

COMMENT L'INTELLIGENCE ARTIFICIELLE TRAITE LES IMAGES ?

Conférence de Jenny BENOIS-PINEAU
Professeure d'informatique - Université de Bordeaux – Laboratoire LABRI
17 janvier 2026

Jenny Benois-Pineau a fait une conférence de haut niveau, dont il est difficile de rendre compte quand on n'a pas une formation mathématique solide. Je vais me contenter de dire ce qui m'est resté accessible et qui a vivement intéressé la littéraire que je suis, en priant les informaticiens de pardonner mes approximations.



Un modèle bio-inspiré

Les informaticiens traitent les informations avec des réseaux de calcul imitant les réseaux de neurones : les flux d'informations sont transmis à un opérateur qui les combine et les pondère pour donner une réponse. Le traitement se fait en plusieurs étapes à travers plusieurs couches de calcul. Mais le système devient instable si on accumule trop de couches.

Dans le modèle du Perceptron Multiniveaux (1958) chaque neurone d'entrée est connecté à tous les autres dans les couches intermédiaires, il en est de même pour la couche de sortie : c'est un réseau complètement

connecté.

Une révolution : les réseaux à convolution (CNN)

Les images, qui comptent souvent des milliers de pixels, sont des fichiers informatiques trop lourds pour être traités de cette manière. En 1998, ont été inventés les réseaux à convolution (en anglais CNN), dans lesquels chaque neurone d'entrée a un champ de réception limité et se focalise sur une région de l'image. Ainsi pour identifier une forme géométrique le système se focalise, par exemple, sur les extrémités, les angles et les intersections, seules informations qui sont conservées dans les couches suivantes de traitement. Le processus de transformation d'une image par un CNN comporte plusieurs phases de concentration (*pooling*) pour ne retenir que les informations les plus pertinentes et aboutir à la décision.

Un exemple d'utilisation du CNN : le corpus Mexculture

Il s'agissait de classer informatiquement 16 000 photos de monuments mexicains selon leur style en quatre catégories : art préhispanique, art colonial, art moderne, inclassables.

Pour accélérer le processus d'identification, des chercheurs ont eu l'idée de fournir au réseau une information déjà sélectionnée par les humains. En enregistrant les mouvements des yeux de 5000 visiteurs regardant les tableaux de la National Gallery, Wooding a établi une carte de l'attention visuelle. En utilisant ces points de saillance visuelle, on peut réduire la zone de travail d'un réseau à convolution. Un des réseaux à convolution utilisant la saillance visuelle a donné 97% de résultats exacts dans le classement de monuments mexicains.

Comme on ne peut mettre des observateurs humains devant chaque image, on s'est demandé comment forcer les réseaux à diriger eux-mêmes leur attention sur les zones pertinentes. Par exemple on a pu introduire dans le programme de traitement des images des cartes d'attention visuelle générées par des modèles de prédiction de l'attention visuelle.

Des expériences ont montré que les réseaux de neurones profonds équipés de cartes d'attention sont aussi attentifs que les humains. Ces questions ont abouti à la création de

nouvelles architectures de réseaux de neurones profonds : les transformeurs (ChatGPT est fondé sur des transformeurs).

Assistance aux porteurs de neuroprothèses

Le laboratoire LABRI est partie prenante d'un autre projet recourant au traitement de l'image. Il s'agit de guider les mouvements d'une prothèse des membres supérieurs à l'aide d'une caméra intégrée aux lunettes du porteur de la prothèse.



Actuellement on utilise des prothèses bioniques qui sont reliées aux nerfs du patient, technique invasive qui exige une opération chirurgicale lourde. Des prothèses commandées électroniquement à partir des images enregistrées par une caméra en temps réel seraient non-invasives et coûteraient beaucoup moins cher.

Il faut créer des logiciels de traitement d'images capables de tenir compte à la fois de la direction du regard du patient, de la forme de l'objet à saisir et de la profondeur de la scène. De nombreux essais ont permis d'obtenir des réglages satisfaisants avec une marge d'erreur très réduite.

Le projet TELLi (Train Léger Innovant)

Il s'agit d'un projet visant à automatiser l'inspection des voies de chemins de fer, grâce à l'analyse d'images enregistrées par des caméras fixées sur les trains qui circulent. Ce projet répond à un besoin crucial, dans un contexte de fréquentation croissante du train et d'usure des infrastructures.

Ce projet recourt à l'IA hybride, c'est-à-dire qu'il utilise les connaissances humaines (les règles du métier) pour valider ou invalider les conclusions de l'IA.

Un détecteur d'objets repère les défauts sur la voirie (par exemple un écrou manquant), les classe et les localise. Mais ces informations doivent être combinées avec celles qui concernent le contexte. En l'occurrence une image montre, à côté de l'écrou manquant, des éclisses et des cés de serrage indiquant que le défaut a déjà été réparé. Un programme transformeur fusionne l'image de l'objet et celle du contexte et on lui demande d'appliquer la règle logique consistant à invalider la détection du défaut « écrou manquant » chaque fois que le contexte montre des cés de serrage. Les résultats obtenus par cette méthode combinant double détection et règles logiques sont fiables à 99%.

Conclusion générale

Beaucoup de travail est encore nécessaire pour gérer ensemble images et informations multimodales (texte, son, signaux).

L'utilisation de l'IA dans le traitement et la création d'images présente des risques mais offre aussi de grandes possibilités.

La multiplication d'images trafiquées grâce à l'IA générative est préoccupante mais d'un autre côté l'IA générative permet de restaurer des images dégradées. L'IA peut être une aide à la création sans pour autant remplacer les artistes. Elle aide à réduire les tâches pénibles, ce qui devra s'accompagner d'une transformation des qualifications professionnelles.

L'un des principaux défis sera d'économiser les calculs pour créer une IA plus « verte ».

Claire Evesque

Lien pour écouter la conférence : <https://youtu.be/91oo2O0hPwg>