

Mélina MOREL
Essai
ARC-6013

Utilisation de l'apprentissage-machine non supervisé dans le cadre d'analyses urbaines

Essai présenté dans le cadre de la maîtrise en design urbain (M. Sc.)
Sous la direction de M. Michael Robert Doyle, professeur adjoint à l'école d'architecture

École d'architecture

Université Laval

21 août 2024

TABLE DES MATIÈRES

1	Introduction	7
2	Contexte	8
2.1	Problématique	8
2.2	Application de l'intelligence artificielle à la ville	10
2.3	Questions de recherche.....	11
3	Méthodologie.....	13
4	Analyse	17
4.1	Présentation des articles	17
4.1.1	Urban morphology meets deep learning: Exploring urban forms in one million cities, town and villages across the planet (Moosavi, 2017)	17
4.1.2	A Novel Road Segmentation Technique from Orthophotos Using Deep Convolutional Autoencoders (Sameen & Pradhan, 2017)	19
4.1.3	Land Use detection with cell phone data using topic models: Case Santiago, Chile (Rios & Munoz, 2017).....	19
4.1.4	Modelling urban networks using Variational Autoencoders (Kempinska & Murcio, 2019).....	20
4.1.5	Unsupervised Deep Learning to Explore Streetscape Factors Associated with Urban Cyclist Safety (Zhao et al., 2019)	21
4.1.6	Using convolutional autoencoders to extract visual features of leisure and retail environments (Comber et al., 2020).....	23
4.1.7	The big picture of cities: Analysing Flickr photos of 222 cities worldwide (Taecharungroj & Mathayomchan, 2020).....	24
4.1.8	NCAE and ELM Based Enhanced Ensemble Optimized Model for Traffic Flow Forecasting (Qi et al., 2020)	25
4.1.9	Data-Driven Approach for Passenger Mobility Pattern Recognition Using Spatiotemporal Embedding (Chao et al., 2021).....	26
4.1.10	Machine learning-based characterisation of urban morphology with the street pattern (Wu et al., 2024)	27
4.2	Analyse des articles	28
4.2.1	Composition de la machine	28
4.2.1.1	Auto-encodeurs et leurs variantes.....	30
4.2.1.2	Réseaux antagonistes génératifs (GAN).....	36
4.2.1.3	Allocation latente de Dirichlet (LDA).....	37

	4.2.1.4 Analyse par regroupement hiérarchique (HCA).....	38
	4.2.2 Éléments observés.....	39
5	Discussion	46
5.1	Opportunités ; hypothèses au Nunavik	46
5.2	Limites.....	56
6	Conclusion.....	58
7	Annexe.....	60
8	Bibliographie	65

REMERCIEMENTS

À mes parents, merci pour le support tout au long de ces nombreuses années d'études et au travers de toutes ces périodes d'incertitudes et de changement. Merci pour votre patience et votre soutien malgré les nombreux changements de parcours.

À ma sœur, fidèle acolyte d'étude. Merci pour les conseils sur nos projets et pour être toujours prête à échanger tant sur le design que sur l'architecture.

À mes amies avec qui j'ai partagé les hauts et les bas de mon cheminement. Merci pour le soutien, les échanges et tous ces merveilleux moments qui ont défini mon parcours universitaire.

Je souhaite également remercier M. Michael Robert Doyle, mon directeur d'essai, pour les conseils pour la rédaction de cet essai.

LISTE DES TABLEAUX ET FIGURES

Tableau 1. Liste des mots-clés utilisés

Tableau 2. Exemple de requêtes entrées dans les bases de recherche

Tableau 3. Liste des textes sélectionnés

Figure 1. La ligne supérieure montre les réseaux de rues d'entrée, alors que la ligne inférieure montre les résultats obtenus par l'auto-encodeur. Certaines rues ou réseaux n'apparaissent pas sur les images reconstruites (p.4, Moosavi, 2017).

Figure 2. Distribution des formes urbaines sur l'ensemble de la planète. Les couleurs similaires indiquent des formes urbaines qui sont similaires (p.7, Moosavi, 2017).

Figure 3. Exemple d'une carte vectorielle extraite de Google Maps API. Les rues sont représentées en noir, l'eau en bleu, les parcs en vert et les lignes de transport en commun en orange (p.158, Zhao et al., 2019)

Figure 4. Synthèse des algorithmes utilisés par les divers auteurs.

Figure 5. Architecture d'un auto-encodeur convolutif (Sameen et Pradhan, 2017)

Figure 6. Le spectre des formes urbaines généré par une carte auto-adaptative (Moosavi, 2017)

Figure 7. Architecture d'un auto-encodeur variable (Kempinska et Murcio, 2019)

Figure 8. Architecture d'un auto-encodeur dispersé (Syoya Zhou, Medium.com)

Figure 9. Architecture d'un auto-encodeur empilé (Yu et al., 2021)

Figure 10. Exemple de l'architecture d'un réseau antagoniste génératif (source : Jason Brownlee, machinelearningmastery.com)

Figure 11. Fonctionnement de l'allocation latente de Dirichlet (Source : Medium.com)

Figure 12. Résultats obtenus (c) en comparaison avec les données d'entrée (a) et les mêmes réseaux de rue sous forme vectorisée (b) (Source : Sameen et Pradhan, 2017).

Figure 13. Échantillons des réseaux de rues évalués par Kempinska et Murcio classés selon leur densité et leur orientation (source : Kempinska & Murcio, 2019)

Figure 14. Représentation des types de réseaux de rues retrouvés dans les villes de Chengdu, Amsterdam, New-York, Seoul, Londres et Houston (source : Wu et al., 2024)

Figure 15. Quelques exemples de translation du domaine sans accident vers le domaine avec accident. De gauche à droite : domaine sans accident, domaine avec accident, résultat de la translation (source : Zhao et al., 2019)

Figure 16. Échantillons d'images de façades qui ont été récoltées pour les groupes A et D (source : Comber et al., 2020)

Figure 17. Quelques images prises à Kuujjuaq. On y observe une dominance de la nature, notamment de prises de vues de la rivière Koksoak. Les images représentent principalement la nature. (Source : Flickr.com)

Figure 18. Quelques images prises à Kangiqsualujjuaq. On note une dominance de la prise en photo d'activités culturelles et communautaires, notamment la cueillette et la pêche (Source : Flickr.com)

Figure 19. Quelques images prises à Inukjuak. Les images représentent plusieurs activités culturelles qui touchent la confection, notamment d'igloos ou encore de canot. Il semble y avoir aussi une importance sur les activités réalisées en communauté (Source : Flickr.com)

Figure 20. Kuujjuaq. À gauche, des lignes de désir au travers des habitations du Vieux-Kuujjuaq. À droite, des lignes de désir qui longent ou traversent le ruisseau Siuralik (Source : Google Earth 2024)

Figure 21. Kangiqsualujjuaq. À gauche, des lignes de désir au travers des habitations. À droite, des lignes de désir reliant deux quartiers (Source : Google Earth 2024)

Figure 22. Inukjuak. À gauche, des lignes de désir au travers des habitations. À droite, des lignes de désir reliant le centre de formation professionnelle Nunavimmi Pigiursavik à la rive (Source : Google Earth 2024)

Figure 23. Kuujjuaq. Une ligne de désir (en pointillé rose) qui est identifiée sur OpenStreetMap. (Source : OpenStreetMap, 2024)

Figure 24. Inukjuak. Des lignes de désir (en pointillé) identifiées à l'extérieur du village (Source : OpenStreetMap, 2024)

Figure 25. Images de bâtiments situés à Kuujjuaq, récupérées sur Google Street View (source : Google Street View 2024)

Figure 26. Images de bâtiments situés à Kangiqsualujjuaq, récupérées sur Google Street View (source : Google Street View 2024)

Figure 27. Images de bâtiments situés à Inukjuak, récupérées sur Google Street View (source : Google Street View 2024)

RÉSUMÉ

Cet essai analyse dix articles scientifiques portant sur l'utilisation des données massives géoréférencées et des algorithmes d'apprentissage machine non supervisé dans le cadre d'analyse urbaines, afin d'en présenter les opportunités pour les professionnels en design urbain et en recherche sur la ville. Ces textes utilisant divers types de données et algorithmes d'apprentissage non supervisé ont été sélectionnés et analysés afin d'en comprendre la composition de leur algorithme et ce que celui-ci observe afin de voir la ville. Par la suite, trois hypothèses d'utilisation de l'apprentissage-machine pour une analyse urbaine réalisée au Nunavik, dans le nord du Québec, sont suggérées afin de faire ressortir les opportunités, mais également les limites, que présentent les algorithmes.

1_INTRODUCTION

Cet essai s'intéressera à l'utilisation de l'intelligence artificielle, ici définie comme le couplage algorithme d'apprentissage-machine et données massives, dans le cadre d'analyses urbaines. Une première section (Contexte, chapitre 2) présentera la problématique et définira brièvement les applications actuelles de l'intelligence artificielle dans un contexte d'urbanisme et de gestion des villes. Ceci mènera aux questions de recherche qui guideront l'analyse de dix articles scientifiques. Le chapitre 3 fera une description de la méthodologie utilisée afin d'identifier les recherches à l'étude, présentant entre autres les critères de sélections. Le Tableau 3 regroupe les dix articles et les résume selon les critères de sélection. Par la suite, le chapitre 4 fera d'abord un résumé de chacune des études (*section 4.1*), puis analysera celles-ci selon la composition des algorithmes utilisés (*section 4.2.1*) et ce qui est regardés par ces derniers afin de voir la ville (*section 4.2.2*). Finalement, le chapitre 5 proposera trois hypothèses d'application de l'apprentissage non supervisé dans le contexte du Nunavik, se concentrant sur les villages de Kuujuaq, Inukjuak et Kangiqsualujjuaq, puis présentera les limites de l'utilisation de l'intelligence artificielle dans le cadre d'analyses urbaines.

2_CONTEXTE

2.1_Problématique

Les villes, entités en constant changement, agissent comme marqueurs des époques et de l'évolution de la culture propre à laquelle elles se rattachent. Celles-ci permettent de comprendre une communauté; comment elle se déplace, se rencontre, vie. L'analyse urbaine par les professionnels en urbanisme et en design urbain permet de comprendre ces milieux vécus et leur complexité. Traditionnellement, les analyses misent sur l'expérience de la ville par la déambulation dans ses rues et les arrêts dans ses commerces ou dans ses parcs afin de s'imprégner de l'ambiance de la ville; ses bruits, ses odeurs, ses mouvements, son architecture par exemple, ainsi que par le dessin ou la prise de photos (Panerai, Depaule, Demorgon, 1999). Par la suite, les analyses se font selon les concepts définis par Kevin Lynch dans son œuvre *The image of the city* parue en 1960. Ces concepts identifient les éléments fondamentaux qui composent les villes, c'est-à-dire les voies (*paths*), éléments permettant le mouvement au travers de la ville, les limites (*edges*), éléments anthropiques qui créent une rupture dans la ville, les quartiers (*districts*), éléments d'une certaine homogénéité qui donnent la sensation d'entrer et de sortir à l'habitant, les nœuds (*nodes*), jonctions de voies qui demandent une prise de décision de direction et de mode de transport, et finalement les points de repères (*landmarks*), éléments ponctuels qui, comme leur nom l'indique, permettent de se repérer dans la ville (Lynch, 1960). En plus de dessin et de la photographie, ces éléments sont traditionnellement analysés à l'aide de schémas ou encore de diagrammes.

L'arrivée de nouvelles technologies dans les dernières décennies amène des changements et de nouvelles possibilités pour l'analyse des villes. Par exemple, le service Google Street View permet à ses utilisateurs de parcourir des rues de leurs propre ville ou d'autres villes au travers de la planète sans avoir à sortir de leur maison, diminuant ainsi largement le temps que peuvent prendre les analyses simplement en retirant le temps de déplacement, et rendant plus accessible l'analyse de villes éloignées. Depuis peu, le développement rapide de l'intelligence artificielle amène de nouvelles opportunités pour les analyses urbaines, notamment en permettant de traiter des quantités de données très élevées. Cet essai s'intéressera donc à l'utilisation de l'intelligence artificielle, plus précisément les algorithmes d'apprentissage-machine non supervisé et les données massives, dans le cadre d'analyses urbaines. Sous forme de revue de littérature, l'essai présentera une sélection de dix études réalisées dans les dix dernières années et utilisant l'apprentissage non supervisé pour regarder la ville, se concentrant principalement sur sa trame et les rues qui la compose. Les méthodes des auteurs seront analysées

selon la composition de l'œil de la machine et ce que celle-ci regarde afin de voir la ville. Finalement, une réflexion sera effectuée sur les opportunités qu'offrent ces nouvelles technologies pour les designers urbains, proposant trois hypothèses d'analyses utilisant l'apprentissage-machine et les données massives dans un contexte d'analyse urbaine dans trois villages du nord du Québec; Kuujuaq, capitale du Nunavik, et Kangiqsualujuaq, tous deux situés dans la baie d'Ungava, ainsi qu'Inukjuak, situé dans la baie d'Hudson. Ces villages, de trois tailles différentes, ont un intérêt pour proposer des hypothèses d'utilisation de l'apprentissage-machine dû à leur localisation éloignée des grands centres. En effet, cet éloignement, ainsi que leur faible densité, donne l'impression qu'ils se retrouvent exclus des données massives. Il devient donc intéressant de proposer des hypothèses d'analyse pouvant s'appliquer à leur contexte.

Cet essai porte sur l'utilisation de l'apprentissage non supervisé pour réaliser des analyses urbaines, notamment sur la rue. Ces dernières, porteuses de la mobilité urbaine, jouent un rôle important dans les villes. Dans *Recombinant urbanisme : Conceptual modeling in architecture, urban design and city theory* (2005), David Grahame Shane définit ces entités comme l'armature des villes. Les rues, qui facilitent le déplacement dans les villes, favorisent également les rencontres entre les gens, ce qui encourage le partage et les échanges sociaux. Ces axes qui forment la trame des villes agissent également comme agents liant entre les différents services, que ce soient les commerces, les lieux de loisir ou les espaces résidentiels. Lynch (1960) associe les rues à son concept de voies (*paths*), les définissant comme des éléments organisateurs du mouvement couvrant l'ensemble de la ville et étant structurées sous forme de réseaux. Il indique que les rues sont l'élément prédominant des villes reconnus par les résidents et visiteurs, ce qui renforce l'aspect d'importance de celle-ci. Pour le designer urbain et l'urbanisme qui souhaitent comprendre la ville et en faire une planification durable, une bonne compréhension de la rue et une analyse poussée de ses caractéristiques nombreuses et complexes sont donc essentielles. Idéalement, une vision globale de la ville serait à prioriser, mais les méthodes d'analyse conventionnelles qui ont été mentionnées plus haut deviennent rapidement fastidieuses et observer une ville entière à l'aide de celles-ci prendrait un temps considérable. De ce fait, les professionnels s'intéressant à la ville réalisent généralement des analyses qui ne touchent que certains quartiers de villes qui sont en général très bien connues. Ce phénomène peut être constaté dans les études de la morphologie de plusieurs grandes villes à travers le monde qui sont présentées dans l'œuvre de Serge Salat, *Les villes et les formes : Sur l'urbanisme durable* (2011). La même chose peut être remarquée chez Lynch (1960) qui étudie seulement des zones spécifiques des villes de Boston, Jersey City et Los Angeles, aux États-Unis. Une opportunité se présente donc pour l'utilisation d'algorithme

d'apprentissage-machine afin d'effectuer des analyses à échelle beaucoup plus grande, voire afin de comparer entre elles des formes urbaines de partout sur la planète, d'autant plus qu'avec Internet et les technologies actuelles, l'accès à des quantités considérables de données de toutes sortes se voit facilité. Pour cet essai, un intérêt est porté notamment sur les données géoréférencées, c'est-à-dire un phénomène auquel des coordonnées géographiques permettant de définir sa position exacte par rapport à un système de référence géodésique lui ont été attribuées¹. Ces données sont variées; il peut s'agir de photos dont la localisation a été partagée (Taecharungroj & Mathayomchan, 2020), des images satellites (Sameen & Pradhan, 2017; Zhao et al., 2019), ou encore des données d'appels téléphoniques (Rios & Munoz, 2017), de points de collecte de tarifs des transports en commun (Chao et al., 2021), et même du trafic routier (Qi et al., 2020), pour ne donner que quelques exemples.

Si cet essai s'intéresse spécifiquement à l'apprentissage non supervisé, il est important de noter qu'il existe également des algorithmes d'apprentissage supervisé. Ces deux méthodes se différencient par le fait que l'apprentissage non supervisé ne requiert pas que les données soient étiquetées (par exemple, une voiture apparaissant dans une image n'a pas besoin d'être identifiée comme étant telle) pour pouvoir identifier les différents groupes de données, au contraire de l'apprentissage supervisé (Zhu & Goldberg, 2009). Cette caractéristique permet ainsi d'éliminer une étape assez importante d'identification des données et permet à l'algorithme de trouver des tendances inattendues ou autrement invisibles à l'œil humain. Cela permet également de traiter plus rapidement des volumes importants de données, rendant donc intéressant l'utilisation des données massives, ou mégadonnées.

2.2_Applications de l'intelligence artificielle à la ville

Une prédominance de l'utilisation de l'intelligence artificielle, donc des algorithmes d'apprentissage-machine, se remarque lorsqu'il s'agit de gestion des services de la ville. Ces villes sont appelées *villes intelligentes* et ce concept se définit comme « ville qui utilise les technologies de l'information et de la communication ainsi que les données issues de capteurs pour améliorer la gestion de ses différents services ainsi que la qualité de vie de ses citoyens tout en respectant les principes de développement durable »². Tant l'apprentissage supervisé que l'apprentissage non supervisé sont utilisés pour traiter et analyser les données qui sont produites par les villes, bien que jusqu'à présent, une préférence semble

¹ Selon l'Office québécois de la langue française (<https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/8362930/georeference>)

² Tel que défini par l'Office québécois de la langue française (<https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26544297/ville-intelligente>)

se porter sur les algorithmes d'apprentissage supervisé³. Ceci découle du fait que la planification s'intéresse davantage à la prévoyance et à la prédiction de phénomènes déjà bien connus. Considérant leur importance, une rapide présentation des différents algorithmes d'apprentissage supervisé peut être réalisée. Dans l'idée d'identifier les méthodes d'apprentissage-machine utilisées dans les villes intelligentes, le texte *When smart cities get smarter via machine learning : An in-depth literature review* (2022) de Band et collaborateurs recensent les diverses études existantes et leur méthode. Les algorithmes qui en ressortent sont les arbres décisionnels, les machines à vecteur de support (SVM) et les réseaux de neurones artificiels. Les auteurs identifient également une catégorie d'algorithmes appartenant à l'apprentissage profond, soit de l'apprentissage non supervisé. Dans un même ordre d'idée, des auteurs s'intéressent à l'identification des méthodes non supervisées qui sont utilisées pour les analyses urbaines. À l'instar de Band, l'étude de Wang et Biljecki, *Unsupervised machine learning in urban studies: A systematic review of applications* (2022), fait une recension des différentes méthodes d'apprentissage non supervisé qui sont utilisées pour observer la ville. Les algorithmes qui en ressortent sont notamment les auto-encodeurs et leurs variantes, l'allocation latente de Dirichlet (LDA), l'analyse par regroupement hiérarchique, les réseaux antagonistes génératifs (GAN), ou encore les cartes auto-adaptatives (SOM ou *Self-Organizing Maps*). Ceux-ci seront définis plus en détail dans la section *Analyse* (Chapitre 4). Finalement, une recherche qui a poussé à explorer les différentes façons dont l'apprentissage non supervisé peut être utilisé pour les analyses urbaines et des réseaux de rues est celle de Vahid Moosavi, *Urban morphology meets deep learning: Exploring urban forms in one million cities, town and villages across the planet* (2017). L'étude fait d'ailleurs partie de la sélection qui sera analysée dans cet essai.

2.3_Questions de recherche

À la lumière de ces considérations, plusieurs questions se posent, notamment à savoir de quelle façon la machine perçoit la rue et de quelle façon elle la regarde. Cette question plutôt large, peut se diviser en sous-questionnements, notamment :

1. De quoi est composé l'œil de la machine? Quel algorithme non supervisé est utilisé et quelle est sa composition ?

Puisque l'intelligence artificielle se développe de plus en plus rapidement et trouve application dans des domaines variés, il est intéressant de se demander quelle est la composition de « l'œil » ou plutôt

³ Tel que constaté lors de la recherche d'articles pouvant être intéressants pour cet essai.

de l'algorithme qu'utilise la machine afin de regarder, ou lire la rue. Comme constaté plus haut, il existe plusieurs types d'algorithmes qui possèdent tous leurs caractéristiques, ou architecture, propre.

2. Que regarde la machine pour voir la rue? Quel est le jeu de données et quels sont ses paramètres ?

Puisque la rue peut être définie par de nombreuses caractéristiques, que ce soit sa forme, sa connectivité, les éléments qui la cadre, l'utilisation qui en est faite ou les services qu'elle dessert (et de nombreuses autres caractéristiques), il devient alors intéressant de savoir pour quels types d'analyses ou de lectures de la rue l'apprentissage-machine est utilisé et, donc, ce qui est regardé lors de ces analyses.

Les méthodes d'analyse urbaine traditionnelles pouvant se révéler comme plutôt longues et fastidieuses, elles limitent l'observation et la caractérisation des villes à des secteurs spécifiques, ce qui ne permet pas nécessairement d'obtenir une vision globale de celles-ci. Lorsque vient le temps de comparer plusieurs formes urbaines entre elles, ces méthodes sont plutôt désuètes. Il devient donc intéressant de se tourner vers l'intelligence artificielle pour analyser de vastes ensembles de données. L'utilisation des algorithmes d'apprentissage-machine non supervisé semble intéressante. Le chapitre suivant présentera donc la méthodologie utilisée afin d'identifier des études qui utilisent l'apprentissage non supervisé pour voir la ville.

3_MÉTHODOLOGIE

Pour cette recherche, un total de dix textes a été retenu pour l'analyse. Ces textes portent tous sur des études qui font l'utilisation de l'apprentissage-machine non supervisé dans le but de relever certaines caractéristiques propres à la rue ou aux réseaux de rues. Les études portent tant sur la forme des réseaux de rues d'un point de vue mondial que sur les caractéristiques de l'environnement bâti des rues qui permettent de les définir ou encore qui peuvent causer des accidents, par exemple. Cette recherche d'articles a été menée avec les moteurs de recherche, notamment Google Scholar⁴ et Web of Science⁵.

Dans un premier temps, pour la recherche d'article à l'aide des bases de données, une identification des termes pertinents a été réalisée. Les thèmes clés, comme *machine learning*, *urban design* ou encore *road analysis*, ont été déclinés en plusieurs variantes afin d'approfondir les recherches. Le tableau 1 dresse la liste de ces thèmes et de leurs variations. À noter que les recherches ont été réalisées en anglais, puisqu'elles offraient une plus grande variété de résultats que lorsque les mots étaient entrés en français.

Tableau 1. Liste des mots-clés utilisés

<i>Types de mots-clés</i>	<i>Variations</i>
<i>Machine-learning</i>	Unsupervised learning, Deep learning
<i>Urban design</i>	Urban analysis, urban planning, urban design, urban morphology
<i>Road analysis</i>	Road/street detection, road/street classification, road/street characterization, road/street segmentation
<i>Algorithms</i>	Auto-encoder, unsupervised neural network, neural network, self-organizing map, Gaussian mixture model, hierarchical clustering
<i>Data</i>	Massive data, georeferenced data, satellite imagery, Open Street Map, Google Earth

Une fois les mots-clés identifiés et entrés dans les moteurs de recherche, la base de données qui s'est retrouvée être la plus concluante, malgré le peu de mots-clés entrés, était Google Scholar. En effet, une problématique rapidement rencontrée, malgré l'utilisation de mots-clés tels que *unsupervised learning*, est qu'en temps général, le peu d'études portant sur l'apprentissage non supervisé et la ville, notamment

⁴ <https://scholar.google.com/>

⁵ <https://www.webofscience.com/wos/>

les trames de rues, se faisaient engouffrer par les études utilisant l'apprentissage supervisé, puisque cette application est plus répandue. Toutefois, la requête entrée dans Google Scholar, et écrite dans le Tableau 2, a permis d'identifier l'étude de Wang et Biljecki : *Unsupervised machine learning in urban studies : A systematic review of applications* (2022), qui fait la revue des différents types d'algorithmes d'apprentissage non supervisé et identifie plusieurs études qui en font l'usage. Puisque l'étude de Wang et Biljecki effectue une recension des écrits portant sur les algorithmes d'apprentissage-machine non supervisé, certains des textes présentés ont été identifiés comme pouvant être intéressants pour cet essai. Dans un premier temps, cet essai ne retient que les articles où les auteurs qui utilisent les algorithmes jugés plus intéressants pour des analyses urbaines portant sur la rue. Ces algorithmes sont les auto-encodeurs, l'allocation latente de Dirichlet, les réseaux antagonistes génératifs et l'analyse par regroupement hiérarchique. Ces algorithmes ont été sélectionnés puisqu'ils permettent des quantités des données qui relèvent des données massives. Certains des autres algorithmes présentés traitaient des quantités de données qui relevaient davantage des super statistiques que de l'intelligence artificielle. Chaque algorithme sera présenté plus en détail dans la section 5.

La recension de Wang et Biljecki a permis d'identifier les textes de Taecharungroj et Mathayomchan (2020) et Comber et collaborateurs (2020). Les autres études ont été trouvées dans les bases de données Web of Science et SOFIA⁶, les requêtes entrées étant inscrites dans le tableau 2.

Tableau 2. Exemple de requêtes entrées dans les bases de recherche

<i>Banque de données</i>	<i>Requête</i>
<i>SOFIA</i>	kw : ("données géoréférencées" OU géoréférencement OU analyse spatiale OU entité géographique) ET kw : ("machine learning" OU apprentissage automatique OU estimation de modèle OU intelligence artificielle) ET kw : ("rues" OU maillage urbain OU mobilité urbaine)
<i>Google Scholar</i>	Unsupervised learning AND urban design AND georeferenced data
<i>Web of Science</i>	Urban planning OR urban design OR regional planning (All fields) and Unsupervised algorithm AND (deep learning OR machine learning) (All fields)

3.2_Critères de sélection

⁶ SOFIA est le moteur de recherche de la bibliothèque de l'Université Laval et offre des services et des ressources pour les étudiants, les chercheurs et le public (<https://www.bibl.ulaval.ca/>)

À la suite des premières étapes de la recherche, plus d'une cinquantaine de textes ont été trouvés et lus de façon sommaire. Des critères de sélection ont dû être établis afin de diminuer le nombre de textes et cibler les textes qui permettent de répondre aux questions de recherche. Dans un premier temps, tous les articles se servant d'algorithmes d'apprentissage-machine, mais dont l'application relevait de l'apprentissage supervisé ont été retirés de la liste. Il était en effet possible que certains textes utilisent des algorithmes non-supervisés, mais que les auteurs passent d'abord par une étape supervisée afin de traiter les données. Ce type a donc été jugé comme moins intéressant pour cette recherche. Les articles restants ont été triés et sélectionnés selon le type de données utilisées, ainsi que la quantité de ces dernières. Par ces critères, l'objectif était d'éliminer toute recherche qui relevait davantage des statistiques, où un modèle théorique du phénomène est imposé, que de l'intelligence artificielle, où le modèle du phénomène est proposé par l'instrument algorithmique.

Il est pertinent de noter que pour une question d'actualité, seuls les articles écrits entre les cinq à dix dernières années ont été retenus, avec une préférence pour les articles les plus récents. En effet, l'intelligence artificielle de façon générale est un domaine en constante évolution. Le tableau 3 dresse la liste des différents critères de sélection observés pour chacun des articles sélectionnés. Dix articles ont finalement été choisis comme les plus pertinents à analyser dans le cadre de cette recherche.

Tableau 3. Liste des textes sélectionnés

<i>Auteurs</i>	<i>Année</i>	<i>Titre</i>	<i>Revue</i>	<i>Forme des données</i>	<i>Taille des données</i>	<i>Algorithme</i>
<i>Moosavi, V.</i>	2017	Urban morphology meets deep learning: Exploring urban forms in one million cities, towns and villages across the planet	arXiv	Images matricielles formées à partir de réseaux de rues vectorisées extraits d'Open Street Map	962 639 images de 256x256x1 pixels	Auto-encodeur convolutif
<i>Sameen, M.I., Pradhan, B.</i>	2017	A Novel Road Segmentation Technique from Orthophotos Using Deep Convolutional Autoencoders	Korean Journal of Remote Sensing	Orthophoto (Image RVB) obtenue d'un balayage laser aéroporté	1024 images de 106x106x3 pixels	Auto-encodeur convolutif
<i>Rios, S. A., Munoz, R.</i>	2017	Land Use detection with cell phone data using topic models: Case Santiago, Chile	Computers, Environment and Urban Systems	Données textuelles de l'activité des antennes d'une compagnie téléphonique non spécifiée	880 000 000 appels téléphoniques	Allocation latente de Dirichlet
<i>Kempinska, K., Murcio, R.</i>	2019	Modelling urban networks using Variational Autoencoders	Applied Network Science	Réseaux de rues vectorisées extraits d'Open Street Map	12 464 images de 64x64x1 pixels	Auto-encodeur variable
<i>Zhao, H. et al.</i>	2019	Unsupervised Deep Learning to Explore Streetscape Factors Associated with Urban Cyclist Safety	Smart Innovation, Systems and Technologies	Images satellites provenant de Google Maps et extraites de Google Maps Static API	3773 images de dimension non spécifiée	Réseau antagoniste génératif
<i>Comber et al.</i>	2020	Using convolutional autoencoders to extract visual features of leisure and retail environments	Landscape and Urban Planning	Images JPEG récoltées par des équipes de la Local Data Company	314 542 images de 224x224x3 pixels	Auto-encodeur convolutif
<i>Taecharungroj, V., Mathayomchan, B.</i>	2020	The big picture of cities: Analysing Flickr photos of 222 cities worldwide	Cities	Images provenant de Flickr.com et extraite de Flickr.com API	222 000 images de dimension non spécifiée	Allocation latente de Dirichlet
<i>Qi, X. et al.</i>	2020	NCAE and ELM Based Enhanced Ensemble Optimized Model for Traffic Flow Forecasting	IEEE	Données du trafic routier obtenues des sites gouvernementaux des États-Unis et du Royaume-Uni	Non spécifiée	Auto-encodeur à contraintes non-négatives
<i>Chao, Y. et al.</i>	2021	Data-Driven Approach for Passenger Mobility Pattern Recognition Using Spatiotemporal Embedding	Journal of Advanced Transportation	Données provenant des points de collecte de billets du métro de Beijing	176,81 millions de données sur les déplacements des passagers	Auto-encodeur empilé
<i>Wu, C. et al.</i>	2024	Machine learning-based characterisation of urban morphology with the street pattern	Computers, Environment and Urban Systems	Réseaux de rues vectorisées extraits d'Open Street Map	1054 zones locales fondées sur la rue et leurs réseaux de rues	Analyse par regroupement hiérarchique

4_ANALYSE

La section suivante fait donc une présentation sommaire de chacun des dix articles retenus (Tableau 3), présentant les auteurs, leurs questions de recherche, l'objet de leur étude et leur institution de rattachement. Par la suite, une analyse plus en profondeur suivant les questions de recherche sera réalisée. Cette dernière section (4.2) fera donc l'analyse de la composition des différents algorithmes utilisés par les auteurs, puis fera une revue des différentes caractéristiques qui sont observées par la machine lors des analyses.

4.1_Présentation des articles

4.1.1_Urban morphology meets deep learning: Exploring urban forms in one million cities, town and villages across the planet (2017)

Écrite par Vahid Moosavi, scientifique en données et chercheur senior à l'École polytechnique fédérale de Zurich (ETH), et publiée en 2017 par la Chair of Computer-Aided Architectural Design, cette étude explore l'utilisation d'un auto-encodeur convolutif profond afin d'identifier des similarités dans les trames urbaines à travers le monde. L'objectif de la recherche consiste à définir le potentiel de l'utilisation de données massives géolocalisées pour les architectes et les designers urbains. L'auteur utilise les données disponibles sur OpenStreetMap.⁷ pour recueillir les données nécessaires. Les coordonnées extraites sont utilisées afin de télécharger les images satellites des lieux sélectionnés depuis Mapbox studio, une application permettant de gérer des données géospatiales et de concevoir des cartes personnalisées. Au total, 962 639 cartes représentant des réseaux de rues de centres de villes et villages ont été téléchargées et ajustées à une taille uniforme de 256x256 pixels. L'auto-encodeur est ensuite utilisé afin de reconstruire des images et les comparer à celles d'origine. À cette étape, Moosavi note que les images reconstruites omettent certains détails fins des images d'entrées à l'échelle du quartier. En effet, tel que démontré sur la Figure 1, certains réseaux de rues plus fins et compacts ont tendance à être flous sur les images reconstruites.

⁷ Cette plateforme fournit des données libres cartographiques qui sont entrées par des cartographes bénévoles (<https://www.openstreetmap.org/>)

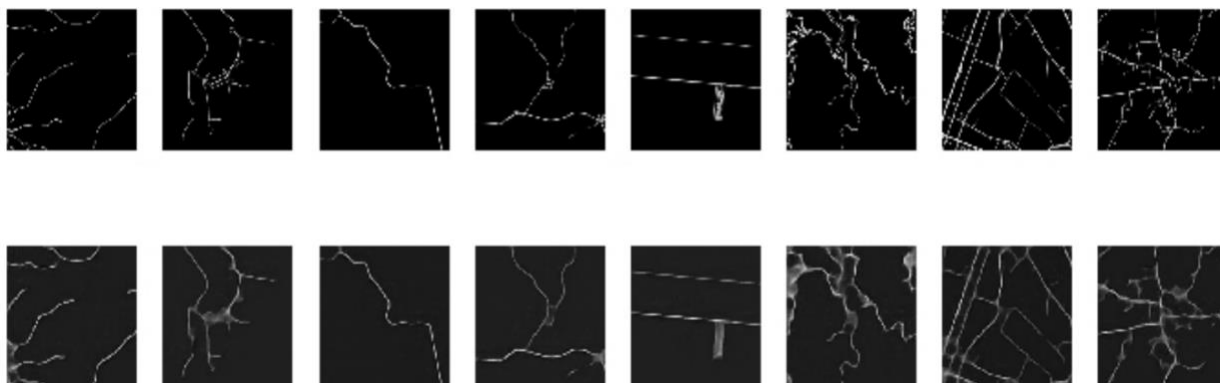


Figure 1. La ligne supérieure montre les réseaux de rues d'entrée, alors que la ligne inférieure montre les résultats obtenus par l'auto-encodeur. Certaines rues ou réseaux n'apparaissent pas sur les images reconstruites (p.4, Moosavi, 2017).

Les résultats obtenus ont été projetés sur une matrice de $8 \times 8 \times 10$, réduisant le nombre de dimensions à 640. En parallèle, une carte auto-adaptative est employée pour projeter les 640 vecteurs dimensionnels des réseaux de rues obtenus sur une grille 2D, permettant ainsi de visualiser les différentes formes urbaines et leur distribution sur une carte de couleur (Figure 2). L'objectif à long terme de cette étude est de développer un 'moteur de recherche pour les villes' permettant aux utilisateurs de découvrir des développements urbains similaires à ceux d'une ville spécifique, et ce où que ce soit sur la planète.

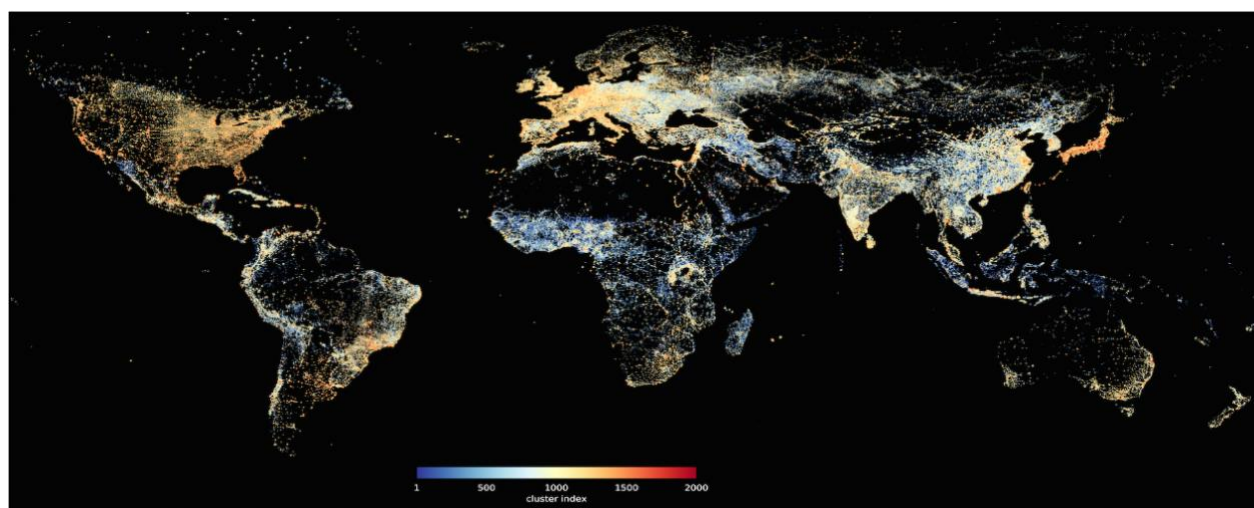


Figure 2. Distribution des formes urbaines sur l'ensemble de la planète. Les couleurs similaires indiquent des formes urbaines qui sont similaires (p.7, Moosavi, 2017).

4.1.2_A Novel Road Segmentation Technique from Orthophotos Using Deep Convolutional Autoencoders (2017)

Cet article, paru en 2017 dans le *Korean Journal of Remote Sensing*, a été rédigé par Maher Ibrahim Sameen, chercheur postdoctoral au département de génie civil de l'Université Putra Malaysia, et Biswajeet Pradhan, directeur fondateur du *Centre for Advanced Modelling and Geo-Spatial Information Systems* (CAMGIS) de l'École des systèmes, de la gestion et du leadership de la faculté d'ingénierie et de technologie de l'information de l'Université de technologie de Sydney. Dans une démarche similaire à Moosavi (2017), les auteurs s'intéressent à l'utilisation de l'apprentissage profond pour l'observation et l'identification de segments de rues à partir d'ortho-photos de haute résolution, et cherchent à évaluer la pertinence de l'utilisation d'un auto-encodeur convolutif pour cette tâche. Ils soulignent le besoin de développer un nouveau modèle d'extraction des réseaux de rues en raison des différentes limites rencontrées par les modèles actuels, notamment le besoin de données d'entraînement. En effet, celles-ci ne sont pas toujours disponibles, ce qui peut restreindre les analyses. Les auteurs démontrent donc leur technique, qui repose sur l'utilisation d'un auto-encodeur convolutif, et proposent trois méthodes d'optimisation des données : la recherche par grille, la recherche aléatoire et la méthode bayésienne. L'ortho-photo utilisée est issue d'un balayage laser aéroporté effectué au-dessus de la vallée de Kinta, en Malaisie. Elle a ensuite été découpée en 1024 images de 106 x 106 pixels. Les auteurs utilisent la recherche par grille pour l'optimisation des données de leur modèle. Après avoir analysé les résultats, ils ont pu conclure que leur méthode était capable d'apprendre automatiquement les caractéristiques essentielles à la segmentation des réseaux routiers, permettant ainsi une réduction du temps d'entrée des données par rapport aux algorithmes d'apprentissage supervisé. Cependant, ils ont observé que leur modèle semble mieux fonctionner avec les rues plus larges que celles plus étroites. Ils attribuent ce phénomène à un potentiel manque de qualité de la résolution des images utilisées ou au nombre limité de données, rendant moins aisée la capacité de l'algorithme à reconnaître les détails des segments de rues plus étroits.

4.1.3_Land Use detection with cell phone data using topic models: Case Santiago, Chile (2017)

Cet article a été publié en 2017 dans la revue *Computers, Environment and Urban Systems* par le Centre de recherche en intelligence économique du département d'ingénierie industrielle de l'Université du Chili. Les auteurs, Sebastian A. Rios et Ricardo Munoz, souhaitent identifier les différents modes d'utilisation du sol dans la ville de Santiago, au Chili. Pour ce faire, ils proposent une méthode qui utilise un algorithme d'allocation latente de Dirichlet, qu'ils appliquent à l'activité des

1183 antennes, ou stations de base émetteur-récepteur, de la ville. Les données utilisées par cette recherche proviennent d'appels téléphoniques passés dans la ville sur une période de 77 jours, pour un total de 880 000 000 appels recensés d'utilisateurs anonymes. Ces données ont été fournies par une grande entreprise de télécommunications qui n'a pas été identifiée par les auteurs, et elles contiennent des informations telles que le jour, l'heure, la durée et les coordonnées de localisation de l'antenne de communication qui a capté l'appel. L'application de cette méthode à l'algorithme d'allocation latente de Dirichlet demande une adaptation de la part des auteurs. Pour y parvenir, ils considèrent que chaque antenne est associée à un thème ou à un mode d'utilisation du sol particulier. Ces modes sont définis par les activités qui sont ici représentées par les appels téléphoniques. La nature de l'activité est interprétée grâce à la géolocalisation de l'appel. Grâce à l'utilisation de l'allocation latente de Dirichlet, les auteurs ont pu découvrir deux regroupements importants qui caractérisent l'utilisation du sol, soit les zones d'emploi et les zones résidentielles, mais ils ont également pu identifier deux nouveaux regroupements, qu'ils ont eux-mêmes interprétés comme des zones de loisirs et commerces et des zones de trafic. Plusieurs autres regroupements sont ressortis de l'analyse, mais les auteurs mentionnent ne pas avoir réussi à les identifier convenablement pour les d comme étant importants. Puisque les données ont été recueillies au fil de plusieurs jours, il a été possible d'identifier quels regroupements sont plus actifs selon les différents jours de la semaine et les différents moments de la journée.

4.1.4_Modelling urban networks using Variational Autoencoders (2019)

Dans cette recherche publiée dans la revue *Applied Network Science* en 2019 et écrite pour le Centre d'analyse spatiale avancée (CASA) du University College de Londres (UCL), Kira Kempinska, chargée de cours en intelligence artificielle géospatiale du CASA, et Roberto Murcio, chercheur senior au Urban Dynamics Lab de l'UCL, s'intéressent à la manière dont les modèles d'apprentissage machine génératifs peuvent accroître la capacité des designers urbains et urbanistes à modéliser des formes urbaines. Pour ce faire, les auteurs utilisent un auto-encodeur variable. Les réseaux de rues sont extraits de l'application OpenStreetMap sous forme de vecteurs et sont enregistrés sous forme d'image. Chaque image représente un réseau de rues d'une dimension de 3 km x 3 km sélectionné à partir du centre des villes choisies, et sont redimensionnées à une taille de 64 x 64 pixels. Au total, l'échantillon contient 12 479 images binaires (en noir et blanc). Une fois traitées par l'encodeur-décodeur, les nouvelles images générées par l'algorithme sont utilisées pour démontrer l'utilité des réseaux de rues vectorisés afin de comparer les similarités entre diverses formes urbaines. Les auteurs ont également utilisé les vecteurs pour identifier les ensembles de formes urbaines similaires. Cette étape a été réalisée à l'aide d'un

algorithme de regroupement *K-means*⁸. Cette analyse a permis entre autres aux auteurs de constater que le décodeur rencontre des difficultés pour reconstruire les détails des segments de rues denses. En effet, les images générées par le décodeur ont une faible résolution, un phénomène qui peut être attribué à un potentiel manque de données selon les auteurs. Malgré la faible résolution des images, l'étude parvient néanmoins à démontrer la capacité de l'auto-encodeur variable à générer des formes urbaines complexes et réalistes qui correspondent aux données observées.

4.1.5_Unsupervised Deep Learning to Explore Streetscape Factors Associated with Urban Cyclist Safety (2019)

Publié dans le rapport *Smart Transportation Systems* 2019 de le volume 149 de la revue Smart Innovation, Systems and Technologies, l'article de Zhao *et al.*, chercheurs en science des données urbaines à l'Université de Melbourne, s'intéresse à la caractérisation des rues dans un objectif de détermination des facteurs de risque qui augmentent la probabilité des accidents chez les cyclistes. L'objectif étant de faire des suggestions aux villes quant aux façons d'aménager le paysage urbain afin de diminuer les accidents et rendre les rues plus sécuritaires. Cette recherche utilise les données d'accidents enregistrées par les autorités routières de la ville à l'étude, soit Melbourne, et plus précisément, la zone métropolitaine du Grand Melbourne. Ces données sont accessibles publiquement via le service VicRoads⁹, une coentreprise gouvernementale de l'état de Victoria, en Australie. Les auteurs ont donc utilisé les données d'emplacement de 5 156 accidents enregistrés par les autorités de la route entre janvier 2010 et décembre 2013. Les emplacements de ces accidents ont ensuite été localisés et extraits sous forme d'images satellites avec le service de carte de Google Maps. Ils ont également été extraits sous forme de cartes vectorielles à l'aide de l'application Google Maps Static API¹⁰ (Figure x).

⁸ Voir *K-means Clustering: Algorithm, Applications, Evaluation Methods and Drawback* (<https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a>)

⁹ <https://www.vicroads.vic.gov.au/>

¹⁰ Google Maps Static API est un service qui permet de générer une carte à partir d'un URL. La carte est générée sous forme d'image. <https://developers.google.com/maps/>

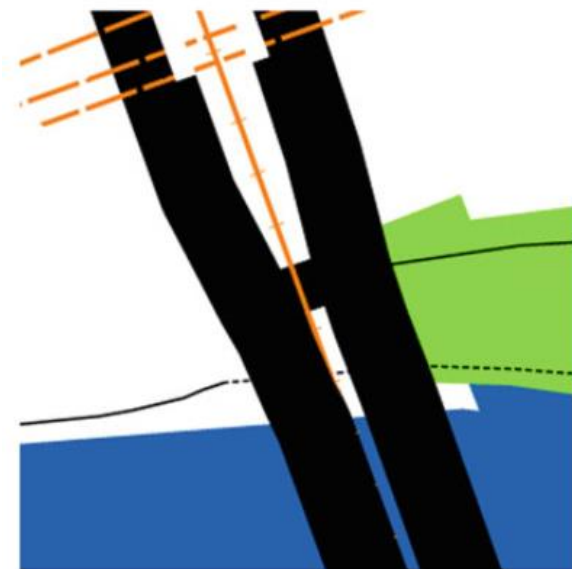


Figure 3. Exemple d'une carte vectorielle extraite de Google Maps API. Les rues sont représentées en noir, l'eau en bleu, les parcs en vert et les lignes de transport en commun en orange (p.158, Zhao et al., 2019)

Au total, l'échantillon de données est composé de 3773 images, soit 1712 images de sites où des accidents se sont produits, et 2061 images de sites où aucun accident n'a été recensé, créant ainsi deux domaines : *sans accident* et *avec accident*. Dans ce cas-ci, l'algorithme d'apprentissage non supervisé utilisé est un réseau antagoniste génératif. Celui-ci est utilisé afin de générer des images d'un domaine dans le second domaine et ainsi comparer les changements qui se sont opérés entre le résultat et l'entrée. Cette traduction d'images à partir du domaine sans accident vers celui avec accident permet aux auteurs d'identifier des facteurs qui augmentent ou diminuent les probabilités d'accidents de la route chez les cyclistes. Par exemple, ils notent que les espaces verts diminuent les risques, alors que les routes marquées par des immeubles de grande hauteur qui forment de grandes ombres ont tendance à augmenter les risques. La génération d'images depuis le domaine avec accident vers le domaine sans accident permet ensuite de confirmer que les facteurs identifiés par la première traduction augmentent ou diminuent les accidents. Finalement, la méthode utilisée permet aux auteurs de faire plusieurs observations; il est soulevé que les voies de tramway, les zones d'immeubles de grande hauteur, les intersections et le stationnement sur rue sont associés à plusieurs accidents, alors que les environnements avec des espaces verts et les rues possédant une bande médiane sont associés à une réduction du nombre d'accident. Les auteurs fournissent également les avantages à ce type de méthode, soit qu'elle aide à mieux comprendre les facteurs de risques pour les cyclistes, tout en permettant

d'établir des suggestions sur des routes qui sont plus sécuritaires et sur des façons de mieux aménager les rues afin de réduire les risques d'accidents cyclistes.

4.1.6_Using convolutional autoencoders to extract visual features of leisure and retail environments (2020)

Publié dans la revue *Landscape and Urban Planning* en 2020, cet article écrit par Sam Comber, Daniel Arribas-Bel, Alex Singleton et Les Dolega, chercheurs au département de géographie et de planification et l'Université de Liverpool, s'intéresse à l'extraction des caractéristiques visuelles des rues marchandes, à l'aide d'images de façades des équipements de détails et de loisirs, pour ultimement s'interroger sur l'utilité de l'apprentissage profond non supervisé pour l'urbanisme et la planification urbaine. Les auteurs utilisent un auto-encodeur convolutif pour réaliser leur analyse et vérifier s'il y a une corrélation entre les caractéristiques visuelles des lieux de loisir et de commerce et les caractéristiques mesurables de l'environnement bâti dans lequel ils s'insèrent. Pour ce faire, 314 542 images provenant de la Local Data Company (LDC)¹¹, un fournisseur de données sur l'emplacement des commerces et loisirs au Royaume-Uni sont récoltées. Il est à noter que ces images sont prises au niveau de la rue et qu'elles montrent les façades des bâtiments. Les auteurs utilisent également des réseaux de rues vectorisés extraits d'OpenStreetMap et couvrant un rayon de 15 minutes de marche autour de chaque établissement sélectionné. Le rayon de marche est déterminé en assumant que la vitesse de marche d'un citoyen dans ce rayon est de 4,5 km/h, ce qui équivaut à une distance équivalente d'environ 1,12 km. Au total, 21 variables sont extraites des deux sources de données, soit des variables en lien à la santé économique, la composition, la taille et la fonction ainsi que des variables sociaux-économiques. Des exemples plus précis de ces variables sont le nombre de lignes d'autobus ou de trains dans chaque zone, la proportion entre commerce de détail et services, la longueur moyenne des rues, ou encore le pourcentage d'unités de logements classées comme appartements dans la zone. Cette recherche se décompose en deux étapes, soit une première d'identification des caractéristiques visuelles des établissements sélectionnés à l'aide de l'auto-encodeur, puis un regroupement des caractéristiques qui présentent des ressemblances. Cette étape est réalisée à l'aide d'une classification ascendante où la typologie initiale composée de 250 groupes sont divisés à l'aide d'un algorithme de regroupement *K-means*. À la suite de l'analyse, cinq types de regroupements ont été identifiés, soit A (comparaison compacte), B (services éparés), C (hôtellerie haut de gamme), D (rues principales diversifiées) et E (achats quotidiens). La méthode a permis aux auteurs de déterminer des modèles de

¹¹ <https://www.localdatacompany.com/>

fonctions et d'activités commerciales, démontrant ainsi que certains environnements bâtis possèdent des caractéristiques visuelles propres. Ils concluent que leur méthode peut être éventuellement utilisée par les urbanistes ou les planificateurs urbains pour déterminer, par exemple, les modes d'usage d'un environnement bâti, ou encore, la situation socio-économique d'un environnement donné.

4.1.7_ The big picture of cities: Analysing Flickr photos of 222 cities worldwide (2020)

L'étude de Virya Taecharungroj et Boonyanit Mathayomchan a été publiée en 2020 par la revue *Cities*. Les auteurs sont chercheurs à l'Université Mahidol International et s'intéressent à l'image de marque et le marketing du lieu, ainsi qu'à la science des données et de l'optimisation. Dans cette recherche, ils utilisent un algorithme d'allocation latente Dirichlet pour identifier les différentes dimensions de l'image de la ville. L'objectif est d'introduire une nouvelle méthode pour identifier les dimensions de l'image de la ville en utilisant l'apprentissage-machine. Les auteurs définissent l'image de la ville comme l'image mentale dont s'en font ses citoyens, mais également ses visiteurs. Les données qu'ils utilisent pour cette étude sont des photos disponibles sur la plateforme de partage de photos Flickr¹². Au total, 222 000 photos provenant de 222 villes, donc 1000 images par ville, ont été extraites de la plateforme. À la suite de l'extraction des images, la première étape réalisée par les auteurs est de les traiter avec le service Google Cloud Vision API¹³ afin de les transformer en données textuelles, notamment grâce à l'option d'identification d'étiquettes (une voiture ou une statue, par exemple). Une fois les données textuelles récupérées, elles ont pu être analysées par l'algorithme afin de déterminer la dimension universelle de l'image (UCID, ou *universal city image dimension*) des 222 villes sélectionnées. De plus, une analyse de la dimension individuelle de l'image des villes de Montréal, Rome, Hong Kong et Bagdad a également été réalisée. En ce qui concerne la dimension universelle de l'image de la ville, les résultats obtenus par les auteurs montrent que celle qui est le plus souvent identifiée est celle de paysage urbain, dans laquelle les mots zone, urbain et ville sont fréquemment retrouvés. Les autres dimensions universelles identifiées sont les suivantes : paysage, transport, architecture et loisirs. En ce qui concerne les dimensions individuelles des villes, les auteurs identifient des variantes dans les quatre villes étudiées. Pour Montréal, celle qui sont identifiées sont, de la plus importante à la moins importante : paysage urbain, paysage, transport, architecture et loisirs; pour Hong Kong, il s'agit de : paysage urbain – nuit, paysage urbain – jour, transport, architecture, paysage; pour Rome : paysage

¹² Application qui permet aux utilisateurs de déposer et partager des photos dans des groupes privés ou de façon publique. Ils peuvent également identifier l'endroit où les photos ont été prises. <https://flickr.com/>

¹³ Service qui permet de collecter, d'analyser et d'extraire des informations textuelles à partir d'images. <https://cloud.google.com/vision/?hl=fr>

urbain, paysage, sculpture, architecture religieuse, architecture ancienne; et pour Bagdad : militaire, paysage, transport, architecture et loisir. De la même façon, les auteurs ont également analysé l'image des villes selon leur classement au niveau de la qualité de vie, ainsi que leur image régionale. Ils découvrent, par exemple, que pour les villes d'Afrique, qui sont moins développées économiquement, la dimension architecturale n'est pas présente, puisqu'il n'y a qu'une différence très faible entre le paysage urbain global et les bâtiments architecturaux, contrairement aux régions plus développées. Finalement, les résultats de l'analyse montrent que ce type d'algorithme permet d'identifier par de multiples façons les différentes dimensions de l'image des villes à travers le monde. L'étude permet également de comparer les villes entre elles et d'identifier les similarités et dissimilarités entre les différents points attractifs des villes.

4.1.8_NCAE and ELM Based Enhanced Ensemble Optimized Model for Traffic Flow Forecasting (2020)

Dans cette recherche publiée en 2020 par l'Institute of Electrical and Electronics Engineers (IEEE), Qi et collaborateurs, qui sont chercheurs au College of Information Science and Engineering de l'Université de Xinjiang, testent un prototype d'algorithme d'apprentissage non-supervisé afin d'analyser et prévoir les flux de circulation routiers. Le modèle proposé par les auteurs utilise une combinaison de plusieurs algorithmes pour des tâches diverses. Ils utilisent entre autres un auto-encodeur à contrainte non-négative, et les données utilisées pour tester la proposition proviennent des autoroutes 15 NB¹⁴ à Portland aux États-Unis, et les segments J6-J7, situés dans la ville de Birmingham, de l'autoroute M6.¹⁵ au Royaume-Uni. Les données sont accessibles sur les sites Internet gouvernementaux de chaque pays. Les données pour l'autoroute 15 NB ont été sélectionnées pour la période entre le 11 septembre 2015 et le 5 octobre 2015. Pour l'autoroute J6-J7, les données ont été récoltées entre le 7 mars 2016 et le 31 mars 2016. Afin de répondre à certaines contraintes, l'algorithme est couplé à un algorithme machine d'apprentissage extrême¹⁶ et d'un algorithme d'ensemble adaptatif amélioré, ce dernier s'étant révélé une méthode efficace pour améliorer la précision ou décomposer des problèmes en sous-problèmes plus simples (Qi et al., 2020). Sans définir exactement les caractéristiques identifiées par leur algorithme à la suite de l'analyse de performance, les auteurs

¹⁴ Les données ont été récupérées à l'emplacement suivant : <http://portal.its.pdx.edu/>

¹⁵ Les données ont été récupérées à l'emplacement suivant : <http://data.gov.uk/data/search>

¹⁶ De façon plus simple, les machines d'apprentissage extrême ont de bonnes capacités de généralisation, et ce à une vitesse extrêmement rapide. Pour plus d'informations, voir *Introduction to extreme learning machines* (<https://towardsdatascience.com/introduction-to-extreme-learning-machines-c020020ff82b>)

parviennent toutefois à obtenir des résultats concluants. En effet, les résultats montrent que l'algorithme parvient à identifier chacune des caractéristiques et même à définir leur périodicité sur une base journalière.

4.1.9_Data-Driven Approach for Passenger Mobility Pattern Recognition Using Spatiotemporal Embedding (2021)

Publié en 2021 dans le *Journal of Advanced Transportation*, l'article de Yu Chao et collaborateurs a pour objectif de proposer une méthode basée sur des données afin d'identifier les différents schémas de mobilité des usagers du métro de Beijing, en Chine. Les auteurs proposent l'utilisation d'un algorithme auto-encodeur empilé, qui est, comme le nom l'indique, une suite de plusieurs auto-encodeurs alignés les uns après les autres. Cette étude exploite des données de collection automatique des tarifs du métro des bornes de collecte géolocalisées. Au total, les données de déplacement de 172,81 millions de passagers ont été acquises entre les mois de septembre et octobre 2018. De plus, afin de caractériser les attributs spatiaux des stations de métro, des données de Points d'Intérêt (POI) ont été récoltées à partir de l'application Amap¹⁷, un fournisseur chinois de contenu cartographique numériques, de solutions de navigation et de services de localisation. Les données géolocalisées permettent de convertir les POI¹⁸ bruts en données pondérées. Trois étapes guident cette recherche, soit une première étape qui consiste en la fusion de données géoréférencées provenant de plusieurs sources dans le but de donner un poids aux POI. La seconde étape consiste en l'utilisation d'un auto-encodeur empilé pour transformer les vecteurs de déplacement à haute dimension en vecteur dense à basse dimension. En effet, le moment entre le temps d'entrée dans la station d'origine et le temps de sortie à la station d'arrivée¹⁹ peut être représenté par un vecteur. Finalement, la troisième étape consiste en l'utilisation de l'algorithme DBSCAN (density-based spatial clustering of applications with noise), un algorithme de partitionnement de données²⁰, afin d'identifier les schémas de mobilité. Pour ce faire, ils ont déterminé une plage de valeurs de [8, 18] et [7,10]. Grâce à cette méthode, les auteurs parviennent à identifier six modèles de mobilité se distinguant par le temps de déplacement, le moment de la journée

¹⁷ <https://www.amap.com/>

¹⁸ les points d'intérêt sont des unités d'information géographique qui permettent d'indiquer l'emplacement des activités humaines (Yu *et al.*, 2021). Ces points peuvent autant être des quartiers résidentiels que des lieux de consommation, ou encore des parcs ou d'emploi

¹⁹ Les heures d'entrée et de sortie sont enregistrées par les bornes de collection automatique des tarifs des stations d'origine et d'arrivée. Celles-ci fournissent également des informations sur le jour de la semaine et le temps de déplacement (Yu *et al.*, 2021)

²⁰ Voir *A Practical Guide to DBSCAN Method* (<https://towardsdatascience.com/a-practical-guide-to-dbscan-method-d4ec5ab2bc99>)

et la raison du déplacement: C1, qui représente des déplacements de longue distance pour le travail dont les départs sont réalisés entre 7 et 8h le matin et dont les temps de déplacements font entre 40 et 80 minutes, C2 et C4, qui représentent la mobilité résidentielle, C3, qui englobe la mobilité associée au divertissement et dont les départs se font entre 9h et 17h, C5 qui révèle la mobilité de courte distance pour le travail entre 7 et 9h le matin, et C6, qui représente des mobilités dont l'objectif est plus difficile à identifier. Ces six modèles ont été identifiés en utilisant les valeurs de paramètres de 16 et 9,5. Finalement, les auteurs découvrent notamment par cette méthode que plus le nombre de regroupements est élevé, plus les modèles ont des caractéristiques détaillées. Toutefois, ils constatent que cela amène également un plus grand nombre de caractéristiques inexplicables.

4.1.10_Machine learning-based characterisation of urban morphology with the street pattern (2024)

Cette étude a été publiée en 2024 dans la revue *Computers, Environment and Urban Systems*. Cai Wu, chercheur à la faculté des sciences de la géo-information et de l'observation de la Terre (ITC) de l'université de Twente aux Pays-Bas, et ses collaborateurs proposent une approche utilisant l'apprentissage-machine afin de répondre au manque de caractéristiques utilisées lors d'analyses plus traditionnelles du tracé des rues, ce qui limite la compréhension de la diversité et de la complexité des réseaux de rues. En effet, en se basant sur une sélection de 9 articles, les auteurs argumentent que les modèles utilisés pour définir les rues reposent sur des caractéristiques visuelles, soit la direction, la courbure des rues, le nombre de rues sans issues, ou encore le nombre de rues connectées par un carrefour, une représentation simplifiée qui ne peut refléter la diversité et la complexité de la morphologie urbaine au travers de la planète (Wu et al., 2024). Ces observations mènent les auteurs, à s'interroger sur l'utilisation d'algorithmes d'apprentissage-machine, et s'intéressent donc à l'utilisation d'un algorithme d'analyse par regroupement hiérarchique pour identifier les modèles de rues de différentes villes à travers le monde et en faire ressortir la diversité et la complexité. Pour ce faire, ils utilisent une unité d'analyse se basant sur des zones locales fondées sur la rue (street-based local area, ou SLA en anglais) dont les données sont définies par quatre caractéristiques fondamentales, soit (1) basées sur les rues, (2) topologiques/configurationnelles, (3) discrètes, et (4) de taille appropriée. Les réseaux de rues qui composent ces zones sont extraits d'OpenStreetMap, pour un total de 1054 zones locales fondées sur la rue qui sont réparties au travers de six villes : Amsterdam, Chengdu, Londres, Séoul, Houston et New-York. Ces villes ont été sélectionnées pour leurs différences historiques, économiques, culturelles et politiques. Cette recherche se divise en trois étapes, soit la représentation

des différentes zones locales retrouvées au travers de la trame urbaine, la caractérisation des différents réseaux de rues au sein de ces zones, et finalement le regroupement des réseaux similaires entre eux à l'aide de l'analyse par regroupement hiérarchique. Afin de caractériser chaque réseau, les auteurs ont générés 13 paramètres d'analyse, qu'ils classent en trois catégories, soit composition, configuration et explicatif. La catégorie *Composition* regroupe les paramètres *Street Length*, *Diameter*, *Circuitry*, *Orientation* et *Entropy*. Ceux-ci font la description physique des réseaux de rue. Les paramètres du groupe *Configuration* (*k_avg*, *Self-Loop*, *L-junction*, *T-junction*, *X-junction*) servent à identifier le potentiel de connectivité des réseaux, et donc les déplacements au sein de la trame urbaine. Les paramètres de la catégorie *Explicatif* (*explanatory* en anglais) comprennent : *Degree Pearson*, *Transitivity*, *Global reaching centrality* et *global efficiency*. À la suite de leur analyse, les auteurs ont identifié quatre types morphométriques majeurs; type I, type II, type III et type IV. S'appuyant sur quatre auteurs, ils associent ces types aux modèles traditionnels utilisés pour les analyses urbaines, soit la grille, la forme organique, déformée, ou le cul-de-sac. Si le type I s'associe visuellement à la grille et le type II à la forme organique, le type III est plutôt un hybride entre ces deux formes. Pour le dernier type, les auteurs l'identifient comme un 'arbre'. Il est à noter que les types I, II et IV englobent également des variantes. Par cette étude, les auteurs ont pu identifier les types morphométriques les plus dominants dans les six villes sélectionnées, et le degré de diversité de leurs trames urbaines. Par exemple, ils soulèvent que Londres possède une trame assez unifiée (avec un degré de diversité plutôt faible) appartenant au type II, alors Houston possède des SLA assez diversifiées (un degré de diversité plutôt élevé).

4.2_Analyse des articles

Cette prochaine section fera l'analyse des articles de façon plus approfondie et se divisera en deux parties principales. La première regardera la composition de l'œil de la machine, ou de chaque algorithme utilisé par les articles sélectionnés, pour ensuite s'intéresser aux éléments qui sont observés par celui-ci. En effet, dans le cadre de cet essai, l'intelligence artificielle se définit comme le couplage des algorithmes d'apprentissage-machine et des données massives nécessaires à leur fonctionnement. Il est donc pertinent d'observer de façon plus approfondie ces deux éléments.

4.2.1_Composition de l'œil de la machine

Un terme qui est utilisé de façon récurrente dans les textes pour décrire la forme des algorithmes est celui d'architecture. Puisque les auteurs utilisent des algorithmes variés, il est pertinent de s'intéresser à cette architecture et d'en ressortir les variantes afin de les comparer entre elles. La première architecture d'algorithme qui sera analysée est celle des auto-encodeurs, soit la méthode non-supervisée

qui s'est révélée être la plus utilisées parmi les auteurs. Les variantes utilisées sont les auto-encodeurs convolutifs (Moosavi, 2017; Sameen et Pradhan, 2017; Comber *et al.*, 2020), les auto-encodeurs variables (Kempinska et Murcio, 2019), les auto-encodeurs à contraintes non-négatives (Qi *et al.*, 2020) et les auto-encodeurs empilés (Chao *et al.*, 2021). Outre les auto-encodeurs, les réseaux antagonistes génératifs (Zhao *et al.*, 2019), l'allocation latente de Dirichlet (Rios et Munoz, 2017; Taecharungroj et Mathayomchan, 2020) et l'analyse par regroupement hiérarchique (Wu *et al.*, 2024) sont des algorithmes qui sont également utilisés et qui seront analysés dans cet essai.

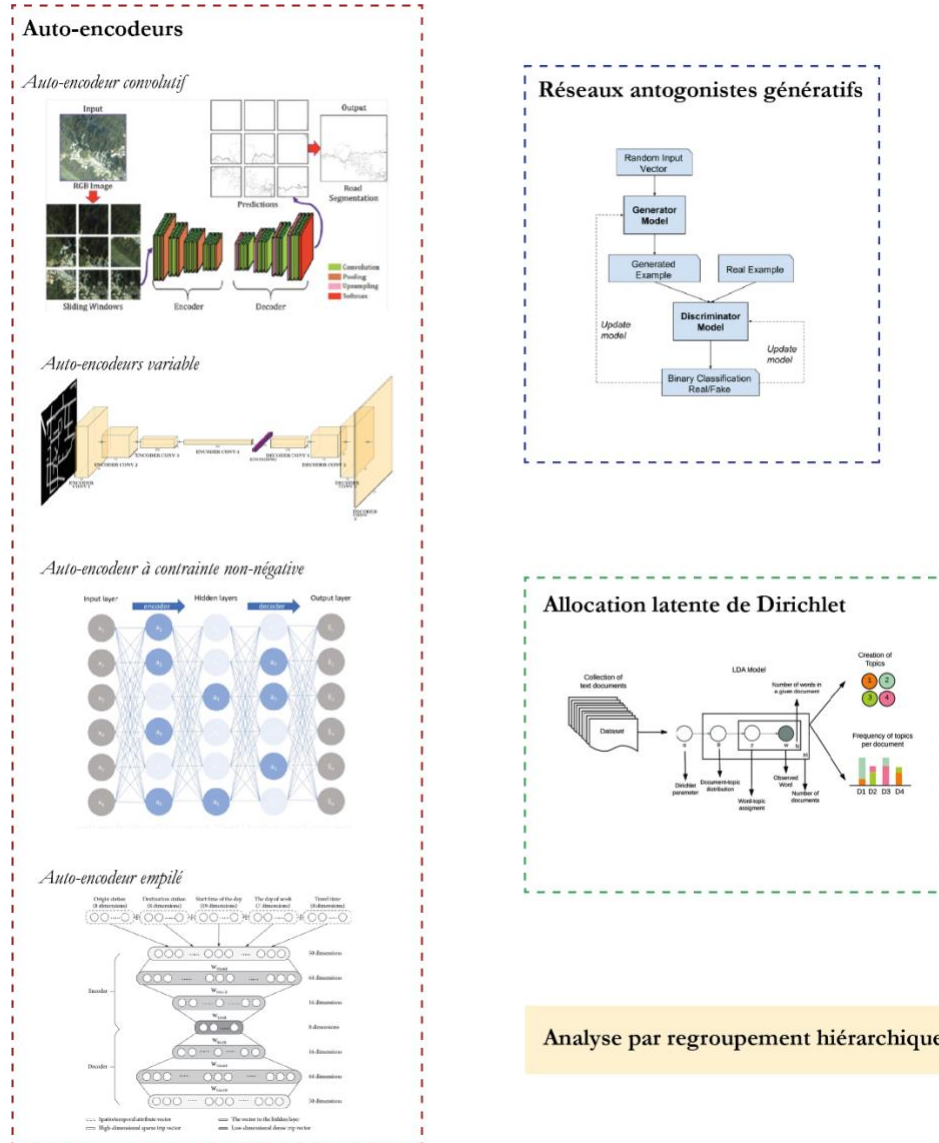


Figure 4. Synthèse des algorithmes utilisés par les divers auteurs.

4.2.1.1 *Auto-encodeurs et leurs variantes*

Un **auto-encodeur** est un réseau neuronal qui apprend à représenter des données d'entrée en les reconstruisant à l'aide des caractéristiques qui ont été apprises au cours d'un processus de compression et de décompression de l'information. Les résultats obtenus à la sortie sont donc une reconstruction des données de départ. La performance du modèle est analysée en comparant ces données. Les auto-encodeurs peuvent également faire des prédictions en générant des données similaires à celles d'origine (Wang et Biljecki, 2022). L'architecture du modèle se présente ainsi : un encodeur, dans lequel les données d'origine sont entrées, d'un espace latent, puis d'un décodeur, duquel sortent les données de sortie reconstruites (Figure 4). L'encodeur et le décodeur sont composés de plusieurs couches. L'encodeur est composé de couches de convolution qui viennent convertir tous les pixels présents dans leur champ de vision en une seule et unique valeur, et de couches de sous-échantillonnage²¹. Celles-ci servent à extraire les caractéristiques importantes des données d'origine. Le décodeur, quant à lui, est composé de couches de déconvolution et de sur-échantillonnage²² qui servent à la reconstruction des données. L'espace latent sert à la représentation en dimension réduite des caractéristiques extraites des données qui ont été traitées par l'encodeur avant d'entrer dans le décodeur. Ce type d'algorithme peut être employé pour regrouper des éléments similaires à partir d'un vaste ensemble de données issues de photographies satellites. Par exemple, Sameen et Pradhan (2017) insèrent dans l'algorithme plusieurs images satellites afin d'en extraire les segments de rues. Ici, l'algorithme sert donc à identifier et reconstruire toutes les données des images qui se rapportent aux réseaux de rues. La figure 4 illustre le processus et l'architecture de leur modèle. On y note la présence d'une dernière couche, nommée Softmax, qui est en fait une couche de classification.

²¹ *Pooling*, en anglais

²² *Upsampling*

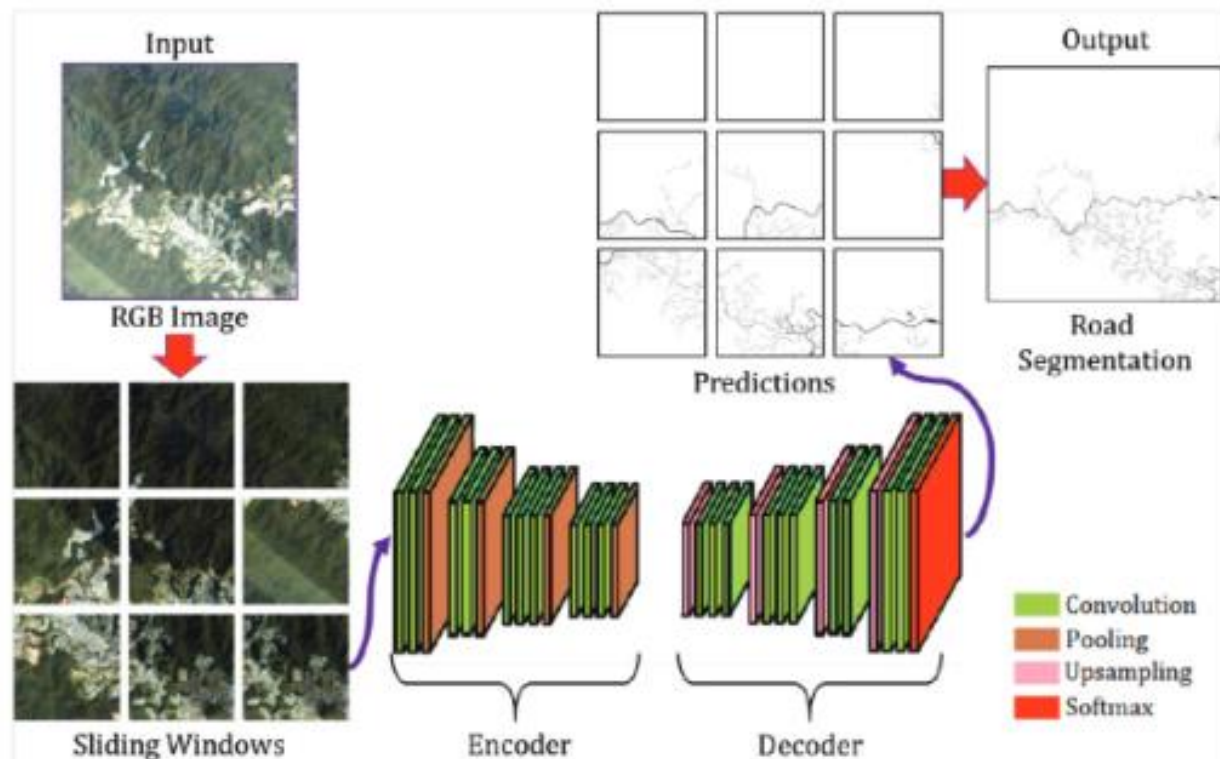


Figure 5. Architecture d'un auto-encodeur convolutif où Input représente l'image d'origine, qui est ensuite divisée en plusieurs images plus petites qui sont entrées dans l'encodeur. Le décodeur utilise les données fournies par l'encodeur pour reconstruire les images qui sont classées par la couche Softmax (p.4, Sameen et Pradhan, 2017)

Dans le cas de Moosavi (2017), l'architecture de l'algorithme est similaire, puisqu'également composée de couches de convolution, mais dans son cas, les couches de regroupement sont remplacées par des intervalles. Ceux-ci s'alternent avec les couches convolutives et permettent de diviser la largeur et la longueur des images en deux après chaque convolution dans l'encodeur puis le décodeur²³ (Moosavi, 2017), permettant à l'algorithme d'identifier les nuances locales de l'image.

En plus d'utiliser un **auto-encodeur convolutif**, Moosavi (2017) fait également usage d'une carte auto-adaptative (*Self-Organizing Map*). Il s'agit d'un réseau neuronal qui s'inspire des réseaux de neurones du cerveau humain. Il se compose d'un réseau de neurones bidimensionnels qui cherche progressivement à se former aux données entrées dans l'algorithme afin d'arriver à une représentation bidimensionnelle fidèle aux données d'entrées à n-dimensions. Chaque cas (chez Moosavi, ce sont des images) est assigné au neurone le plus proche avec les autres cas similaires. Le réseau neuronal peut être visualisé comme

²³ Le décodeur est ici seulement utilisé pour tester le modèle. Moosavi extrait les données de l'espace latent pour l'étape suivante de son étude.

une grille à $n \times n$ dimensions qui affiche la distribution des cas selon leur similarité. Dans le cas de Moosavi, par exemple, la visualisation bidimensionnelle du réseau neuronal de la SOM révèle le degré de nuances que l'algorithme a pu apprendre sur les images de trames urbaines encodées en amont par l'auto-encodeur (Figure 5).

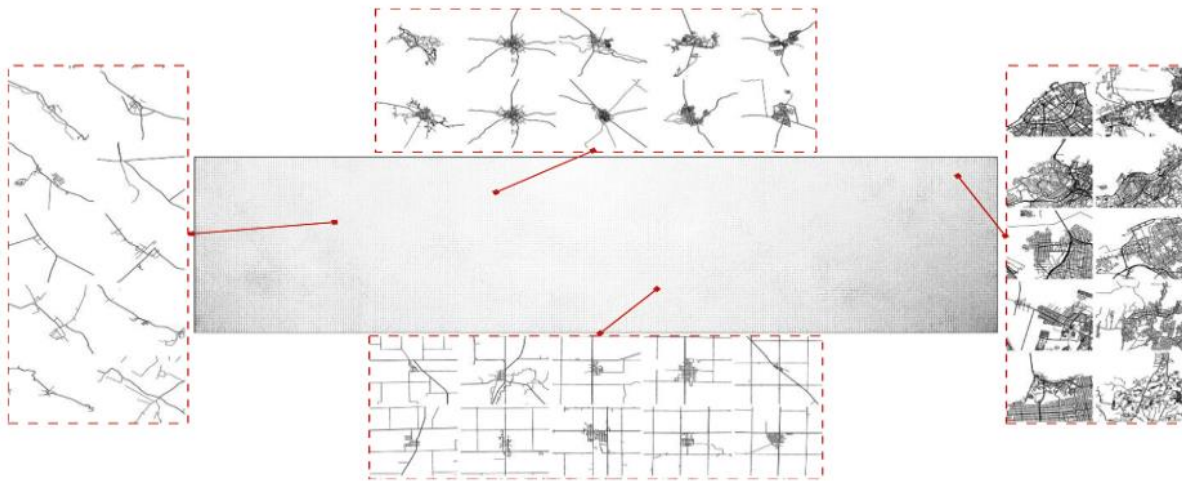


Figure 6. Le spectre des formes urbaines généré par une carte auto-adaptative (p.6, Moosavi, 2017)

Une seconde variante des auto-encodeurs est celle des **auto-encodeurs variables**. Cette variante est utilisée par Kempinska et Murcio (2019), qui s'intéressent à la manière dont les modèles d'apprentissage machine génératifs peuvent accroître la capacité des designers urbains et urbanistes à modéliser des formes urbaines. Simplement, un auto-encodeur variable fonctionne comme un auto-encodeur classique, soit à l'aide d'un encodeur-décodeur, entre lesquels s'y trouve un espace latent. Cet algorithme est entraîné pour minimiser les erreurs de reconstruction des données d'entrée. La différence avec les auto-encodeurs classique vient du fait que l'encodeur convertit les données d'entrées en une distribution de points dans l'espace latent, au lieu de les coder comme un seul point. Cela permet d'éviter un surajustement des données, garantissant ainsi que l'espace latent possède toutes les caractéristiques favorables à la génération de nouvelles données. À la différence de l'auto-encodeur convolutif, chaque couche de convolution est suivie d'une couche d'unité linéaire rectifiée (ReLU), ce qui introduit la non-linéarité, ou flexibilité, dans le modèle (Kempinska & Murcio, 2019). Cette couche d'activation permet d'assurer un entraînement plus rapide de l'algorithme, ainsi qu'un apprentissage

plus efficace. Moosavi (2017) utilise également une couche ReLU dans l'architecture de son auto-encodeur.

La figure 6 illustre l'architecture de l'auto-encodeur variable utilisé par Kempinska et Murcio. Les blocs jaunes représentent les différentes couches de convolutions et les couches d'unité linéaire rectifiée, et la dimension de ces blocs représentent la dimension de sortie. Le bloc violet, entre l'élément encodeur et l'élément décodeur, représente l'intégration qui a été apprise — l'encodage. L'encodage est ce que l'algorithme a appris sur l'ensemble de données sur lequel il est entraîné. Comme démontré sur la figure, l'image est insérée dans l'algorithme et passe d'abord par plusieurs couches de convolutions dans l'encodeur. C'est au niveau de l'intégration, ou espace latent, que l'algorithme va produire plusieurs distributions probables. Finalement, le décodeur, qui prend en compte la variabilité, vient reconstruire les données. Avec un tel algorithme, les auteurs arrivent à assurer une fidélité (un minimum de perte d'information) entre l'étape d'encodage et l'étape de décodage, ce qui permet de générer des réseaux de rues plausibles.

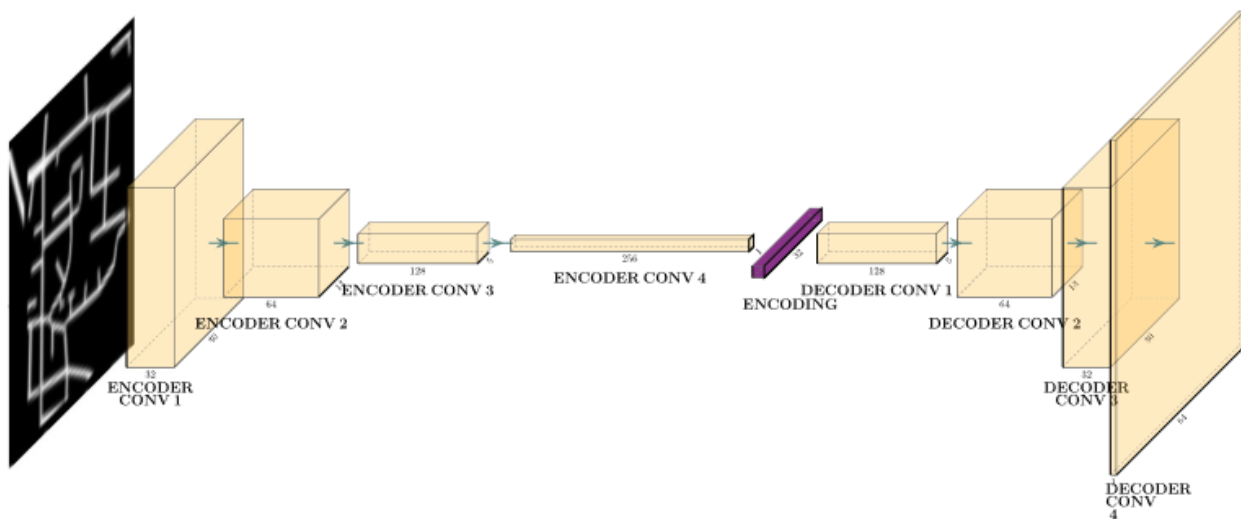


Figure 7. Architecture d'un auto-encodeur variable (p.4, Kempinska et Murcio, 2019)

Une autre variante des auto-encodeurs est celle utilisée par Qi *et al.* (2020), soit un **auto-encodeur à contrainte non-négative**, ou auto-encodeur dispersé. Ce type d'auto-encodeur limite le nombre de nœuds qui sont activés dans chacune de ses couches, à la différence d'un auto-encodeur classique. Cela permet à l'algorithme d'apprendre les caractéristiques les plus importantes et significatives des données d'origine et limite la présence d'information moins pertinentes ou utiles pour ce qui est souhaité d'être

analysé. En d'autres termes, c'est un peu comme s'il était souhaité que l'algorithme se spécialise sur un seul type de donnée et, pour l'aider, seuls les nœuds contenant des données pertinentes lui sont montrés. La figure 7, tirée d'un article écrit par Syoya Zhou et publié sur le site Medium.com, illustre le processus. Les pastilles bleu foncé représentent les nœuds activés, alors que celles qui sont plus pâles représentent l'information qui est désactivée. L'algorithme considère donc uniquement l'information des pastilles bleu foncé. Dans le cas de l'étude menée par Qi et collaborateurs (2020), il est souhaité par les auteurs de développer un modèle pouvant assurer la prédiction du mouvement du trafic de façon précise dans le but de répondre à la problématique du trafic. Les caractéristiques significatives sur lesquelles il est souhaité que l'algorithme s'arrête sont les périodes où la circulation est plus dense (les moments où il y a des ralentissements). Toutes les autres caractéristiques non pertinentes à ce qui est souhaité d'être regardé sont donc désactivées, ou rendues invisibles à la machine. Seules les informations pertinentes lui sont visibles, lui permettant alors d'être plus précise dans ses résultats.

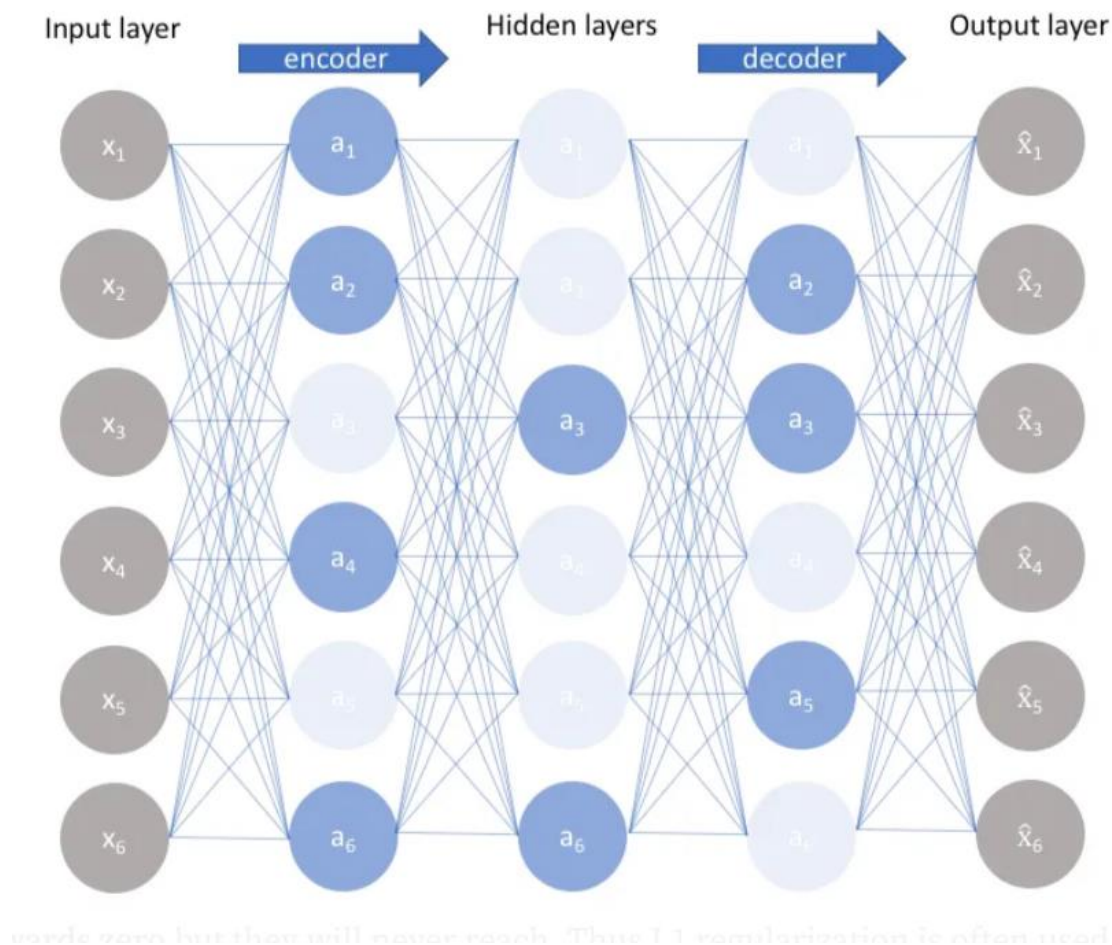


Figure 8. Architecture d'un auto-encodeur dispersé (Syoya Zhou, Medium.com)

La dernière variante qui est utilisée est celle de l'**auto-encodeur empilé**, tel que présenté par Chao et collaborateurs (2021). L'architecture reste simple : un encodeur et un décodeur, et leurs couches respectives. Une machine qui utilise un algorithme d'auto-encodeur empilé est donc composée de plusieurs auto-encodeurs. Les données d'entrées seront fournies au premier auto-encodeur, puis les données de sorties deviendront les données d'entrée du second, et ainsi de suite, selon le nombre d'auto-encodeurs qui sont utilisés.²⁴ Ce type d'auto-encodeur est utilisé dans les cas où un auto-encodeur simple ne parvient pas à réduire la dimension des données d'entrées, dû à leur relation complexe entre chaque caractéristiques.²⁵ L'architecture spécifique de l'algorithme utilisé par Chao et collaborateurs est illustrée à la figure 8. Les données, soit la station d'origine, la station de destination, le moment de départ, le jour de la semaine et le temps de déplacement, sont toutes liées entre elles et donc complexes. Leur traitement dans l'auto-encodeur permet d'en réduire la dimensionalité et donc d'en réduire la complexité.

²⁴Voir *Stacked autoencoders: extract important features from data using deep learning* (<https://towardsdatascience.com/stacked-autoencoders-f0a4391ae282>)

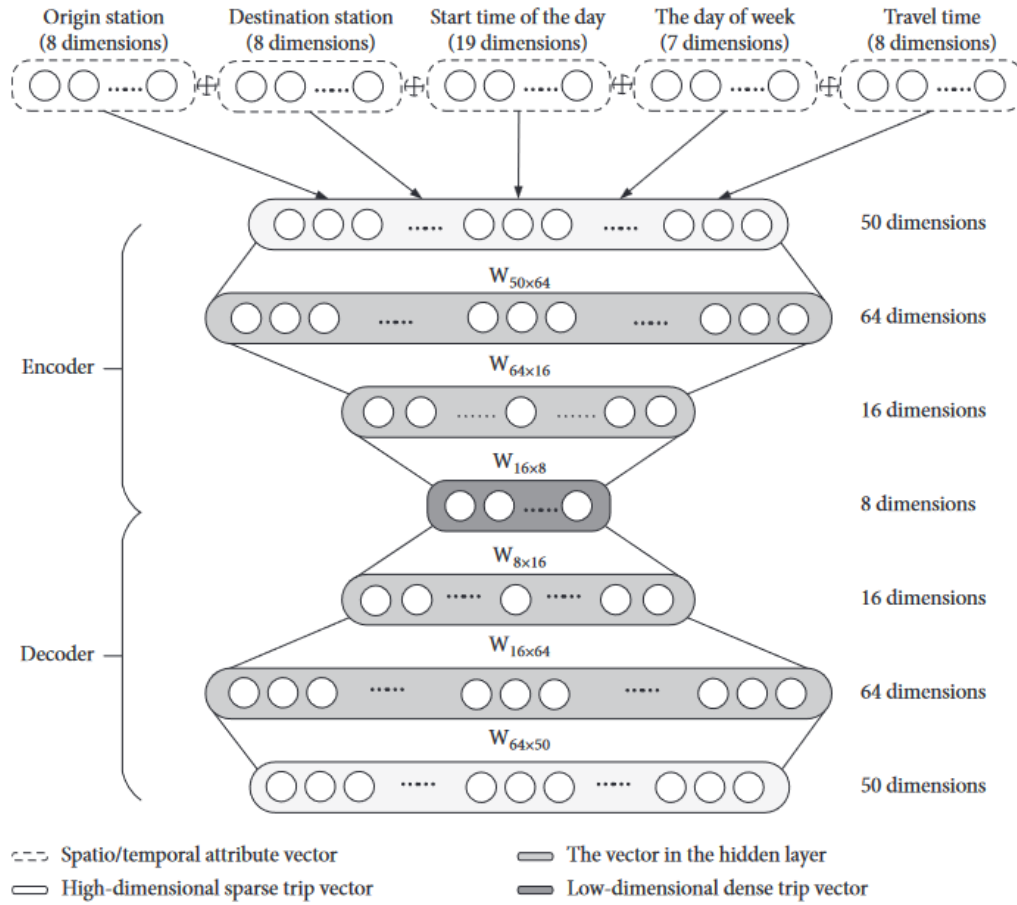


Figure 9. Architecture d'un auto-encodeur empilé (p.7, Yu et al., 2021)

4.2.1.2_Réseaux antagonistes génératifs (GAN)

Les réseaux antagonistes génératifs utilisent un discriminateur, soit une image réelle provenant de l'ensemble de données traitées, dans le but de produire des signaux d'auto-supervision pour affiner les résultats du générateur, soit des images aléatoires, ou bruit (Wang et Biljecki, 2022). Le générateur crée donc de fausses images, alors que le discriminateur utilise les données réelles dans le but de faire la distinction entre information fausse et information juste. Les deux composantes s'entraînent entre elles en tentant, d'un côté, de produire des images fausses qui tentent au mieux de répliquer la réalité et, d'un autre côté, en tentant d'identifier les images réelles sans se tromper. Le but du générateur est donc de créer des images si réalistes que le discriminateur ne peut plus les identifier comme étant fausses. La figure 9 illustre ce processus. Zhao et collaborateurs (2019) soulèvent que, lorsqu'un réseau antagoniste génératif est bien entraîné, il lui est possible de faire des translations d'image, soit d'insérer

une image d'un domaine afin de la reconstruire dans son propre domaine, mais aussi d'en générer une nouvelle dans un autre domaine. En prenant leur étude comme exemple, une image d'un fragment de rue où des accidents cyclistes se sont produits peut être insérée dans le générateur et le résultat fait une translation vers le domaine de segments de rues sans accident. Cela permet aux auteurs de comparer les différences entre chaque image avant et après la translation et donc, dans ce cas-ci, identifier les éléments pouvant causer des accidents cyclistes.

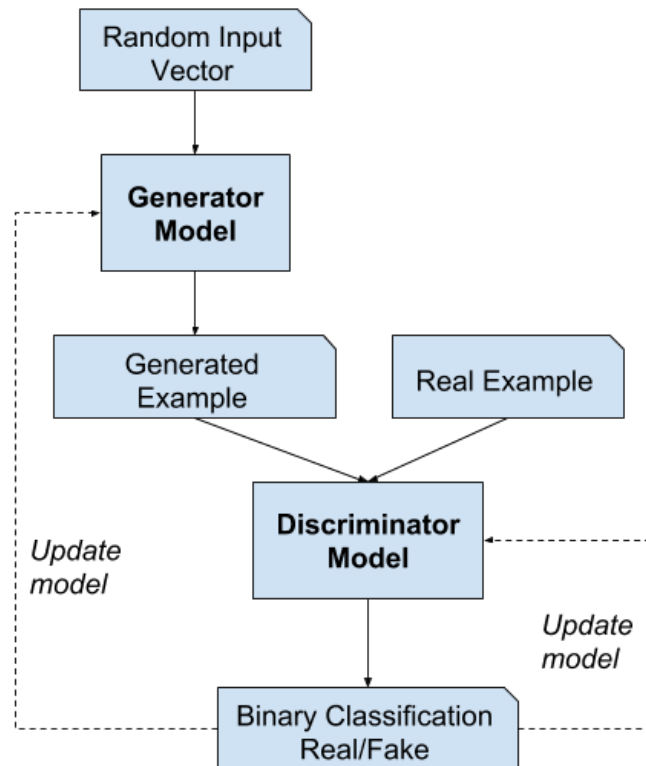


Figure 10. Exemple de l'architecture d'un réseau antagoniste génératif (source : Jason Brownlee, machinelearningmastery.com)

4.2.1.3 Allocation latente de Dirichlet (LDA)

L'allocation latente de Dirichlet permet d'identifier des sujets dans de grands ensembles de données entrées sous forme de texte. Par exemple, en prenant un livre d'urbanisme, les thèmes pouvant en ressortir pourraient être l'aménagement ou encore la mobilité. L'algorithme vient donc faire une analyse de tous les documents qu'on y insère et tente d'en ressortir les thèmes et les différents mots qui y sont associés. Cette première étape permet à l'algorithme d'apprendre les mots qui sont les plus intéressants selon les différents thèmes et, éventuellement, d'identifier le sujet de nouveaux documents. La figure 9 schématise ce processus. En prenant pour exemple l'étude de Taecharungroj et

Mathayomchan (2020), qui tente d'identifier les dimensions de l'image des villes, le fonctionnement de l'algorithme peut être expliqué. Dans un premier temps, les auteurs utilisent Google Cloud Vision API pour transformer les images récupérées sur Flickr.com en données textuelles. Ce sont ces données qui sont entrées dans l'algorithme (*dataset* sur la figure). L'algorithme va ensuite associer chaque mot à un sujet, créant ainsi un nombre x de sujets qui sont retrouvés dans les données. Il identifiera également la fréquence de chaque sujet dans chaque document. Les documents, dans le cas de cette étude, peuvent être représentés par les villes, et chaque ville se verra associer un certain nombre de sujets. Par exemple, les auteurs identifient les sujets *cityscape*, *landscape*, *transport*, *architecture* et *recreation*.

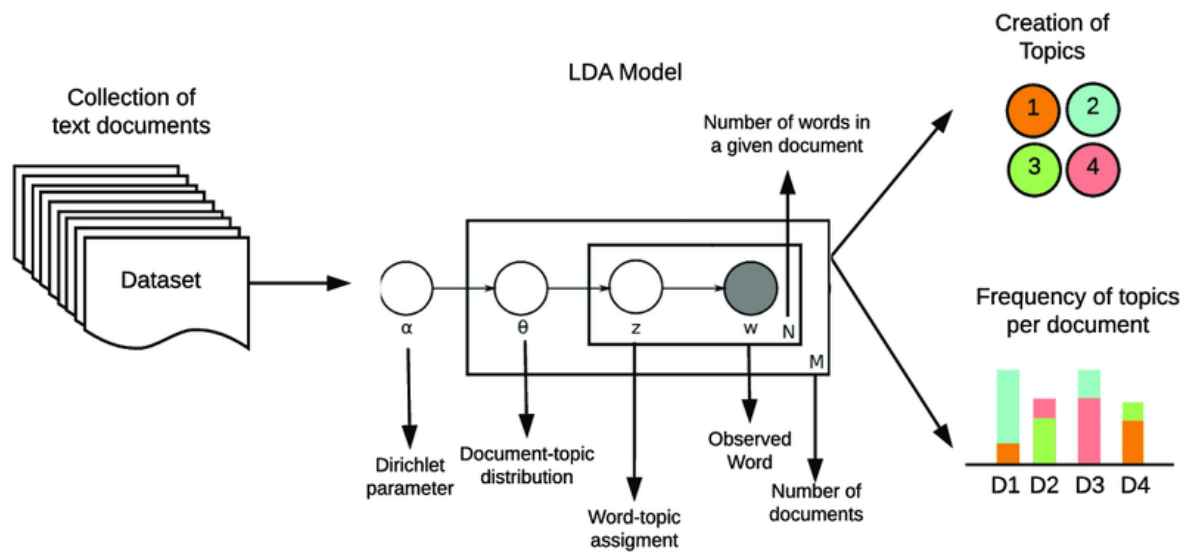


Figure 11. Fonctionnement de l'allocation latente de Dirichlet (Source : Medium.com)

4.2.1.4 _Analyse par regroupement hiérarchique (HCA)

Finalement, l'analyse par regroupement hiérarchique se compare à un dendrogramme, ou une structure arborescente. Comme son nom l'indique, cet algorithme vient organiser les données de façon hiérarchique. Ainsi, les données qui sont similaires sont regroupées entre elle, ce qui vient donc former plusieurs groupes et sous-groupes de données présentant des similarités, un peu comme si des objets étaient classés selon leur couleur et leur forme, ou encore leur fonction. Un grand groupe peut donc être divisé en plusieurs sous-groupes selon le nombre de caractéristiques différentes observées. Dans le cas de Wu et collaborateurs (2024), ce type d'algorithme a été utilisé afin de regrouper entre elles les trames urbaines similaires dans six villes, soit Chengdu, Amsterdam, New York, Séoul, Londres et Houston. Quatre grands types, soit des trames sous forme de grille, organique, déformée ou en cul-de-sac ont été identifiés dans chacune des villes, ainsi que plusieurs sous-groupes pour chaque type.

Les auteurs peuvent ensuite identifier les types qui ont tendance à se retrouver dans toutes les villes et ceux qui sont plus spécifiques à certaines villes en particulier.

4.2.2_Éléments observés

Plusieurs formes de données peuvent être utilisées par les chercheurs afin de produire des analyses urbaines grâce à l'apprentissage-machine. Les textes présentés utilisent notamment des images tant satellites que prises au niveau de la rue, des réseaux de rues vectoriels transformés en images, ou encore des données provenant de téléphones portatifs ou de points de collecte de billets.

Dans le cas d'images satellites, la machine peut faire l'identification des routes, des zones agricoles ou urbaines, des zones forestières ou encore des zones hydriques, comme les ruisseaux ou les lacs, par exemple. La figure suivante, provenant de la recherche de Sameen et Pradhan (2017) montre les résultats obtenus à la suite de l'insertion d'images provenant d'une ortho photo. Dans ce cas-ci, l'œil de la machine fait l'identification de pixels de couleurs (RVB) afin de voir les fragments de rues.

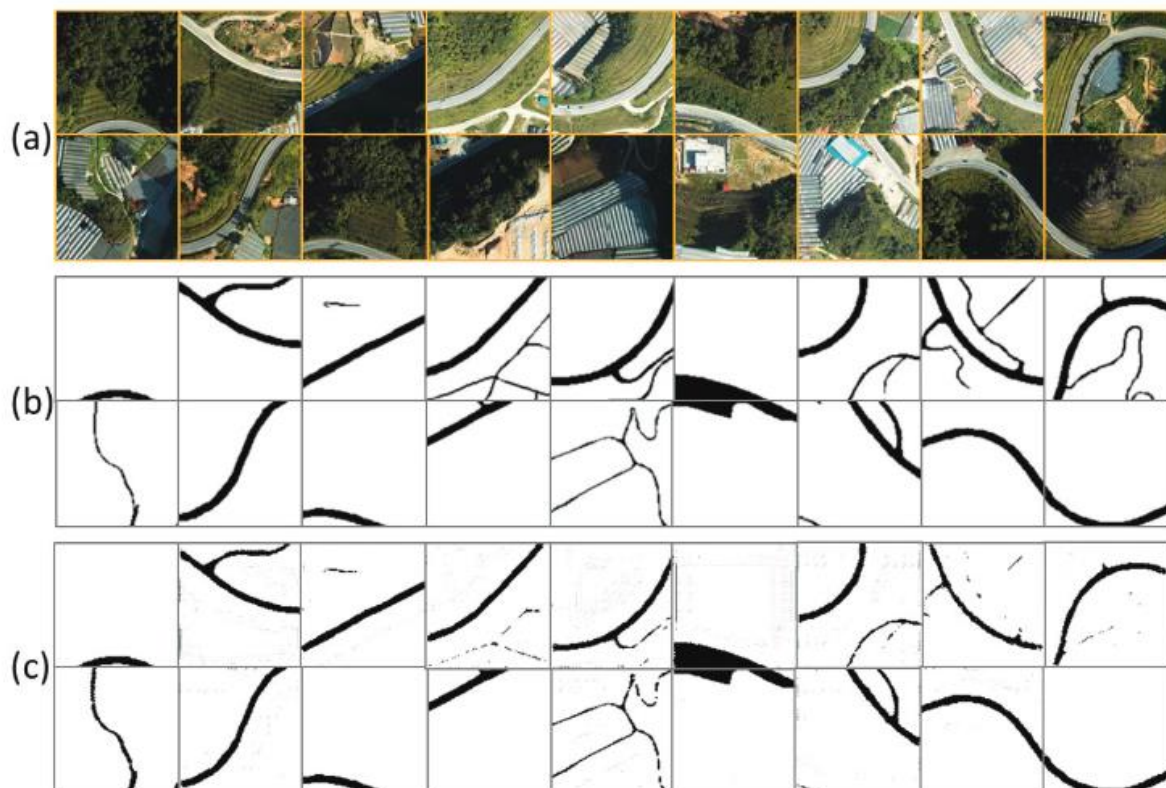


Figure 12. Résultats obtenus (c) en comparaison avec les données d'entrée (a) et les mêmes réseaux de rues de référence obtenus par association de fragments d'images dans GIS par digitalisation manuelle (b) (p.9, Sameen et Pradhan, 2017).

La grille (a) montre les images des segments de rues, alors que (b) montre les mêmes segments de rues extraits par digitalisation manuelle de GIS. Ces dernières ont été créées afin de servir de référence pour comparer les images obtenues à l'aide de l'algorithme. La dernière grille, (c), montre les résultats obtenus par l'algorithme. Dans ce cas-ci, la machine regarde donc la présence de segments de rues sur les données d'entrée (a). Les résultats obtenus par l'algorithme montrent toutefois que celui-ci, bien qu'il observe spécifiquement les segments de rues, a davantage de difficulté à voir les rues qui sont plus étroites. Cela se constate lors de la comparaison avec les segments vectorisés; les résultats fournis par l'auto-encodeur donnent des segments légèrement effacés ou inexistantes lorsqu'il s'agit de rues plus étroites. Les auteurs expliquent toutefois que ce phénomène peut être dû à un nombre insuffisant de données ou encore à une résolution trop faible des images.

De façon similaire, si, comme le font Moosavi (2017), Kempinska et Murcio (2019) et Wu et collaborateurs (2024), les images insérées dans la machine sont des réseaux de rues vectorisés, celle-ci regardera les réseaux de rues comme une distribution de pixels blancs sur un fond noir (Figure 12). L'organisation de ces pixels est également regardée afin de classer les réseaux de rues selon leurs similarités. À la figure 12, tirée de Kempinska et Murcio, l'algorithme utilisé par les auteurs classifie les réseaux de rues selon leur densité, plus précisément la nombre de pixels blancs retrouvés sur chaque image. Si la forme de la trame n'était pas le principal élément observé, force est de constater que certaines densités sont définies par des trames spécifiques, alors que, dans le cas de basses densités, la forme des rues semble plus aléatoire et variée.



Figure 13. Échantillons des réseaux de rues évalués par Kempinska et Murcio classés selon leur densité et leur orientation (p.5, Kempinska & Murcio, 2019)

Dans le cas de Wu et collaborateurs (figure 12), la machine observe plutôt l'organisation et la forme de la trame urbaine, la classant sous quatre grands types, dans lesquels on retrouve également des sous-types. Pour ce faire, l'algorithme regarde l'organisation d'ensembles de pixels noirs sur fond blanc. Tel que mentionné dans la section 4.1.10, une fois les formes urbaines classifiées par l'algorithme, les auteurs ont associé visuellement les types obtenus à des formes urbaines conventionnelles, soit la grille, la forme organique, et la forme d'arbre. Ils définissent donc que le type I comme une grille, le type II comme organique, le type III comme étant un hybride entre les types I et II, et le type IV comme une forme d'arbre. Leur algorithme observe donc les façons dont les rues se connectent entre elles et s'organisent.

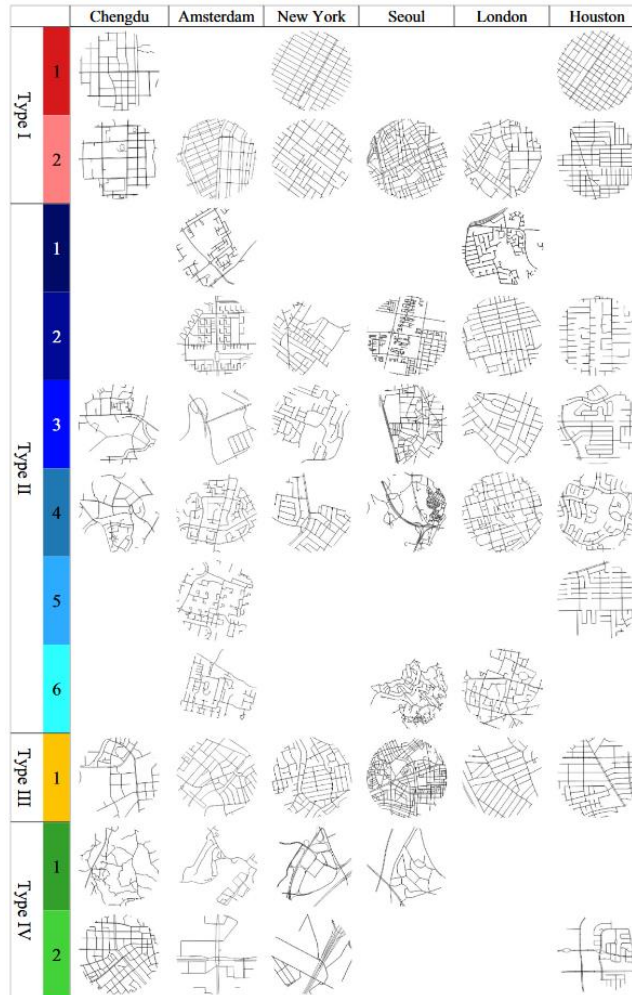


Figure 14. Représentation des types de réseaux de rues retrouvés dans les villes de Chengdu, Amsterdam, New-York, Seoul, Londres et Houston (p.8, Wu et al., 2024)

Zhao et collaborateurs (2019), qui tentent de déterminer les facteurs de risque qui augmentent la probabilité d'accidents chez les cyclistes (section 4.1.5), utilisent quant à eux des images satellites provenant de Google Maps, ainsi que des cartes vectorielles sorties via l'application Google Maps Static API. Pour ces images, les auteurs ont choisi de conserver toutes représentation de parcs, rues, lignes de transports en commun et leurs stations, et masses d'eau. Les éléments (a), (g) et (h) de la figure 10 sont des exemples d'images issues de Google Maps Static API qui ont été utilisées. Dans tous les cas, la troisième image est celle qui a été produite par la translation du domaine sans accident vers le domaine avec accident. Puisque l'objectif de la recherche est d'identifier les caractéristiques qui définissent des segments de rues sur lesquels des accidents cyclistes sont survenus, l'algorithme fait donc l'observation et l'identification des caractéristiques qui reviennent le plus fréquemment dans le cas d'accident, comme la présence de lignes de transport en commun ou de stationnement sur rue, par

exemple. La figure 14 montre les images d'entrée et le résultat obtenu par la translation des images du domaine sans accident vers le domaine avec accident. Dans le cas de (a), par exemple, l'ajout de lignes de tramway et la suppression des espaces verts sont étudiés. L'image ressortissante de la translation montre que ces opérations augmentent les risques d'accidents chez les cyclistes.

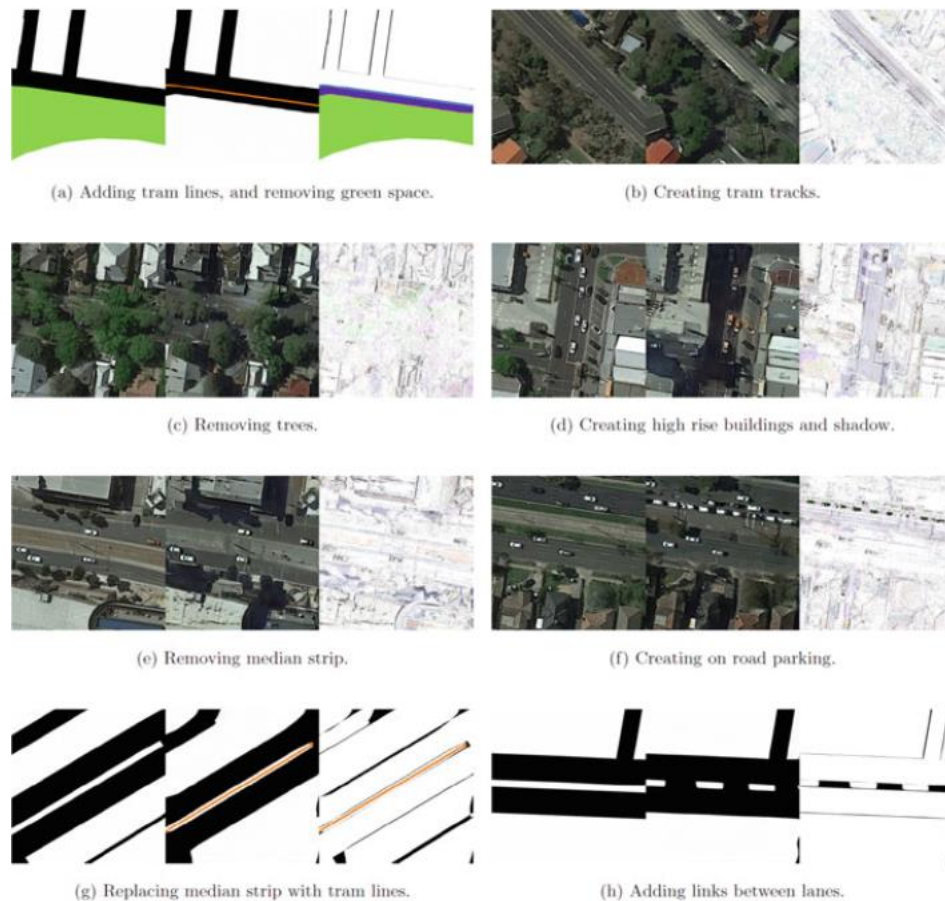


Figure 15. Quelques exemples de translation du domaine sans accident vers le domaine avec accident. De gauche à droite : domaine sans accident, domaine avec accident, résultat de la translation (p.159, Zhao et al., 2019)

Dans leur cas, Comber et collaborateurs (2020) et Taecharungroj et Mathayomchan (2020) font l'usage d'images prises au niveau de la rue. Comber utilise des images de façades de bâtiments commerciaux et de loisir issues de Google Street View. L'auto-encodeur regarde ici les composantes architecturales des bâtiments qui sont visibles depuis la rue et donc qui animent et caractérisent cette dernière, soit les ouvertures, la matérialité, la couleur ou encore les types d'enseignes. Une observation de la présence de véhicules sur les images est également réalisée, celle-ci permettant d'identifier les secteurs qui se

prêtent davantage à l'usage de véhicules. Un exemple tiré de l'étude est illustré par la figure 14, où deux groupes sont comparés l'un à l'autre. Le groupe D montre une plus grande diversité de commerces et semble plus favorable à l'utilisation de la voiture que le groupe A.

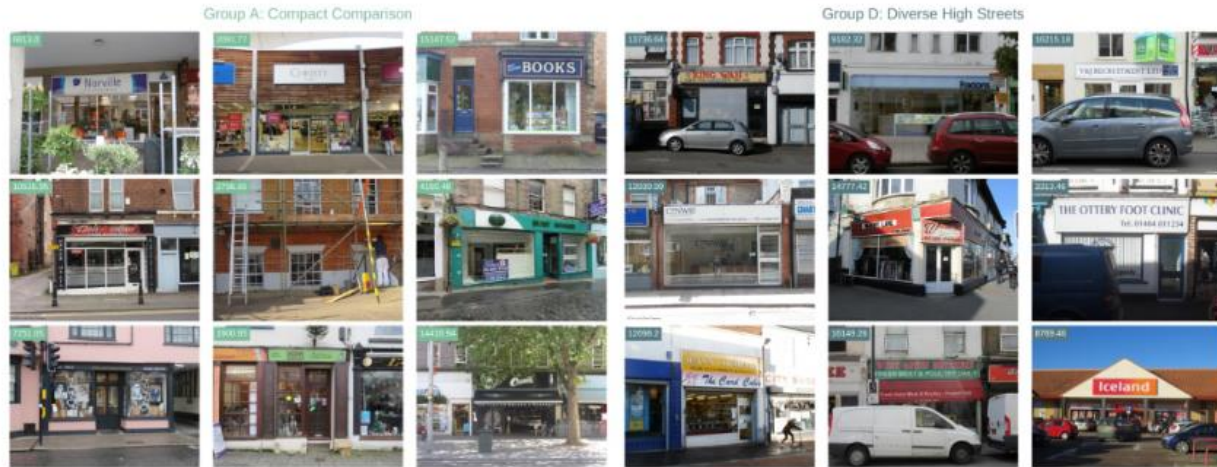


Figure 16. Échantillons d'images de façades qui ont été récoltées pour les groupes A et D (p.9, Comber et al., 2020)

Taecharungroj et Mathayomchan, quant à eux, utilisent des photos prises par les usagers et publiées sur l'application Flickr.com. Ces photos capturent différents aspects des villes étudiées, que ce soient des places publiques, des monuments, des activités culturelles, des paysages ou encore tout éléments pouvant capturer l'essence de la ville et l'intérêt de ses habitants et visiteurs. Puisque les auteurs utilisent un algorithme d'allocation latente de Dirichlet, une première étape a été d'identifier le contenu de chacune des 220 000 images à l'aide de l'application Google Cloud Vision API, une plateforme permettant d'extraire des informations sous forme de texte à partir d'images. Les étiquettes associées au contenu de chaque image ont ensuite été traitées via l'apprentissage machine dans l'objectif d'identifier ce que les auteurs nomment les dimensions de l'image de la ville universelle (UCIDs), ainsi que les dimensions de la ville individuelle (ICIDs). Puisque l'allocation latente de Dirichlet utilise des mots comme données d'analyse, la machine perçoit donc la ville comme une liste de mots. Certains exemples d'étiquettes qui ont été identifiées par Google Cloud Vision sont : human, settlement, architecture, river, nature, train, railway, bus, window, sculpture, street, facade, art, light, tourism, people, crowd.

Rios et Munoz (2017) font eux aussi usage de l'allocation latente de Dirichlet, mais ils appliquent cet algorithme aux données qui sont captées par les antennes qui assurent la réception des signaux provenant des téléphones cellulaires, dans un objectif d'identification des modes d'utilisation du sol. Puisque les appels téléphoniques sont géolocalisés, les auteurs peuvent faire des suppositions sur

l'activité réalisée par les utilisateurs selon leur localisation dans la ville. Chaque antenne se rattachant à un certain nombre de données, celles-ci sont considérées comme des documents, qui contiennent des mots (les données). Un document équivaut à une semaine d'activité. En traitant les données de cette façon, les auteurs parviennent donc à adapter l'activité des antennes à l'algorithme, qui traite des données textuelles. Ainsi, il peut être défini que la machine observe des mots qui représentent des activités réalisées pendant une semaine, et ce sur une période totale de 77 jours. Chaque semaine, ou document, contient des thèmes récurrents qui, finalement, permettent de définir l'utilisation du sol qui est faite à proximité de chaque antenne.

Plusieurs algorithmes ont été présentés, ainsi que différentes méthodes qui les utilisent, offrant plusieurs possibilités intéressantes pour les professionnels en urbanisme et en design urbain. Afin de rendre plus concrètes ces méthodes, il est intéressant d'évaluer lesquelles pourraient être pertinentes à utiliser dans le cadre d'une analyse urbaine spécifique. Le prochain chapitre tentera de proposer des hypothèses d'utilisation de l'intelligence artificielle dans un contexte urbain bien précis du Québec.

5_DISCUSSION

À la suite de l'analyse des dix textes à l'étude, plusieurs algorithmes d'apprentissage non-supervisé ont été mis de l'avant pour observer la rue. S'il se trouve que plusieurs des articles sélectionnés font l'usage des auto-encodeurs (Moosavi, 2017; Sameen et Pradhan, 2017; Kempinska et Murcio, 2019; Comber et al., 2020; Qi et al, 2020; Chao et al., 2021), d'autres auteurs utilisent différents algorithmes, soit les réseaux antagonistes génératifs (Zhao et al., 2019), l'allocation latente de Dirichlet (Rios et Munoz, 2017; Taecharungroj et Mathayomchan, 2020), ou encore l'analyse de regroupement hiérarchique (Wu et al., 2024). De plus, plusieurs façons d'observer la rue ont été identifiées; la forme des réseaux de rues, l'identité visuelle des villes, l'utilisation du sol, la prédiction du trafic, la détection des zones à risque d'accidents, ou encore la mobilité des utilisateurs des transports en commun. Les algorithmes arrivent donc à voir et identifier de nombreuses caractéristiques qui sont également observées par les urbanistes et designers urbains lors d'analyses traditionnelles du tissu urbain.

L'apprentissage non-supervisé se présente donc comme une méthode pouvant être intéressante pour la réalisation d'analyses urbaines par des urbanistes ou des designers urbains. Si, dans un premier temps, cela permet d'observer un nombre de données beaucoup plus imposant que les méthodes traditionnelles, et de façon plus rapide, cela peut également offrir l'opportunité de faire l'analyse de formes urbaines d'autres pays sans nécessairement avoir à se déplacer sur les lieux. Cet aspect devient ainsi intéressant lorsque le déplacement n'est pas possible pour diverses raisons. Comme constaté au travers de plusieurs des études, cela peut également permettre de comparer et analyser plus rapidement plusieurs villes entre elles, un travail pouvant se montrer fastidieux lorsque réalisé à l'aide de méthodes traditionnelles telles que la photographie, le dessin, les schémas ou les diagrammes résultant d'observations sur le terrain.

5.1 Opportunités ; hypothèses au Nunavik

Tel que mentionné précédemment, ces méthodes sont de façon générale réalisées dans de grandes villes et seulement dans des secteurs précis, donc elles ne permettent pas une vision globale de la ville. Si les algorithmes d'apprentissage machine permettent l'observation d'un nombre de données beaucoup plus important, il a été remarqué au travers des textes étudiés que, bien que les villes soient analysées dans leur totalité, les auteurs optent encore en général pour l'étude de métropoles, ou des villes avec une population assez importante. Cependant, puisque l'apprentissage machine permet d'utiliser des données géoréférencées. Et donc des données de partout sur la planète, une opportunité se crée pour l'observation de villes et villages de taille moins importante, ou dont la localisation est

plus éloignée des grands centres, mais où les analyses urbaines sont tout aussi pertinentes pour y réaliser des interventions, ou simplement comprendre l'organisation. Un exemple concret pour un tel exercice peut être celui des villages inuit du Nunavik. Situés au-delà du 55° parallèle dans le nord du Québec, ces villages font face à divers enjeux, tels que la pénurie de logements ou le manque d'accès à certains équipements, qui peuvent intéresser les designers urbains et les pousser à penser et proposer des projets pour ces communautés. Bien que ces villages soient plus reculés, de la documentation et des données sont de plus en plus disponibles sur ceux-ci, rendant ainsi potentiellement intéressant des exercices d'utilisation de l'apprentissage non supervisé. Pour illustrer les hypothèses possibles, un exemple peut être fait d'une équipe souhaitant faire une intervention dans trois villages du Nunavik, soit deux étant situés dans la baie d'Ungava; Kuujjuaq, la capitale, et Kangiqsualujjuaq, et finalement un troisième village situé sur l'autre rive, dans la baie d'Hudson; Inukjuak. Cet exercice se veut donc une discussion sur ce que pourraient apporter les algorithmes d'apprentissage-machine non supervisés aux analyses urbaines se portant sur des villes et villages en situation éloignée. Sans procéder à des tests réels des divers algorithmes, trois hypothèses se basant sur les dix études analysées précédemment sont proposées afin d'illustrer ce que pourraient identifier les algorithmes.

Une première proposition de l'utilisation de l'apprentissage machine dans ce cadre est celui de l'analyse de l'image de la ville, ou du village dans ce cas-ci, de façon similaire à ce que Taecharungroj et Mathayomchan (2020) ont réalisé pour 222 villes sur la planète, utilisant des photographies publiées par les utilisateurs du site Flickr.com. De la même façon que les auteurs, cette approche peut être utilisée pour définir les dimensions universelles et individuelles des quatorze villages du Nunavik, ou encore, dans le cas d'une sélection de trois villages, pour définir l'image spécifique de Kuujjuaq, Kangiqsualujjuaq et Inukjuak. Lorsqu'une recherche est effectuée sur la carte du monde de Flickr.com²⁶, il est possible de constater que plusieurs photographies prisent dans chacun des villages sont disponibles, ce qui rend l'exercice possible. Pour Kuujjuaq, 421 éléments géotaggés sont accessibles, alors que Kangiqsualujjuaq en présente 206, et Inukjuak 230. Les figures qui suivent présentent des échantillons d'images qui peuvent être retrouvées pour chacun des trois villages.

²⁶ <https://www.flickr.com/map>



Figure 17. Quelques images prises à Kuujuaq. On y observe une dominance de la nature, notamment de prises de vues de la rivière Koksoak. Les images représentent principalement la nature. (Source : Flickr.com)



Figure 18. Quelques images prises à Kangisualijuaq. On note une dominance de la prise en photo d'activités culturelles et communautaires, notamment la cueillette et la pêche (Source : Flickr.com)



Figure 19. Quelques images prises à Inukjuak. Les images représentent plusieurs activités culturelles qui touchent la confection, notamment d'igloos ou encore de canot. Il semble y avoir aussi une importance sur les activités réalisées en communauté (Source : Flickr.com)

Les quelques images sélectionnées présentent à première vue quelques similarités. En effet, une tendance parmi celles-ci semble être les paysages et les vastes étendues de territoire que le Nunavik a à offrir. L'hiver est également un thème récurrent. Toutefois, certaines particularités ressortent; si les images de Kuujuaq (figure 16) sont d'une certaine diversité, montrant tant le village que le paysage et les activités, les images prises à Kangiqsualujuaq (figure 17) semblent toucher davantage les activités de subsistance, dont la pêche et la cueillette de champignons. Dans le cas d'Inukjuak (figure 18), les activités culturelles et communautaires, comme la pêche sur glace, la construction d'igloos et de kayaks, et les sports, semblent davantage misent de l'avant. Il s'agit bien évidemment de quelques échantillons d'images et une supposition à première vue, mais si toutes les images disponibles pour chacun des trois villages étaient entrées dans un algorithme, celui-ci pourrait en ressortir les grands thèmes et ainsi permettre d'identifier la vision globale qui est faite du Nunavik à travers les photos publiées sur Flickr, mais également celle de chacun des villages. Considérant que Taecharungroj et Mathayomchan (2020) utilisent l'allocation latente de Dirichlet (LDA), si le même algorithme est utilisé pour les trois villages, les images devraient donc être passées dans une application permettant de sortir sous forme de mots les informations présentes sur chacune d'elles; Google Cloud Vision API, par exemple, en reprenant la méthode des auteurs. Les étiquettes identifiées seraient ensuite entrées dans l'algorithme LDA qui

pourra alors définir les thèmes récurrents. Dans ces cas-ci, les thèmes observés pourraient être hiver, neige, paysage, plage, eau, nature, animal, véhicule, bâtiment, pêche, sport, humain, pour ne donner que quelques exemples.

Une seconde approche pour réaliser une analyse urbaine des villages du Nunavik peut être l'utilisation d'images satellites, ou encore de réseaux de rues vectorisés qui ont été transformé en image, comme le font Moosavi (2017), ou encore Kempinska et Murcio (2019). La première méthode peut se révéler intéressante pour l'observation et l'identification des lignes de désir. Ces dernières se définissent selon l'office québécois de la langue française comme des lignes qui relient les centres de deux zones et dont l'épaisseur est proportionnelle au déplacement entre ces deux zones.²⁷ Ainsi, plus l'un de ces parcours est utilisé, plus il sera marqué et donc visible tant au niveau du sol que sur des images satellites. S'il est possible de constater ces lignes de désir un peu partout dans les différents villes et villages, une caractéristique des villages inuit est que ces lignes traversent également les îlots d'habitation, et donc les terrains des autres villageois. Les inuit se déplacent également beaucoup en véhicule tout terrain ou encore en skidoo, ce qui vient créer de nouveau sentiers en plus de la trame de rue existante dans les villages. L'observation de ces chemins peut être intéressante pour voir la manière dont ceux-ci rejoignent les routes existantes et jouent un rôle dans la mobilité inuit, ou encore pour repérer les lieux et les espaces les plus importants pour la communauté. Les trois figures suivantes montrent des lignes de désir qui peuvent être aperçues sur les photos satellites tirées de Google Earth de Kuujjuaq, Kangisualujjuaq et Inukjuak.

²⁷ Définition tirée de la Vitrine Linguistique de l'Office québécois de la langue française (<https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/17065328/ligne-de-desir>)



Figure 20. Kuujuaq. À gauche, des lignes de désir au travers des habitations du Vieux-Kuujuaq. À droite, des lignes de désir qui longent ou traversent le ruisseau Siuvalik (Source : Google Earth 2024)



Figure 21. Kangiqsualujuaq. À gauche, des lignes de désir au travers des habitations. À droite, des lignes de désir reliant deux quartiers (Source : Google Earth 2024)



Figure 22. Inukjuak. À gauche, des lignes de désir au travers des habitations. À droite, des lignes de désir reliant le centre de formation professionnelle Nunavimmi Pigiursavik à la rive (Source : Google Earth 2024)

Les images satellites présentées aux figures 19, 20 et 21 montre ainsi diverses appropriations du sol par les inuit à travers des lignes de désir. Il est possible de constater que ceux-ci coupent les îlots et les terrains pour se rendre plus rapidement d'une destination ou à une autre, ou encore traversent des ruisseaux ou des espaces naturels, ou portions de territoire, afin de se rendre d'une partie du village à une autre, ou encore pour accéder à la rive. En insérant plusieurs images des villages, voir même du territoire à proximité, dans un auto-encodeur tels que le font Sameen et Pradhan (2017), cela pourrait faire ressortir les connexions que font ces lignes de désir entre elles, mais également avec la trame de rue existante. Cependant, des difficultés pourraient être rencontrées quant à l'identification de ces lignes de désirs. En effet, Sameen et Pradhan (2017) notent les difficultés que peut avoir l'algorithme pour identifier les routes étroites. Puisqu'il s'agit de sentiers créés par des déplacements piétonniers ou par petits véhicules tout terrain, il y a un risque que les lignes de désirs soient trop étroites pour être correctement identifiées et reconstruites par l'encodeur-décodeur.

Dans l'idée que l'exercice peut être réalisable malgré la faible largeur des sentiers, il peut être intéressant d'utiliser une technique similaire à Moosavi (2017), ou encore Kempinska et Murcio (2019), soit en sortant d'Open Street Map les réseaux de rue vectorisés des villages et les transformant en image pour les entrer dans un auto-encodeur. Dans le cas de Sameen et Pradhan, les réseaux de rues vectorisés ont notamment été utilisé pour comparer les résultats de l'auto-encodeur à la suite de l'entrée des images satellites.

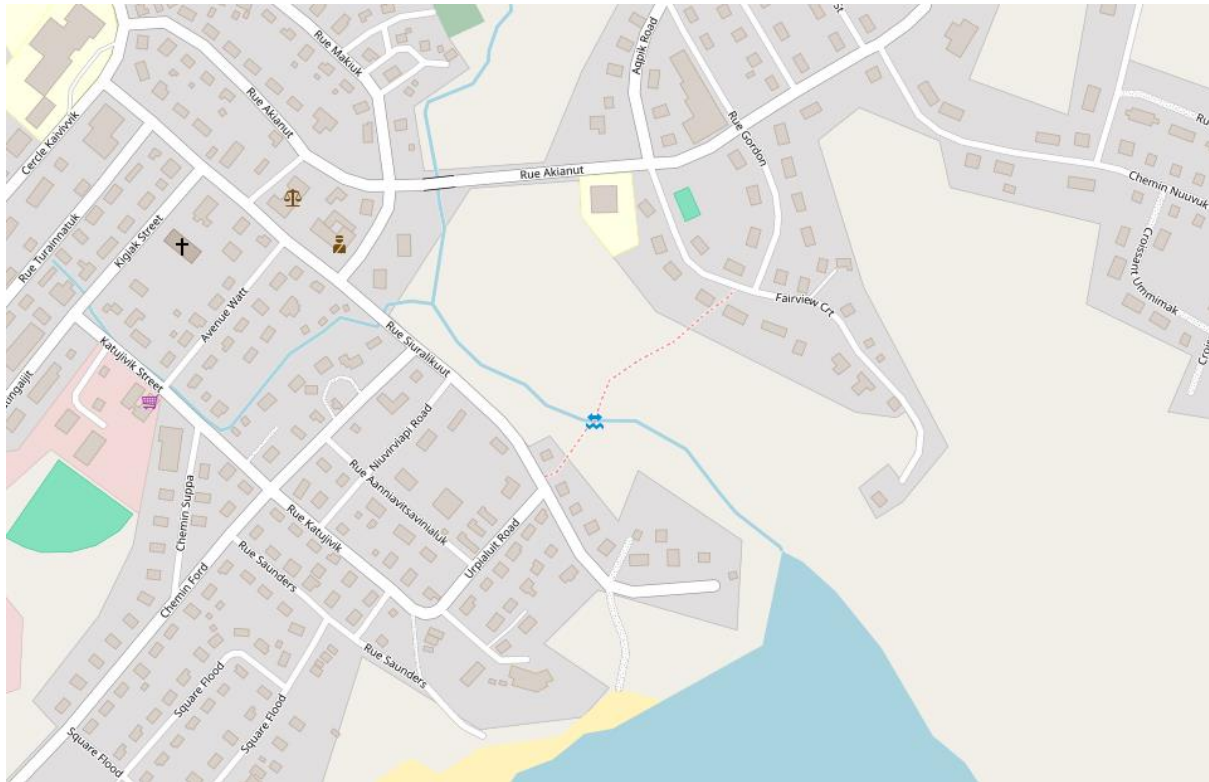


Figure 23. Kuujuuaq. Une ligne de désir (en pointillé rose) qui est identifiée sur OpenStreetMap. (Source : OpenStreetMap, 2024)



Figure 24. Inukjuak. Des lignes de désir (en pointillé) identifiées à l'extérieur du village (Source : OpenStreetMap, 2024)

Peu de lignes de désir ont été identifiées et entrées dans les bases de données d'OpenStreetMap (OSM), ce qui devient limitant pour l'observation de celles-ci via des réseaux vectorisés. Puisqu'il s'agit d'une plateforme de cartographie communautaire, il serait intéressant d'aller y ajouter les sentiers informels identifiables afin qu'ils apparaissent sur les cartes vectorielles. Il est à noter que la ligne de désir qui est entrée dans l'application pour Kuujuaq correspond à celle qui est retrouvée sur l'image de droite à la figure 15, ce qui peut révéler l'importance de ce sentier pour la communauté. La supposition peut être faite qu'il s'agit non seulement d'un sentier emprunté à la marche, mais également par les véhicules tout terrain ou les skidoo en hiver. Malgré le manque de recension des lignes de désir, les réseaux de rues des différents villages du Nunavik sont tout de même recensés sur OSM. Une alternative serait donc d'utiliser ces réseaux et de les entrer dans un algorithme afin de comparer les trames urbaines du Nunavik, ou encore de les comparer avec des trames d'autres villages sur la planète, à l'image du travail de Moosavi (2017). L'application de cette méthode pourrait permettre de révéler une trame urbaine différente, et donc une connectivité plus importante, ou non. Cela permettrait de révéler des lieux forts de rencontre qui sont peut-être différents des espaces publics qui sont habituellement associés aux rencontres et aux échanges communautaires. Cependant, tel que mentionné précédemment, il est possible de rencontrer des difficultés quant à l'identification de ces sentiers, dû à l'étroitesse de ceux-ci.

Une troisième opportunité est l'utilisation d'images présent au niveau de la rue, suivant la méthode utilisée par Comber *et al.* (2020) pour observer les façades des bâtiments commerciaux et de loisir qui consiste en l'utilisation d'un auto-encodeur convolutif et d'images présent au niveau de la rue des façades de bâtiments commerciaux et de loisirs. Une majorité des rues des villages du Nunavik, notamment les trois villages qui sont utilisés ici en exemple, ont été parcourus par les services de Google Street View, rendant ainsi possible l'exploration de ces rues via les applications de Google, et rendant donc par le fait même possible une étude telle que réalisée par Comber et collaborateurs. S'il est possible de procéder comme les auteurs et analyser les façades de bâtiments commerciaux et de loisirs, cette méthode peut également être utilisée pour observer la relation qu'ont les bâtiments avec la rue, ou encore la façon dont les inuit occupent et s'approprient les terrains. Les figures 24, 25 et 26 montrent des images provenant de Kuujuaq, Kangqsualujuaq et Inukjuak et pouvant être récupérées depuis Google Street View.



Figure 25. Images de bâtiments situés à Kunijuaq, récupérées sur Google Street View (source : Google Street View 2024)



Figure 26. Images de bâtiments situés à Kangiqsualujuaq, récupérées sur Google Street View (source : Google Street View 2024)



Figure 27. Images de bâtiments situés à Inukjuak, récupérées sur Google Street View (source : Google Street View 2024)

À l'image de la méthode utilisée par Comber, qui consiste à extraire les images de façades, ainsi que les réseaux de rues dans un rayon de 15 minutes de marche pour localiser les commerces, les réseaux de rues vectorisés des trois villages pourraient également être extraits, et les bâtiments sélectionnés localisés sur ces cartes, les données extraites pouvant fournir des variables caractérisant les habitations, les bâtiments publics, ou l'occupation des espaces par exemple, selon ce qui est souhaité d'être analysé. L'auto-encodeur convolutif est l'algorithme qui, comme pour les auteurs de la recherche de référence, servirait à identifier les caractéristiques visuelles des bâtiments d'entrée. Ces caractéristiques peuvent être la présence ou non d'ouvertures, de rampes d'accès, de véhicules, de garages, ou encore la couleur ou la matérialité des bâtiments, ou leur hauteur et la forme. L'identification de celles-ci, ajouté à la localisation des bâtiments, pourrait permettre d'identifier les différents types d'occupation du sol, s'il y a lieu.

5.2_Limites

Si les algorithmes d'apprentissage-machine non supervisé permettent de réaliser des analyses urbaines à échelle beaucoup plus grande et de façon plus rapide que les méthodes traditionnelles, ceux-ci ne

permettent cependant pas de transmettre certaines qualités que possèdent, par exemple, les observations sur le terrain. En effet, tel que mentionné précédemment, les algorithmes étant des machines, ils ne possèdent pas la capacité de comprendre que les éléments qui sont observés appartiennent à une culture spécifique. De plus, les étiquettes ne représentent pas toujours les particularités culturelles. Ainsi, si l'objectif est d'identifier les nuances culturelles, l'architecture de l'algorithme devra être conçue pour identifier les nuances culturelles là où elles pourraient être visibles au sein des données. Une limite est donc que les algorithmes vont identifier les éléments qu'ils peuvent voir sans les associer, par exemple, à la culture inuit. Cela peut toutefois être un avantage lors de la comparaison de plusieurs cultures au travers de la planète – des caractéristiques qui semblent propres à une culture pourraient se révéler comme étant partagées entre plusieurs autres modes d'établissements. Finalement, les méthodes traditionnelles comme la photographie ou le dessin, par exemple, permettent une expérience sensorielle de la ville, apportant ainsi une autre dimension à l'analyse qui ne se retrouve pas avec l'utilisation d'algorithmes. De plus, comme mentionné par Panerai, Depaule, Demorgon (1999), les observations sur le terrain permettent de s'imprégner de l'ambiance de la ville. En effet, les bruits, les sons, les odeurs font partie intégrante de l'expérience de la ville et sont des éléments qui permettent de la définir, et les algorithmes ne permettent pas de ressentir cette dimension sensorielle et humaine.

6_CONCLUSION

Le développement rapide de l'intelligence artificielle offre de nouvelles opportunités dans divers domaines, notamment dans le cadre du design urbain et des analyses urbaines. Si les algorithmes d'apprentissage-machine supervisé sont déjà largement utilisés dans le cadre de la prévoyance et la prédiction de phénomènes connus, et donc se révèlent intéressants pour la facilitation de la gestion des villes intelligentes, les algorithmes d'apprentissage-machine non supervisés présentent eux aussi un certain potentiel. S'intéressant à ces derniers, cet essai a permis de présenter et d'analyser dix articles scientifiques qui utilisent l'apprentissage non supervisé pour réaliser des analyses urbaines. Plusieurs algorithmes sont identifiés, notamment les auto-encodeurs, les cartes auto-adaptatives (SOM), les réseaux antagonistes génératifs (GAN), l'allocation latente de Dirichlet (LDA) et l'analyse par regroupement hiérarchique (HCA). L'identification détaillée de la composition de ces différents algorithmes, ainsi que l'identification des éléments qu'ils observent, a permis de proposer trois hypothèses d'application de l'apprentissage-machine non supervisé dans un contexte nordique, soit dans les villages de Kuujuaq, Inukjuak et Kangiqsualujuaq, au Nunavik. L'objectif de cette portion de l'essai était de démontrer l'utilité de l'intelligence artificielle (algorithme et données massives) dans des contextes plus reculés qui semblent moins bien documentés.

S'il y a perte de l'expérience sensorielle des villes qu'on retrouve dans les analyses urbaines traditionnelles, l'utilisation de l'apprentissage-machine pour regarder les villes demeure intéressante pour comparer de larges groupes de données entre eux, et donc de comparer plusieurs formes urbaines entre elles et en identifier les similarités et les différences. L'analyse des dix textes sélectionnés le démontre bien. En effet, les applications des algorithmes par les auteurs sont variées et donnent une idée de ce que l'intelligence artificielle peut faire dans le cadre d'analyses urbaines. Moosavi (2017), Sameen et Pradhan (2017), Kempinska et Murcio (2019), ainsi que Wu *et al.* (2024) s'intéressent à l'utilisation de l'apprentissage-machine pour l'identification des différentes formes urbaines au travers de la planète. D'autres auteurs utilisent les algorithmes pour définir les types d'usage du sol (Rios et Munoz, 2017), déterminer les facteurs à risque d'accidents chez les cyclistes (Zhao *et al.*, 2019), faire ressortir les éléments visuels de quartiers aux activités précises (Comber *et al.*, 2020) ou encore de différentes villes (Taecharungroj et Mathayomchan, 2020), ainsi que réaliser des analyses portant sur le trafic routier (Qi *et al.*, 2020) et la mobilité (Chao *et al.*, 2021). Les auteurs utilisent des algorithmes variés, tels que les auto-encodeurs, l'allocation latente de Dirichlet, les réseaux antagonistes génératifs ou encore l'analyse par regroupement hiérarchique, ce qui démontre les différentes options qui

s'offrent aux designers urbains souhaitant utiliser l'intelligence artificielle dans le cadre d'analyses urbaines.

Si cet essai fait l'analyse de dix études et présente les algorithmes utilisés ainsi que leur fonctionnement, tentant ainsi de rendre plus accessible les algorithmes d'apprentissage-machine non supervisé aux designers urbains, urbanistes et architectes, il ouvre également des portes pour de futurs essais en design urbain. En effet, il agit comme base pour mieux comprendre les différents types d'algorithmes et peut guider sur la sélection d'un algorithme précis dans le cadre d'une recherche particulière, en donnant plusieurs exemples concrets d'études déjà réalisées, et fictifs lorsqu'on regarde les hypothèses d'applications au Nunavik.

7_ANNEXE

- Ameen, M., & Szu, H. (1999). Early vision image analyses using ICA in unsupervised learning ANN. *IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339)*, 2, 1022-1027. <https://doi.org/10.1109/IJCNN.1999.831095>
- Ammar, A., Koubaa, A., Ahmed, M., Saad, A., & Benjdira, B. (2021). Aerial Images Processing for Car Detection using Convolutional Neural Networks : Comparison between Faster R-CNN and YoloV3. *Electronics*, 10(7), 820. <https://doi.org/10.3390/electronics10070820>
- Benali, A., Dermeche, H., Belhadj, S., Adnane, A., & Hanifi Elhachemi Amar, R. (2014). Buildings extraction of very high spatial resolution satellite images. *2014 International Conference on Multimedia Computing and Systems (ICMCS)*, 277-282. <https://doi.org/10.1109/ICMCS.2014.6911229>
- Bergado, J. R., Persello, C., & Gevaert, C. (2016). A deep learning approach to the classification of sub-decimetre resolution aerial images. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 1516-1519. <https://doi.org/10.1109/IGARSS.2016.7729387>
- Block, J., Yazdani, M., Nguyen, M., Crawl, D., Jankowska, M., Graham, J., DeFanti, T., & Altintas, I. (2017). An Unsupervised Deep Learning Approach for Satellite Image Analysis with Applications in Demographic Analysis. *2017 IEEE 13th International Conference on E-Science (e-Science)*, 9-18. <https://doi.org/10.1109/eScience.2017.13>
- Campos-Taberner, M., Romero-Soriano, A., Gatta, C., Camps-Valls, G., Lagrange, A., Le Saux, B., Beaupere, A., Boulch, A., Chan-Hon-Tong, A., Herbin, S., Randrianarivo, H., Ferecatu, M., Shimoni, M., Moser, G., & Tuia, D. (2016). Processing of Extremely High-Resolution LiDAR and RGB Data : Outcome of the 2015 IEEE GRSS Data Fusion Contest–Part A: 2-D Contest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(12), 5547-5559. <https://doi.org/10.1109/JSTARS.2016.2569162>
- Canty, M. J. (2009). Boosting a fast neural network for supervised land cover classification. *Computers & Geosciences*, 35(6), 1280-1295. <https://doi.org/10.1016/j.cageo.2008.07.004>
- Celik, T. (2009). Multiscale Change Detection in Multitemporal Satellite Images. *IEEE Geoscience and Remote Sensing Letters*, 6(4), 820-824. <https://doi.org/10.1109/LGRS.2009.2026188>
- Chen, Y., Jiang, H., Li, C., Jia, X., & Ghamisi, P. (2016). Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10), 6232-6251. <https://doi.org/10.1109/TGRS.2016.2584107>
- Cheng, G., Wang, Y., Xu, S., Wang, H., Xiang, S., & Pan, C. (2017). Automatic Road Detection and Centerline Extraction via Cascaded End-to-End Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing*, 55(6), 3322-3337. <https://doi.org/10.1109/TGRS.2017.2669341>
- Cheng, G., Zhou, P., & Han, J. (2016). Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12), 7405-7415. <https://doi.org/10.1109/TGRS.2016.2601622>
- Christodoulou, C. I., Michaelides, S. C., Pattichis, C. S., & Kyriakou, K. (2001). Classification of satellite cloud imagery based on multi-feature texture analysis and neural networks. *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, 1, 497-500. <https://doi.org/10.1109/ICIP.2001.959062>

- Defferrard, M., Bresson, X., & Vandergheynst, P. (s. d.). *Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering*.
- Del Frate, F., Petrocchi, A., Lichtenegger, J., & Calabresi, G. (2000). Neural networks for oil spill detection using ERS-SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5), 2282-2287. <https://doi.org/10.1109/36.868885>
- Feng, Z., Peng, J., & Wu, J. (2020). Using DMSP/OLS nighttime light data and K-means method to identify urban-rural fringe of megacities. *Habitat International*, 103, 102227. <https://doi.org/10.1016/j.habitatint.2020.102227>
- Giannakopoulos, T., & Perantonis, S. (2018). *Recognition of Urban Sound Events using Deep Context-Aware Feature Extractors and Handcrafted Features* [Preprint]. MATHEMATICS & COMPUTER SCIENCE. <https://doi.org/10.20944/preprints201811.0509.v1>
- Grbčić, L., Lučin, I., Kranjčević, L., & Družeta, S. (2020). A Machine Learning-based Algorithm for Water Network Contamination Source Localization. *Sensors*, 20(9), 2613. <https://doi.org/10.3390/s20092613>
- Grinias, I., Panagiotakis, C., & Tziritas, G. (2016). MRF-based segmentation and unsupervised classification for building and road detection in peri-urban areas of high-resolution satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 145-166. <https://doi.org/10.1016/j.isprsjprs.2016.10.010>
- Gupta, D., Bhurchandi, K., Murala, S., Raman, B., & Kumar, S. (Éds.). (2023). *Computer Vision and Image Processing : 7th International Conference, CVIP 2022, Nagpur, India, November 4–6, 2022, Revised Selected Papers, Part II* (Vol. 1777). Springer Nature Switzerland. <https://doi.org/10.1007/978-3-031-31417-9>
- Hammadi-Mesmoudi, F. (1995). An unsupervised neural network classifier and its application in remote sensing. *Fifth International Conference on Image Processing and Its Applications*, 1995, 236-240. <https://doi.org/10.1049/cp:19950656>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition* (arXiv:1512.03385). arXiv. <http://arxiv.org/abs/1512.03385>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). *MobileNets : Efficient Convolutional Neural Networks for Mobile Vision Applications* (arXiv:1704.04861). arXiv. <http://arxiv.org/abs/1704.04861>
- Ibrahim, M. R., Haworth, J., & Cheng, T. (2021). URBAN-i : From urban scenes to mapping slums, transport modes, and pedestrians in cities using deep learning and computer vision. *Environment and Planning B: Urban Analytics and City Science*, 48(1), 76-93. <https://doi.org/10.1177/2399808319846517>
- Jiang, K., Wang, Z., Yi, P., & Jiang, J. (2018a). A Progressively Enhanced Network for Video Satellite Imagery Superresolution. *IEEE Signal Processing Letters*, 25(11), 1630-1634. <https://doi.org/10.1109/LSP.2018.2870536>
- Kit, O., Lüdeke, M., & Reckien, D. (2012). Texture-based identification of urban slums in Hyderabad, India using remote sensing data. *Applied Geography*, 32(2), 660-667. <https://doi.org/10.1016/j.apgeog.2011.07.016>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90. <https://doi.org/10.1145/3065386>
- Lan, M., Zhang, Y., Zhang, L., & Du, B. (2020). Global context based automatic road segmentation via dilated

convolutional neural network. *Information Sciences*, 535, 156-171. <https://doi.org/10.1016/j.ins.2020.05.062>

Liu, Z.-G., Zhang, Z.-W., Pan, Q., & Ning, L.-B. (2022). Unsupervised Change Detection From Heterogeneous Data Based on Image Translation. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-13. <https://doi.org/10.1109/TGRS.2021.3097717>

Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 645-657. <https://doi.org/10.1109/TGRS.2016.2612821>

Mahjourian, R., Wicke, M., & Angelova, A. (2018). *Unsupervised Learning of Depth and Ego-Motion from Monocular Video Using 3D Geometric Constraints* (arXiv:1802.05522). arXiv. <http://arxiv.org/abs/1802.05522>

Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions on Geoscience and Remote Sensing*, 42(8), 1778-1790. <https://doi.org/10.1109/TGRS.2004.831865>

Moosavi, V. (2017). *Urban morphology meets deep learning : Exploring urban forms in one million cities, town and villages across the planet* (arXiv:1709.02939). arXiv. <http://arxiv.org/abs/1709.02939>

Paoletti, M. E., Haut, J. M., Plaza, J., & Plaza, A. (2018). A new deep convolutional neural network for fast hyperspectral image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 120-147. <https://doi.org/10.1016/j.isprsjprs.2017.11.021>

Peng, B., Huang, Q., Vongkusolkrit, J., Gao, S., Wright, D. B., Fang, Z. N., & Qiang, Y. (2021). Urban Flood Mapping With Bitemporal Multispectral Imagery Via a Self-Supervised Learning Framework. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 2001-2016. <https://doi.org/10.1109/JSTARS.2020.3047677>

Peng, C., Li, Y., Jiao, L., & Shang, R. (2021). Efficient Convolutional Neural Architecture Search for Remote Sensing Image Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7), 6092-6105. <https://doi.org/10.1109/TGRS.2020.3020424>

Ruf, B., Thiel, L., & Weinmann, M. (2018). DEEP CROSS-DOMAIN BUILDING EXTRACTION FOR SELECTIVE DEPTH ESTIMATION FROM OBLIQUE AERIAL IMAGERY. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-1, 125-132. <https://doi.org/10.5194/isprs-annals-IV-1-125-2018>

Saboori, M., Homayouni, S., Shah-Hosseini, R., & Zhang, Y. (2022). Optimum Feature and Classifier Selection for Accurate Urban Land Use/Cover Mapping from Very High Resolution Satellite Imagery. *Remote Sensing*, 14(9), 2097. <https://doi.org/10.3390/rs14092097>

Samany, N. N. (2019). Automatic landmark extraction from geo-tagged social media photos using deep neural network. *Cities*, 93, 1-12. <https://doi.org/10.1016/j.cities.2019.04.012>

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition* (arXiv:1409.1556). arXiv. <http://arxiv.org/abs/1409.1556>

Song, X., Wang, Z., Liang, J., Zhang, B., Du, Y., Zeng, Z., & Liu, M. (2022). Automatic Extraction of the Basal Channel Based on Neural Network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 5013-5023. <https://doi.org/10.1109/JSTARS.2022.3184156>

Sublime, J., Troya-Galvis, A., & Puissant, A. (2017). Multi-Scale Analysis of Very High Resolution Satellite Images Using Unsupervised Techniques. *Remote Sensing*, 9(5), 495. <https://doi.org/10.3390/rs9050495>

- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2014). *Going Deeper with Convolutions* (arXiv:1409.4842). arXiv. <http://arxiv.org/abs/1409.4842>
- Taecharunroj, V., & Mathayomchan, B. (2020). The big picture of cities : Analysing Flickr photos of 222 cities worldwide. *Cities*, 102, 102741. <https://doi.org/10.1016/j.cities.2020.102741>
- Tahir, H., Shahbaz Khan, M., & Owais Tariq, M. (2021). Performance Analysis and Comparison of Faster R-CNN, Mask R-CNN and ResNet50 for the Detection and Counting of Vehicles. *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 587-594. <https://doi.org/10.1109/ICCCIS51004.2021.9397079>
- Umaselvi, M., Kumar, S. S., & Athithya, M. (2013). Color and wavelet based identification of urban and agricultural area using texture features in satellite images. *2013 International Conference on Green High Performance Computing (ICGHPC)*, 1-4. <https://doi.org/10.1109/ICGHPC.2013.6533914>
- Verma, D., Jana, A., & Ramamritham, K. (2019). Classification and mapping of sound sources in local urban streets through AudioSet data and Bayesian optimized Neural Networks. *Noise Mapping*, 6(1), 52-71. <https://doi.org/10.1515/noise-2019-0005>
- Wu, C., Peng, N., Ma, X., Li, S., & Rao, J. (2020). Assessing multiscale visual appearance characteristics of neighbourhoods using geographically weighted principal component analysis in Shenzhen, China. *Computers, Environment and Urban Systems*, 84, 101547. <https://doi.org/10.1016/j.compenvurbsys.2020.101547>
- Wuest, B., & Zhang, Y. (2009). Region based segmentation of QuickBird multispectral imagery through band ratios and fuzzy comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1), 55-64. <https://doi.org/10.1016/j.isprsjprs.2008.06.005>
- Xu, D., Zhou, D., Wang, Y., Meng, X., Chen, W., & Yang, Y. (2020). Temporal and spatial variations of urban climate and derivation of an urban climate map for Xi'an, China. *Sustainable Cities and Society*, 52, 101850. <https://doi.org/10.1016/j.scs.2019.101850>
- Yi-Chung Rau, Comiso, J. C., & Lure, F. Y. M. (1994). Application of neural networks for identification of sea ice coverage and movements from satellite imagery. *Proceedings of IGARSS '94 - 1994 IEEE International Geoscience and Remote Sensing Symposium*, 3, 1407-1409. <https://doi.org/10.1109/IGARSS.1994.399453>
- Zagoruyko, S., & Komodakis, N. (2017). *Wide Residual Networks* (arXiv:1605.07146). arXiv. <http://arxiv.org/abs/1605.07146>
- Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., & Atkinson, P. M. (2018a). An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sensing of Environment*, 216, 57-70. <https://doi.org/10.1016/j.rse.2018.06.034>
- Zhang, C., Wei, S., Ji, S., & Lu, M. (2019). Detecting Large-Scale Urban Land Cover Changes from Very High Resolution Remote Sensing Images Using CNN-Based Classification. *ISPRS International Journal of Geo-Information*, 8(4), 189. <https://doi.org/10.3390/ijgi8040189>
- Zhang, F., Du, B., & Zhang, L. (2015). Saliency-Guided Unsupervised Feature Learning for Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4), 2175-2184. <https://doi.org/10.1109/TGRS.2014.2357078>

Zhao, X., Tao, R., Li, W., Li, H.-C., Du, Q., Liao, W., & Philips, W. (2020). Joint Classification of Hyperspectral and LiDAR Data Using Hierarchical Random Walk and Deep CNN Architecture. *IEEE Transactions on Geoscience and Remote Sensing*, 58(10), 7355-7370. <https://doi.org/10.1109/TGRS.2020.2982064>

Zhou, G., Sohn, K., & Lee, H. (s. d.). *Online Incremental Feature Learning with Denoising Autoencoders*.

8_BIBLIOGRAPHIE

- Band, S. S., Ardabili, S., Sookhak, M., Chronopoulos, A. T., Elnaffar, S., Moslehpour, M., Csaba, M., Torok, B., Pai, H.-T., & Mosavi, A. (2022). When Smart Cities Get Smarter via Machine Learning : An In-Depth Literature Review. IEEE Access, 10, 60985-61015. <https://doi.org/10.1109/ACCESS.2022.3181718>
- Comber, S., Arribas-Bel, D., Singleton, A., & Dolega, L. (2020). Using convolutional autoencoders to extract visual features of leisure and retail environments. Landscape and Urban Planning, 202, 103887. <https://doi.org/10.1016/j.landurbplan.2020.103887>
- Deltombe, M., & Deudon, L. (2020). Repenser les sources de l'histoire environnementale grâce aux outils numériques : Le cas de la vallée de l'Escaut (France). Cahiers d'histoire, 37(1), 67-97. <https://doi.org/10.7202/1067957ar>
- Hirschberg, U., Hovestadt, L., & Fritz, O. (Éds.). (2020). Atlas of digital architecture : Terminology, concepts, methods, tools, examples, phenomena. Birkhauser.
- Kempinska, K., & Murcio, R. (2019). Modelling urban networks using Variational Autoencoders. Applied Network Science, 4(1), 114. <https://doi.org/10.1007/s41109-019-0234-0>
- Lynch, K. (1960). *The Image of the City*. Cambridge [Mass.] Technology Press.
- Mahrez, Z., Sabir, E., Badidi, E., Saad, W., & Sadik, M. (2022). Smart Urban Mobility : When Mobility Systems Meet Smart Data. IEEE Transactions on Intelligent Transportation Systems, 23(7), 6222-6239. <https://doi.org/10.1109/TITS.2021.3084907>
- Moosavi, V. (2017). Urban morphology meets deep learning : Exploring urban forms in one million cities, town and villages across the planet (arXiv:1709.02939). arXiv. <http://arxiv.org/abs/1709.02939>
- Neves, F., Finamore, A. C., Madeira, S. C., & Henriques, R. (2022). Mining Actionable Patterns of Road Mobility From Heterogeneous Traffic Data Using Biclustering. IEEE Transactions on Intelligent Transportation Systems, 23(7), 6430-6445. <https://doi.org/10.1109/TITS.2021.3057240>
- Panerai, P., Depaule, J.-C., & Demorgon, M. (1999). Analyse urbaine. Éditions Parenthèses.
- Qi, X., Qin, X., Jia, Z., Lin, M., & Liu, Y. (2020). NCAE and ELM Based Enhanced Ensemble Optimized Model for Traffic Flow Forecasting. IEEE Access, 8, 200486-200499. <https://doi.org/10.1109/ACCESS.2020.3034763>
- Qu, X., Zhen, L., Howlett, R. J., & Jain, L. C. (Éds.). (2019). Smart Transportation Systems 2019 (Vol. 149). Springer Singapore. <https://doi.org/10.1007/978-981-13-8683-1>
- Ríos, S. A., Muñoz, R. (2017). Land Use detection with cell phone data using topic models : Case Santiago, Chile. Computers, Environment and Urban Systems, 61, 39-48. <https://doi.org/10.1016/j.compenvurbsys.2016.08.007>
- Salat, S., Labbé, F., & Nowacki, C. (2011). Les villes et les formes : Sur l'urbanisme durable. Hermann CSTB, Laboratoire des morphologies urbaines.

- Sameen, M. I., Pradhan, B. (2017). A Novel Road Segmentation Technique from Orthophotos Using Deep Convolutional Autoencoders. *Korean Journal of Remote Sensing*, 33(4), 423-436. <https://doi.org/10.7780/KJRS.2017.33.4.8>
- Shane, D. G. (2005). *Recombinant urbanism : Conceptual modeling in architecture, urban design, and city theory*. Wiley.
- Taecharungroj, V., Mathayomchan, B. (2020). The big picture of cities : Analysing Flickr photos of 222 cities worldwide. *Cities*, 102, 102741. <https://doi.org/10.1016/j.cities.2020.102741>
- Wang, J., Biljecki, F. (2022). Unsupervised machine learning in urban studies : A systematic review of applications. *Cities*, 129, 103925. <https://doi.org/10.1016/j.cities.2022.103925>
- Wu, C., Wang, J., Wang, M., & Kraak, M.-J. (2024). Machine learning-based characterisation of urban morphology with the street pattern.
- Yu, C., Li, H., Xu, X., Liu, J., Miao, J., Wang, Y., & Sun, Q. (2021). Data-Driven Approach for Passenger Mobility Pattern Recognition Using Spatiotemporal Embedding. *Journal of Advanced Transportation*, 2021, 1-21. <https://doi.org/10.1155/2021/5574093>
- Zhu, X., & Goldberg, A. B. (2009). *Introduction to semi-supervised learning*. Morgan & Claypool.