

# François Labastie

Cas Pratique E3

## Comptage d'un groupe de personnes grâce à l'intelligence artificielle

Décembre 2020

Titre professionnel "Développeur en intelligence artificielle"

de niveau 6 enregistré au RNCP sous le n°34757

Passage par la voie de la formation -

parcours de 19 mois achevé le 20 octobre 2020

---

Introduction et choix de thématique .....	2
I Documents .....	2
1. Méthodologie et outils de recherche	
2. Sources & analyse de pertinence	
II Synthèse .....	7
1. Tendances des solutions en Intelligence Artificielle	
2. Conséquences pour la pratique professionnelle	
Conclusion .....	8

# Introduction et choix de thématique

L'école SalsaNueva - située à Paris - souhaite compter de manière automatique les participants à ses cours de danse: A l'entrée des studios, les danseurs déposent leurs cartes d'abonnements qui servent à décompter leurs présences et facturer les cours. L'employée en charge de cette action doit ensuite comparer le nombre de cartes déposées avec le nombre de danseurs et danseuses présent(e)s dans la salle.

Pourtant, le comptage de personnes en mouvement dans un lieu clos n'est pas une tâche évidente. C'est pourquoi il nous est demandé d'étudier une solution de détection automatique, à l'aide d'une webcam par exemple.

## I. Documents

### 1.1 Méthodologie et outils de recherche

Afin d'identifier de manière efficace les documents relatifs au sujet qui nous concerne, notre stratégie de recherche porte en premier lieu sur les mots-clés représentatifs du domaine. Nous en proposons une série - principalement en français - mais nous testerons également leurs traductions anglaises. Ces mots-clés seront appliqués dans les moteurs de recherche choisis, ainsi que sur d'autres outils.

#### **Mots-clés**

- *comptage de formes grâce à intelligence artificielle*
- *counting shapes using artificial intelligence*
- *reconnaissance de formes intelligence artificielle*
- *artificial intelligence pattern recognition*
- *reconnaissance des formes : méthodes et applications.*
- *solutions reconnaissance de formes intelligence artificielle*

Notre méthode repose sur des recherches partant du général jusqu'au particulier, en employant des outils généralistes puis de plus en plus ciblés et spécialisés. Nos requêtes seront alors reformulées selon les résultats obtenus afin d'atteindre nos objectifs. C'est donc une méthode itérative "recherche → analyse → recherche" qui nous anime.

## Outils

### Moteurs de recherche généralistes

- [google.com](https://www.google.com/) <sup>1</sup> Moteur de recherche le plus visité au monde
- [scholar.google.fr](https://scholar.google.fr/) <sup>2</sup> Articles et publications scientifiques
- [qwant.com](https://www.qwant.com/) <sup>3</sup> Moteur de recherche généraliste

### Moteurs de recherche scientifiques

- [worldwidescience.org](https://worldwidescience.org/) <sup>4</sup> Portail mondial d'information scientifique
- [bib.cnrs.fr](https://bib.cnrs.fr/) <sup>5</sup> Ressources documentaires du CNRS.
- [hal.archives-ouvertes.fr](https://hal.archives-ouvertes.fr/) <sup>6</sup> Hyper articles en ligne (ou HAL); plateforme en ligne développée par le Centre pour la communication scientifique directe (CCSD) du CNRS.
- [https://hal.inria.fr](https://hal.inria.fr/) <sup>7</sup> Ensemble des publications de l'INRIA sur HAL

### Revues et sites

- [medium.com](https://medium.com/) <sup>8</sup> Plateforme web de blog
- [stackoverflow.com](https://stackoverflow.com/) <sup>9</sup> Plateforme "questions-réponses" informatique
- [youtube.com](https://www.youtube.com/) <sup>10</sup> Hébergement de vidéos & média social
- [cite-sciences.fr](https://cite-sciences.fr) Catalogue de la bibliothèque Cité des Sciences

En appliquant notre série de mots-clés à ces différents outils sélectionnés, nous évaluons les réponses retournées selon notre objectif de recherche. Les sites et plateformes généralistes nous permettent de naviguer de manière globale autour du thème pour cibler ensuite des documents plus pertinents à analyser.

---

<sup>1</sup> <https://www.google.com/>

<sup>2</sup> <https://scholar.google.fr/>

<sup>3</sup> <https://www.qwant.com/>

<sup>4</sup> <https://worldwidescience.org/>

<sup>5</sup> <https://bib.cnrs.fr/>

<sup>6</sup> <https://hal.archives-ouvertes.fr/>

<sup>7</sup> <https://hal.inria.fr/>

<sup>8</sup> <https://medium.com/>

<sup>9</sup> <https://stackoverflow.com/>

<sup>10</sup> <https://www.youtube.com/>

## 1.2 Sources & analyse de pertinence

Une première recherche générale avec le moteur Google et les mots clés "*comptage de formes grâce à intelligence artificielle*", retourne une série de documents dont nous sélectionnons les suivants:

- [\*Analyse et comptage via Deep Learning\*](#)<sup>11</sup>  
pixalarm.net/fr - Site web de la société InPixal
- [\*Comptage et analyse de flux - NomoSense\*](#)<sup>12</sup>  
nomosense.com - Site web de la société nomosense
- [\*Une startup automatise le comptage des clients dans les magasins\*](#)<sup>13</sup>  
geeko.lesoir.be - Article de Charlotte Rabatel - 22 avril 2020

Nous en retenons que certaines sociétés - dont une belge et une autre française - commercialisent des solutions pour la régulation de fréquentations de magasins ou bien le comptage de personnes dans une file d'attente par exemple.

Une nouvelle étape de recherche sur Google, avec maintenant les mots-clés "*reconnaissance de formes intelligence artificielle*", nous propose une série de documents davantage orientés scientifiques. La raison de cette différence proviendrait du terme "reconnaissance". Nous retenons les documents suivants:

- [\*Intelligence artificielle et reconnaissance des formes, vision, apprentissage pour la robotique\*](#)<sup>14</sup>  
ISIR - Institut des Systèmes Intelligents et de Robotique  
Article de Isabelle Bloch et al · 2014
- [\*Intelligence artificielle : la reconnaissance d'images prend de l'ampleur\*](#)<sup>15</sup>  
medium.com · Unitec · Veille et Prospective · 15 mars 2019

Le premier article décrit quelques problèmes et méthodes associant intelligence artificielle, reconnaissance des formes, vision par ordinateur et apprentissage. Il s'agit d'un texte universitaire inclus dans un ouvrage qui dresse un panorama des recherches effectuées en IA allant de travaux fondamentaux aux applications.

---

<sup>11</sup> <http://www.pixalarm.net/fr/analyse-et-comptage-via-deep-learning/>

<sup>12</sup> <https://www.nomosense.com/comptage-et-analyse-de-flux/>

<sup>13</sup> <https://geeko.lesoir.be/2020/04/22/une-startup-automatise-le-comptage-des-clients-dans-les-magasins/>

<sup>14</sup> <https://dumas.ccsd.cnrs.fr/ISIR/hal-00995039>

<sup>15</sup> <https://medium.com/@veilleunitec/intelligence-artificielle-la-reconnaissance-dimages-prend-de-l-ampleur-9836dfddf82>

Le deuxième article - issu du site *medium.com* - est moins technique, et fait un balayage des applications en reconnaissance d'images grâce à l'intelligence artificielle. Il nous informe que les grands acteurs mondiaux proposent des API "sur étagères":

- Amazon → [Rekognition](#)
- Microsoft → [Azure cognitive services](#)
- IBM → [Watson Visual Recognition](#)
- Google → [IA Vision](#)
- Intel → [OpenVINO Toolkit](#)
- Apple → [Vision](#)

Nous explorons les présentations de ces solutions, via leurs sites web:

**[Amazon Rekognition](#)** est une plateforme de vision par ordinateur SaaS (Software as a Service) basée sur le cloud et utilisant une technologie de deep learning. Elle permet d'identifier des objets, des personnes, du texte, des scènes et des activités dans des images et des vidéos.

**[Azure Cognitive Services](#)** est une famille complète de services d'intelligence artificielle et d'API cognitives pour aider à créer des applications intelligentes. Par exemple, son service d'IA "Vision par ordinateur" analyse le contenu dans les images et vidéos.

**[Watson Visual Recognition](#)** est un service qui permet l'étiquetage, la classification et la recherche rapide et précise de contenu visuel à l'aide de l'apprentissage automatique.

**[Google Cloud Vision](#)** est un service d'analyse d'images via API qui utilise un modèle pré-entraîné formé sur un vaste ensemble de données d'images, similaire aux modèles utilisés pour alimenter Google Photos.

**[Intel OpenVINO](#)** est un kit d'outils téléchargeable gratuitement qui permet aux développeurs et aux scientifiques des données de développer facilement des capacités de vision par ordinateur. Ce framework se décline en deux versions, dont l'une est maintenue par la communauté open source sous licence Apache.

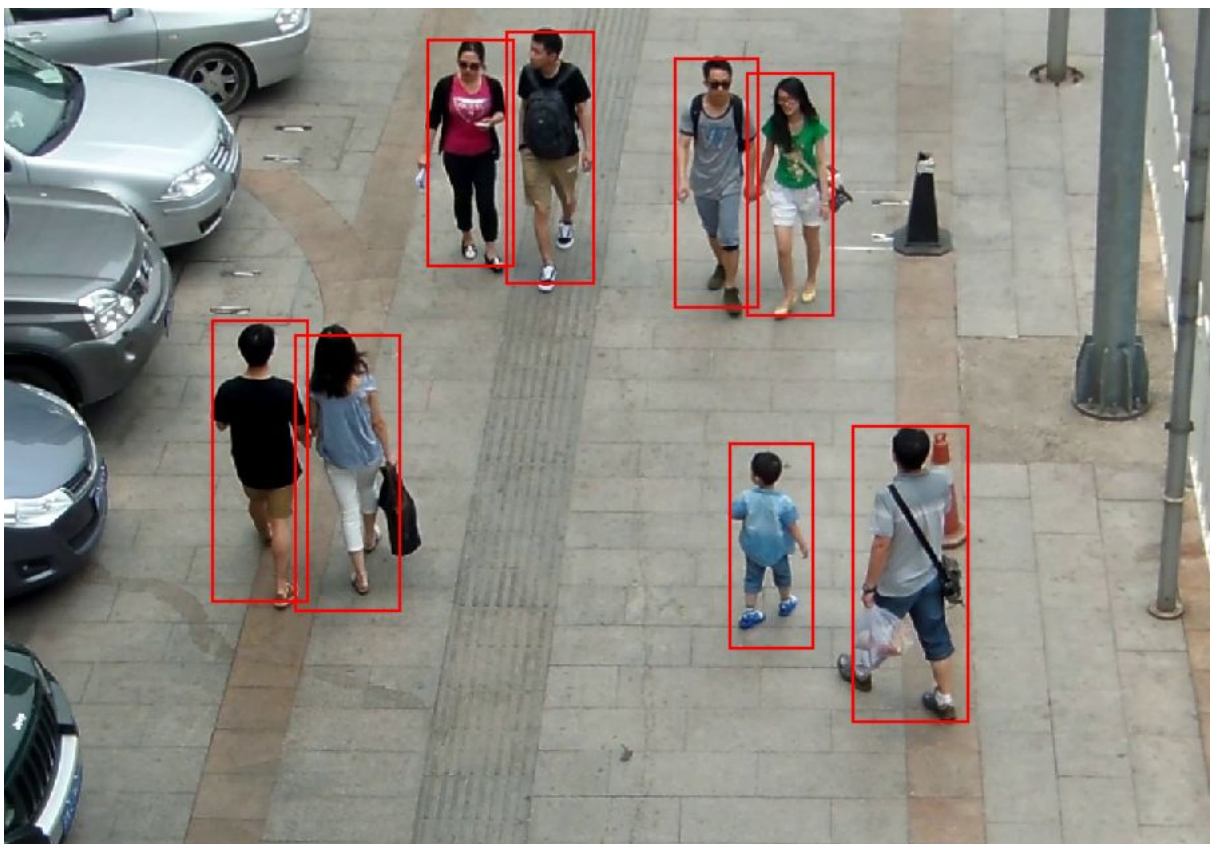
**[Apple vision](#)** est un framework pour le développement d'applications iOS de vision par ordinateur. Cet outil permet par exemple la détection de

torses humains <sup>16</sup>, ce qui nous conduit à lancer une recherche sur Google avec les mots-clés suivants : "*human torso detection*". Et le résultat nous mène vers d'autres outils (Keras, TensorFlow, etc.)

Le framework opensource OpenVINO nous semble le plus adapté pour notre projet. Et une recherche avec les mots-clés "*openvino person detection*" sur Google puis Youtube nous renvoie les résultats suivants :

- [person-detection-retail-0013](#) <sup>17</sup> OpenVINO™ Toolkit Documentation
- [Pedestrian Tracking Demo | OpenVINO™ toolkit](#) <sup>18</sup> Youtube, 2019

Ces documents nous renseignent sur les possibilités d'OpenVINO pour la détection et le comptage de personnes dans un lieu complexe. Des cas d'utilisations codés en C++ et en Python sont proposés par Intel.<sup>19</sup>



Exemple de détections de piétons avec Intel OpenVINO

<sup>16</sup> <https://developer.apple.com/videos/play/wwdc2020/10653/>

<sup>17</sup> <https://docs.openvino toolkit.org/>

<sup>18</sup> <https://www.youtube.com/watch?v=OKH57mvO9k0>

<sup>19</sup> <https://software.intel.com/content/www/us/en/develop/tools/openvino-toolkit.html>

## II. Synthèse

### 2.1 Tendances des solutions en Intelligence Artificielle

Les solutions des grands acteurs tels que Amazon, Azure, IBM, Google présentées précédemment, sont des plateformes interrogeables via API, et qui permettent à l'utilisateur d'introduire ses propres données et payer les algorithmes ou les ressources de calcul lors de leur utilisation.

Une recherche "*vision par ordinateur tendance 2020*" sur Google nous guide vers des articles généralistes confirmant cette idée:

- [Tendances retail 2020 : vision par ordinateur](https://www.usine-digitale.fr) <sup>20</sup> usine-digitale.fr
- [La vision par ordinateur](https://www.numericube.com/) <sup>21</sup> · numericube.com · 2020

Ces articles nous apprennent que la vision par ordinateur n'est plus cantonnée à la surveillance et à la sécurité comme c'était le cas il y a quelques années. Elle est présente également dans l'imagerie médicale, le manufacturing, la recherche scientifique ou l'informatique de gestion.

Aussi, la tendance globale est à la généralisation des algorithmes sur la base de dispositifs de prise de vue courants (notamment smartphone). Et les avancées récentes de la vision par ordinateur se sont donc concentrées tant sur la partie matérielle (amélioration des dispositifs de prise de vue) que sur la partie logicielle.

### 2.2 Conséquences pour la pratique professionnelle

Ces évolutions entraînent des facilités d'utilisation pour les développeurs et les entreprises qui souhaitent commercialiser des solutions en intelligence artificielle. Les outils et méthodes mis en œuvre peuvent être matériels (caméras, dispositifs d'imagerie) et logiciels (algorithmes, ou intégration de programmes « prêts à l'emploi ») avec parfois des modèles pré-entraînés permettant une démocratisation de l'IA et une mise en pratique plus accessible des outils.

---

<sup>20</sup> <https://www.usine-digitale.fr/>

<sup>21</sup> <https://www.numericube.com/>

## Conclusion

Dans le cas d'utilisation qui nous concerne - le comptage de personnes en mouvement dans une salle de danse - plusieurs possibilités semblent disponibles, avec leurs avantages et inconvénients. Soit l'une des solutions via API (Amazon, Azure, IBM, Google), soit la conception d'une application spécifique à l'aide d'un framework.

C'est l'orientation que nous envisageons : l'utilisation du framework OpenVINO sur un [raspberry pi](https://www.raspberrypi.org/)<sup>22</sup> associé à une webcam, et couplé à un mini [écran](https://www.adafruit.com/product/1601)<sup>23</sup> par exemple.

Une recherche Google "*openvino raspberry*" nous confirme la compatibilité de ce framework avec l'OS Raspian. NVidia propose également un stick accélérateur d'IA pour le raspberry.

- [Install OpenVINO™ toolkit for Raspbian\\* OS](https://docs.openvinotoolkit.org/)<sup>24</sup>  
OpenVINO Documentation
- [Turning a Raspberry Pi 3B+ into a powerful object recognition edge server with Intel Movidius NCS2](https://towardsdatascience.com/turning-a-raspberry-pi-3b-plus-into-a-powerful-object-recognition-edge-server-with-intel-movidius-ncs2/)<sup>25</sup> towardsdatascience.com

Mais connaissant également la carte embarquée [Jetson Nano](https://www.nvidia.com/fr-fr/autonomous-machines/embedded-systems/jetson-nano/)<sup>26</sup> de NVidia - qui dispose d'un GPU intégré - cette solution associée à une webcam et mini écran nous semble bien plus attractive pour un développement à base de OpenVINO, en C++ ou en Python.

Vu les retours d'expériences diffusés et les documentations disponibles nous pouvons alors nous engager sur un développement d'application en C++, avec OpenVINO et Jetson Nano pour un comptage de personnes en mouvement dans un studio de danse.

---

<sup>22</sup> <https://www.raspberrypi.org/>

<sup>23</sup> <https://www.adafruit.com/product/1601>

<sup>24</sup> <https://docs.openvinotoolkit.org/>

<sup>25</sup> <https://towardsdatascience.com/>

<sup>26</sup> <https://www.nvidia.com/fr-fr/autonomous-machines/embedded-systems/jetson-nano/>