

# NOTE de VEILLE



MARS 2019

Tous les mois, retrouvez à travers les notes de veille, un diptyque Tendances globales / Actualités locales sur l'innovation et les mondes numériques décryptés par Unitec

## IA : LA RECONNAISSANCE D'IMAGES PREND DE L'AMPLEUR

En ce début d'année 2019, l'OMPI (Organisation mondiale de la propriété intellectuelle) a [publié un rapport](#) sur le dépôt de brevets à travers le monde portant sur les technologies d'intelligence artificielle. Selon lui, nous assistons depuis le début des années 2000 et encore plus depuis 2013 à un véritable déferlement de dépôts de brevets. Pour preuve, depuis les années 1950, près de 340.000 brevets portant sur des technologies liées à l'intelligence artificielle ont été déposés dont plus de la moitié depuis 2013, avec un passage clair de la recherche théorique à l'industrialisation de produits et de services faisant appel à l'intelligence artificielle (8 publications pour 1 brevet en 2010 contre 3 pour 1 en 2016). « *L'intelligence artificielle (IA) correspond à un ensemble de technologies qui permet de simuler l'intelligence et d'accomplir automatiquement des tâches de perception, de compréhension et de prise de décision. Ces techniques font particulièrement appel à l'utilisation de l'informatique, de l'électronique, des mathématiques (notamment statistiques), des neurosciences et des sciences cognitives.*<sup>1</sup> » Selon l'OMPI, 49% des brevets déposés dans le cadre de l'intelligence artificielle concerne la vision par ordinateur, ce qui en fait l'application la plus populaire avec quelques 21000 brevets déposés pour

l'année 2016<sup>2</sup> loin devant le traitement du langage naturel (14%) et traitement de la voix (12%). Le succès de la vision par ordinateur est notamment dû au spectre très large de marchés et de solutions qu'elle adresse : l'imagerie médicale, l'industrie, la réalité augmentée, la télédétection, la robotique, etc. A l'occasion du Forum NAIA qui se tiendra le 19 mars 2019 à Bordeaux, nous souhaitons faire le point sur une technologie qui fait couler beaucoup d'encre.

### 1 - Qu'est-ce que la vision par ordinateur ?

La vision par ordinateur peut être définie comme la science des machines, des robots, des systèmes informatiques et de l'intelligence artificielle qui analysent des images et des vidéos, reconnaissent des objets et agissent en conséquence. Il s'agit donc d'une branche de l'intelligence artificielle dont l'objectif est d'analyser des informations contenues dans des images ou des vidéos.

Les applications de la vision par ordinateur peuvent être divisées en grands domaines technologiques : la reconnaissance des caractères, la détection des visages (par la segmentation d'images/vidéo et la biométrie), la détection et le [suivi d'objets](#), la

détection d'événements (ou compréhension de scènes) et enfin la navigation. Au niveau des brevets, la détection des visages (+31%/an) et la compréhension de scènes (+28%/an) sont les deux domaines les plus actifs. Quel que soit le domaine, cette technique n'a pas pour but de reproduire la vision humaine mais cherche à construire des modèles algorithmiques puissants (des patterns) permettant de décrire précisément les propriétés d'images dans le cadre de cas d'usage bien précis : classification d'images (de bas niveau [détection de contours, suppression des bruits] et de haut niveau [interprétation d'images, reconstitution 2D/3D]), reconnaissance faciale, détection d'objets (fixes ou en mouvement), reconnaissance de caractères, etc. L'analyse d'image est une tâche complexe pour une machine et le recours aux techniques relatives au Machine Learning – ML ([Deep Learning](#), réseaux de neurones – dont les [réseaux neuronaux convolutifs](#), [apprentissage par renforcement](#), [apprentissage supervisé/non supervisé](#), etc.) permet des avancées majeures à des niveaux de complexité jamais atteints. De nombreux centres de recherche et des entreprises travaillent d'arrache-pied à améliorer l'efficacité des algorithmes (à l'Université de Bordeaux, le centre de recherche du LaBRI possède [plusieurs](#)

<sup>1</sup> Définition issue du [Rapport Pipame](#), février 2019. Le rapport montre très bien, chiffres à l'appui, que l'accélération récente est due, entre autres, à l'augmentation de la puissance de calcul des machines et à la baisse des coûts de stockage des données, elles-mêmes de plus en plus nombreuses. Cette note de veille ne développera pas cet aspect de l'intelligence artificielle.

<sup>2</sup> Les données de l'OMPI s'arrêtent à 2016 pour tenir compte du délai entre la demande de dépôt et la publication officielle 18 mois plus tard.



avec le soutien de



équipes qui travaillent spécifiquement sur le sujet, tandis que l'INRIA a lancé le projet Pimiento pour améliorer le suivi des cancers grâce à l'imagerie médicale), et les grands acteurs mondiaux proposent des API sur étagères: Amazon propose Rekognition, Microsoft Azure, Apple Vision, IBM Watson tandis que Google offre Cloud Vision et Intel OpenVino.

## 2 - Pour quels usages concrets ?

L'intérêt pour les entreprises est de faire appel à des solutions clé en main, puissantes et éprouvées à des coûts d'accès raisonnables. A titre d'exemple, la société Butterfleye a utilisé l'API Rekognition d'Amazon pour doter sa caméra de sécurité intelligente (Nero1) d'une fonction de reconnaissance faciale. Cloudinary, solution BtoB de stockage et de gestion de banques d'images, s'est appuyé sur l'API Face d'Azure Microsoft pour offrir à ses clients de nouvelles fonctionnalités comme l'Advanced Facial Attributes Detection qui permet d'automatiser et de traiter très rapidement la détection d'attributs physiques (le genre, l'âge, l'expression de sentiment) sur des photos ajoutant ainsi une couche de services supplémentaires aux gestionnaires d'images pour catégoriser et classer ces dernières. Ces deux premiers exemples montrent que la vision par ordinateur propose déjà des solutions opérationnelles<sup>3</sup>. La robotique s'est emparée de cette problématique de la vision par ordinateur. En dotant les robots de capacités visuelles performantes, l'idée est de les rendre autonomes : à titre d'exemple, une équipe du MIT a réussi à doter un robot de capacités visuelles (couplées au ML) et de capacités physiques (le toucher, la préhension, la capacité à aligner des pièces) lui permettant de jouer au jeu de Jenga. Si cette avancée peut

sembler futile, elle ouvre cependant des perspectives dans l'industrie manufacturière où les robots seront capables, sans programmation préalable, de reconnaître, de saisir et d'assembler parfaitement des pièces quelle que soit leur taille avec plus de précision et de rapidité qu'un travailleur humain. L'équipe AUCTUS de l'INRIA Bordeaux travaille en collaboration avec l'entreprise Aerospline à la captation et la modélisation de l'activité humaine afin d'améliorer les cobots industriels et par-là le confort de l'ouvrier. Autre marché en pleine explosion, celui de la voiture autonome. Afin d'être autonome<sup>4</sup>, la voiture doit être capable de percevoir parfaitement son environnement afin d'anticiper toute modification brusque des conditions dans lesquelles elle évolue. La vision par ordinateur pour la conduite autonome repose alors sur des images provenant de capteurs, le traitement d'images et l'apprentissage en profondeur pour transformer les données en actions appropriées. La voiture doit être capable de comprendre la scène qui se déroule devant « ses yeux » et identifier tout ce qui l'entoure : piétons, mobilier urbain, infrastructure et autres véhicules. La technologie LiDAR<sup>5</sup> et plus particulièrement celle développée par Luminar, adoptée par Volvo, permet ceci.

La médecine et plus particulièrement le diagnostic de maladie est un domaine dans lequel la vision par ordinateur est très présente : l'analyse d'images permet d'identifier des maladies rares caractérisées par des traits particuliers du visage et peut ainsi permettre de détecter un potentiel cancer du pancréas par la caméra d'un smartphone ou de déterminer avec une très bonne fiabilité (de l'ordre de 95%) la présence d'un mélanome. Ou encore de mesurer la pression artérielle en temps réel grâce à

un casque HoloLens. Les images médicales constituent un enjeu particulièrement important dans le champs de l'analyse par vision artificielle : le Bordelais Deski, accompagné par Unitec, associe imagerie médicale et deep learning dans le cadre de l'étude des maladies cardiaques et cardio-vasculaires. Nvidia de son côté continue d'entraîner ses algorithmes d'intelligence artificielle en créant, grâce à la technique des réseaux antagonistes génératifs<sup>6</sup> (GANs), de fausses IRM de cerveaux alimentant alors la base d'images utilisée par le réseau de neurones pour identifier les anomalies et les maladies.

Si les cas d'usages sont nombreux et s'inscrivent dans un véritable progrès pour la société, il n'en reste pas moins que certaines initiatives posent questions : la Chine généralise la reconnaissance faciale dans l'espace public ou à l'école, la Ville de Nice va tester (auprès de 1000 volontaires) un système de reconnaissance faciale lors du Carnaval, notamment pour réduire les risques d'attentat. Mais la manipulation des images et des vidéos rendu possible par l'IA peut également conduire à renforcer la propagation des fausses nouvelles en trompant le spectateur, devenant ainsi un véritable sujet de société.

**Auteur : Alexandre Bertin**  
**Responsable Veille et Prospective**

<sup>3</sup> Même si la recherche continue à améliorer l'efficacité des algorithmes dans la robotique, le cinéma, la santé entre autres.

<sup>4</sup> On parle ici des quatre 3.4 et 5.

<sup>5</sup> Pour Light Detection And Ranging systems. Le lidar a de nombreuses autres applications dans des domaines variés comme en topographie, sismologie, météorologie, agronomie, archéologie, régulation du trafic aérien, défense, etc...

<sup>6</sup> Les GAN font partie, avec l'apprentissage par renforcement et l'apprentissage par transfert des techniques qui montent en matière de Deep Learning.

Unitec accompagne les entreprises du secteur #IA comme Aerospline, Askovet, Cibler, Deski, Flowtide, GAIAR, InHeart, Lucine, Mieuxplacer.com, Nurea, Quorum ou Wozza.  
Apprenez-en plus sur [www.unitec.fr/intelligence-artificielle-ia-data/](http://www.unitec.fr/intelligence-artificielle-ia-data/), et candidatez en ligne pour être accompagné par Unitec !

Vous pouvez nous suggérer des thèmes que vous souhaiteriez voir traités dans une prochaine Note (ou Dossier) de Veille  
veille@unitec.fr | communication@digital-aquitaine.com



avec le soutien de  
**DIGITAL**  
**AQUITAINE**