent, la polarimétrie pourrait avoir un avantage sur le modèle RGB. Étant donné que l’ensemble de données est acquis dans un seul type de conditions météorologiques, le RGB peut avoir une avance sur l’autre modèle, ce qui peut expliquer ces résultats.

Cependant, le modèle de polarimétrie donne une précision convaincante dans toutes les classes. Plus précisément, lors de la segmentation de zones telles que les fenêtres, les voitures et les bâtiments, le modèle obtient une grande différence positive par rapport au RGB. La segmentation de la fenêtre est presque deux fois plus performante avec le modèle polarimétrique que le modèle RGB. En effet, ces résultats peuvent s’expliquer par l’état de polarisation de telles zones.

5 Conclusion

Dans cet article, nous avons proposé l’introduction de la polarimétrie dans le champ de segmentation pixélique des scènes routières. À notre connaissance, il n’existait aucun jeu de données contenant des scènes en extérieur capturées par la polarimétrie. Nous avons donc créé notre propre jeu de données. Cet ensemble de données étant composé

de plusieurs modalités, l’idée principale était d’avoir une mesure de comparaison. Nous avons été contraint de développer une méthode d’augmentation permettant de préserver les propriétés physiques de cette modalité. Cette approche définit les transformations possibles et fournit les formules nécessaires pour une rotation ou une symétrie. Nous avons ensuite utilisé notre jeu de données enrichi en tant qu’entrée du réseau SegNet pour estimer les résultats. Après avoir comparé SegNet appliqué à la polarimétrie et au RGB, nous pouvons en déduire que la polarimétrie offre un avantage considérable par rapport au RGB. En effet, les zones réfléchissantes sont mieux détectées tout en maintenant ou en améliorant les performances de segmentation des autres zones. Nous pouvons en conclure que la polarimétrie peut fournir un nouveau type d’informations utiles dans de nombreux domaines tels que la vision par ordinateur, la robotique mobile ou les voitures autonomes. Nous envisageons d’utiliser un réseau plus élaboré afin de comparer les résultats de manière plus large.

Remerciements

Ce travail a été soutenu par l’ANR ICUB (ANR-17-CE22-0011). Nous remercions sincèrement NVIDIA Corporation pour le soutien en faisant un don de GPU utilisé pour ce travail.

TABLE 3 – Précision par classe et différence de précision.

	Ciel	Eau	Fenêtre	Route	Voiture	Bâtiment	Aucun	Moyenne
Polarimétrie	75.34 %	75.70 %	82.85 %	77.82 %	71.40 %	87.69 %	78.95 %	78.54 %
RGB	89.57 %	78.61 %	44.50 %	78.45 %	48.48 %	67.84 %	83.4 %	69.83 %
Difference	-14.23%	-3.51 %	38.35 %	-0.63 %	22.92 %	19.85 %	-4.45 %	8.71 %

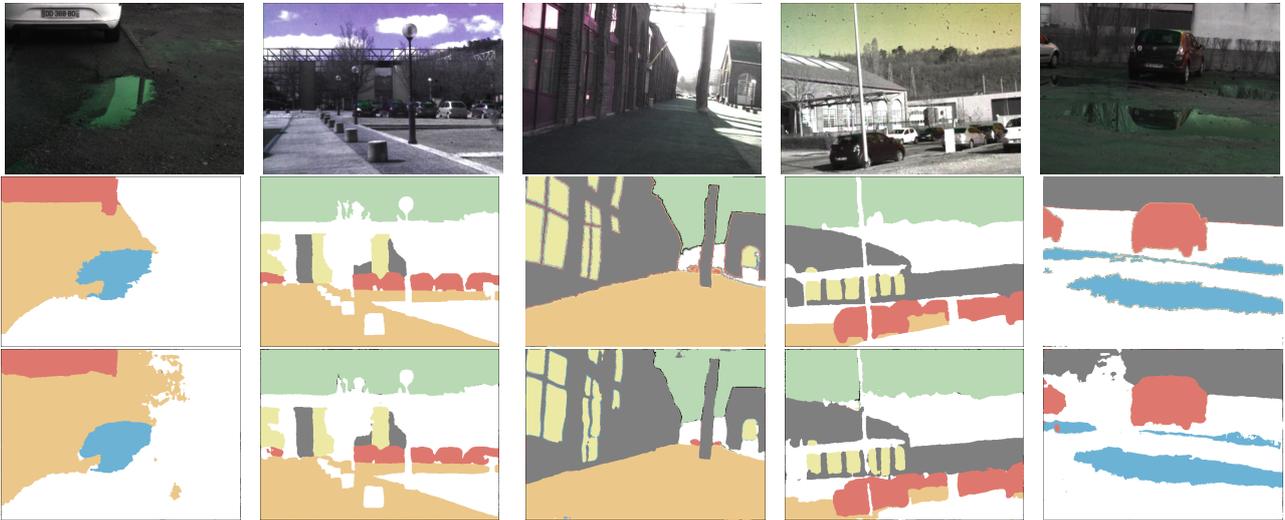


FIGURE 4 – Résultats de polarimétrie - Sortie du set de tests. La rangée supérieure correspond à l'image HSL en entrée. La rangée du milieu est la vérité terrain segmentée manuellement. La rangée du bas représente les prévisions émises par le réseau SegNet polarimétrie.

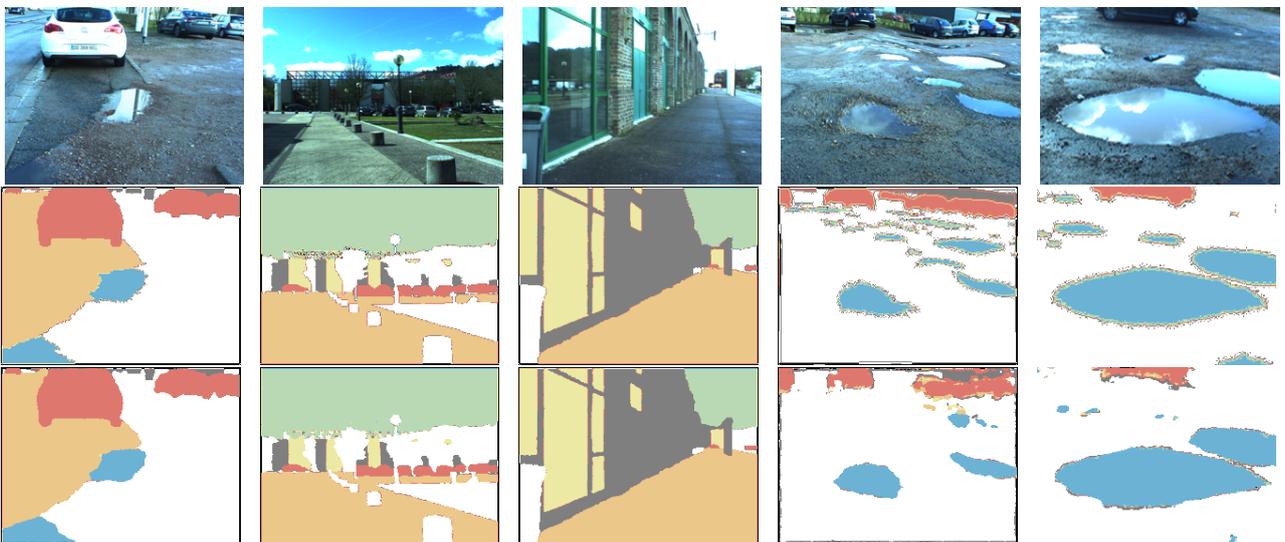


FIGURE 5 – Résultats RGB - Sortie de test. La rangée supérieure est l'image RGB d'entrée. La rangée du milieu est la vérité terrain segmentée manuellement. La rangée inférieure représente la sortie de prédiction du SegNet RGB.

Références

- [1] Vijay Badrinarayanan, Ankur Handa, and Roberto Cipolla. Segnet : A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. *CoRR*, abs/1505.07293, 2015.
- [2] Kai Berger, Randolph Voorhies, and Larry H Matthies. Depth from stereo polarization in specular scenes for urban robotics. In *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pages 1966–1973. IEEE, 2017.
- [3] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Semantic image segmentation with deep convolutional nets and fully connected crfs.

- In *ICLR*, 2015.
- [4] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab : Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *arXiv :1606.00915*, 2016.
 - [5] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab : Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4) :834–848, 2018.
 - [6] Camille Couprie, Clément Farabet, Laurent Najman, and Yann LeCun. Indoor semantic segmentation using depth information. *arXiv preprint arXiv :1301.3572*, 2013.
 - [7] Saurabh Gupta, Ross Girshick, Pablo Arbeláez, and Jitendra Malik. Learning rich features from rgb-d images for object detection and segmentation. In *European Conference on Computer Vision*, pages 345–360. Springer, 2014.
 - [8] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *arXiv preprint, 2017*.
 - [9] Diederik P Kingma and Jimmy Ba. Adam : A method for stochastic optimization. *arXiv preprint arXiv :1412.6980*, 2014.
 - [10] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.
 - [11] James Millerd, Neal Brock, John Hayes, Michael North-Morris, Brad Kimbrough, and James Wyant. Pixelated phase-mask dynamic interferometers. In *Fringe 2005*, pages 640–647. Springer, 2006.
 - [12] Olivier Morel, Fabrice Meriaudeau, Christophe Stolz, and Patrick Gorria. Polarization imaging applied to 3d reconstruction of specular metallic surfaces. In *Machine Vision Applications in Industrial Inspection XIII*, volume 5679, pages 178–187. International Society for Optics and Photonics, 2005.
 - [13] Chuong V Nguyen, Michael Milford, and Robert Mahony. 3d tracking of water hazards with polarized stereo cameras. In *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pages 5251–5257. IEEE, 2017.
 - [14] Gregory P Nordin, Jeffrey T Meier, Panfilo C Deguzman, and Michael W Jones. Diffractive optical element for stokes vector measurement with a focal plane array. In *Polarization : Measurement, Analysis, and Remote Sensing II*, volume 3754, pages 169–178. International Society for Optics and Photonics, 1999.
 - [15] Gregory P Nordin, Jeffrey T Meier, Panfilo C Deguzman, and Michael W Jones. Micropolarizer array for infrared imaging polarimetry. *JOSA A*, 16(5) :1168–1174, 1999.
 - [16] Gabriel L Oliveira, Wolfram Burgard, and Thomas Brox. Efficient deep models for monocular road segmentation. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pages 4885–4891. IEEE, 2016.
 - [17] Sinno Jialin Pan, Qiang Yang, et al. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10) :1345–1359, 2010.
 - [18] Charles R Qi, Hao Su, Kaichun Mo, and Leonidas J Guibas. Pointnet : Deep learning on point sets for 3d classification and segmentation. *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 1(2) :4, 2017.
 - [19] Stefan Rahmann and Nikos Canterakis. Reconstruction of specular surfaces using polarization imaging. In *null*, page 149. IEEE, 2001.
 - [20] Arturo Rankin and Larry Matthies. Daytime water detection based on color variation. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 215–221. IEEE, 2010.
 - [21] Arturo L Rankin and Larry H Matthies. Passive sensor evaluation for unmanned ground vehicle mud detection. *Journal of Field Robotics*, 27(4) :473–490, 2010.
 - [22] Bradley M Ratliff, Charles F LaCasse, and J Scott Tyo. Interpolation strategies for reducing ifov artifacts in microgrid polarimeter imagery. *Optics express*, 17(11) :9112–9125, 2009.
 - [23] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv :1409.1556*, 2014.
 - [24] Lisa Torrey and Jude Shavlik. Transfer learning. In *Handbook of Research on Machine Learning Applications and Trends : Algorithms, Methods, and Techniques*, pages 242–264. IGI Global, 2010.
 - [25] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. High-resolution image synthesis and semantic manipulation with conditional gans. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, page 5, 2018.
 - [26] Lawrence B Wolff and Andreas G Andreou. Polarization camera sensors. *Image and Vision Computing*, 13(6) :497–510, 1995.
 - [27] Sze Hon Yan. Water body detection using two camera polarized stereo vision. *International Journal of Research in Computer Engineering & Electronics*, 3(3), 2014.