

**UNIVERSITÉ LIBRE DE BRUXELLES – FACULTÉ DES SCIENCES**  
**DÉPARTEMENT D'INFORMATIQUE**

*Ilyass OUALI ALAMI, Nicolas NGANDO NGENA, Malcom OGOUBY-KOMLAN-DIOGO, Szymon JANUSZEWSKI, Cyril SCHETGEN*

**Question Scientifique :**

Les systèmes actuels de vision artificielle ont la capacité de reconnaître des objets avec une très grande efficacité. Mais être performant suffit-il pour parler d'intelligence ?

Dans ce projet, nous comparons un humain et un modèle d'IA sur une tâche de reconnaissance visuelle, c'est-à-dire la capacité à décider, à partir d'une image, si un objet est présent ou non. Ici, la question posée est simple en apparence : **l'image contient-elle un chat, ou non ?** Cette comparaison permet de mettre en évidence une différence essentielle : l'humain interprète une image avec le contexte et le sens, alors que l'IA apprend surtout à associer des motifs visuels à partir d'exemples. On peut ainsi mieux voir dans quels cas l'IA paraît très forte, et dans quels cas elle révèle ses limites.

**Description de l'expérience:**

Notre expérience repose sur un **affrontement** en 3 épreuves **entre l'homme et l'ia**, dans lequel il faut déterminer si un chat est présent ou non dans la photo.

**Épreuve 1:** Le participant doit décider très vite si un chat est présent ou non dans l'image. Cette première étape met surtout en évidence qu'une IA peut sembler presque parfaite lorsque les images sont nettes, classiques et proches de celles vues à l'entraînement.

**Épreuve 2:** L'image est progressivement dégradée pour tester la robustesse du modèle. Cette étape montre qu'une IA peut perdre beaucoup en fiabilité dès que l'information visuelle se détériore, alors qu'un humain continue parfois à reconnaître l'objet grâce à sa perception globale.

**Épreuve 3:** On présente des images ambiguës ou ressemblantes, comme des dessins, des peluches ou des animaux proches d'un chat. Cette étape sert à montrer qu'un modèle peut se tromper avec une forte confiance, car il associe surtout des motifs visuels appris plutôt qu'une véritable notion abstraite de "chat".

**Pourquoi cette comparaison est-elle intéressante ?**

Cette comparaison est intéressante parce qu'elle ne consiste pas seulement à mesurer qui répond le mieux. Elle permet surtout de comparer deux façons très différentes de traiter une image. L'humain mobilise le contexte, l'expérience et le sens, tandis que l'IA transforme l'image en nombres, repère des motifs visuels appris, puis calcule la réponse la plus probable. En confrontant humain et IA aux mêmes images, on peut donc mieux comprendre non seulement leurs performances, mais aussi leurs limites, leurs types d'erreurs et leur manière de réagir face à l'ambiguïté ou à des situations nouvelles.

**Que faut-il en retenir ?**

Les résultats obtenus montrent qu'un système de vision artificielle peut être très performant pour reconnaître certaines images, sans pour autant posséder une compréhension comparable à celle d'un humain. Une IA de vision apprend à partir d'exemples, transforme les images en valeurs numériques, puis calcule la réponse la plus probable à partir de motifs visuels repérés pendant son entraînement. Cela explique pourquoi elle peut réussir dans des situations familières, mais aussi pourquoi elle peut devenir fragile lorsque l'image est dégradée, ambiguë ou différente de ce qu'elle a appris. La performance seule ne suffit donc pas pour parler d'intelligence au sens humain du terme.

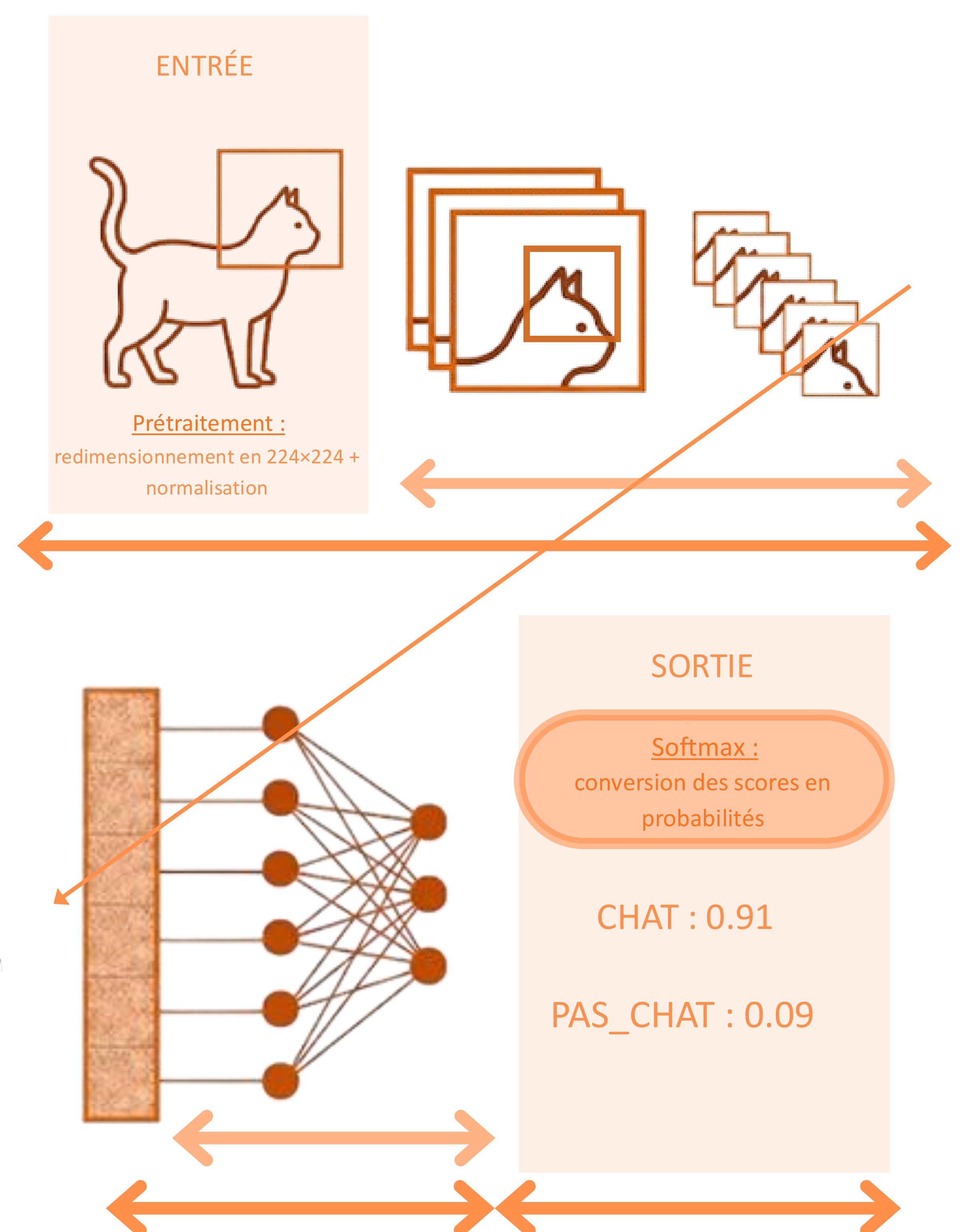
**Que se passe-t-il dans l'algorithme ?**

Notre système utilise un réseau de neurones convolutionnel, plus précisément un modèle ResNet-18.

Lorsqu'une image est envoyée au programme, elle est d'abord préparée : elle est redimensionnée, normalisée, puis convertie en valeurs numériques. Le réseau analyse ensuite cette image par étapes successives afin d'en extraire des caractéristiques visuelles, comme des contours, des motifs, des textures ou des formes.

Enfin, la dernière couche du modèle attribue un score à chaque réponse possible, puis le programme retient la plus probable : "chat" ou "pas chat".

**RÉSEAU DE NEURONES CONVOLUTIFS :**



**Pourquoi "convolutionnel" ?**

Un réseau convolutionnel analyse l'image localement. Les premières couches détectent des motifs simples, comme des bords, des contrastes ou des textures. Les couches plus profondes combinent ces motifs pour former des représentations plus complexes. La couche finale attribue ensuite un score à chaque classe possible.