

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la
Propriété Intellectuelle
Bureau international



(10) Numéro de publication internationale
WO 2020/165544 A1

(43) Date de la publication internationale
20 août 2020 (20.08.2020)

(51) Classification internationale des brevets :
G06K 9/00 (2006.01) G06K 9/62 (2006.01)
G06K 9/46 (2006.01)

(21) Numéro de la demande internationale :
PCT/FR2020/050268

(22) Date de dépôt international :
13 février 2020 (13.02.2020)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
1901448 13 février 2019 (13.02.2019) FR

(71) Déposant : SAFRAN [FR/FR] ; 2, boulevard du Général
Martial Valin, 75015 PARIS (FR).

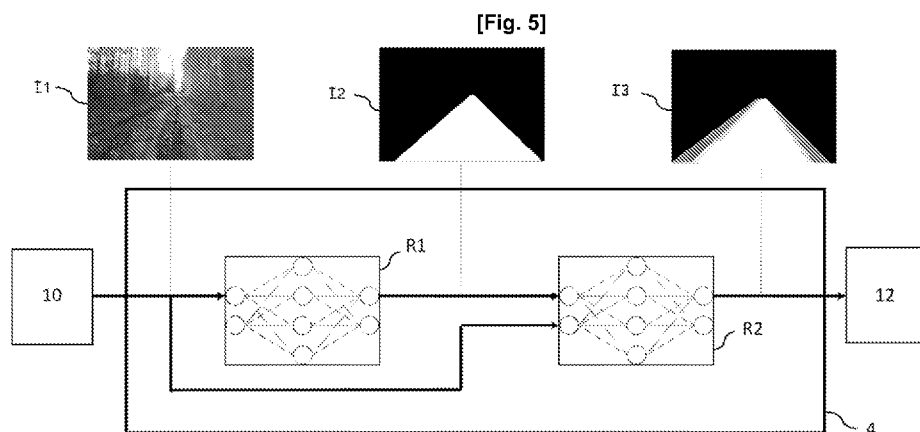
(72) Inventeurs : **PICARD, Sylvaine** ; c/o SAFRAN CEPI
Rond-Point René Ravaud - Réau, 77550 MOISSY-CRA-
MAYEL (FR). **BURSUC, Andrei** ; c/o SAFRAN CEPI
Rond-Point René Ravaud - Réau, 77550 MOISSY-CRA-
MAYEL (FR).

(74) Mandataire : **REGIMBEAU** ; 20, rue de Chazelles, 75847
PARIS CEDEX 17 (FR).

(81) États désignés (sauf indication contraire, pour tout titre de
protection nationale disponible) : AE, AG, AL, AM, AO,
AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA,
CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ,
EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR,
HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR,
KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,

(54) Title: IDENTIFICATION OF DRIVABLE AREAS WITH CONSIDERATION OF THE UNCERTAINTY BY A DEEP LEARNING METHOD

(54) Titre : IDENTIFICATION DE ZONES ROULABLES AVEC PRISE EN COMPTE DE L'INCERTITUDE PAR UNE MÉTHODE D'APPRENTISSAGE PROFOND



(57) **Abstract:** The present invention relates to a method for identifying drivable areas in at least one image of a terrain, characterised in that it comprises the implementation of the following steps in a processing unit (4): - supervised learning of at least one neural network (R2), from a learning database comprising terrain learning images having been annotated by at least two operators in order to identify drivable areas, - segmentation by means of the at least one neural network (R2), of an input image (I1) in order to obtain a probabilistic map (I3) of the presence of at least one drivable area in said input image (I1), said segmentation comprising an estimation of a probability of presence of a drivable area in the input image (I1), so as to obtain a probabilistic map (I3) defined by a matrix, each element of which corresponds to a probability of presence of at least one drivable area on a pixel of the input image (I1).

(57) **Abrégé :** La présente invention concerne un procédé d'identification de zones roulables dans au moins une image d'un terrain, caractérisé en ce qu'il comprend l'implémentation des étapes suivantes dans une unité de traitement (4) : - apprentissage supervisé d'au moins un réseau de neurones (R2), à partir d'une base de données d'apprentissage comprenant des images d'apprentissages de terrain ayant été annotées par au moins deux opérateurs en vue d'identifier des zones roulables, - segmentation au moyen de l'au moins un réseau de neurones (R2), d'une image d'entrée (I1) pour obtenir une carte probabiliste (I3) de présence d'au moins une zone roulable

SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT,
TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) États désignés (*sauf indication contraire, pour tout titre de protection régionale disponible*) : ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, RU, TJ, TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Publiée:

— avec rapport de recherche internationale (Art. 21(3))

dans ladite image d'entrée (I1), ladite segmentation comprenant une estimation d'une probabilité de présence d'une zone roulable dans l'image d'entrée (I1), de manière à obtenir une carte probabiliste (I3) définie par une matrice dont chaque élément correspond à une probabilité de présence d'au moins une zone roulable sur un pixel de l'image d'entrée (I1).

Identification de zones roulables avec prise en compte de l'incertitude par une méthode d'apprentissage profond

DOMAINE DE L'INVENTION

- 5 La présente invention concerne le domaine des véhicules autonomes, et en particulier le domaine des véhicules autonomes terrestres, ainsi que les procédés de contrôle de tels véhicules.

ETAT DE LA TECHNIQUE

10 La conception et le développement de véhicules de plus en plus autonomes et connectés posent de nombreux défis. En effet, afin qu'un véhicule puisse naviguer sans pilote, il faut qu'il soit capable de percevoir son environnement, au moyen, par exemple de nombreux capteurs, puis d'analyser et interpréter les données reçues, afin de prendre des décisions sur la conduite du véhicule.

Afin de pouvoir se déplacer sans risque pour des usagers et d'adapter convenablement sa conduite, le véhicule autonome doit en outre être capable de détecter et d'identifier tous les
15 objets fixes ou mobiles de son environnement.

Pour atteindre ces objectifs, des techniques d'intelligence artificielle sont actuellement utilisées. En particulier, ce sont des techniques d'apprentissage automatique qui donnent les meilleurs résultats, notamment des méthodes basées sur l'apprentissage à partir de données à
20 l'aide de réseaux de neurones profonds, et notamment des réseaux de neurones convolutionnels (en anglais « Convolutional Neural Network » (CNN)), particulièrement adaptés à la segmentation d'images ou de vidéos.

Toutefois, l'utilisation de ces techniques d'intelligence artificielle n'est pas sans difficulté. En effet, entraîner un système d'intelligence artificielle à la conduite automobile est un processus
25 particulièrement long et compliqué, qui nécessite notamment une base de données d'apprentissage importante, afin de prédire un fonctionnement sur les multiples scénarios auxquels peuvent être confrontés les véhicules.

De nombreuses bases de données existent déjà, par exemple la base de données Cityscapes™ (voir à ce titre le document M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M.
30 Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016), qui dispose de plusieurs milliers d'images sur lesquelles sont identifiées différentes classes d'objets (routes, usagers, véhicules, bâtiments, ...). Cependant, la

construction de telles bases de données annotées précisément nécessite une identification manuelle des objets par un opérateur humain et est par conséquent fastidieuse.

La **figure 1** représente un exemple d'image annotée par un opérateur humain, destinée à entraîner un réseau de neurones à la segmentation d'images en plusieurs classes d'objets. La figure 1 est une scène d'extérieur urbain, dans laquelle cinq catégories d'éléments ont été annotées. Dans cet exemple, la zone roulable a1 est une route goudronnée, définie sans incertitude. Les trottoirs a2, les êtres humains (piétons ou cyclistes) a3, les bâtiments a4, le fond et la végétation a5 ont été également segmentés.

A partir de la base de données annotées, il est possible de construire un algorithme d'apprentissage profond (en anglais, « deep learning »), par exemple un réseau de neurones profonds, permettant d'identifier automatiquement les différentes classes d'objet, et notamment d'identifier une zone « roulable ».

Toutefois, en dehors d'une zone urbaine, dans un contexte de chemins ou de pistes non goudronnées par exemple, un problème de définition de ce que l'on considère comme une zone roulable se pose, même pour un opérateur humain chargé d'annoter une image d'une base de données d'apprentissage. Ainsi, pour une image de chemin donnée, il n'est pas rare que différents annotateurs proposent une annotation différente de la présence du chemin. En effet, dans le cas des chemins l'appréciation de la présence du chemin est assez subjective par exemple :

- est-ce qu'une flaque doit être considérée comme faisant partie du chemin ?
- est-ce que les ornières en bord de chemin font partie du chemin ? etc.

Cette subjectivité est en fait intéressante à observer. Mais ce n'est pas le cas dans les méthodes connues. Ainsi, le contexte des scènes urbaines ou de routes goudronnées reste privilégié.

De plus, pour un système de contrôle d'un véhicule autonome, il est souhaitable d'intégrer la connaissance de l'incertitude des informations, et notamment de compléter l'information de détection d'un chemin dans l'environnement du véhicule par une estimation de probabilité, afin notamment de minimiser un risque, d'embourbement par exemple.

Cependant, les algorithmes existants n'intègrent pas la connaissance de l'incertitude des informations, en particulier pour la détection de chemin, lors de circulation en dehors de routes, et ne tiennent en particulier pas compte du problème d'incertitude dans la définition de la présence ou non d'un chemin en un point donné d'une image. Ainsi l'apprentissage du modèle, par exemple du réseau de neurones convolutionnels (CNN) est sous optimal ainsi que son utilisation pratique pour le contrôle du véhicule.

EXPOSE DE L'INVENTION

Un but de l'invention est de remédier au moins en partie aux inconvénients précités et de proposer selon un premier aspect un procédé d'identification de zones roulables sur une image, permettant
5 notamment de détecter de façon probabiliste des zones roulables sur des images de terrain. Selon un autre aspect est proposé un procédé de contrôle d'un véhicule terrestre, mettant en œuvre un tel procédé d'identification de zones roulables, permettant de définir des lois de contrôle d'un véhicule, par exemple une trajectoire ou une vitesse, tenant compte de l'incertitude de zones roulables identifiées.

10 A cet effet, l'invention concerne, selon un premier aspect, un procédé d'identification de zones roulables dans au moins une image d'un terrain, caractérisé en ce qu'il comprend l'implémentation des étapes suivantes dans une unité de traitement :

E1 : apprentissage supervisé d'au moins un réseau de neurones, à partir d'une base de données d'apprentissage comprenant des images d'apprentissages de terrain ayant été annotées
15 par au moins deux opérateurs en vue d'identifier des zones roulables,

E2 : segmentation au moyen de l'au moins un réseau de neurones, d'une image d'entrée pour obtenir une carte probabiliste de présence d'au moins une zone roulable dans ladite image d'entrée, ladite segmentation comprenant une estimation d'une probabilité de présence d'une zone roulable dans l'image d'entrée, de manière à obtenir une carte probabiliste (I3) définie par
20 une matrice dont chaque élément correspond à une probabilité de présence d'au moins une zone roulable sur un pixel de l'image d'entrée.

L'invention est avantageusement complétée par les caractéristiques suivantes, prises seules ou en une quelconque de leur combinaison techniquement possible :

- le procédé comprend avant l'étape d'apprentissage, une étape de calcul d'au moins une vérité terrain définie par une carte probabiliste pour au moins une image associée de la base de données d'apprentissage, à partir d'au moins deux annotations, la vérité terrain étant utilisée pour
25 l'étape d'apprentissage.

- la base de données d'apprentissage comprend un ensemble d'images, associé à un ensemble de vérités terrain associé, chaque vérité terrain étant une matrice dont chaque élément p correspond à un pixel de l'image associée, chaque élément p étant calculé de la manière
30 suivante :

$$p(pixel) = \sum_{i=1}^K \frac{annotation(pixel, i)}{K}$$

avec K le nombre d'opérateurs ayant annoté l'image, et annotation (pixel, i) = 1 si un opérateur i a considéré le pixel comme appartenant à une zone roulable, 0 sinon.

- l'étape d'apprentissage comprend les étapes :

5 S1 : apprentissage supervisé d'un premier réseau de neurones à partir de la base de données d'apprentissage, de sorte que le premier réseau de neurones entraînés soit configuré pour segmenter une image entre zone roulable et zone non roulable ;

S2 : apprentissage supervisé d'un second réseau de neurones à partir de la base de données d'apprentissage et de l'au moins une vérité terrain, de sorte que le second réseau de neurones entraîné soit configuré pour segmenter l'image selon une carte probabiliste de présence
10 de zone roulable.

- l'étape de segmentation de l'image d'entrée comprend les étapes de :

S10 : calcul d'une segmentation de l'image d'entrée, par le premier réseau de neurones entraîné,

15 S20 : calcul d'une carte probabiliste de présence de zones roulables de l'image d'entrée par le second réseau de neurones entraîné, à partir de l'image d'entrée et de la segmentation issue du premier réseau de neurones.

L'invention concerne selon un deuxième aspect, un procédé de contrôle d'un véhicule terrestre, comprenant les étapes suivantes :

20 C1 : acquisition d'images par un capteur situé sur le véhicule ;

C2 : traitement des images acquises par une unité de traitement, par un procédé de d'identification selon le premier aspect de l'invention, afin d'obtenir au moins une carte probabiliste de présence de zones roulables ;

C3 : calcul d'une consigne comprenant les sous-étapes de :

25 - C31 : calcul d'une probabilité de présence de zone roulable par l'unité de traitement (4') à partir de la carte probabiliste de présence de zones roulables ;

- C32 : calcul d'une trajectoire et/ou d'une consigne en vitesse du véhicule par l'unité de traitement à partir de la probabilité de présence de zone roulable.

30 - le capteur acquiert une vidéo, et dans lequel l'identification au cours de l'étape de traitement des images acquises est effectuée de manière séquentielle.

- l'étape de calcul de la consigne dépend d'une valeur seuil prédéterminée, la valeur seuil étant un niveau de risque à prendre correspondant à une probabilité qu'un pixel d'une image acquise appartienne à une zone roulable.

- la consigne en vitesse augmente lorsque la probabilité de présence de zone roulable
5 augmente.

- la consigne en trajectoire est calculée à partir de la carte probabiliste de zones roulables de sorte à maximiser la probabilité de présence de zone roulable.

- la consigne en trajectoire est calculée à partir de la carte probabiliste de zones roulables de sorte à maximiser la probabilité de présence de zone roulable et à minimiser selon un niveau
10 d'exigence k , prédéfini ou choisi dynamiquement, les zones au caractère roulable faiblement probables.

- la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones roulables finale issue de la fusion dite conservative de M cartes probabilistes successives, dont chaque élément p_{final} est calculé de la manière suivante :

$$p_{final}(pixel) = \min (p_t(pixel), p_{t-1}(pixel) \dots p_{t-M+1}(pixel))$$

- la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones roulables finale issue de la fusion dite robuste de M cartes probabilistes successives, dont chaque élément p_{final} est calculé de la manière suivante :

$$p_{final}(pixel) = \text{mediane} (p_t(pixel), p_{t-1}(pixel) \dots p_{t-M+1}(pixel))$$

- la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones roulables finale issue de la fusion dite agressive de M cartes probabilistes successives, dont chaque élément p_{final} est calculé de la manière suivante :

$$p_{final}(pixel) = \max (p_t(pixel), p_{t-1}(pixel), \dots p_{t-M+1}(pixel))$$

- la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones
25 roulables finale issue de la fusion dite moyenne de M cartes probabilistes successives, dont chaque élément p_{final} est calculé de la manière suivante :

$$p_{final}(pixel) = \frac{p_t(pixel) + p_{t-1}(pixel) + \dots + p_{t-M+1}(pixel)}{M}$$

- la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones roulables finale issue de d'une fusion de M cartes probabilistes successives, et d'une
30 détection d'obstacles réalisée par un capteur d'obstacle.

DESCRIPTION DES FIGURES

D'autres caractéristiques, buts et avantages de l'invention ressortiront de la description qui suit, qui est purement illustrative et non limitative, et qui doit être lue en regard des dessins annexés sur lesquels :

[Fig. 1] La figure 1 représente un exemple de segmentation d'une image de zone urbaine, telle qu'il en existe dans la base de données Cityscapes™.

[Fig. 2] La figure 2 représente un exemple de segmentation d'une image prenant en compte une incertitude de la définition de la zone roulable.

[Fig. 3] La figure 3 représente un exemple d'architecture d'implémentation du procédé d'identification de zones roulables.

[Fig. 4] La figure 4 représente schématiquement les étapes d'un procédé de d'identification de zones roulables sur au moins une image de terrain selon un premier aspect de l'invention.

[Fig. 5] La figure 5 représente un premier exemple d'architecture globale d'un procédé de segmentation d'une image d'entrée selon le premier aspect de la présente invention.

[Fig. 6] La figure 6 représente un deuxième exemple d'architecture globale d'un procédé de segmentation d'une image d'entrée selon le premier aspect de la présente invention.

[Fig. 7] La figure 7 représente le deuxième exemple d'architecture globale d'un procédé de segmentation d'une vidéo d'entrée selon le premier aspect de la présente invention.

[Fig. 8] La figure 8 représente schématiquement un véhicule terrestre permettant de mettre en œuvre un procédé de contrôle selon un deuxième aspect de la présente invention.

[Fig. 9] La figure 9 représente schématiquement les étapes d'un procédé de contrôle de véhicule terrestre selon le deuxième aspect de l'invention.

[Fig. 10] La figure 10 représente schématiquement une première image issue d'un capteur d'un véhicule tel qu'illustré sur la figure 8, et une première carte de probabilité associée.

[Fig. 11] La figure 11 représente schématiquement une deuxième image issue d'un capteur d'un véhicule tel qu'illustré sur la figure 8, et une deuxième carte de probabilité associée.

[Fig. 12] La figure 12 représente schématiquement une troisième image issue d'un capteur d'un véhicule tel qu'illustré sur la figure 8, et une troisième carte de probabilité associée, utilisée pour émettre une consigne en trajectoire et en vitesse pour le véhicule.

[Fig. 13] La figure 13 représente un exemple de segmentations d'une image prenant en compte une incertitude de la définition de la zone roulable par deux annotateurs différents.

DESCRIPTION DETAILLEE DE L'INVENTION

Dans ce qui suit, deux principaux aspects vont être décrits :

- Un procédé d'identification d'au moins une zone roulable par au moins un réseau de neurones, de préférence par deux réseaux de neurones, comprenant
5 une première étape d'apprentissage de l'au moins un réseau de neurones, afin que ledit au moins un réseau de neurones entraîné puisse ensuite, lors d'une deuxième étape de segmentation, segmenter une image ou une vidéo pour obtenir une carte probabiliste permettant d'identifier des zones roulables en prenant en compte une incertitude associée ;
- 10 - Un procédé de contrôle d'un véhicule équipé de capteurs mettant en œuvre le procédé de segmentation précédent afin d'établir une consigne permettant d'avoir une vitesse et/ou une trajectoire adaptée(s) à la qualité de la zone roulable.

Procédé d'identification d'au moins une zone roulable (E)

15 Le **figure 4** illustre schématiquement des étapes du procédé d'identification (E) de zones roulables dans un exemple de réalisation.

Le procédé d'identification E de zones roulables met en œuvre un procédé de création de base d'apprentissage E0, un procédé d'apprentissage E1, et un procédé de segmentation E2 d'une image d'entrée, mis en œuvre par au moins un réseau de neurones entraîné au cours du
20 procédé d'apprentissage E1.

L'homme du métier comprendra qu'une fois que l'au moins un réseau de neurones entraîné est obtenu au cours du procédé d'apprentissage E1, le procédé d'identification E de zones roulables peut comprendre une répétition du procédé de segmentation d'images E2 à de nombreuses reprises, sans nécessiter la répétition du procédé d'apprentissage E1.

25 Par segmentation d'une image d'entrée, on entendra une classification de chaque pixel de l'image, c'est-à-dire la prédiction d'un label pour chaque pixel de l'image, chaque label étant associé à un objet réel à identifier, par exemple une route ou plus généralement une zone roulable. Ainsi, l'image est segmentée en une pluralité d'ensembles de pixels comprenant le même label, chaque ensemble de pixels définissant un objet. Comme expliqué précédemment,
30 l'image de la **figure 1** est segmentée en cinq classes, la zone roulable a1, les trottoirs a2, les usagers a3, les bâtiments a4, le fond et la végétation a5.

Par la suite, on utilisera de manière équivalente les termes « piste », « chemin » ou « zone roulable », pour désigner une surface au sol sur laquelle un véhicule terrestre est capable de rouler.

De manière générale, une image est une matrice de pixels de taille donnée, par exemple 321 x 321 pixels. On parlera de manière générale d'image d'un terrain pour désigner une image d'une zone extérieure, ne comprenant pas nécessairement une zone roulable. On pourra désigner par « la segmentation » une image segmentée, ou une matrice dont les éléments sont des labels correspondant à un objet ou à une probabilité.

Architecture

La **figure 3** illustre un exemple d'architecture dans laquelle les étapes des procédés ci-dessus peuvent être implémentées.

Le procédé d'apprentissage (E1) peut être mis en œuvre sur un premier serveur 1 et le procédé de segmentation (E2) peut être mis en œuvre sur un deuxième serveur 2. Chaque serveur 1, 2 comprend avantageusement au moins une unité de traitement 3, 4, par exemple un processeur ayant de préférence une puissance de calcul élevé.

Chaque serveur 1, 2 peut également comprendre une interface de communication 7, 8 permettant une communication entre eux. Une telle communication peut par exemple être mise en œuvre au moyen d'une liaison filaire ou sans fil par l'intermédiaire de tout type de réseau de communication par exemple le réseau Internet. S'agissant du premier serveur 1, il peut avoir accès par l'intermédiaire de l'interface de communication 7 à une base de données d'apprentissage distante (non représentée).

De manière avantageuse également, chaque serveur 1, 2 peut comprendre une unité de stockage 5,6, par exemple un disque dur. Typiquement, le premier serveur d'apprentissage 1 peut stocker dans l'unité de stockage 5 une base de données d'apprentissage ou avoir accès à une base de données d'apprentissage.

L'homme du métier comprendra sans difficulté que d'autres architectures pour la mise en œuvre des procédés d'apprentissage (E1) et de segmentation (E2) sont possibles, les deux serveurs 1,2 pouvant notamment être fusionnés.

Procédé de segmentation (étape E2)

Selon un premier aspect, le procédé d'identification E de zone(s) roulable(s) comprend la mise en œuvre d'un procédé de segmentation d'image ou de vidéo (E2). La segmentation d'image ou de vidéo est réalisée au moyen d'au moins un réseau de neurones préalablement entraînés, au cours d'un procédé d'apprentissage (E1) qui sera détaillé par la suite.

De manière générale, un réseau de neurones comprend une pluralité de neurones (de type perceptron ou de type convolutionnel) sur plusieurs couches connectées entre elles..

L'au moins un réseau de neurones peut être implémenté selon un grand nombre d'architectures différentes. De manière non limitative, il est possible d'utiliser un réseau selon l'une
5 des architectures connues de l'homme du métier de type ConvNet comme par exemple : ResNet (réseau de neurones convolutionnels (CNN) permettant d'optimiser l'apprentissage de réseaux très profonds (plus de 150 couches)), DenseNet (réseau comprenant plusieurs « block » dont les couches sont reliées par des connexions plus courtes, SegNet (réseau de neurones
entièrement convolutifs présentant une architecture encodeur-décodeur symétrique. L'encodeur
10 est une succession de couches convolutives suivies par une normalisation par batch et des fonctions de transfert non linéaires, et le décodeur est une symétrie de l'encodeur (Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. *SegNet : A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation*. November 2015),. On pourra également s'inspirer
d'une de ces architectures est l'adapter au contexte des véhicules autonomes.

15 Par la suite, on se placera dans un premier temps dans le cas où la donnée d'entrée est une image, qu'on désignera généralement par « image d'entrée » ou « image de terrain d'entrée », avant de s'intéresser à la segmentation d'une vidéo ou d'une séquence d'images

Dans un premier exemple de réalisation illustré sur la **figure 5**, le procédé de segmentation (E2) est réalisé au moyen d'un premier réseau de neurones profonds R1 et d'un
20 second réseau de neurones profonds R2, et préalablement entraînés.

Le premier réseau de neurones profonds R1 reçoit en entrée une image I1 pour effectuer une première segmentation de la zone roulable (étape S10) afin d'obtenir une image I2 segmentée. L'image I1 est de préférence fournie au premier réseau de neurones profonds R1 par
l'intermédiaire de l'interface de communication 8 connectée à une unité 10 distante du serveur 2
25 (par exemple un capteur) ou bien est stockée dans l'unité de stockage 6 du serveur 2.

Typiquement, comme expliqué précédemment, il peut s'agir de classifier chaque pixel de l'image en prédisant pour chaque pixel de l'image I1 un label, parmi un premier label correspondant à la classe « zone roulable » et un deuxième label correspondant à la classe
« zone non roulable ». Le premier réseau de neurones R1 produit en sortie une image segmentée
30 I2.

Dans un exemple de réalisation, il peut s'agir d'une binarisation de l'image I1 fournie en entrée, l'image segmentée I2 obtenue en sortie comprenant des pixels blancs et des pixels blancs

uniquement, les pixels blancs correspondant par exemple à la zone roulable identifiée, comme illustré schématiquement sur la **figure 5**.

Au moins l'image segmentée I2 en sortie du premier réseau de neurones R1 est donnée en entrée du second réseau de neurones R2. Dans l'exemple d'architecture illustré sur la **figure 5** par exemple, le second réseau de neurones R2 prend en entrée l'image d'entrée I1 image segmentée I2 issue du premier réseau de neurones R1.

Le second réseau de neurones R2, est configuré, c'est-à-dire entraîné, pour fournir en sortie une estimation de la probabilité de la présence de zone roulable dans l'image d'entrée I1. L'estimation de la probabilité de la présence de zone roulable (étape S20) dans l'image d'entrée I1 peut ensuite être transmise à une unité 12, par l'intermédiaire de l'interface de communication 8 ou bien être stockée dans l'unité de stockage 6 du serveur 2.

De manière générale, le second réseau de neurones R2 calcule l'estimation de la probabilité de la présence de zone roulable sous la forme d'une carte probabiliste I3 de présence d'au moins une zone roulable dans l'image d'entrée I1. La carte probabiliste I3 est une matrice de deux dimensions $L \times l$, dont chaque élément ou valeur est compris entre 0 et 1 et correspond à une probabilité de présence d'au moins une zone roulable sur un pixel de l'image d'entrée I1.

On parlera de manière indépendante de carte probabiliste ou de carte de probabilité pour définir l'estimation d'une probabilité de présence d'une zone roulable sur une image.

Chaque valeur ou élément de la carte probabiliste I3 est destiné à exprimer la probabilité pour un pixel correspondant de l'image d'entrée I1 d'appartenir à une zone roulable. Par exemple, on associe à un pixel situé dans une zone identifiée comme non roulable le chiffre 0 (noir), et dans une zone identifiée comme roulable de façon certain le chiffre 1 (blanc). On associe aux pixels situés dans des zones présentant une incertitude des valeurs intermédiaires, typiquement 0,2 pour un pixel d'une zone probablement pas roulable, ou 0,8 un pixel d'une zone probablement roulable.

Dans un exemple de réalisation, la sortie du second réseau de neurones R2 peut être une matrice de même dimension que l'image d'entrée I1, comprenant par exemple 321×321 valeurs. Une telle représentation des cartes probabilistes est également illustrée sur les **figures 10 à 12**.

Plus généralement, les dimensions de la matrice représentant la carte de probabilité I3 peuvent être inférieures à celles de l'image d'entrée I1, par exemple suite à une compression de l'image d'entrée I1. Cela implique une perte d'informations mais permet d'améliorer la vitesse de calcul de la carte de probabilité I3.

De manière alternative, les dimensions de la matrice représentant la carte de probabilité I3 peuvent être supérieures à celles de l'image d'entrée I1, par exemple en réalisant des interpolations entre les pixels de l'image d'entrée I1.

Dans un deuxième exemple de réalisation illustré sur la **figure 6**, le procédé de segmentation (E2) peut être réalisé au moyen d'un unique réseau de neurones profonds R3.

A partir de l'image I1 donnée en entrée le réseau de neurones profonds R3 est configuré pour fournir en sortie une estimation de la probabilité de la présence de zone roulable dans l'image d'entrée I1.

Selon ce deuxième exemple de réalisation, le réseau de neurones R3 ne comprend pas d'étape de segmentation de l'image d'entrée I1 pour obtenir une image segmentée I2. Comme cela sera détaillé par la suite, le réseau de neurones R3 est configuré lors du procédé d'apprentissage (E1) pour directement calculer l'estimation de la probabilité de la présence de zone roulable sous la forme d'une carte probabiliste I3.

Le réseau de neurones R3 peut comprendre une ou plusieurs « skip connexion(s) » (31, 32), c'est-à-dire que l'information de l'image d'entrée I1 peut être redonnée en entrée au niveau d'un ou plusieurs point(s) du réseau de neurones R3, comme cela est visible sur la **figure 6**.

La carte probabiliste I3 calculée par le réseau de neurones R3 peut être une matrice semblable à celle obtenue en sortie du second réseau de neurones R2, et décrite précédemment.

A partir des deux exemples de réalisations décrits précédemment, il est également possible de manière analogue de segmenter une vidéo.

De manière générale, une vidéo est une séquence de T images, dépendant de la longueur de la vidéo, numérotées de 1 à T . En d'autres termes, une image quelconque d'une vidéo correspond à une image t , avec $t \in \llbracket 1; T \rrbracket$.

On considèrera par la suite que la donnée d'entrée est une vidéo d'entrée c'est-à-dire une séquence de T images.

Dans le premier exemple de réalisation, le procédé de segmentation E2 mis en œuvre dans le procédé d'identification E de zones roulables est effectué par deux réseaux de neurones (R1, R2).

Le premier réseau de neurones R1 peut être configuré pour segmenter chaque image de la vidéo d'entrée. Dans ce cas, la sortie du premier réseau de neurones R1 peut être la séquence d'images de la vidéo d'entrée segmentées, c'est-à-dire la vidéo d'entrée segmentée, par exemple par binarisation.

De même, le second réseau de neurones R2 peut être configuré pour calculer une carte de probabilité de présence pour chaque image segmentée obtenue en sortie du premier réseau de neurones R1.

De préférence, il est possible d'effectuer un séquençage de la vidéo d'entrée avant la segmentation par le premier réseau de neurones R1 et/ou avant le calcul de cartes probabilistes par le second réseau de neurones R2.

Par exemple, le premier réseau de neurones R1 peut être configuré pour ne calculer de segmentation d'une image qu'à une certaine fréquence, par exemple de ne calculer la segmentation que d'une image sur deux, ou d'une image sur trois de la vidéo d'entrée.

Lors d'une segmentation séquentielle, le résultat en sortie du premier réseau de neurones R1 est un ensemble de matrices d'images segmentées dont le nombre est inférieur au nombre d'images de la vidéo d'entrée.

De préférence, la vidéo est divisée en séquences de taille $T' < T$, la taille de la séquence étant choisie de sorte que, pour chaque séquence, le résultat en sortie du premier réseau de neurones R1 peut être une unique matrice représentant la classification en zone roulant/non roulant d'une image suffisamment représentative de la séquence. Typiquement, il peut s'agir de la dernière image de la séquence. La valeur de T' peut dépendre de la vitesse d'acquisition du capteur fournissant la vidéo d'entrée, par exemple d'une caméra. La valeur de T' peut dépendre de la taille des réseaux de neurones (R1,R2,R3), et/ou de la capacité de traitement de l'unité de traitement (3,4).

De préférence, la taille d'une séquence T' est inférieure à dix images. Cela permet d'assurer la fiabilité d'une succession de cartes de probabilités, notamment dans une mise en œuvre dans un procédé de contrôle d'un véhicule terrestre, comme il sera détaillé par la suite.

De manière avantageuse, il est possible de calculer séquentiellement la première segmentation puis de redonner en entrée au deuxième réseau de neurones R2 la séquence d'images segmentées et toutes les images de la vidéo d'entrée.

De même, le second réseau de neurones R2 peut être configuré pour ne calculer une carte probabiliste d'une image segmentée par le premier réseau de neurones R1 qu'à une certaine fréquence, par exemple de ne calculer la carte probabiliste que d'une image segmentée sur deux ou sur trois.

Ne pas segmenter toutes les images de la vidéo d'entrée et/ou ne pas calculer toutes les cartes de probabilité associées aux images de la vidéo d'entrée permet avantageusement de diminuer la puissance de calcul nécessaire au traitement de la vidéo d'entrée, et donc le temps

de traitement. Cela est particulièrement pertinent dans le cas où les cartes de probabilités sont utilisées pour prendre une décision de manière dynamique, comme par exemple dans le procédé de contrôle décrit dans le deuxième aspect de l'invention.

5 Dans un deuxième exemple de réalisation, le procédé de segmentation E2 mis en œuvre dans le procédé d'identification E de zones roulables est effectué par l'unique réseau de neurones R3.

De même que précédemment, le réseau de neurones R3 peut être configuré pour ne pas calculer des cartes probabilistes pour toutes les images de la vidéo d'entrée.

10 Un exemple schématique est illustré sur la **figure 7**. La vidéo d'entrée, fournie en entrée par l'unité 10 à l'unité de traitement 4 sur laquelle est implémentée le réseau de neurones R3, est dans cet exemple divisée en séquences de taille trois ($T'=3$) par l'unité de traitement 4. Chaque séquence de la vidéo d'entrée comprend par exemple les images ($I1, I1', I1''$).

Le réseau de neurones R3 peut être configuré pour ne calculer qu'une carte de probabilité pour chaque séquence de la vidéo d'entrée, de sorte que la carte de probabilité soit
15 représentative de l'estimation de la probabilité de la présence de zone roulable dans les images ($I1, I1', I1''$) de la séquence d'entrée. Par exemple, le réseau de neurones R3 ne peut calculer que la carte de probabilité associée à la première image $I1$ de la séquence ou à la dernière image $I1''$ de la séquence ou à une image centrale $I1'$ de la séquence. La carte de probabilité $I3$ représentative de la séquence d'images ($I1, I1', I1''$) peut ensuite être transmise à l'unité 12.

20 Il est possible de réaliser ce traitement en une seule opération en utilisant une ou plusieurs « skip connexions » comme cela a été détaillé précédemment. Tout ou partie des images ($I1, I1', I1''$) de la séquence de la vidéo d'entrée peuvent être redonnées en entrée en un ou plusieurs points du réseau de neurones R3.

L'homme du métier comprendra que lors d'une utilisation du procédé d'identification E
25 de zones roulable pour le traitement d'une vidéo, le procédé E2 de segmentation d'image peut être répétée indépendamment du procédé d'apprentissage E1, qui va maintenant être décrit.

Procédé d'apprentissage (E1)

De manière générale et connue de l'homme du métier, un procédé d'apprentissage supervisé d'un réseau de neurones correspond à mettre en œuvre un algorithme consistant à
30 modifier les pondérations des neurones du réseau de neurones, jusqu'à ce que la sortie du réseau de neurone corresponde à la sortie attendue.

Un procédé d'apprentissage nécessite une base de données d'apprentissage comprenant un ensemble de données d'entrée exemple, et un ensemble de sorties attendues

associées. Par exemple, la base de données d'apprentissage pour un réseau de neurones destiné à la segmentation d'images comprend une pluralité d'images (ou plus généralement de vidéos) et une pluralité de segmentation probabiliste des images.

Par exemple, l'algorithme peut être configuré pour initialiser les pondérations à des valeurs quelconques, puis pour présenter successivement toutes les données d'entrée exemple de la base de données d'apprentissage au réseau de neurones. Pour chaque donnée d'entrée exemple, l'algorithme peut être configuré pour comparer la sortie du réseau de neurones à la sortie attendue associée à la donnée d'entrée exemple, et pour ajuster les pondérations selon que la sortie du réseau de neurones est « proche » de la sortie attendue. L'algorithme peut être configuré pour s'arrêter lorsque toutes les données d'entrée exemple ont été présentées sans modification d'aucune pondération.

De nombreux algorithmes d'apprentissage peuvent être utilisés, à titre illustratif, un algorithme d'apprentissage par rétropropagation du gradient, ou par exemple un algorithme du gradient stochastique (en anglais « Stochastic Gradient Decent »), ou d'estimation adaptative de moment (en anglais ADAM ou « adaptive moment estimation »). L'homme du métier comprendra que de nombreux algorithmes d'apprentissage, tels que connus dans l'art antérieur peuvent être utilisés pour entraîner le réseau de neurones.

De préférence, le procédé d'apprentissage E1 peut être mis en œuvre par un algorithme d'apprentissage implémenté dans l'unité de traitement 3 du serveur d'apprentissage 1. L'étape d'apprentissage est réalisée à partir d'un ensemble d'images et de sorties attendues associées, contenu dans une base de données d'apprentissage, qui peut être stockée dans l'unité de stockage 5, ou accessible via l'interface de communication 7.

On considère d'abord un premier exemple de réalisation, dans lequel le procédé d'identification E de zones roulables est mis en œuvre au moyen du premier réseau de neurones profonds R1 et du second réseau de neurones profonds R2.

Dans un exemple de réalisation de procédé d'apprentissage, comme illustré par exemple sur la **figure 4**, le procédé d'apprentissage E1 comprend les étapes suivantes :

- S1 : apprentissage supervisé du premier réseau de neurones R1 à partir de la base de données d'apprentissage, de sorte que le premier réseau de neurones R1 entraîné soit configuré pour segmenter une image entre zone roulant et zone non roulant ;
- S2 : apprentissage supervisé du second réseau de neurones R2 à partir de la base de données d'apprentissage et d'au moins une vérité terrain, de sorte que le second réseau

de neurones R2 entraîné soit configuré pour segmenter l'image selon une carte probabiliste de présence de zone roulable.

Il est avantageux de distinguer deux étapes S1, S2 pour entraîner les deux réseaux de neurones distincts (R1, R2) indépendamment. L'algorithme d'apprentissage peut être plus simple à mettre en place, et le procédé d'apprentissage E1 peut être plus rapide à implémenter. Il peut être plus rapide d'entraîner le premier réseau de neurones R1 et le deuxième réseau de neurones R2 successivement.

Dans un exemple de réalisation préférentiel, lors de la première étape d'apprentissage S1, on entraînera le premier réseau de neurones R1 en utilisant une fonction de coût adaptée telle que la fonction de coût à entropie croisée (en anglais « Cross Entropy »), par exemple utilisée dans l'article « SegNet: A deep Convolutional Encoder-Decoder Architecture for Image Segmentation ».

Dans un exemple de réalisation préférentiel, lors de la deuxième étape d'apprentissage S2, on entraînera le deuxième réseau de neurones R2 en utilisant une fonction de coût adaptée telle que la valeur absolue ou le carré des différences entre la matrice de sortie du réseau de neurones R2 et la vérité terrain associée, en chaque pixel de la matrice de sortie.

De façon avantageuse on pourra utiliser des méthodes de régularisation connues de l'homme du métier, par exemple la méthode de dégradation des pondérations (en anglais « weight decay ») pour la définition de la fonction de coût.

Pour mettre en œuvre la première étape S1 d'apprentissage supervisé du premier réseau de neurones R1, permettant la segmentation d'images ou de vidéos, et en particulier la détection de zones roulables, une première étape préliminaire (étape E0) peut comprendre une construction de la base de données d'apprentissage B1 adaptée.

La base de données d'apprentissage B1 peut comprendre un ensemble d'images ou de vidéos de scènes d'extérieur, de préférence plusieurs centaines, de préférence encore plusieurs milliers. Les scènes d'extérieur des images de terrain ou des vidéos ne représentent pas nécessairement toutes de zones roulables, cependant, un ensemble important des scènes d'extérieur présentent au moins un chemin, une piste, ou une zone roulable.

La base de données d'apprentissage B1 comprend également une annotation de l'ensemble des images ou des vidéos. Une annotation d'une image de terrain est une segmentation de l'image (c'est-à-dire une classification de chaque pixel de l'image de terrain), réalisée par un opérateur humain. L'opérateur humain, désigné également par « annotateur » définit un label pour chaque pixel de l'image, correspondant au label attendu pour chaque pixel

de l'image en sortie du premier réseau de neurones R1 entraîné. L'annotation associée à une image est donc le résultat attendu en sortie du réseau de neurones R1 entraîné, s'il recevait en entrée ladite image.

Par exemple, dans l'exemple de réalisation où le premier réseau de neurones R1
5 entraîné est destiné à binariser une image d'entrée entre une classe « zone roulable » et une classe « zone non roulable », l'annotateur devra attribuer à chaque pixel un label parmi deux labels (un premier label correspondant à une zone roulable, et un second label correspondant à une zone non roulable), selon qu'il considère que le pixel appartient à une zone roulable ou pas.

De préférence, chaque image de terrain de la base de données d'apprentissage B1 peut
10 être associée à une annotation d'un annotateur correspondant à la perception de l'annotateur, par exemple à la distinction par l'annotateur en une zone roulable et une zone non roulable.

Dans le cas où les zones roulables sur une image de terrain sont mal définies, l'appréciation dépendant de l'opérateur humain ou annotateur réalisant l'annotation, il est également possible d'associer plusieurs annotations différentes à l'image de terrain, afin de
15 multiplier les données de la base de données d'apprentissage B1.

A partir des images de terrain et des annotations associées, contenus dans la base de données d'apprentissage B1, l'algorithme d'apprentissage peut entraîner le premier réseau de neurones R1, de sorte qu'il soit capable de segmenter des images. Par exemple, de sorte que le premier réseau de neurones R1 entraîné soit capable de binariser des images entre « zone
20 roulable » et « zone non roulable ».

Dans un autre exemple de réalisation, la première étape d'apprentissage S1 n'est pas nécessaire, et le réseau de neurones profonds R1 utilisé lors de l'étape E2 de segmentation est un réseau de neurones déjà entraîné à identifier des zones roulables.

Pour mettre en œuvre la deuxième étape S2 d'apprentissage supervisé du deuxième
25 réseau de neurones R2, permettant le calcul de cartes probabilistes de présence de zones roulables, il convient d'adapter la base de données d'apprentissage B1, en une base de données d'apprentissage B2.

A l'issue de la deuxième étape d'apprentissage S2, le deuxième réseau de neurones R2 peut avoir appris à estimer la probabilité de présence d'un chemin, piste, zone roulable à partir
30 d'une image d'une scène d'extérieur et/ou d'une image segmentée pouvant être issue du premier réseau de neurones R1.

La base de données d'apprentissage B2 peut comprendre l'ensemble des images de terrain de la base de données d'apprentissage B1.

La base de données d'apprentissage B2 comprend également un ensemble de cartes probabilistes associées aux images de terrain. Chaque carte probabiliste correspond à une vérité terrain, c'est-à-dire à l'estimation de présence de zone roulable attendue en sortie du deuxième réseau de neurones R2 entraîné, si l'image associée était donnée en entrée.

5 L'étape E0 peut comprendre une étape de calcul d'au moins une vérité terrain S200 définie par une carte probabiliste pour au moins une image associée de la base de données d'apprentissage B1. L'au moins une vérité terrain calculée est utilisée pour l'étape d'apprentissage S2 illustrée sur la **figure 4**.

De préférence, les vérités terrain ont un format semblable aux cartes probabilistes qui
10 sont calculées au cours du procédé de segmentation E2.

Comme expliqué précédemment, dans le cadre de zones roulables mal définies, il est avantageux d'ajouter une information sur l'incertitude de la zone roulable à identifier, et notamment sur l'incertitude de l'existence de la zone roulable.

Pour construire une vérité terrain associée à une image de la base de données B1, il est
15 possible d'utiliser des annotations de l'image réalisées par des annotateurs différents, de manière indépendante, lors d'une étape d'annotation d'images S100. De préférence, la vérité terrain de l'image est construite à partir d'au moins deux annotations différentes de l'image, c'est-à-dire que l'image peut avoir été préalablement annotée par plus de deux annotateurs, de préférence encore par au moins cinq annotateurs indépendants.

20 La **figure 2** illustre un exemple d'image de terrain dans laquelle il existe une incertitude dans la définition de la zone roulable. Par exemple, la zone la plus claire (zones z1) a été identifiée comme roulable par les cinq annotateurs de l'image, et correspond donc à une voie roulable identifiée avec une forte certitude, tandis que la zone la plus sombre (ciel) correspond à une zone n'ayant été identifiée comme roulable par aucun des annotateurs. Les zones intermédiaires
25 correspondent à des zones n'ayant pas été identifiées par tous les annotateurs comme des zones roulables. Ainsi, la même zone a été annotée par plusieurs annotateurs et peut être annotées différemment d'un annotateur à l'autre. Ceci est illustré sur la figure 13 où une image conduit à deux annotations différentes de la zone roulable. Un annotateur aura identifié une zone roulable z1' plus grande (en bas à gauche) que l'autre annotateur (en bas à droite), la zone non roulable
30 z'5 étant aussi différente d'un annotateur à l'autre.

Ainsi, plus le nombre d'annotateurs différents est élevé, plus on dispose de perceptions différentes de la zone roulable éventuelle, et il sera simple d'estimer une incertitude dans la définition même de la capacité à rouler sur la zone roulable éventuelle.

Pour le calcul de la vérité terrain, on peut considérer qu'un consensus des annotations de tous les annotateurs représente une zone roulable certaine, associée par exemple à la valeur 1, tandis que la divergence des annotations porte l'information d'une probabilité de chemin plus faible. De manière générale, plus un nombre élevé d'annotateurs différents identifie une zone

d'une image comme zone roulable, plus l'incertitude de la capacité à rouler sur la zone diminuera. De préférence, on peut construire une vérité terrain pour chacune des images de la base de données d'apprentissage B1. L'homme du métier comprendra que, dans une certaine mesure, plus le nombre d'images de la base de données d'apprentissage B2 augmente, plus la fiabilité du réseau de neurones R2 entraîné augmentera.

Selon un exemple de réalisation, le calcul de la vérité terrain S200 consiste à additionner pour chaque pixel d'une image annotée de la base de données B1 le nombre de fois qu'un annotateur a considéré que ledit pixel fait partie d'une zone roulable puis l'unité de traitement 3 divise cette somme par le nombre d'annotateurs ayant annoté l'image. On obtient un chiffre entre 0 et 1 qui représente la probabilité pour un pixel d'appartenir à un chemin (piste, zone roulable...).

Ainsi, l'ensemble de vérités terrain correspond à un ensemble de matrices de la taille des images de la base de données d'apprentissage B1, chaque matrice représentant une moyenne des annotations de l'image correspondante ou associée dans la base de données B1.

Dans l'exemple d'une binarisation, où l'annotation d'un pixel vaut 1 si l'annotateur i parmi les K annotateurs de l'image a considéré que le pixel appartient à une zone roulable, ou 0 sinon, chaque élément p de la vérité terrain est calculé de la manière suivante :

$$p(pixel) = \sum_{i=1}^K \frac{annotation(pixel, i)}{K}$$

On obtient ainsi une carte probabiliste de présence de zones roulables dans un ensemble d'images de la base de données d'apprentissage B1.

Par exemple, la vérité terrain associée à l'image de terrain de la **figure 2** comprend une zone $z1$ annotée comme « zone roulable » par tous les annotateurs de l'image, et une zone $z5$ annotée par aucun des annotateurs. Les pixels de la zone $z1$ sont associés à la valeur 1 et les pixels de la zone $z5$ sont associés à la valeur 0. Les pixels de la zone $z2$, annotée par quatre annotateurs sur les cinq annotateurs de l'image peuvent être associés à la valeur 0,8. Les pixels de la zone $z3$, annotée par trois annotateurs, peuvent être associés à la valeur 0,6. Les pixels de la zone $z3$, annotée par un annotateur, peuvent être associés à la valeur 0,2.

A partir de la deuxième base de données d'apprentissage B2, il est possible d'entraîner le réseau de neurones R2. Comme expliqué précédemment, il s'agit de redéfinir itérativement les pondérations du deuxième réseau de neurones R2 de sorte à faire converger les cartes de probabilité en sortie du réseau de neurones R2 vers les cartes de probabilités correspondant à la vérité terrain calculée précédemment.

Dans le deuxième exemple de réalisation décrit, le procédé d'identification E de zones roulables est mis en œuvre au moyen d'un seul réseau de neurones R3 pouvant comprendre des « skip connexions ».

Dans cet exemple de réalisation, l'apprentissage peut être réalisé de bout en bout (« end-to-end »), c'est-à-dire les deux étapes d'apprentissages S1 et S2 sont réalisées conjointement.

L'algorithme d'apprentissage du réseau de neurones R3 peut prendre directement en entrée la base de données d'apprentissage B2 comprenant les images de la base de données B1, brutes c'est-à-dire sans annotation, avec l'ensemble des vérités terrain associé.

Procédé de contrôle d'un véhicule utilisant la segmentation probabiliste de zones roulables

On décrit ici un procédé de contrôle d'un véhicule terrestre en vitesse et en trajectoire, mettant en œuvre le procédé d'identification E de zones roulables ci-dessus décrit.

La **figure 8** illustre schématiquement un véhicule terrestre V, équipé d'au moins un capteur 10' embarqué et d'une unité de traitement 4'. Le capteur 10' est destiné à fournir une image d'entrée à l'unité ou moyen de traitement 4', sur laquelle peut être implémenté au moins l'étape de segmentation d'image E2 par l'au moins un réseau de neurones entraîné.

De préférence, le moyen de traitement 4' peut comprendre au moins un processeur, par exemple un processeur graphique de type GPU (en anglais, « *Graphics Processing Unit* »), destiné à traiter les cartes probabilistes calculées par l'au moins un réseau de neurones entraîné, et communiquer avec un calculateur C du véhicule terrestre V.

Le véhicule terrestre V comprend des actionneurs A, par exemple un système motorisé relié à une transmission. Le calculateur C du véhicule V peut être configuré pour contrôler le fonctionnement de l'actionneur A.

De préférence, le calculateur C reçoit des informations en provenance de capteurs situés au niveau de l'actionneur A, et est capable d'adapter le contrôle du fonctionnement de l'actionneur A à ces données. Par exemple, le calculateur C peut permettre de contrôler une direction assistée, ou une vitesse d'un axe de transmission.

Comme cela est illustré sur la **figure 9**, le procédé de contrôle du véhicule terrestre V peut comprendre plusieurs étapes :

- C1 : acquisition d'images par l'au moins un capteur 10' situé sur le véhicule V ;
- 5 - C2 : traitement des images acquises par l'unité de traitement 4', selon un procédé de segmentation E2 décrit précédemment, afin d'obtenir au moins une carte probabiliste de présence de zones roulables ;
- C3 : calcul d'une consigne comprenant les sous-étapes de :
 - 10 C31 : calcul d'une probabilité de présence de zone roulable par le moyen de traitement 4' à partir de la carte probabiliste de présence de zones roulables ;
 - C32 : calcul d'une trajectoire et/ou d'une consigne en vitesse du véhicule V par le moyen de traitement à partir de la probabilité de présence de zone roulable.

15 Dans un exemple de réalisation, le procédé de contrôle peut comprendre une étape C4 de transmission de la consigne au calculateur C du véhicule V.

Afin de pouvoir mettre en œuvre la première étape C1 du procédé de contrôle, l'au moins un capteur 10' est configuré pour l'acquisition de données nécessaires au contrôle du véhicule V. Le capteur 10' peut être un capteur permettant l'acquisition d'images ou de vidéos par exemple une caméra, située de préférence à l'avant ou sur le toit du véhicule V.

20 Sur l'exemple illustré à la **figure 8**, l'unité de traitement de données 4' est un système embarqué comprenant un processeur par exemple. De manière alternative, l'unité de traitement 4' peut être située sur un serveur à distance tel que le serveur 2 décrit précédemment, communiquant avec les capteurs 10' via l'interface de communication 8 et avec le calculateur C.

 Sur l'unité de traitement de données ou unité de traitement 4' est implémenté le procédé
25 de segmentation E2 du procédé d'identification de zones roulables et illustré dans différents exemples de réalisation sur les **figures 5 à 7**. L'unité de traitement de données 4' détermine à partir des données collectées par l'au moins un capteur 10' une segmentation de l'environnement prenant en compte l'incertitude de la capacité du véhicule V à pouvoir rouler sur les différentes zones détectées. Une telle segmentation de l'environnement est typiquement une carte de
30 probabilité ou carte probabiliste de présence de zone roulable telle que décrite précédemment pour une image de l'environnement du véhicule V.

De préférence, la segmentation de l'étape de traitement C2 des images acquises lors de l'étape C1 est effectuée de manière séquentielle, comme cela a été expliqué lorsque la donnée

d'entrée de du procédé de segmentation E2 du procédé d'identification E de zones roulables est une vidéo par exemple. Cela permet de garantir des temps de calculs plus courts, et une mise en œuvre plus efficace et fiable du procédé de contrôle du véhicule terrestre V.

A partir de cartes probabilistes de présence de zones roulables, il est possible de piloter le véhicule. L'étape C3 peut comprendre une première sous-étape d'analyse C31 de la segmentation de l'environnement obtenue lors de l'étape C2.

On définit un premier seuil (Seuil1) représentant le niveau de risque que l'on souhaite prendre sur le niveau de probabilité nécessaire d'appartenance à une zone roulant. Le premier seuil de risque Seuil1 peut dépendre par exemple du type de véhicule, de l'utilisateur, ou de l'environnement. Le premier seuil de risque Seuil1 peut être une donnée préprogrammée sur le moyen de traitement de données 4' ou réglée par un utilisateur du véhicule V.

Lors de la première sous-étape d'analyse C31, l'unité de traitement 4' peut calculer à partir de la segmentation des grandeurs prédéfinies pertinentes pour établir une loi de contrôle du véhicule V.

On note P une grandeur représentant la probabilité moyenne des pixels considérés comme appartenant à une zone roulant de la segmentation d'une image de l'environnement. L'unité de traitement 4' est configurée pour calculer la somme sur tous les pixels considérés comme appartenant à la zone roulant au premier seuil de risque Seuil1 de la probabilité des pixels, puis diviser la somme obtenue par la taille N de la segmentation.

Soit :

$$P = \sum_{\substack{pixel \\ p(pixel) > Seuil1}} \frac{p(pixel)}{N}$$

avec p(pixel) la probabilité pour qu'un pixel appartienne à un chemin, selon la carte de probabilité obtenue par le procédé de segmentation d'image à l'étape C2.

Cette grandeur P donne une probabilité moyenne d'un pixel de l'image de l'environnement d'appartenir à une zone roulant.

On peut également définir P' une grandeur représentant la probabilité moyenne des pixels considérés comme appartenant à une zone roulant par rapport à la zone considérée comme roulant au seuil Seuil1. L'unité de traitement 4' calcule la somme sur tous les pixels considérés comme appartenant à la zone roulant au premier seuil de risque Seuil1 de la probabilité des pixels, puis divise la somme obtenue par la taille N' correspondant à la taille du chemin au seuil Seuil1, soit

$$N' = \sum_{pixel} \chi_{\{p(pixel) > Seuil1\}}(pixel)$$

avec χ la fonction caractéristique de l'ensemble
 $\{pixel / p(pixel) > Seuil1\}$.

On a donc :

5

$$P' = \sum_{\substack{pixel \\ p(pixel) > Seuil1}} \frac{p(pixel)}{N'}$$

Cette grandeur donne un indicateur de la qualité du chemin :

- Si $P' \rightarrow 1$, alors la zone roulable est de bonne qualité : toute la zone identifiée comme roulable au seuil Seuil1 est roulable de façon certaine.
- 10 - Si $P' \rightarrow Seuil1$, alors la zone roulable est de mauvaise qualité : toute la zone identifiée comme roulable au seuil Seuil1 est estimée roulable au seuil de risque Seuil1 uniquement.

Par exemple, dans le cas où l'on choisit de prendre un risque élevé, le seuil de risque Seuil1 est faible, par exemple 0,1.

Sur l'exemple illustré sur la **figure 10**, le réseau de neurones, par exemple le second
 15 réseau de neurones R2 entraîné ou le réseau de neurones R3, a segmenté l'image de 321 x 321 pixels en une zone inférieure roulable de manière certaine de probabilité 1, correspondant en réalité à une piste, et une zone non roulable de manière certaine, correspondant en réalité à un fond montagneux. Dans cet exemple, la moitié inférieure de l'image d'entrée est considérée comme un chemin au seuil de risque Seuil1, soit $N' = 160 \times 321$ pixels. Comme l'ensemble de la
 20 zone considéré comme roulable au seuil de risque Seuil1 l'est de façon certaine, on a $P' = 1$ et $P = 0,5$.

Sur l'exemple illustré sur la **figure 11**, le deuxième réseau de neurones R2 entraîné ou le réseau de neurones R3 entraîné a segmenté l'image de 321 x 321 pixels en une zone inférieure roulable de manière non certaine, correspondant en réalité à une piste, et une zone non roulable
 25 de manière certaine, correspondant en réalité à un fond de forêt. Parmi les pixels de la moitié inférieure de l'image, une zone centrale d'environ 160 x 160 pixels est roulable de façon certaine ($p=1$), et les zones latérales de la moitié inférieure, représentant 160 x 161 pixels sont roulables de façon probable ($p=0.5$), et correspondent en réalité à une piste recouverte de feuilles. On obtient dans cet exemple $P' = (160 \times 160 \times 1 + 160 \times 161 \times 0.5) / 160 \times 321 = 0.75$ et $P = (160 \times$
 30 $160 \times 1 + 160 \times 161 \times 0.5) / 321 \times 321 = 0.375$.

Ainsi, plus P' et P sont grands plus la présence d'un chemin ou zone roulable sur l'image et donc dans l'environnement du véhicule V est certaine.

On définit également une grandeur S représentant la taille de la plus grande composante connexe, c'est-à-dire la plus grande zone de la segmentation considérée comme roulable au seuil de risque $Seuil1$. Elle permet de vérifier si la segmentation du chemin est morcelée, et donc mesure la qualité de la zone roulable.

$$S = \max(\text{surface en pixel de composante connexe } \{pixel \mid p(pixel) > Seuil1\})$$

S représente la surface de la plus grande composante connexe composée de pixels présentant une probabilité supérieure à $Seuil1$.

Par exemple, la grandeur S sera maximale dans l'exemple illustré sur les figures 10 et 11 puisque la zone identifiée comme roulable au seuil de risque $Seuil1$ est constituée d'un seul tenant.

En revanche, une piste comprenant par exemple des nids de poules, des ornières ou des obstacles divers ne sera pas identifiée comme une zone roulable d'un seul tenant, et aura donc une valeur S plus faible.

Dans l'exemple illustré sur la **figure 12**, l'image de l'environnement acquise par le capteur lors de l'étape C1 comprend des obstacles (flaques, piétons, véhicules), identifiés comme des zones non roulables de manière certaine ou probablement non roulables par le réseau de neurones lors de l'étape C2. Selon le seuil de risque choisi, la taille de la composante connexe maximale varie.

Lors de la sous-étape d'analyse C31, l'unité de traitement 4' peut également calculer d'autres grandeurs intéressantes pour le contrôle du véhicule V .

A partir des grandeurs calculées lors de la première sous-étape d'analyse C31 de la qualité du chemin de l'étape C3, l'unité de traitement 4' peut calculer, lors d'une deuxième sous-étape C32, une commande ou consigne en vitesse du véhicule terrestre V .

De manière générale, il est connu de l'art antérieur des procédés de contrôle de véhicule (en anglais « path planning » par exemple :

- Daniel Harabor and Alban Grastien. An Optimal Any-Angle Pathfinding Algorithm. Proceedings of the Twenty-Third International Conference on Automated Planning and Scheduling.
- Sinyukov, Dmitry A.; Padiş, Taskin (May–June 2017). "CWave: High-Performance Single-Source Any-Angle Path Planning on a Grid". Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA). 2017 IEEE International Conference on

Robotics and Automation (ICRA). Singapore: IEEE. pp. 6190–6197), permettant de piloter une trajectoire d'un véhicule, notamment en amont, à partir de données de navigation GPS (en anglais, « Global Positioning System ») du véhicule et de données concernant le réseau routier, puis à partir de la perception de l'environnement direct du véhicule par des capteurs, et de modèles prédictifs de comportement.

Dans le cas où le véhicule V effectue une trajectoire « off-road », sur des pistes ou des zones non cartographiées notamment, un procédé de contrôle de véhicule tel que décrit par la suite est particulièrement adapté.

De manière générale, la commande ou consigne en vitesse est définie comme une fonction des grandeurs calculées lors de la première sous-étape C31, et notamment de la probabilité de l'existence d'une zone roulable. De manière générale, on a :

$$Vitesse = f(\theta)$$

avec θ comprenant au moins un paramètre calculé lors de la sous-étape C31, par exemple $\theta \in \{P, P', S, (P, S), (P', S), (P, P', S)\}$, et f une fonction continue.

Il est particulièrement avantageux que la vitesse du véhicule V dépende de la probabilité de l'existence de zone roulable P ou P', puisque cette probabilité est un indicateur de la qualité de la zone roulable.

Par exemple, le moyen de traitement 4' peut établir une consigne en vitesse selon la formule :

$$Vitesse = V_0 \times \frac{S}{N'}$$

avec V_0 la vitesse cible du véhicule.

De préférence, la consigne en vitesse peut tendre à ralentir le véhicule terrestre V si la qualité de la zone roulable n'est pas bonne.

De préférence, la consigne en vitesse peut augmenter lorsque la probabilité de présence de zone roulable P ou P' augmente.

De nombreuses fonctions de commande de vitesse peuvent être définies, par exemple

$$Vitesse = V_0 \times P'.$$

On a alors

$$Vitesse \rightarrow V_0$$

lorsque le chemin identifié est une zone roulable de manière certaine.

De manière générale, la commande de vitesse calculée à la sous-étape C32 peut dépendre de la carte de probabilité établie lors de l'étape C2 et du seuil de risque autorisé *Seuil1*, qui peut être prédéfini par l'utilisateur.

$$Vitesse = f(\text{carte de probabilité}, \text{Seuil1})$$

5 A partir des grandeurs calculées lors de la sous-étape d'analyse de la qualité du chemin, le moyen de traitement 4' peut également, lors de la deuxième sous-étape C32, établir une commande de la trajectoire du véhicule terrestre.

De préférence, les capteurs 10' peuvent comprendre un système de détection de distance de type ultrason, laser, LIDAR. Des données de distance permettent avantageusement de compléter les cartes de probabilités obtenues lors de l'étape C2, en fournissant une information sur la profondeur de la zone roulable identifiée au seuil de risque *Seuil1*, et sur la distance à des obstacles.

De préférence, les capteurs 10' peuvent comprendre un système de détection de position de type GPS. Des données de position permettent avantageusement de compléter les cartes de probabilités obtenues lors de l'étape C2. Par exemple, des données de position peuvent être utilisées lorsque le véhicule V circule sur des zones de routes bien définies. Des données de position peuvent aussi être utilisées pour calculer dynamiquement la vitesse du véhicule V, et notamment identifier des dérapages, afin d'ajuster la trajectoire du véhicule V.

De préférence, la consigne en trajectoire est calculée à partir de la carte probabiliste de zones roulables obtenue lors de l'étape C2 de sorte à maximiser la probabilité de présence de zone roulable.

Par exemple, la consigne de la trajectoire peut être calculée de sorte à maximiser l'intégrale des probabilités des pixels empruntés par la trajectoire prévue pour le véhicule V.

Par exemple, il est possible à partir des images issues des capteurs 10' de définir une trajectoire en trois dimensions (3D) par interpolation. On peut paramétrer par exemple une trajectoire par une fonction d'interpolation polynomiale de type *spline* représentant les trajectoires possibles paramétrées par un ensemble de paramètres $\theta 1$, par exemple

$$\theta 1 \in \{P, P', S, (P, S), (P', S), (P, P', S)\}.$$

On écrira de manière générale :

$$30 \quad Trajectoire = \operatorname{argmax}_{\theta 1} \int p(\text{Spline}(\theta 1))$$

avec *Spline* la projection des points de passage du véhicule V dans l'image.

En particulier, on peut choisir de contraindre la trajectoire à ne pas passer par les zones où le chemin est le moins probable. La trajectoire peut être alors calculée:

$$Trajectoire = \underset{\theta_1}{argmax} \frac{\int_{s \in Spline(\theta_1)} p(s) - k \times (1 - p(s)) \times (1 - p(s)) ds}{1}$$

Avec k un niveau d'exigence ou coefficient de pénalisation pouvant être un facteur prédéfini dans l'unité de traitement 4' ou déterminé dynamiquement par l'unité de traitement 4'. Le coefficient k permet de choisir un degré de pénalisation associé au passage dans une zone roulable de manière incertaine.

On peut aussi choisir de contraindre la longueur du chemin emprunté par le véhicule V. La trajectoire peut être alors calculée :

$$Trajectoire = \underset{\theta_1}{argmax} \frac{\int_{s \in Spline(\theta_1)} p(s) ds}{\int_{s \in Spline(\theta_1)} 1 ds}$$

De manière plus générale, l'homme du métier comprendra qu'il est possible de définir, à partir des images segmentées et des cartes de probabilité différentes paramétrisations de la trajectoire, par exemple des portions de droites projetées dans l'image de l'environnement.

Comme décrit ci-dessus, le calcul de la trajectoire peut se faire dans l'image puis projetée dans le référentiel de contrôle du véhicule V. Par exemple, cela peut être fait en appliquant à l'image des probabilités de chemin une homographie qui lie le plan de l'image issue du capteur 10', par exemple l'image caméra, à un plan de référence du véhicule V. On peut également projeter la carte de probabilité de présence de zone roulable dans le référentiel de contrôle et configurer l'unité de traitement 4' pour effectuer tous les calculs dans ce référentiel.

Il est possible de fusionner ou consolider les informations disponibles, acquises par le capteur 10' ou calculées par l'unité de traitement 4', pour obtenir un meilleur contrôle du véhicule V.

A un instant t, il est possible d'obtenir une connaissance du déplacement du véhicule V à partir d'au moins deux cartes de probabilités, par exemple entre une première carte de probabilité à l'instant t, et une deuxième carte de probabilité à l'instant t-1. De manière générale, on obtient la connaissance du déplacement du véhicule V à partir de M cartes probabilistes successives, c'est-à-dire de la première carte de probabilité à l'instant t, et des cartes de probabilités antérieures aux instants t-1, t-2... t-M+1.

La connaissance du déplacement du véhicule V peut être calculée par l'unité de traitement 4', par exemple par une technique d'odométrie, par des données de navigation (GPS),

ou par une centrale inertielle comprenant un système de mesure du mouvement et de cartographie de l'environnement (aussi appelée « inertie vision »).

A partir de la connaissance de déplacement du véhicule V, il est possible d'effectuer une consolidation temporelle d'au moins deux cartes de probabilité :

- 5 - comme expliqué ci-dessus, l'unité de traitement 4' peut projeter et recalculer les au moins deux cartes de probabilité (t , $t-1$) dans le repère de contrôle du véhicule V ;
- L'unité de traitement 4' calcule une carte de probabilité finale correspondant à la fusion des au moins deux cartes de probabilités des zones roulables, en appliquant une règle de consolidation parmi une des règles de consolidation suivantes :
 - 10 • Consolidation par fusion conservatrice, dans laquelle chaque élément p_{final} de la matrice finale est calculé selon : $p_{final}(\text{pixel}) = \min(p_{t-1}(\text{pixel}), p_t(\text{pixel}))$;
 - consolidation par fusion agressive, dans laquelle chaque élément p_{final} de la matrice finale est calculé selon : $p_{final}(\text{pixel}) = \max(p_{t-1}(\text{pixel}), p_t(\text{pixel}))$;
 - 15 • consolidation par fusion moyenne, dans laquelle chaque élément p_{final} de la matrice finale est calculé selon : $p_{final}(\text{pixel}) = \frac{p_{t-1}(\text{pixel}) + p_t(\text{pixel})}{2}$;
 - consolidation par fusion robuste, dans laquelle chaque élément p_{final} de la matrice finale est calculé selon les éléments p_{t-2}, p_{t-1}, p_t des cartes de probabilités aux instants $t-2$, $t-1$ et t respectivement, selon : $p_{final}(\text{pixel}) =$
 - 20 mediane ($p_{t-2}(\text{pixel}), p_{t-1}(\text{pixel}), p_t(\text{pixel})$) ;

Les consolidations précédentes peuvent être généralisées dans le cas où $M > 2$, en appliquant des formules semblables aux M cartes de probabilité. Il est aussi possible d'utiliser d'autres règles de consolidation utilisant les données des éléments p des M cartes de probabilités antérieures.

- 25 Il est également possible d'effectuer une consolidation comprenant une détection d'obstacle externe, dans le cas où les capteurs 10' comprennent un capteur d'obstacle, par exemple un capteur de télédétection par laser (LIDAR, en anglais «light detection and ranging »), un capteur par radar ou un capteur par ultrasons. Dans un tel exemple de réalisation, il est possible de configurer l'unité de traitement 4' pour que chaque élément p soit mis à zéro de
- 30 manière forcée dans le cas d'une détection d'obstacle. Ainsi, la probabilité de présence de zone roulable est automatiquement nulle quand il y a présence d'un obstacle dans la zone. Dans cet

exemple de réalisation, le niveau d'exigence ou coefficient de pénalisation k utilisé pour le calcul de trajectoire peut être prédéfini à une valeur élevée, typiquement supérieure à 100.

Selon l'exemple de la **figure 12**, la trajectoire calculée à la sous-étape C32 est définie pour que le véhicule V reste sur les zones considérées comme roulable de manière certaine, en passant par les zones ayant la plus grande probabilité d'être néanmoins roulable. Un exemple de consigne est également schématisé, dans lequel la vitesse de consigne diminue lorsque le véhicule V atteint le point P1, où il entre dans la zone dont la probabilité d'être roulable est plus faible, et la vitesse de consigne augmente lorsque le véhicule V atteint le point P2, où il entre dans une zone dont la probabilité d'être roulable est élevée. La zone entre P1 et P2 identifiée sur l'image de l'environnement du véhicule correspond ici à une portion de route boueuse par exemple.

Dans un exemple de réalisation, la commande en vitesse et/ou en trajectoire calculée à la sous-étape C32 peuvent être transmises à un calculateur C du véhicule V. Le calculateur C peut être de préférence capable de modifier la vitesse du véhicule V de sorte à atteindre la commande en trajectoire, par un système de type « motion planner » par exemple.

De préférence, le calculateur C peut comprendre un contrôleur de type « feedback control », capable de modifier la commande en vitesse et/ou en trajectoire du véhicule V en fonction des données issues de capteurs présents sur l'actionneur A, par exemple en cas de dérapage du véhicule V.

Les étapes C1 à C3 du procédé de contrôle du véhicule V pourront être réitérées de manière quasi continue, c'est-à-dire avec une fréquence élevée, afin de permettre un contrôle dynamique du véhicule V.

Le lecteur aura compris que de nombreuses modifications peuvent être apportées sans sortir matériellement des nouveaux enseignements et des avantages décrits ici. Par conséquent, toutes les modifications de ce type sont destinées à être incorporées à l'intérieur de la portée du procédé de contrôle présenté.

REVENDEICATIONS

1. Procédé d'identification (E) de zones roulables dans au moins une image d'un terrain, caractérisé en ce qu'il comprend l'implémentation des étapes suivantes dans une unité de traitement (3,4) :

E0 : construction d'une base de données d'apprentissage (B1) comprenant un ensemble d'images, ladite construction comprenant une étape de calcul d'un ensemble d'au moins une vérité terrain associé à l'ensemble d'images de ladite base de données d'apprentissage (B1), chaque vérité terrain étant définie par une carte probabiliste et est obtenue à partir d'au moins deux annotations d'une même image ;

E1 : apprentissage supervisé d'au moins un réseau de neurones (R2,R3), à partir d'une base de données d'apprentissage (B1) comprenant des images d'apprentissages de terrain ayant été annotées par au moins deux opérateurs en vue d'identifier des zones roulables,

E2 : segmentation au moyen de l'au moins un réseau de neurones (R2,R3), d'une image d'entrée (I1) pour obtenir une carte probabiliste (I3) de présence d'au moins une zone roulable dans ladite image d'entrée (I1), ladite segmentation comprenant une estimation d'une probabilité de présence d'une zone roulable dans l'image d'entrée (I1), de manière à obtenir une carte probabiliste (I3) définie par une matrice dont chaque élément correspond à une probabilité de présence d'au moins une zone roulable sur un pixel de l'image d'entrée (I1).

2. Procédé d'identification (E) selon la revendication précédente, dans lequel la base de données d'apprentissage (B1) comprend un ensemble d'images, associé à un ensemble de vérités terrain associé, chaque vérité terrain étant une matrice dont chaque élément p correspond à un pixel de l'image associée, chaque élément p étant calculé de la manière suivante :

$$p(pixel) = \sum_{i=1}^K \frac{annotation(pixel, i)}{K}$$

avec K le nombre d'opérateurs ayant annoté l'image, et annotation (pixel, i) = 1 si un opérateur i a considéré le pixel comme appartenant à une zone roulable, 0 sinon.

3. Procédé d'identification (E) selon l'une quelconque des revendications précédentes, dans lequel l'étape d'apprentissage (E1) comprend les étapes :

S1 : apprentissage supervisé d'un premier réseau de neurones (R1) à partir de la base de

données d'apprentissage (B1), de sorte que le premier réseau de neurones (R1) entraînés soit configuré pour segmenter une image entre zone roulable et zone non roulable ;

S2 : apprentissage supervisé d'un second réseau de neurones (R2) à partir de la base de données d'apprentissage et de l'au moins une vérité terrain, de sorte que le second réseau de neurones (R2) entraîné soit configuré pour segmenter l'image selon une carte probabiliste de présence de zone roulable.

4. Procédé de d'identification (E) de zones roulables selon la revendication précédente, dans lequel l'étape de segmentation (E2) de l'image d'entrée (I1) comprend les étapes de :

S10 : calcul d'une segmentation (I2) de l'image d'entrée (I1), par le premier réseau de neurones (R1) entraîné,

S20 : calcul d'une carte probabiliste (I3) de présence de zones roulables de l'image d'entrée (I1) par le second réseau de neurones (R2) entraîné, à partir de l'image d'entrée (I1) et de la segmentation (I2) issue du premier réseau de neurones (R1).

5. Procédé de contrôle d'un véhicule terrestre (V), comprenant les étapes suivantes:

C1 : acquisition d'images par un capteur (10') situé sur le véhicule (V) ;

C2 : traitement des images acquises par une unité de traitement (4'), par un procédé de d'identification (E2) selon l'une quelconque des revendications 1 à 4, afin d'obtenir au moins une carte probabiliste de présence de zones roulables ;

C3 : calcul d'une consigne comprenant les sous-étapes de :

- C31 : calcul d'une probabilité de présence de zone roulable par l'unité de traitement (4') à partir de la carte probabiliste de présence de zones roulables ;

- C32 : calcul d'une trajectoire et/ou d'une consigne en vitesse du véhicule (V) par l'unité de traitement (4') à partir de la probabilité de présence de zone roulable.

6. Procédé de contrôle d'un véhicule (V) selon la revendication précédente, dans lequel le capteur (10') acquiert une vidéo, et dans lequel l'identification (E2) au cours de l'étape de traitement des images acquises (C2) est effectuée de manière séquentielle.

7. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 ou 6, dans lequel l'étape de calcul de la consigne (C3) dépend d'une valeur seuil prédéterminée, la valeur

seuil étant un niveau de risque à prendre correspondant à une probabilité qu'un pixel d'une image acquise appartienne à une zone roulable.

8. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 à 7, dans lequel la consigne en vitesse augmente lorsque la probabilité de présence de zone roulable augmente.

9. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 à 8, dans lequel la consigne en trajectoire est calculée à partir de la carte probabiliste de zones roulables de sorte à maximiser la probabilité de présence de zone roulable.

10. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 à 8, dans lequel la consigne en trajectoire est calculée à partir de la carte probabiliste de zones roulables de sorte à maximiser la probabilité de présence de zone roulable et à minimiser selon un niveau d'exigence k , prédéfini ou choisi dynamiquement, les zones au caractère roulable faiblement probables.

11. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 à 10, dans lequel la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones roulables finale issue de la fusion dite conservative de M cartes probabilistes successives, dont chaque élément p_{final} est calculé de la manière suivante :

$$p_{final}(pixel) = \min (p_t(pixel), p_{t-1}(pixel) \dots p_{t-M+1}(pixel))$$

12. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 à 10, dans lequel la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones roulables finale issue de la fusion dite robuste de M cartes probabilistes successives, dont chaque élément p_{final} est calculé de la manière suivante :

$$p_{final}(pixel) = \text{mediane} (p_t(pixel), p_{t-1}(pixel) \dots p_{t-M+1}(pixel))$$

13. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 à 10, dans lequel la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de

zones roulables finale issue de la fusion dite agressive de M cartes probabilistes successives, dont chaque élément p_{final} est calculé de la manière suivante :

$$p_{final}(pixel) = \max (p_t(pixel), p_{t-1}(pixel), \dots p_{t-M+1}(pixel))$$

- 5 14. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 à 10, dans lequel la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones roulables finale issue de la fusion dite moyenne de M cartes probabilistes successives, dont chaque élément p_{final} est calculé de la manière suivante :

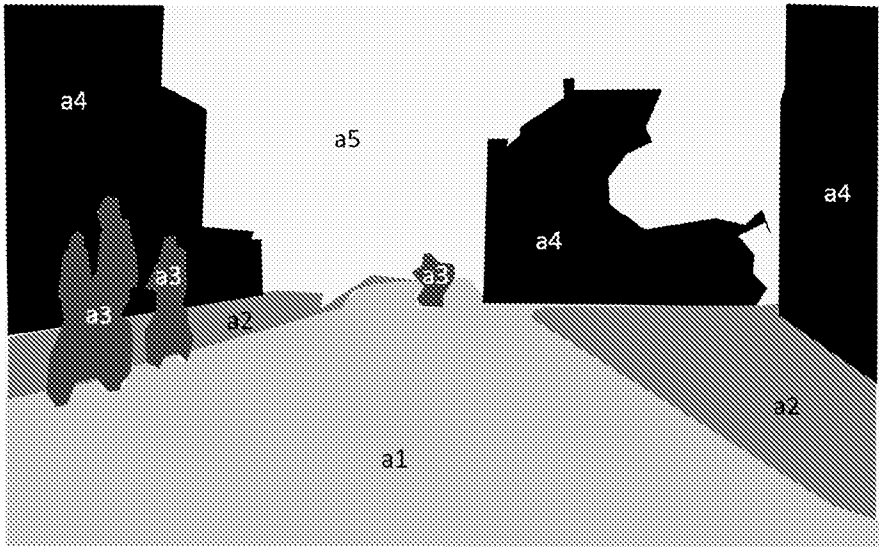
$$p_{final}(pixel) = \frac{p_t(pixel) + p_{t-1}(pixel) + \dots + p_{t-M+1}(pixel)}{M}$$

10

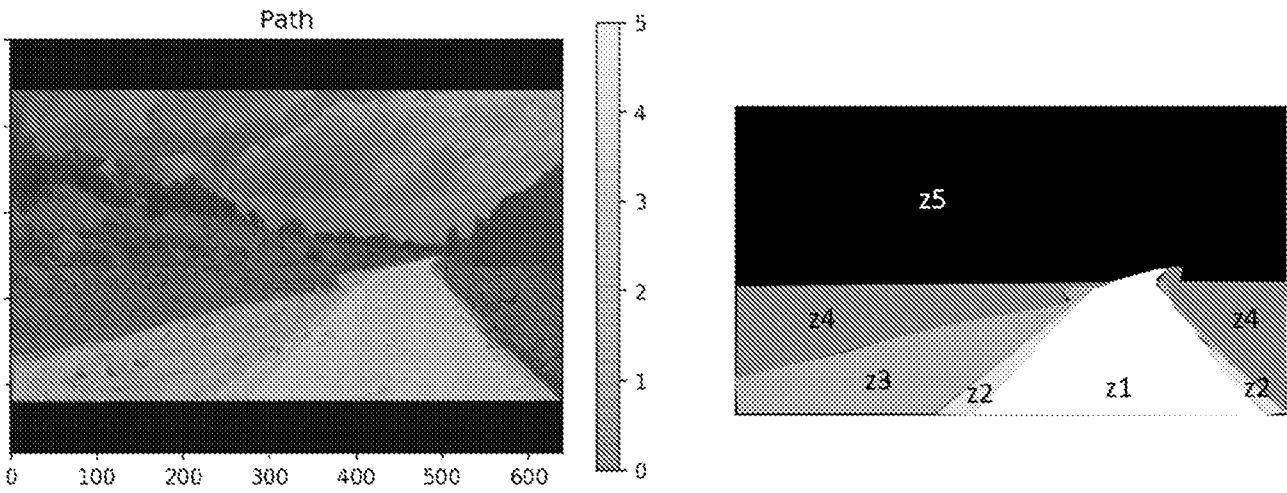
15. Procédé de contrôle d'un véhicule (V) selon l'une quelconque des revendications 5 à 14, dans lequel la consigne de trajectoire est calculée à partir de la carte de probabilité de présence de zones roulables finale issue de d'une fusion de M cartes probabilistes successives, et d'une détection d'obstacles réalisée par un capteur d'obstacle.

15

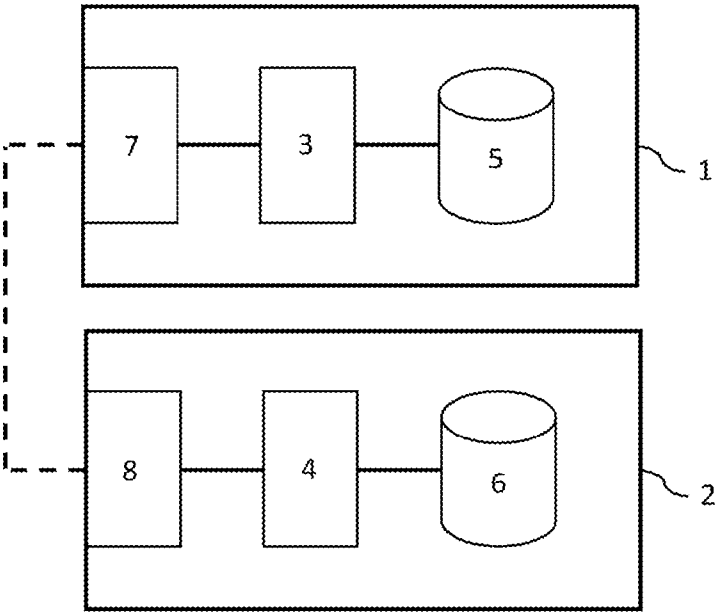
[Fig. 1]



[Fig. 2]

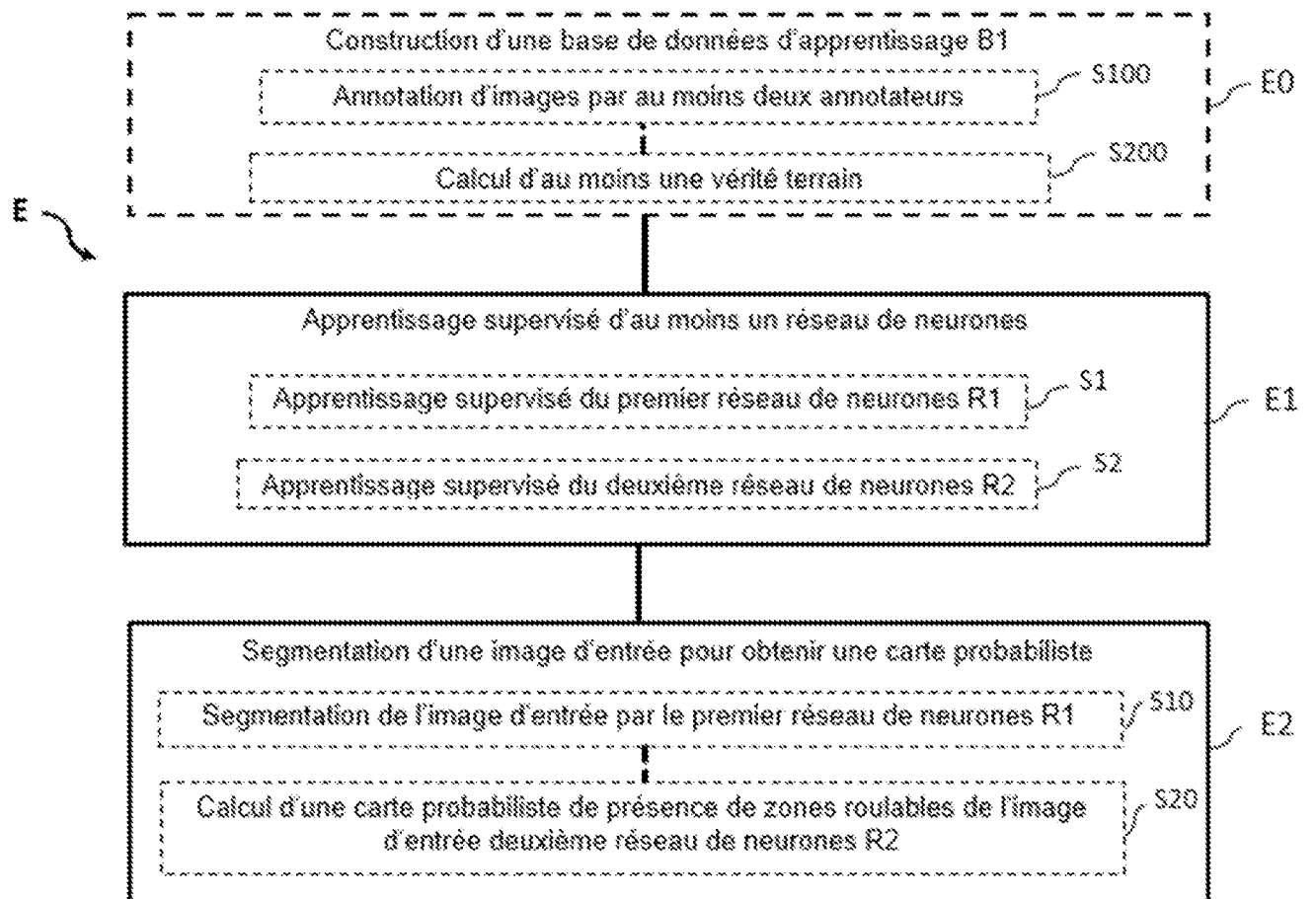


[Fig. 3]



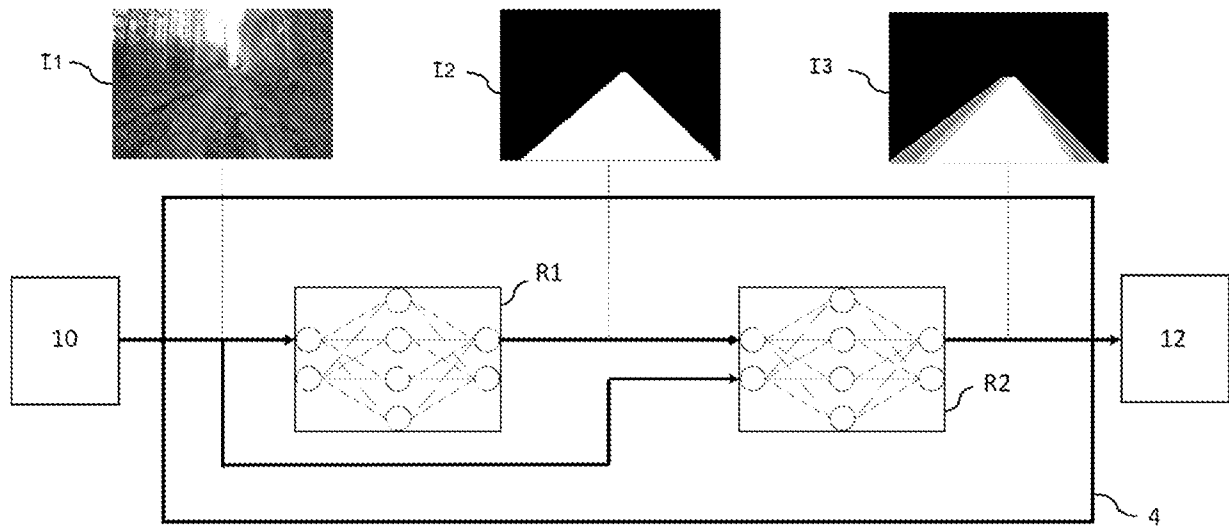
3/8

[Fig. 4]

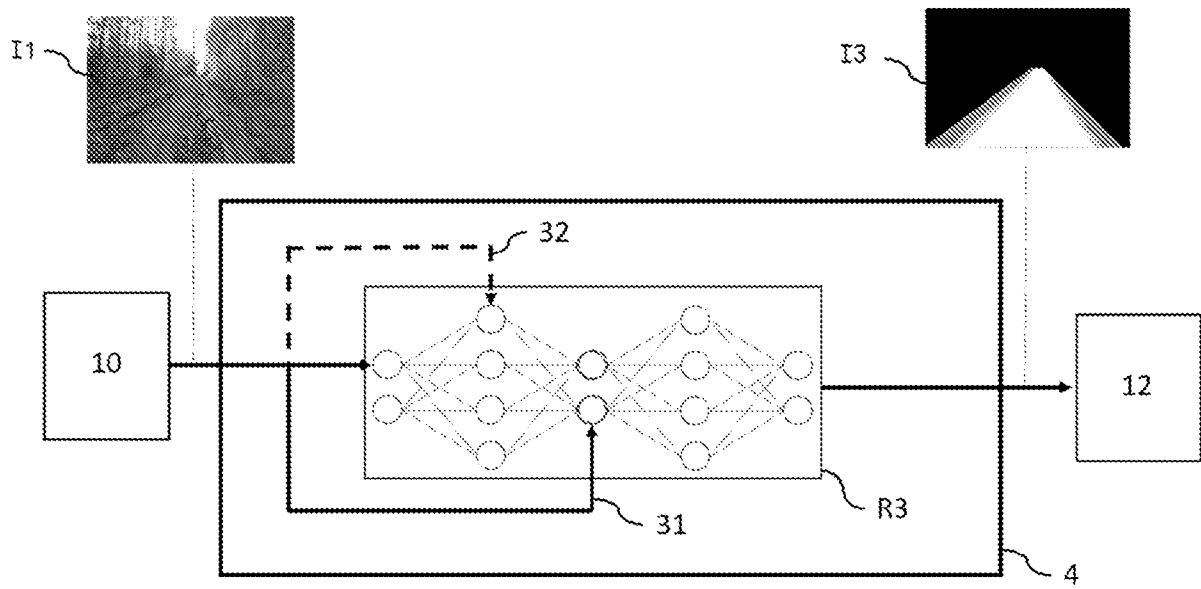


4/8

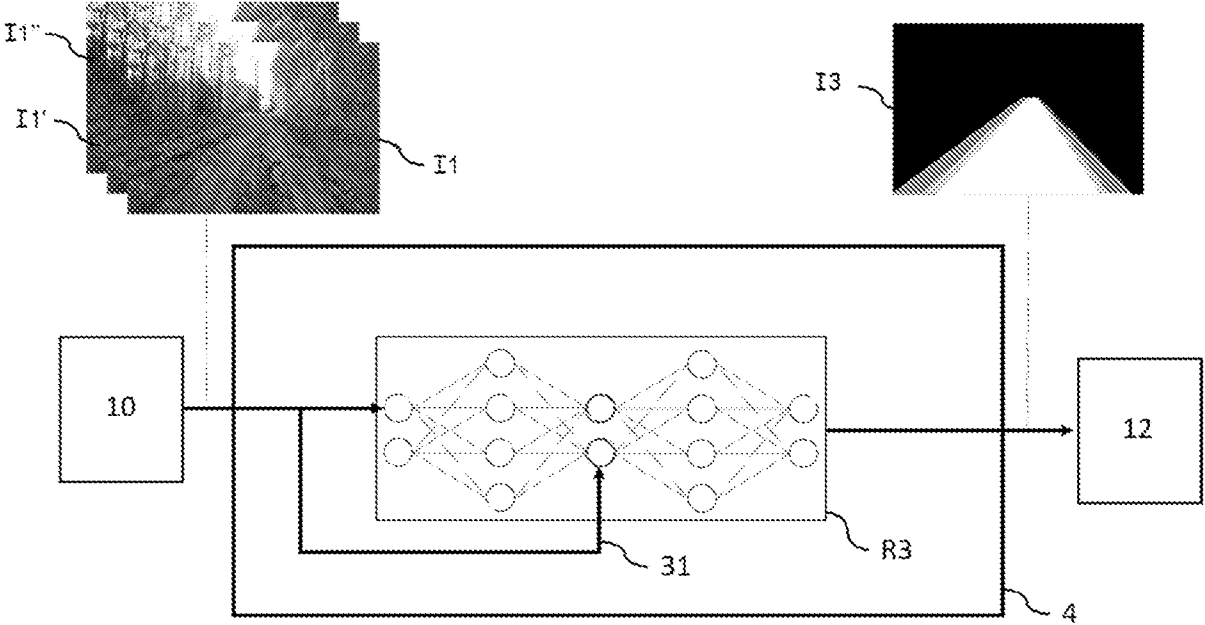
[Fig. 5]



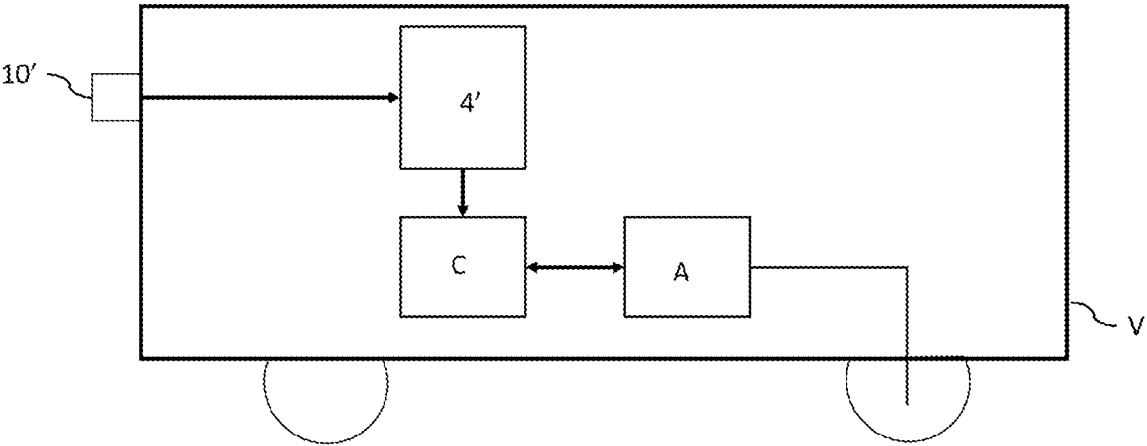
[Fig. 6]



[Fig. 7]

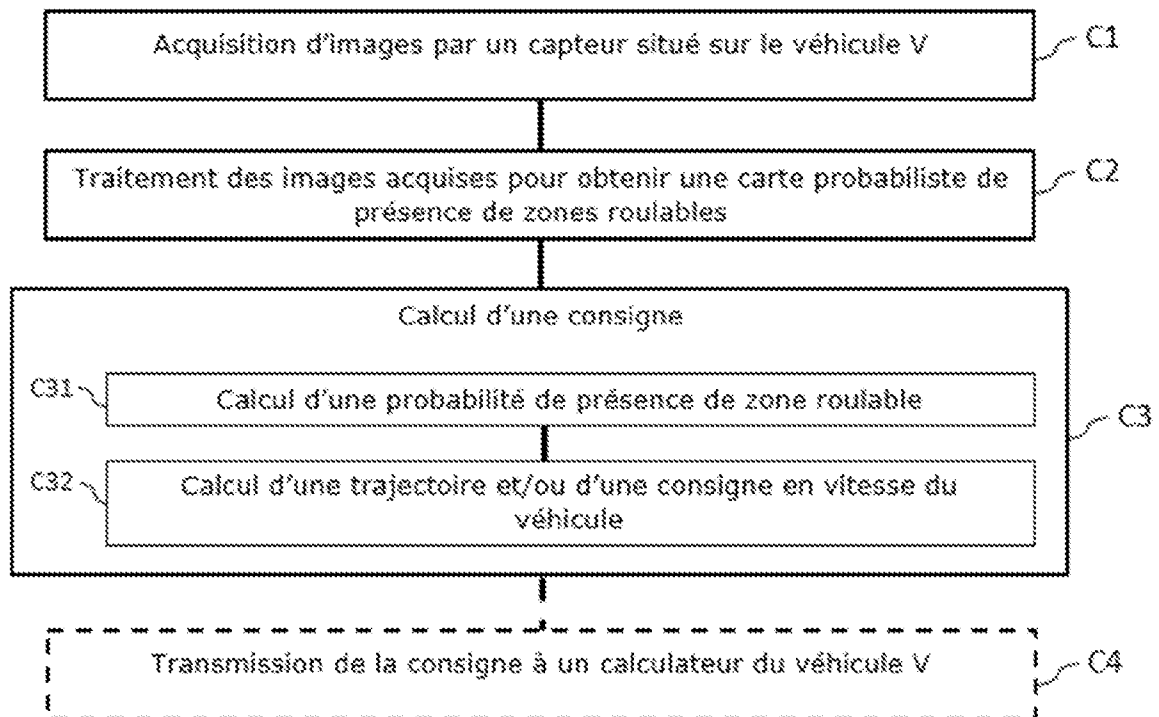


[Fig. 8]



6/8

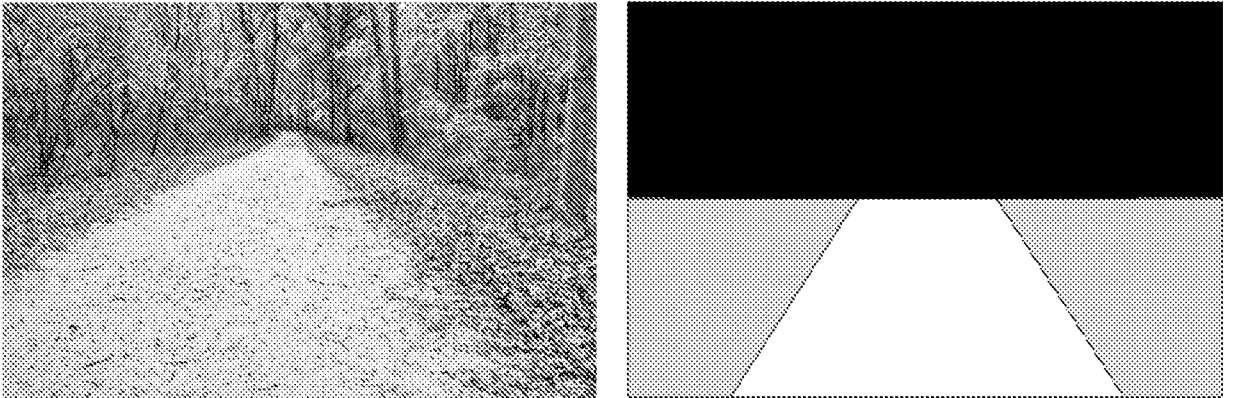
[Fig. 9]



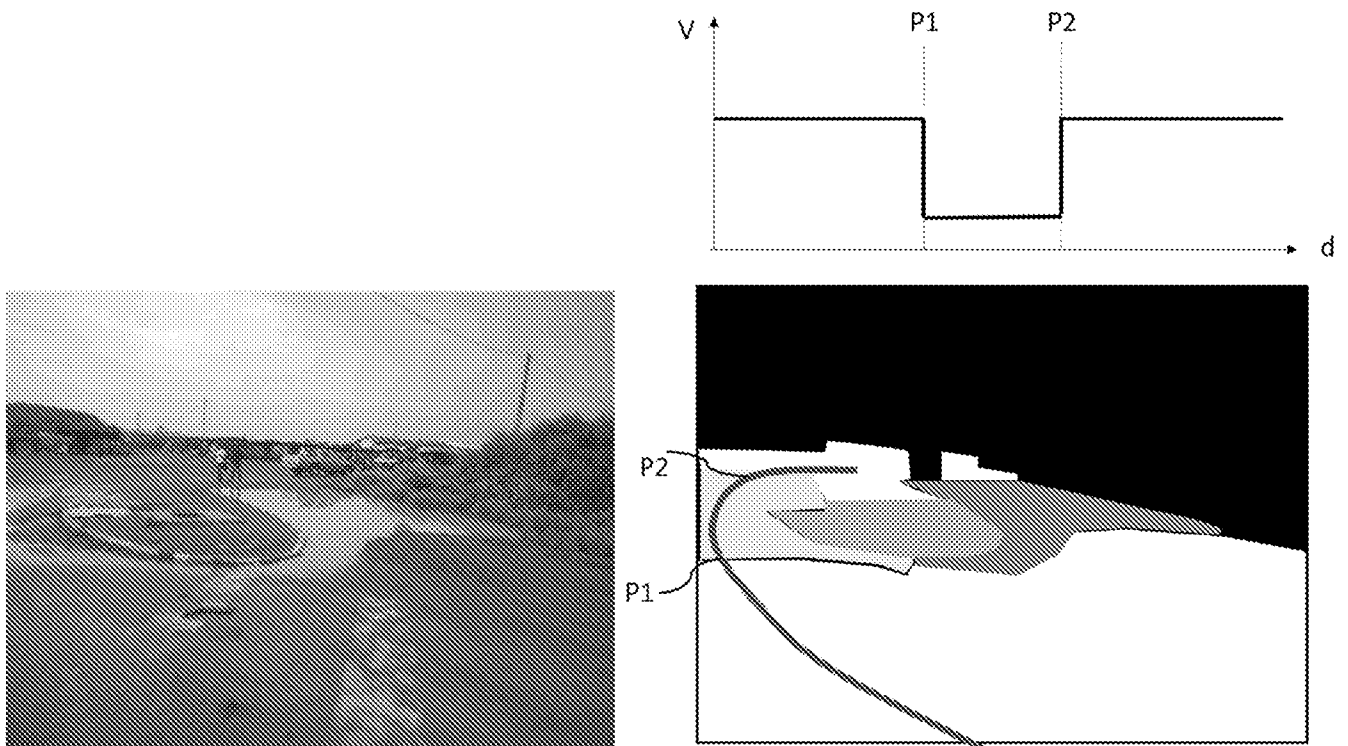
[Fig. 10]



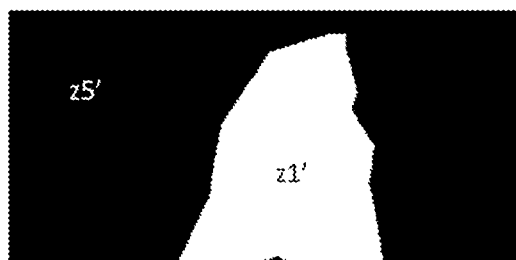
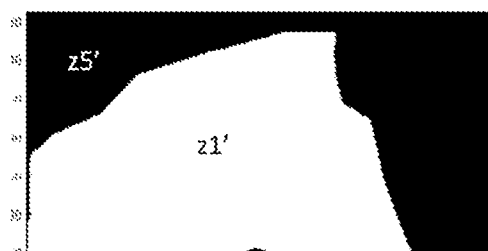
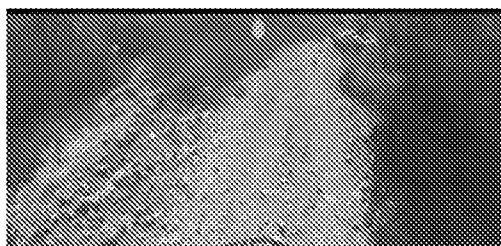
[Fig. 11]



[Fig. 12]



[Fig. 13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR2020/050268

A. CLASSIFICATION OF SUBJECT MATTER**G06K 9/00**(2006.01)i; **G06K 9/46**(2006.01)i; **G06K 9/62**(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K; G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, COMPENDEX, INSPEC, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KIM JIMAN ET AL. "End-To-End Ego Lane Estimation Based on Sequential Transfer Learning for Self-Driving Cars" <i>2017 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW)</i> , IEEE, 21 July 2017 (2017-07-21), pages 1194-1202 DOI: 10.1109/CVPRW.2017.158 XP033145900 the whole document	1-15
X	US 2018157972 A1 (HU RUI [US] ET AL) 07 June 2018 (2018-06-07) figure 2 figure 4 figure 5	1-15
A	US 10176388 B1 (GHAFARIANZADEH MAHSA [US] ET AL) 08 January 2019 (2019-01-08) figure 1B figure 2	1-15

☒ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

04 May 2020

Date of mailing of the international search report

14 May 2020

Name and mailing address of the ISA/EP

European Patent Office
p.b. 5818, Patentlaan 2, 2280 HV Rijswijk
Netherlands

Telephone No. (+31-70)340-2040

Facsimile No. (+31-70)340-3016

Authorized officer

Martinière, Anthony

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/FR2020/050268

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NELSON FERNANDEZ. "Two-stream convolutional networks for end-to-end learning of self-driving cars" <i>ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853</i> , 13 November 2018 (2018-11-13), XP081050083 the whole document	1-15
A	MOHAMMAD HAJIZADEH SAFFAR ET AL. "Semantic Video Segmentation: A Review on Recent Approaches" <i>ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853</i> , 16 June 2018 (2018-06-16), XP081238558 paragraph [2.1.6.]	1-15
A	YUNPENG PAN ET AL. "Agile Off-Road Autonomous Driving Using End-to-End Deep Imitation Learning" <i>ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853</i> , 21 September 2017 (2017-09-21), XP080822298 the whole document	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/FR2020/050268

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)		
US	2018157972	A1	07 June 2018	NONE			
US	10176388	B1	08 January 2019	US 10176388 B1	08 January 2019		
				US 2019138826 A1	09 May 2019		

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/FR2020/050268

A. CLASSEMENT DE L'OBJET DE LA DEMANDE INV. G06K9/00 G06K9/46 G06K9/62 ADD.		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE Documentation minimale consultée (système de classification suivi des symboles de classement) G06K G06N		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés) EPO-Internal, COMPENDEX, INSPEC, WPI Data		
C. DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	KIM JIMAN ET AL: "End-To-End Ego Lane Estimation Based on Sequential Transfer Learning for Self-Driving Cars", 2017 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), IEEE, 21 juillet 2017 (2017-07-21), pages 1194-1202, XP033145900, DOI: 10.1109/CVPRW.2017.158 le document en entier	1-15
X	----- US 2018/157972 A1 (HU RUI [US] ET AL) 7 juin 2018 (2018-06-07) figure 2 figure 4 figure 5 ----- -/-	1-15
☒ Voir la suite du cadre C pour la fin de la liste des documents ☒ Les documents de familles de brevets sont indiqués en annexe		
* Catégories spéciales de documents cités: "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent "E" document antérieur, mais publié à la date de dépôt international ou après cette date "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée) "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée "T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier "&" document qui fait partie de la même famille de brevets		
Date à laquelle la recherche internationale a été effectivement achevée 4 mai 2020		Date d'expédition du présent rapport de recherche internationale 14/05/2020
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Fonctionnaire autorisé Martinière, Anthony

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	US 10 176 388 B1 (GHAFARIANZADEH MAHSA [US] ET AL) 8 janvier 2019 (2019-01-08) figure 1B figure 2 -----	1-15
A	NELSON FERNANDEZ: "Two-stream convolutional networks for end-to-end learning of self-driving cars", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 13 novembre 2018 (2018-11-13), XP081050083, le document en entier -----	1-15
A	MOHAMMAD HAJIZADEH SAFFAR ET AL: "Semantic Video Segmentation: A Review on Recent Approaches", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 16 juin 2018 (2018-06-16), XP081238558, alinéa [2.1.6.] -----	1-15
A	YUNPENG PAN ET AL: "Agile Off-Road Autonomous Driving Using End-to-End Deep Imitation Learning", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 21 septembre 2017 (2017-09-21), XP080822298, le document en entier -----	1-15

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°

PCT/FR2020/050268

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 2018157972 A1	07-06-2018	AUCUN	
US 10176388 B1	08-01-2019	US 10176388 B1	08-01-2019
		US 2019138826 A1	09-05-2019