

Classification crédibiliste d'objets LIDAR en monde ouvert, par apprentissage profond

Édouard CAPELLIER^{1,2}, Franck DAVOINE², Véronique CHERFAOUI²

¹Renault S.A.S, 1 av. du Golf, 78288 Guyancourt, France.

²Sorbonne Universités, Université de technologie de Compiègne, CNRS, HeuDiaSyc, Centre de recherche Royallieu, CS 60319, 60 203 Compiègne cedex, France.

{edouard.capellier, franck.davoine, veronique.cherfaoui}@hds.utc.fr

Résumé – Nous proposons un système de classification crédibiliste d'objets LIDAR par apprentissage profond. L'algorithme proposé est basé sur une reformulation crédibiliste d'un perceptron multi-couche, ainsi que sur un mécanisme de filtrage statistique simple. Le système, qui n'est entraîné que sur une base d'usagers routiers, est cependant capable de classifier des objets incohérents vis-à-vis de ce jeu d'entraînement comme étant des objets inconnus.

Abstract – We propose an evidential end-to-end deep neural network to classify LIDAR objects. The system is capable of classifying ambiguous and incoherent objects as unknown, while only having been trained on vehicles and vulnerable road users. This is achieved thanks to an evidential reformulation of generalized logistic regression classifiers, and an online filtering strategy based on statistical assumptions.

1 Introduction

Le modèle standard 4D/RCS [1] recommande de réaliser la tâche de perception au sein d'un véhicule autonome terrestre en deux temps : d'abord, des objets d'intérêt doivent être détectés ; ensuite, un algorithme doit classifier chacun d'entre eux. Il est donc naturel d'utiliser ce type de traitement sur des données LIDAR, afin de permettre une perception en trois dimensions de l'environnement routier.

Cette indépendance entre les étapes de détection, et de classification, suppose que tout objet d'intérêt appartient à un type d'objet reconnu par l'algorithme de classification. Cela est difficile à garantir en pratique, notamment lorsque les algorithmes en question doivent être déployés au sein d'un système robotique évoluant en environnement ouvert.

Une solution possible serait de proposer des algorithmes de classification qui seraient entraînés à reconnaître un nombre restreint de classes d'objets pertinentes, tout en étant capable de classifier des objets éloignés de ces classes pertinentes comme étant inconnus. Un exemple de comportement attendu est présenté par la figure 1. Le cadre évidentiel, dans lequel le fait de ne pas savoir est représenté explicitement, paraît adapté à cette approche. Cependant, les algorithmes évidentiels de classification s'appuient, généralement, sur l'hypothèse d'un monde fermé [2], ce qui limite leur applicabilité.

Nous proposons un classifieur évidentiel d'objets LIDAR. Un perceptron multi-couche est utilisé comme un classifieur logistique généralisé multi-tâche, afin de reconnaître les objets détectés via un capteur LIDAR. Nous nous affranchissons de l'hypothèse du monde fermé par un mécanisme de filtrage sta-

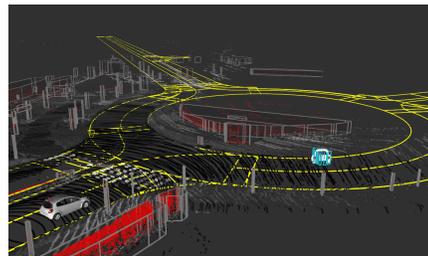


FIGURE 1 – L'objet bleu est classifié comme véhicule, et les objets gris comme inconnus.

tistique simple. Le système n'est entraîné que sur une base de données représentant des véhicules et des usagers vulnérables de la route. Il est néanmoins capable de classifier d'autres types d'objets comme étant inconnus.

2 Classification logistique binaire et théorie des fonctions de croyance

Nous reprenons les observations faites par T Denoeux dans [3]. Soit un problème de classification binaire. $X = (x_1, \dots, x_k)$ est un vecteur d'entrée de dimension d , et $Y \in \Theta$ une variable de classe. Soit $p_1(x)$ la probabilité que $Y = \theta_1$, étant donné que $X = x$. Soit w la sortie d'un classifieur logistique binaire traitant le problème sus-mentionné. Alors $p_1(x)$ est telle que :

$$p_1(x) = S(w) = S\left(\sum_{i=1}^d \beta_i \phi_i(x) + \beta_0\right) \quad (1)$$

S étant la fonction sigmoïde; w correspond parfaitement à la sortie d'un perceptron.

La fonction sigmoïde est strictement croissante. Donc, dans l'équation 1, plus w est grand, plus $p_1(x)$ est élevée et $p_2(x)$ petite. En outre, en posant $\sum_{j=1}^d \alpha_j = \beta_0$, w peut être réécrit sous forme de somme :

$$w = \sum_{j=1}^d w_j = \sum_{j=1}^d (\beta_j \phi_j(x_c) + \alpha_j) \quad (2)$$

Chaque w_j peut ainsi être vu, selon son signe, comme une preuve vers l'hypothèse θ_1 ou θ_2 . Imaginons qu'une fonction de masse crédibiliste simple, notée m_j , puisse être construite à partir de w_j . Soit $w_j^+ = \max(0, w_j)$ la partie positive de w_j , et $w_j^- = \max(0, -w_j)$ sa partie négative.

Supposons que la fonction de masse simple m_j soit égale à $\{\theta_1\}^{w_j^+} \oplus \{\theta_2\}^{w_j^-}$, \oplus étant l'opérateur de Dempster-Shafer.

Soient $w^+ = \sum_{j=1}^d w_j^+$ et $w^- = \sum_{j=1}^d w_j^-$. En supposant que toutes les fonctions de masse m_j soient indépendantes, l'opérateur de Dempster-Shafer peut être utilisé pour les fusionner. La fonction de masse résultante, notée m_{LR} est telle que :

$$K = [1 - \exp(-w^+)] [1 - \exp(-w^-)] \quad (3a)$$

$$m_{LR}(\theta_1) = \frac{[1 - \exp(-w^+)] (\exp(-w^-))}{1 - K} \quad (3b)$$

$$m_{LR}(\theta_2) = \frac{[1 - \exp(-w^-)] (\exp(-w^+))}{1 - K} \quad (3c)$$

$$m_{LR}(\Theta) = \frac{\exp(-w^+ - w^-)}{1 - K} \quad (3d)$$

Or, à partir de la transformation proposée par Cobb et. al. [4], $m_{LR}(\theta_1)$ peut être transformée en une probabilité $p_{m_{LR}}$ telle que :

$$p_{m_{LR}}(\theta_1) = S(w) \quad (4)$$

On retrouve exactement l'expression de $p_1(x)$, dans l'équation 1. Cela signifie que tout classifieur logistique généralisé binaire peut être converti en classifieur évidentiel.

3 Filtrage statistique et optimisation en ligne d'un classifieur logistique binaire

Dans le modèle original proposé par T. Denoeux [3], les paramètres α_i sont calculés après que les autres paramètres du classifieur aient été estimés, de sorte que les fonctions de masse résultantes soient les moins informatives possibles. Cela mène au problème de minimisation suivant :

$$\min f(\alpha) = \sum_{i=1}^n \sum_{j=1}^d (\beta_j \phi_j(x_i) + \alpha_j)^2 \quad (5)$$

$\{(x_i, y_i)\}_{i=1}^n$ étant le jeu d'entraînement, et $\alpha = (\alpha_1, \dots, \alpha_d)$.

Au lieu de ce traitement a posteriori, il peut être observé qu'une solution approchée peut être obtenue durant l'entraînement initial du système. Supposons que la dernière couche

du classifieur, sous forme de perceptron, effectue une normalisation par lot (*Batch Normalization*) sur chaque sortie de la pénultième couche. Soit $v(x_c) = (v_1(x_c), \dots, v_d(x_c))$ l'application modélisée par toutes les couches consécutives du classifieur, sauf la dernière. L'équation 1 devient alors :

$$p_1(x) = S(w) = S\left(\sum_{j=1}^d \left(\beta_j \frac{v_j(x_c) - \bar{v}_j}{\sqrt{\sigma(v_j)^2 + \epsilon}}\right) + \sum_{j=1}^d \alpha_j\right) \quad (6)$$

En supposant que ϵ soit négligeable, on obtient alors :

$$\min f(\alpha) = n \sum_{j=1}^d \beta_j^2 + n \sum_{j=1}^d \alpha_j^2 \quad (7)$$

Cela signifie que, si la dernière couche d'un perceptron, entraîné comme un classifieur logistique généralisé binaire, applique une telle normalisation par lot, alors une simple dégradation des pondérations (*weight decay*) linéaires de la couche finale permet de créer des fonctions de masse crédibilistes prudentes. En outre, cette normalisation par lot a un second avantage : elle peut servir de base à un filtrage statistique, en ligne, des sorties du classifieur.

Soit $z(v_j(x_i)) = \frac{v_j(x_i) - \bar{v}_j}{\sigma(v_j)}$ le score Z de $v_j(x_i)$. En supposant que v_j peut être assimilée à une variable aléatoire de loi normale, un simple seuillage permet de définir des intervalles de confiance sur sa pertinence : plus $z(v_j(x_i))$ est grand, plus $v_j(x_i)$ est éloignée, en nombre d'écart-types, de \bar{v}_j . De plus, le théorème central limite affirme que la somme de variables aléatoires indépendantes peut être modélisée par une distribution normale [5]. Si le classifieur repose sur un ensemble d'opérations linéaires, les $v_j(x_i)$ peuvent être vues comme des sommes de valeurs aléatoires. Des $v_j(x_i)$ statistiquement éloignés du jeu d'entraînement peuvent alors être rejetés, simplement en observant leur score Z.

L'équation 2 peut être interprétée comme :

$$w_j \approx \beta_j * z(v_j(x_c)) + \alpha_j \quad (8)$$

Soit $ZMax$ un seuil à définir. Pendant la classification, nous proposons d'introduire le comportement suivant pour le calcul des w_j :

$$w_j = \begin{cases} 0, & \text{si } \left| \frac{v_j(x_i) - \bar{v}_j}{\sqrt{\sigma(v_j)^2 + \epsilon}} \right| > ZMax \\ \beta_j * \frac{v_j(x_i) - \bar{v}_j}{\sqrt{\sigma(v_j)^2 + \epsilon}} + \alpha_j, & \text{sinon} \end{cases} \quad (9)$$

Ainsi, en suivant l'équation 3, lorsque $w_j = 0$, la fonction de masse m_j indique une ignorance totale. En filtrant ainsi les fonctions de masse m_j , ces dernières n'influencent pas la fonction de masse finale décrite dans l'équation 3, dans les cas où elles ne peuvent pas être capables de produire des résultats pertinents. L'hypothèse du monde fermé peut ainsi être levée.

4 Classification d'objets Lidar

4.1 Génération des données d'entraînement

Pour tester le modèle évidentiel précédemment proposé, un jeu d'objets LIDAR a été collecté sur routes ouvertes grâce à

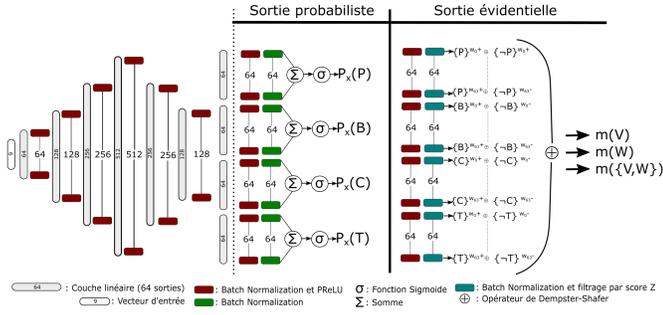


FIGURE 2 – L'architecture multi-tâche proposée

un capteur Velodyne VLP32C, et labellisé. Les données brutes ont été acquises en des lieux et à des dates différentes. Un algorithme de détection temps réel [6] a ensuite été utilisé pour générer les objets LIDAR à classifier. Seuls les objets à moins de quarante-cinq mètres des véhicules ont été conservés. Tous les objets associés par suivi via un filtre de Kalman étendu ont été labellisés ensemble comme étant "voiture", "camion", "vélo", "piéton" ou "inconnu". Au total, 91297 voitures, 9713 camions, 3461 piétons, 946 vélos et 10492 objets inconnus composent l'ensemble d'apprentissage.

Une boîte englobante était ensuite obtenue pour chaque objet, grâce à l'algorithme de minimisation de variance présenté dans [7]. Chaque objet LIDAR est ainsi représenté, in fine, par neuf attributs :

- Distance entre le centroïde de la boîte et le capteur ;
- Longueur, largeur, hauteur de la boîte englobante ;
- Distance moyenne, et écart-type correspondant, entre les points d'un objet et le centroïde de la boîte englobante ;
- Les valeurs propres obtenues par une analyse en composantes principales sur les coordonnées des points ;

4.2 Modèle

Le perceptron multi-couche entraîné sur le jeu de données ainsi obtenu est présenté par la figure 2. Le comportement multi-tâche est nécessaire car le modèle défini dans la section 3 n'est valide que pour des classifieurs binaires. Pour chaque objet x , le perceptron multi-tâche prédit quatre probabilités : $P_x(P)$ la probabilité que l'objet soit un piéton ; $P_x(B)$, la probabilité qu'il soit un vélo ; $P_x(C)$, la probabilité qu'il soit une voiture ; et $P_x(T)$, la probabilité qu'il soit un camion. D'après l'équation 3, ces probabilités peuvent être converties en fonctions de masse crédibilistes. Soit $\Omega = \{V, W\}$ un cadre de discernement, V représentant le fait qu'un objet soit un véhicule, W représentant le fait qu'un objet soit un usager vulnérable. Les fonctions de masse crédibilistes sur Ω peuvent être construites comme suit :

$$m_p(V) = 0, m_p(W) = m(P), m_p(\{V, W\}) = 1 - m(P) \quad (10a)$$

$$m_b(V) = 0, m_b(W) = m(B), m_b(\{V, W\}) = 1 - m(B) \quad (10b)$$

$$m_c(V) = m(C), m_c(W) = 0, m_c(\{V, W\}) = 1 - m(C) \quad (10c)$$

$$m_t(V) = m(T), m_t(W) = 0, m_t(\{V, W\}) = 1 - m(T) \quad (10d)$$

L'opérateur de Dempster-Shafer permet d'obtenir la fonction de masse finale générée par le perceptron multi-couche :

$$m = m_p \oplus m_b \oplus m_c \oplus m_t \quad (11)$$

4.3 Entraînement du modèle

Soit D_p l'ensemble des objets pertinents du jeu de données, c'est à dire le jeu de données privé des objets étiquetés comme inconnus. Un jeu d'entraînement D_{pt} a été créé à partir de soixante-dix pourcents de D_p , et un jeu de validation D_{pv} a été créé à partir de trente pourcents de D_p . Le nombre d'objets de chaque classe de D_{pt} a été rééquilibré sur le nombre de camions présents, grâce à de l'échantillonnage aléatoire et à l'algorithme SMOTE [8]. Ce jeu d'entraînement raffiné est noté D'_{pt} . Le perceptron a été optimisé via l'algorithme d'optimisation ADAM, et un taux d'apprentissage de 0,0001. Une dégradation des pondérations de $1e-5$ était appliquée sur les paramètres linéaires de la dernière opération de normalisation par lot. L'algorithme a optimisé le système en itérant quatre-cents fois sur D'_{pt} . La fonction de coût utilisée est une somme d'entropies croisées, calculées sur chacune des classes possibles pour chaque objet, et à partir des sorties probabilistes.

4.4 Evaluation du comportement évidentiel

Le comportement évidentiel est évalué vis-à-vis du cadre de discernement Ω , et des fonctions de masse crédibilistes générées à partir des équations 3, 10 et 11. La règle ID [3], est utilisée pour classifier les objets selon les fonctions des masse sur V et W prédites par le perceptron. Cependant, pour lever l'hypothèse de monde fermé, on considère que si un objet n'est pas classé comme étant V ou W par la règle ID, alors il doit être considéré comme un objet inconnu au sens large. Le seuil $ZMax$ dans l'équation 9 doit toujours être choisi. Les seuils habituels pour détecter des anomalies statistiques à partir d'un score Z sont 2.58, 1.96, and 1.65 [5]. Le perceptron multi-tâche est donc testé avec ces trois valeurs pour $Zmax$.

Les décisions générées depuis les fonctions de masse crédibilistes sont comparées avec des décisions prises à partir de quatre séparateurs à vastes marges mono classes, entraînés sur notre jeu de données (*one-class SVMs*) [9]. Les SVMs mono classes sont des solutions classiques pour reconnaître les objets inconnus. Chaque SVM est entraîné individuellement à reconnaître soit les piétons, soit les vélos, soit les voitures, soit les camions, à partir des échantillons de chaque classe présents dans D_{pt} . Une règle de classification simple est utilisée pour classifier un objet partir des quatre séparateurs, ou des sorties probabilistes du perceptron :

- Si un objet est classifié comme vélo, piéton, ou les deux, alors l'objet est classifié comme usager vulnérable (W) ;
- Si un objet est classifié comme voiture, camion, ou les deux, alors l'objet est classifié comme véhicule (V) ;
- Sinon, l'objet est considéré inconnu (Ω) ;

Pour simuler un test en conditions réelles, les séparateurs et le perceptron sont testés sur $D_{rc} = D_u \cup D_{pv}$, l'union du l'en-

Method	IoU	Précision	score F1 (V)	score F1(W)	score F1(Ω)
$ZMax = \infty$	0.320	0.733	0.890	0.412	0.388
$ZMax = 2.58$	0.558	0.825	0.938	0.458	0.675
$ZMax = 1.96$	0.682	0.872	0.945	0.570	0.786
$ZMax = 1.65$	0.725	0.897	0.929	0.661	0.825
SVMs mono classes [9]	0.507	0.661	0.672	0.556	0.660

TABLE 1 – Résultats sur les données de validation (objets connus et inconnus)

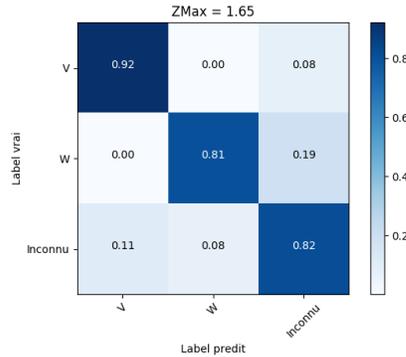


FIGURE 3 – Matrice de confusion normalisée pour $ZMax=1.65$ sur D_{rc}

Précision	score F1 (V)	score F1(W)
0.914	0.959	0.890

TABLE 2 – Résultats de la classification crédibiliste sur D_{pv} , pour $ZMax = 1.65$

semble de validation et de l'ensemble des objets inconnus. La table 1 présente les résultats.

D'après les indices de Jaccard (IoU) de chaque méthode D_{rc} , l'approche la plus performante est la classification crédibiliste avec une valeur de $ZMax$ égale à 1,65. Cette version est également la meilleure sur tous les indicateurs, mis-à-part le score F1 sur V . Ainsi, l'intérêt du filtrage des sorties du perceptron en fonction de leurs scores Z , avec une interprétation crédibiliste des sorties, est visible. Le score F1 sur Ω est égal à 0.825 pour un $ZMax$ égal 1.65, alors que le système n'a jamais été entraîné sur des objets inconnus. Les usagers vulnérables restent cependant difficiles à classer correctement. Cela peut s'expliquer du fait que l'ensemble d'entraînement original contenait beaucoup moins de piétons et de vélos que de voitures et de camions. Comme le montre la figure 3, tous les objets pertinents (véhicules et usagers vulnérables) mal classifiés sont classifiés comme inconnus, lors d'une classification crédibiliste avec $ZMax = 1.65$. De plus, 81% des usagers vulnérables (W) sont correctement classifiés : le faible score F1 pour cette classe s'explique donc par les 19% qui sont classifiés comme inconnus, et les 8% d'objets inconnus qui sont classifiés comme utilisateurs vulnérables. La table 2, indique les précisions et les scores F1 uniquement calculés sur D_{pv} , l'ensemble de validation, dans le cas d'une classification crédibiliste avec $ZMax = 1.65$. Le score F1 pour W est alors égal à 0.890, ce qui est plus satisfaisant.

5 Conclusions

Nous avons proposé un algorithme de classification crédibiliste d'objets LIDAR. Le système, qui a été entraîné en tant que classifieur probabiliste puis converti en classifieur évidentiel, est efficace pour la détection d'objets LIDAR inconnus arbitraires. La classification crédibiliste peut se faire sous contraintes temps-réel : les 30190 objets LIDAR de l'ensemble D_{rc} peuvent être classifiés en 400ms, sur une carte graphique NVidia TitanX Pascal. Les prochaines étapes de ce travail consisteront donc à intégrer ce classifieur au sein d'un système de fusion.

Ce travail a été réalisé au sein de SIVALab, le laboratoire commun entre Renault S.A.S, le CNRS et HeuDiaSyc.

Références

- [1] J. S. ALBUS, "4D/RCS : a reference model architecture for intelligent unmanned ground vehicles", in *Unmanned Ground Vehicle Technology IV*, International Society for Optics et Photonics, t. 4715, 2002, p. 303-311.
- [2] P. SMETS, "What is Dempster-Shafer's model", *Advances in the Dempster-Shafer theory of evidence*, p. 5-34, 1994.
- [3] T. DENOEU, "Logistic regression, neural networks and Dempster-Shafer theory : a new perspective", *arXiv preprint arXiv :1807.01846*, 2018.
- [4] B. R. COBB et P. P. SHENOY, "On the plausibility transformation method for translating belief function models to probability models", *International Journal of Approximate Reasoning*, t. 41, n° 3, p. 314-330, 2006.
- [5] D. J. RUMSEY, *U Can : statistics for dummies*. John Wiley & Sons, 2015.
- [6] Y. LI, C. LE BIHAN, T. RISTORCELLI, J. IBANEZ-GUZMAN et K. BOUZIANE, "Fast Coarse-to-Fine 3D Point Cloud Segmentation in Spherical Coordinates for Autonomous Driving", *Submitted to (ICRA)*, 2019.
- [7] X. ZHANG, W. XU, C. DONG et J. M. DOLAN, "Efficient L-shape fitting for vehicle detection using laser scanners", in *Intelligent Vehicles Symposium (IV)*, IEEE, 2017, p. 54-59.
- [8] N. V. CHAWLA, K. W. BOWYER, L. O. HALL et W. P. KEGELMEYER, "SMOTE : synthetic minority over-sampling technique", *Journal of artificial intelligence research*, t. 16, p. 321-357, 2002.
- [9] B. SCHÖLKOPF, J. C. PLATT, J. SHAWE-TAYLOR, A. J. SMOLA et R. C. WILLIAMSON, "Estimating the support of a high-dimensional distribution", *Neural computation*, t. 13, n° 7, p. 1443-1471, 2001.