L'intelligence artificielle souffre d'illusions d'optique

VÉHICULES AUTONOMES La technique d'apprentissage de la machine doit s'affiner

Un véhicule autonome semblable à celui de Han-sur-Lesse fera la navette vers la butte du lion à Waterloo. © D. DUCHESNES. L.TH.

▶ Alors que l'avènement des véhicules 100 % autonomes se profile, des expériences scientifiques révèlent une faiblesse de l'intelligence artificielle qui devrait les piloter.
▶ Un risque de confusion des panneaux de signalisation routière n'est pas exclu.
▶ Ce jeudi, une navette autonome est lancée à Waterloo.

Deux expériences de véhicules autonomes ont cours sur des voies publiques wallonnes. La première transporte des touristes sur 500 m à Han-sur-Lesse ; la deuxième en emmène désormais du parking au pied de la butte du lion à Waterloo. Un voyage d'un kilomètre à bord d'un véhicule piloté par une intelligence artificielle (IA). Alors que le bureau d'études McKinsey prévoit qu'à l'horizon 2030, jusqu'à 15 % des véhicules neufs vendus seront entièrement autonomes ; on peut s'interroger sur la fiabilité des IA qui joueront aux pilotes. Les exemples d'illusions d'optique fleurissent en effet dans la littérature scientifique. Sans conteste, la perception de l'environnement est l'un des défis majeurs des véhicules autonomes.

Afin de se passer de pilote humain, les industriels et chercheurs explorent le deep learning. Il s'agit d'une méthode d'apprentissage automatique basée sur la conception d'un programme informatique dont l'architecture mime un réseau de neurones. C'est-à-dire un ensemble de cellules de calculs reliées entre elles et empilées en couches successives. Ensuite, il convient de l'entraîner en l'alimentant de millions de clichés de circulation routière.

Les hallucinations de l'IA

Si le discours médiatique dominant clame que la technologie est au point, le monde scientifique met en garde contre les faiblesses de l'IA. Il peut en effet lui arriver de souffrir d'hallucinations. Au point de confondre des panneaux de signalisation routière.

Lors d'une expérimentation menée aux Etas-Unis, des chercheurs ont recouvert un panneau « stop » d'un filtre algorithmique transparent ayant la caractéristique d'en modifier la valeur de quelques pixels. Alors que l'humain continuait d'y voir un panneau « stop » classique, le robot aidé de sa caméra l'identifiait désormais comme un panneau de… priorité.

Concrètement, les risques de sécurité associés à un tel piratage visuel sont-ils importants ? *« Dans le cas spécifique des véhicules autonomes, je n'en suis pas sûr. Comme d'autres chercheurs l'ont souligné, il est beaucoup plus facile de renverser physiquement un panneau « stop » que de générer un autocollant qui peut, ou non, tromper l'algorithme de vision. Cela est néanmoins un sujet de débat parmi les chercheurs,* explique Dr Jonathan Peck, spécialiste des mathématiques des réseaux de neurones à l'Université de Gand. *Indépendamment du fait qu'il présente ou non un risque de sécurité réel, cet exemple souligne que les modèles d'apprentissage automatique de deep learning que l'on applique actuellement à l'intelligence artificielle ne se généralisent pas aussi bien que nous le souhaiterions. »*

Test de piratage

Dans la littérature scientifique, il y a plusieurs autres exemples d'hallucinations de l'IA. L'un d'eux a été réalisé par des chercheurs de Google. Ils ont subtilement modifié la valeur de quelques pixels sur une photo de panda. Résultat ? Alors que l'œil humain continue d'y voir le même ursidé, l'algorithme, quant à lui, identifie un singe gibbon avec un degré de confiance de 99,3 % !

Si cela peut prêter à sourire, ça devient nettement moins drôle quand on se rappelle que ces algorithmes sont voués à envahir notre quotidien. Qu'adviendra-t-il quand ils prendront un feu rouge pour un feu vert, une cellule cancéreuse pour un organe sain, une école pour une cible militaire, le visage d'un quidam pour celui d'un terroriste ? *« De nombreuses méthodes ont été développées pour tenter d'éviter ces illusions d'optique, mais leurs performances sont loin d'être satisfaisantes pour le moment »*, déplore le Dr Peck.

A noter toutefois qu'aucun véhicule autonome actuellement sur le marché n'est totalement basé sur l'apprentissage par deep learning. *« Ils sont l'intégration de plusieurs sources de données provenant du GPS, de caméras et de capteurs embarqués*, explique Pr Gilles Louppe, spécialiste des réseaux de neurones à l'ULiège. *Actuellement, la méthode de deep learning n'est utilisée que pour analyser en direct le flux de caméras embarquées. A partir de celui-ci, les portions dans l'image qui correspondent à la route, un trottoir, des personnes ou d'autres véhicules sont détectées. Ensuite, un programme informatique classique intègre tous ces morceaux d'informations pour prendre une décision. »* ■

LAETITIA THEUNIS

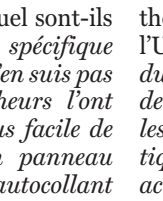
ENTRETIEN

« Pas encore de solutions 100 % fiables »

Le Pr Gilles Louppe est spécialiste des réseaux de neurones à l'ULiège.

Les risques associés aux illusions d'optique de l'IA sont-ils vraiment importants ?

A l'heure actuelle, ce type d'illusions, aussi appelées « attaques adversaires », est un réel problème pour tout système basé sur un réseau de neurones. Les recherches concernant les « défenses » sont nombreuses et la communauté étudie ce problème très sérieusement. Il existe plusieurs solutions plus ou moins efficaces, mais on assiste pour l'instant davantage à un jeu du chat et de la souris. Dès qu'une nouvelle méthode de défense est proposée, de nouvelles attaques sont alors observées et vice-versa. Il n'existe pas encore de solution permettant d'éviter ces attaques de façon 100 % fiable.



Y a-t-il d'autres limites à l'usage de l'IA ?

Quand on construit un réseau de neurones, il y a toute une étape d'entraînement qui consiste à lui présenter beaucoup d'exemples et à lui dire telle image, c'est une voiture, telle autre, c'est une banane etc. Quand on passe ensuite en mode prédictif, on déploie le système dans la réalité : si les images présentées au réseau de neurones sont un peu différentes des images utilisées lors de l'entraînement, il n'y a pas encore beaucoup de garanties théoriques. Quand les différences sont trop fortes, les performances peuvent baisser assez vite. C'est une des grosses limitations. Aussi, il peut y avoir des biais dans les bases de données d'apprentissage. Imaginons que dans les images montrées, tous les chiens sont sur du gazon tandis que tous les chats sont sur un divan. Le réseau de neurones va apprendre à différencier un chien d'un chat, mais il pourrait le faire sur base de la présence ou non de pelouse dans l'image…

Les défis de l'IA risquent-ils de freiner l'avènement, par exemple, des véhicules autonomes ?

Il y a eu énormément d'avancées ces 5 dernières années. Je pense que ça va continuer de la sorte. C'est normal qu'il y ait des difficultés techniques sur le parcours. Elles ne sont pas, selon moi, insurmontables.

L.TH.

CONDUITE

Catégories d'autonomie

La liste qui suit montre que l'autonomie des véhicules se situe actuellement au niveau 3.

0 = Conduite (C) manuelle

1 = Conduite assistée

Le conducteur peut être dispensé de la commande par les pédales. A cela s'ajoutent le freinage ABS et le régulateur de vitesse

2 = Conduite partiellement automatisée

Grâce au centrage automatique sur la chaussée et à l'assistance à la trajectoire, le conducteur peut ponctuellement lever les mains du volant

3 = Conduite conditionnellement automatisée

Les trajectoires et les manœuvres sont automatisées et disposent d'un système d'alerte. Le conducteur peut ponctuellement ne pas regarder la route. Les meilleurs véhicules autonomes actuellement sur le marché se classent dans cette catégorie.

4 = Conduite hautement automatisée

Par la conduite et la surveillance de l'environnement automatisées, le conducteur n'aura plus à intervenir sur la route

5 = Conduite totalement automatisée

Ce véhicule 100 % autonome sera donc piloté exclusivement par une IA. Il pourra circuler sur tout type de voies pendant que l'humain embarqué dormira sur ses deux oreilles.

WATERLOO

Un trajet d'un kilomètre

Ce jeudi, une navette autonome emmènera les touristes du parking jusqu'au pied de la butte du lion à Waterloo. Ce voyage d'un kilomètre se fera à bord d'un véhicule piloté exclusivement par une intelligence artificielle (IA).

Ce sera la deuxième expérience de ce type sur la voie publique en terre wallonne. En juin dernier, une navette semblable avait été présentée par le ministre de la Mobilité François Bellot et le premier ministre Charles Michel à Han-sur-Lesse.

Cette fois, le parcours sera *« plus difficile et cinq fois plus long qu'à Han-sur-Lesse »*, précise le communiqué de presse.

François Bellot avait alors annoncé que les premières voitures autonomes allaient faire leur apparition sur les routes belges dès cette deuxième moitié d'année. *« Il y a déjà une quinzaine de tests qui sont prévus en 2018 et 2019 en Belgique »*, expliquait-il.

L. TH.