



HAL
open science

Comment l'IA transforme l'analyse des images de la ville

Benjamin Beaucamp

► **To cite this version:**

Benjamin Beaucamp. Comment l'IA transforme l'analyse des images de la ville. Culture et recherche, 2024, 147, pp.87-89. hal-04800985

HAL Id: hal-04800985

<https://hal.science/hal-04800985v1>

Submitted on 24 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Comment l'IA transforme l'analyse des images de la ville

Se promener en ville peut offrir des expériences variées : certaines rues paraissent vivantes et agréables, tandis que d'autres sont ennuyeuses, voire oppressantes. Les ambiances urbaines, liées en partie à l'aspect visuel de la ville, agissent sur la qualité de vie de ses habitants : c'est pourquoi qualifier la perception visuelle de l'espace urbain par le piéton est un enjeu pour les collectivités territoriales et les acteurs de l'urbanisme.

Comprendre comment les piétons perçoivent et interagissent avec leur environnement est primordial pour créer des espaces non seulement fonctionnels, mais aussi psychologiquement et socialement attrayants et stimulants pour les habitants. Grâce à l'Intelligence artificielle (IA), et notamment *via* des techniques de traitement d'image, de nouvelles méthodes émergent pour évaluer, de manière automatique, le ressenti des piétons dans n'importe quelle ville.

Évaluer l'expérience du piéton

L'évaluation des qualités urbaines est généralement chronophage, car elle nécessite de réaliser des études *in situ* avec des experts et des enquêtes auprès d'habitants.

Pour pallier ce problème, nous travaillons sur des modèles d'IA permettant de qualifier automatiquement l'espace urbain à l'échelle d'une ville entière, à partir de corpus d'images de rues. L'objectif est de quantifier ces qualités subjectives pour « mesurer le non-mesurable¹ » et aider les experts dans leur prise de décision.

Tandis que les modèles classiques se basent sur des quantités mesurables comme la largeur d'une rue, la hauteur des bâtiments, le ratio de ciel vu, la présence de sites historiques ou la densité de végétation, la particularité des modèles d'IA est qu'ils sont entraînés directement à partir des ressentis des piétons.

L'IA permet en effet d'analyser des millions d'images prises au niveau de la rue et de les associer à des évaluations humaines, collectées *via* des questionnaires en ligne. Les images de rue sont présentées aux utilisateurs pour qu'ils se familiarisent avec le lieu de l'étude, puis ils évaluent les qualités du lieu, sur une échelle de 1 à 5 : sentiment de sécurité, caractère dynamique ou paisible, etc. Les chercheurs développent ainsi des modèles capables de prédire comment les piétons perçoivent différents espaces urbains.

En 2016, le MIT (Massachusetts Institute of Technology) a publié un jeu de données qui fait référence, nommé Place Pulse 2. Il contient 110 000 images provenant de 28 pays et plus de 1,5 million de votes d'utilisateurs du monde entier. Ce jeu de données offre un socle commun pour que les chercheurs puissent développer et évaluer de nouvelles techniques sur une grande variété d'environnements urbains².

Les modèles d'IA utilisés sont des réseaux de neurones convolutionnels, qui sont tout particulièrement adaptés pour traiter des images. Une fois entraînés, ces modèles d'IA peuvent analyser de nouvelles images et prédire des indicateurs qui nous renseignent sur la manière dont les piétons sont susceptibles de percevoir ces lieux, au sens d'un comportement « moyen ».

À grande échelle, ces indicateurs permettent par exemple de créer une cartographie sensible de la ville. Cette approche peut aussi être utilisée pour prévoir les effets d'un nouvel élément de mobilier urbain ou pour évaluer l'avant et l'après d'un projet de rénovation d'un lieu public.

Les avantages de l'IA pour l'analyse urbaine

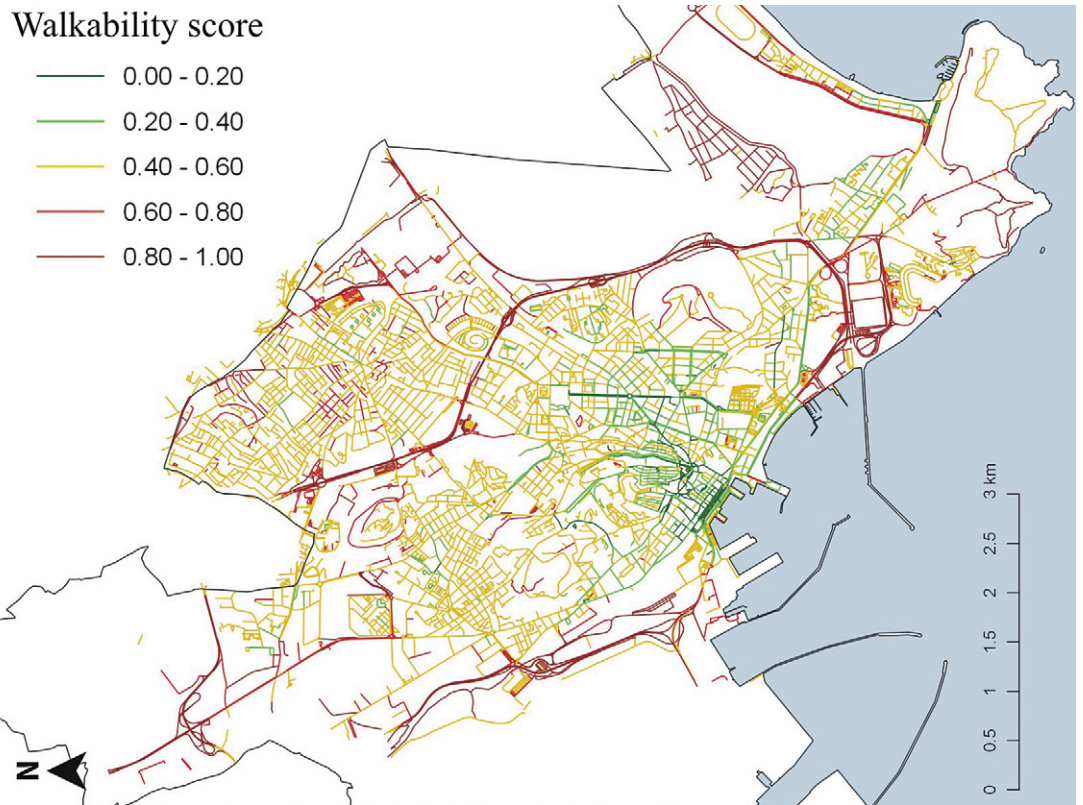
L'utilisation de l'IA permet de traiter de grandes masses de données, ce qui présente plusieurs avantages significatifs : un échantillon beaucoup plus large et diversifié de personnes, capturant ainsi un vaste éventail de points de vue ; la possibilité d'évaluer rapidement des villes entières, fournissant une vue d'ensemble impossible à obtenir avec les méthodes d'enquête habituelles. De plus, nous pouvons nous concentrer sur les ressentis individuels des résidents et des utilisateurs réels des espaces, qui sont des données difficiles à traiter. Cela nous permet de compléter les évaluations plus classiques qui reposent sur des attributs physiques des lieux. En outre, elle aide à identifier des relations entre les éléments perçus et

BENJAMIN BEAUCAMP

Post-doctorant, ingénieur de recherche à l'Université Gustave Eiffel, chercheur associé UMR CNRS Ambiances Architectures Urbanités – Centre de recherche nantais architectures urbanités (AAU-CRENAU)

1. Reid Ewing et Susan Handy, « Measuring the unmeasurable: Urban design qualities related to walkability », *Journal of Urban Design*, vol. 14, n° 1, 2009, p. 65-84. <https://doi.org/10.1080/13574800802451155> (site consulté le 26 juillet 2024).

2. Abhimanyu Dubey, Nikhil Naik, Devi Parikh, Ramesh Raskar et César A. Hidalgo, « Deep learning the city: Quantifying urban perception at a global scale », dans Bastian Leibe *et al.* (dir.), *Computer Vision – ECCV 2016*, Springer, p. 196-212. https://doi.org/10.1007/978-3-319-46448-0_12 (site consulté le 26 juillet 2024).



Carte de marchabilité de la ville de Cagliari en Italie.
 © Ivan Blečić, Arnaldo Cecchini et Giuseppe A. Trunfio, « Towards automatic assessment of perceived walkability », dans Osvaldo Gervasi et al. (dir.), *Computational Science and its Applications – ICCSA 2018*, Springer, 2018, 10962 LNCS, p. 351-365. https://doi.org/10.1007/978-3-319-95168-3_24 (site consulté le 26 juillet 2024).

donne donc des explications qui pourraient échapper aux observateurs humains. Cette approche vise à démocratiser le processus d'évaluation, en intégrant l'expertise ordinaire d'un large éventail de citoyens grâce à des méthodes innovantes de collecte de données telles que le *crowdsourcing* (« collecte de données participatives »).

Améliorer la représentation virtuelle des espaces urbains

Les chercheurs travaillent à l'amélioration de cette approche, notamment en se concentrant sur l'obtention de recommandations basées sur l'expérience des piétons, élément essentiel des approches d'apprentissage automatique.

Un facteur majeur pour obtenir des résultats fiables est de s'assurer que les images utilisées représentent fidèlement l'expérience des piétons. Les images de rue, largement utilisées en raison de leur couverture dense des villes, ont traditionnellement été produites par des plateformes offrant principalement des perspectives centrées sur les voitures. Cette utilisation

prépondérante de l'imagerie automobile peut fausser la représentation visuelle des espaces urbains, en particulier celles des piétons et des cyclistes. Par exemple, on peut observer un même lieu photographié depuis la route avec le point de vue d'un automobiliste et depuis le trottoir avec celui d'un piéton. Comme l'image est l'unique information communiquée aux participants sur le lieu, ce changement de perspective a un impact fort sur la manière dont il est évalué. Pour remédier à ce problème, les chercheurs peuvent utiliser des images produites directement par les piétons avec leur smartphone ou des caméras embarquées, afin de multiplier les points de vue. Des plateformes comme Mapillary³ ou Panoramax⁴ permettent le partage de ces images.

Un autre aspect influant sur notre capacité à nous projeter dans un lieu est l'utilisation d'images à 360 degrés. Comparées aux images « plates », ces images sont une manière de mieux appréhender l'espace urbain environnant et permettent aux participants de mieux comprendre l'espace. La figure ci-contre, en haut, montre l'évolution de la perception d'un lieu en fonction de la direction d'orientation de

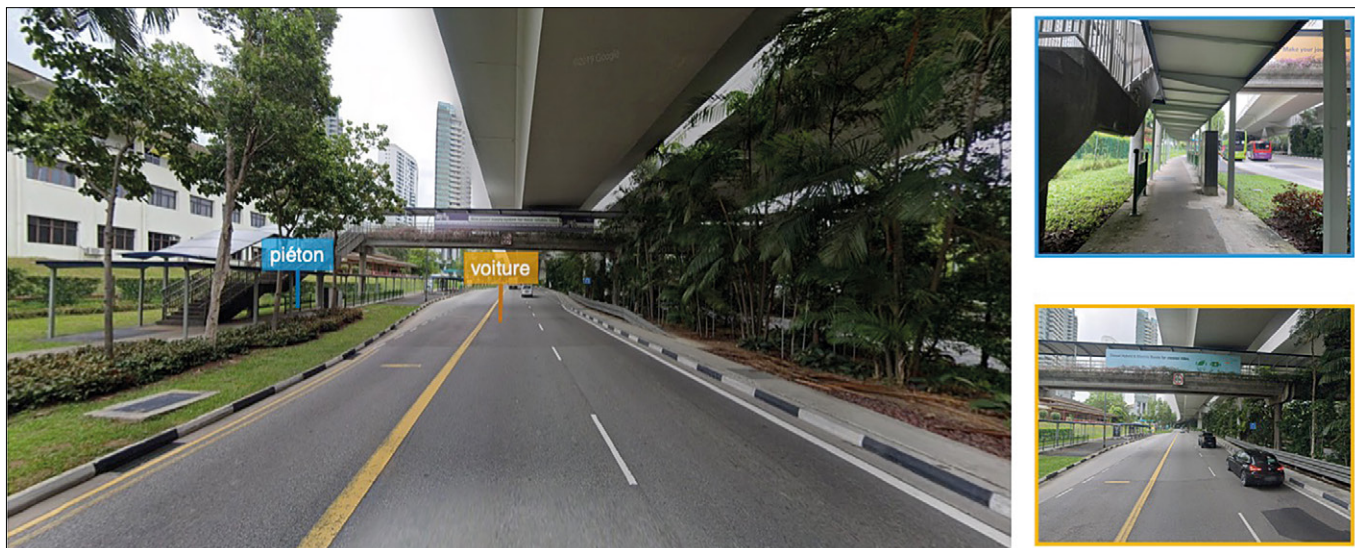
3. <https://www.mapillary.com/> (site consulté le 26 juillet 2024).

4. <https://panoramax.fr> (site consulté le 26 juillet 2024).



Capture d'écran de la plateforme UP2 (*urban panoramic perception*), développée par les auteurs pour collecter l'avis de participants sur des images *Street View* plates ou à 360°.

© plateforme UP2 (*urban panoramic perception*)



© Google

la caméra. Le simple fait de regarder devant ou derrière soi peut amener à des expériences différentes. Cette nouvelle technique d'évaluation des espaces urbains devrait correspondre davantage à l'expérience humaine du lieu. Cela contribue à obtenir des données qui se rapprochent de l'expérience réelle du piéton, et qui sont donc plus fiables.

Défis

Malgré le potentiel de cette approche, des questions demeurent. La collecte de données auprès d'un échantillon véritablement représentatif de la population reste un défi. Les enquêtes en ligne peuvent toucher rapidement de nombreux participants, mais elles risquent de ne pas saisir toute la diversité des résidents urbains, de leurs activités et de leurs expériences.

Par ailleurs, cette approche basée sur l'IA ne restitue pas tous les aspects de la perception urbaine, tels que le son, l'odeur et les souvenirs personnels qui jouent un rôle dans notre expérience des villes. Ainsi, les modèles d'IA visuels apportent une perspective qui doit être complétée avec d'autres méthodes d'analyse.

Enfin, les modèles basés sur l'IA sont souvent difficiles à expliquer. Bien que les réseaux puissent prédire les différentes qualités subjectives d'un lieu, il est difficile de comprendre pourquoi ce lieu possède ces qualités. Des approches complémentaires basées sur l'analyse du contenu visuel peuvent aider à corréliser des éléments tels que la présence de végétation, de voitures ou d'autres piétons, avec des qualités perceptuelles élevées ou faibles.

Perspectives

Plusieurs pistes sont possibles pour la suite de ce travail. La priorité est de continuer à améliorer la qualité des données récoltées *via* les questionnaires en ligne. Cela doit se faire notamment en proposant des expériences virtuelles plus proches de l'expérience réelle, sur site. Par exemple, il serait intéressant de restituer le caractère dynamique des expériences urbaines. Une rue est ainsi perçue différemment selon qu'on la parcourt ou qu'on l'observe à l'aide d'une image statique. D'une part, des vidéos ou des séquences d'images pourraient être présentées aux

participants pour créer une situation plus réaliste. D'autre part, les modèles de prédiction futurs pourraient utiliser ces vidéos pour mieux représenter ce mouvement.

Une autre perspective est le développement de modèles capables d'évaluer la perception à différentes échelles spatiales – d'un point à une rue, jusqu'à des quartiers entiers. Cette prise en compte du contexte pourrait nous aider à comprendre comment les différents éléments de l'environnement urbain interagissent pour façonner notre expérience globale d'un lieu.

Enfin, les modèles actuels sont conçus pour prédire l'avis de la « foule », sans distinction des spécificités de chaque citoyen, ou de leur activité précise (trajet domicile-travail, promenade...). Prendre en compte ces éléments pourrait aider à mieux nous renseigner sur la manière dont un lieu est perçu.

L'objectif de cette recherche est de compléter le jugement humain ou les méthodes traditionnelles utilisées en urbanisme, en fournissant de nouveaux outils et perspectives pour aider à créer des villes plus vivables, agréables et répondant aux besoins de leurs habitants. En exploitant les outils de traitement d'image et les expériences collectives des citoyens, nous ouvrons de nouvelles possibilités pour comprendre et façonner les espaces urbains que nous habitons, tout en mettant en avant la perspective unique qu'a le piéton en ville. ■

À gauche, positions de caméra dans une rue à Singapour pour une voiture et un piéton ; à droite, images prises du point de vue du piéton et d'une voiture.

Image 360° d'une rue à Nantes (en projection équivalente).



© « nantes.jpg » : Google