



**HAL**  
open science

# Représentation et combinaison de l'information géographique pour l'apprentissage profond

Azelle Courtial, Guillaume Touya

► **To cite this version:**

Azelle Courtial, Guillaume Touya. Représentation et combinaison de l'information géographique pour l'apprentissage profond. GAST – Gestion et Analyse de données Spatiales et Temporelles, Jan 2022, Blois, France. pp.54-65. hal-03719234

**HAL Id: hal-03719234**

**<https://hal.science/hal-03719234v1>**

Submitted on 11 Jul 2022

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

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



Distributed under a Creative Commons Attribution 4.0 International License

# Représentation et combinaison de l'information géographique pour l'apprentissage profond

Azelle Courtial\*, Guillaume Touya\*

\* LASTIG Univ. Gustav Eiffel, ENSG, IGN, Champs-sur-Marne  
prenom.nom@ign.fr,

**Résumé.** L'apprentissage profond permet maintenant de générer des cartes transformées à partir d'images d'autres cartes. Mais contrairement aux méthodes traditionnelles de prédiction de carte qui reposent sur des couches de données vectorielles stockées dans des bases de données géographiques, l'image ne transmet qu'une vue limitée des informations contenues dans la version vectorielle des données. Dans cet article, nous nous intéressons à la représentation de l'information géographique sous forme de tenseurs pour améliorer la génération de cartes par apprentissage profond. Nous proposons d'abord une stratégie alternative pour la création des données d'apprentissage : un ensemble de masques où chacun décrit les formes et positions d'un type d'objet géographique sur une même portion de carte (bâtiments, routes, ...). Nous étudions ensuite comment combiner de l'information géographique additionnelle dans les mécanismes d'apprentissage pour améliorer l'abstraction des cartes générées.

## 1 Introduction

L'apprentissage profond est de plus en plus utilisé pour générer de l'information géographique et notamment des cartes, afin d'accélérer les processus de production. Le problème de génération de cartes est souvent traité comme un problème de traduction d'image à image comme par exemple passer d'une image aérienne vers une carte (Isola et al., 2017). Néanmoins, de telles approches ne permettent pas à proprement parler de générer des cartes, mais plutôt des images dans le style d'une carte ou ressemblant à une carte. En effet, la création d'une carte inclut la symbolisation de l'information géographique, mais il s'agit surtout de l'abstraction d'une portion de terrain à une certaine échelle pour répondre à un certain objectif. Les méthodes de génération de cartes doivent apprendre à la fois la combinaison, l'adaptation à l'échelle cible (généralisation) et la représentation de l'information géographique dans l'image cible. Pour cela, l'image initiale doit permettre de discerner l'ensemble des informations nécessaires à ces processus.

Ces informations incluent non seulement la forme et la localisation des objets, mais aussi des caractéristiques implicites telles que leurs importances, leurs rôles dans le paysage, leurs relations spatiales, ... L'acquisition de ces connaissances et leur utilisation dans un processus automatique, sont des défis majeurs de la généralisation automatique (Sester, 2005; Weibel et al., 1995). Lorsque nous nous intéressons à la génération de cartes à partir d'images de

## Représentation et combinaison de l'information géographique pour l'apprentissage profond

cartes plutôt que de données vectorielles, l'accès à ces informations est compliqué. Si certaines de ces informations sont implicitement visibles dans les exemples d'apprentissage (sinuosité d'une route, régularité d'un groupe de bâtiments, ...), il reste à s'assurer qu'elles sont comprises et utilisées correctement par le réseau d'apprentissage. Et surtout une part importante des informations est masquée soit par la représentation initiale des données (fonction des bâtiments, sens de circulation des tronçons routiers, ...), soit par le découpage de l'espace en tuiles (densité d'un îlot urbain, place d'un arc routier dans le réseau routier, ...), soit par le format image (limite des objets adjacents non-visibles, unicité de valeur de pixel en cas d'intersection d'objets géographiques, ...). Cette inadéquation entre la forme des données d'apprentissage et l'objectif d'apprentissage est la principale limite des premières expérimentations pour générer des nouvelles cartes par apprentissage (Courtial et al., 2021b).

Nous proposons donc dans cet article de tester comment une bonne combinaison de l'information géographique peut améliorer la génération de cartes en permettant d'intégrer des informations supplémentaires non contenues dans l'image, et/ou de mettre en valeur les relations spatiales entre objets géographiques. Cet article est organisé comme suit : d'abord, la partie 2 est un l'état de l'art qui décrit les limites des différentes approches de cartographie usant d'apprentissage profond, et comment la combinaison d'informations peut être utilisée pour améliorer la génération de cartes. Puis nous présentons nos expérimentations dans la section 3. Ensuite, les sections 4 et 5 présentent nos propositions pour (a) la structuration et la représentation des données initiales, (b) la combinaison d'informations additionnelles. Enfin, nous concluons sur l'importance de la représentation et de la combinaison des données initiales en apprentissage profond dans la section 6.

## 2 État de l'art

### 2.1 Apprentissage profond et cartographie

La traduction d'image aérienne vers une tuile dans le style de Google Maps figure parmi les nombreux exemples d'usage des premiers réseaux antagonistes de génération d'images (GAN) (Isola et al., 2017; Zhu et al., 2017). Toutes ces expérimentations, menées par des non-cartographes, démontrent la capacité de ces réseaux à reproduire les principales caractéristiques d'une carte, préserver l'information géographique saillante dans l'image initiale et reproduire une symbolisation. Plus tard, des variations de ces réseaux, spécialisées pour la cartographie tentent d'améliorer les résultats dans ce cas et s'intéressent particulièrement à la préservation de l'information géographique initiale. Par exemple, Chen et al. (2020) proposent une fonction d'apprentissage qui utilise le gradient des images pour décrire leurs structures et s'assurer de leurs préservations lors de la génération. Cette fonction peut être apparentée à une contrainte de préservation des relations topologiques de la carte. De façon similaire, Fu et al. (2019) proposent une fonction d'apprentissage qui s'assure de la consistance géométrique de la prédiction, c'est-à-dire qui applique une transformation géométrique de l'image en entrée et compare la prédiction pour cette image avec la transformation de l'image produite depuis l'image initiale. Cette caractéristique est une contrainte spatiale importante sur les données géographiques à une échelle fixe. Ganguli et al. (2019) proposent quant à eux un réseau plus proche du transfert de style, évaluant à la fois la ressemblance de l'image produite avec le style désiré et avec l'image initiale. Li et al. (2020) complètent un GAN avec un classifieur qui

doit apprendre ce qu'est une carte pour améliorer le rendu des prédictions. L'objectif de tous ces travaux reste la représentation des données et l'abstraction est laissée de côté : la fonction et l'information mise en valeur par la carte finale est exactement similaire à celle de la carte initiale. Il ne s'agit pas d'une nouvelle carte, mais de la même carte avec un style différent.

Au contraire, d'autres études travaillent à montrer que ces réseaux sont capables d'abstraction via la génération d'informations géographiques généralisées (à plus petite échelle que les données initiales). Elles ne traitent pas du tout de la représentation et se concentrent sur un type d'objet particulier, par exemple : les bâtiments (Feng et al., 2019), les traits de côte (Du et al., 2021), et les routes de montagne (Courtial et al., 2020). Ces approches n'abordent pas la représentation de l'information et mènent à la prédiction d'images assez éloignées d'une image de carte. Mais elles démontrent la capacité des réseaux d'apprentissage à sélectionner, simplifier, déplacer, déformer, abstraire de l'information géographique pour répondre à un objectif de lisibilité à une échelle cible, et par conséquent la capacité des réseaux d'apprentissage profond à construire une carte (et non lui appliquer un style). Pour étayer ces conclusions Jenny et al. (2020) apprend le positionnement d'ombrage sur une carte à partir d'un Modèle Numérique de Terrain.

Quelques expérimentations visent à la fois l'abstraction et la représentation de l'information initiale en une carte généralisée pour une échelle cible. Courtial et al. (2021b) tentent d'apprendre à généraliser l'information d'une carte topographique pour une représentation qui est également une carte topographique, mais à une échelle intermédiaire. Tandis que Kang et al. (2019) tente à partir de données OSM (donc avec un niveau de détail hétérogène) de produire une carte homogène de type Google Map. Ces premières expérimentations ont montré que la représentation des données d'apprentissage sous forme d'une rasterisation des données vecteurs symbolisées est source de limitations importantes :

- La symbolisation doit être adaptée pour éviter la superposition des symboles et permettre l'apprentissage. L'ajout de contour permet de définir les limites des objets (Touya et al., 2019) mais cela élargit les formes et ajoute de la complexité à la symbolisation.
- La symbolisation choisie doit permettre de différencier sans ambiguïté les pixels représentant des objets différents. Les couleurs trop proches sont donc à éviter au sens de distance entre les valeurs de pixels comme au sens visuel. Les attributs des objets ne sont accessibles que s'ils sont représentés dans l'image via des variables visuelles, telles que la couleur ou la taille de l'objet, cela permet de représenter par exemple la nature des bâtiments ou l'importance des routes. Mais le nombre d'attributs représentés par objet est très rapidement limité.
- Enfin le découpage de l'espace d'étude en tuile peut amener à créer des effets de bord qui perturberont l'apprentissage. Par exemple, des tuiles presque vides en bordure de zone perturbent l'apprentissage, la généralisation des objets étant influencée par leur contexte géographique.

En effet, les productions de ces deux expérimentations sont prometteuses, mais ne tirent pas entièrement parti des informations initiales. Finalement, Chen et al. (2021) tente d'apprendre la traduction d'image aérienne en une carte à plusieurs échelles en tirant parti de la similarité entre les échelles pour apprendre ce qu'est une carte. Mais là encore les résultats sont limités par les données initiales qui ne permettent pas une lisibilité suffisante de l'information géographique du terrain.

Représentation et combinaison de l'information géographique pour l'apprentissage profond

## 2.2 Apprentissage profond et fusion d'information.

La fusion de données est une utilisation possible de l'apprentissage profond, il s'agit d'utiliser plusieurs images ou tenseurs en entrée du réseau de neurones pour une tâche unique de classification ou de segmentation d'image. En cartographie ces méthodes pourront servir pour intégrer de l'information géographique additionnelle dans l'image de carte initiale. Mais les usages actuels de ces réseaux ne sont pas orientés pour la cartographie. Par exemple, Hazırbas et al. (2016) proposent la fusion d'information de profondeur pour améliorer la segmentation de photographies de scènes en intérieur. L'architecture proposée repose sur l'encodage et la fusion progressive des informations des deux images initiales puis la création d'une segmentation. Ainsi, différents niveaux de relations entre l'image en couleurs et celle de profondeur peuvent être encodés. Guo et al. (2019) propose d'utiliser un réseau de fusion d'information comme générateur d'un réseau antagoniste génératif (GAN). Les GAN sont des réseaux d'apprentissage constitué de deux réseaux de convolutions : un générateur qui réalise à proprement parler la tâche cible et un discriminateur qui classe les prédictions du générateur. Ces deux réseaux apprennent de façon antagoniste : le générateur à tromper le discriminateur et le discriminateur à améliorer sa classification, pour produire des images plus réalistes (qui ressemblent à une image de cible). Zhang (2021) réalise un comparatif des stratégies de fusion d'image pour apprendre à combiner des images avec différentes parties floues en une image nette. Finalement, Zhang et al. (2020) propose d'améliorer l'apprentissage de la génération de cartes à partir d'image aérienne avec de l'information supplémentaire telle que des traces GPS.

## 3 Expérimentations

Dans cette section, nous présentons nos expérimentations : dans un premier temps notre cas d'étude, puis les modèles d'apprentissage profond employés.

### 3.1 Cas d'étude

Nous nous intéressons au problème de génération de cartes topographiques d'espace urbain à une échelle intermédiaire autour du 1 :50k, à partir de donnée détaillées produite pour des cartes à 1 :25k. Notre cas d'étude inclut la combinaison et la symbolisation des objets géographiques suivants : routes, bâtiment, et rivières. La densité importante de ces objets en milieu urbain, requiert une abstraction importante de l'information initiale malgré le saut d'échelle faible : outre la simplification des formes individuelles (rectangularisation et grossissement de bâtiment, lissage des lignes), les bâtiments et le réseau routier doivent être sélectionnés et déplacés voire déformés, de sorte à conserver la structure globale et les agencements remarquables (i.e. les alignements de bâtiments, réseau en étoile, ...), mais réduire l'encombrement et produire une carte lisible. Cette tâche a été choisie pour nos premières expérimentations, car elle demande une abstraction importante, et la carte topographique répond à une symbolisation, un objectif et une fonctionnalité clairement définis quel que soit le niveau de familiarité de l'utilisateur avec le lieu représenté, ce qui facilitera l'évaluation de la carte en sortie. Notre jeu d'exemple est produit par tuilage régulier des données dans l'aire urbaine de Saint-Jean-de-Luz. On obtient alors 3 700 paires d'images initiales et cibles, on en garde 91 pour l'évaluation et le reste sert à l'entraînement du modèle.

### 3.2 Modèle de génération

Les précédentes expérimentations visant à la symbolisation d'information géographique, ou à la génération de cartes ont montré le potentiel des réseaux antagonistes, c'est pourquoi nous avons concentré nos expérimentations sur ce type de réseaux. Les GANs sont des réseaux qui apprennent grâce à l'opposition de deux réseaux d'apprentissage : un générateur qui produit une image la plus proche possible du domaine cible à partir de l'image initiale et un discriminateur qui prédit la crédibilité de la prédiction. Ces deux réseaux fonctionnent en opposition : le générateur apprend à tromper le discriminateur en générant des images toujours plus réalistes tandis que le discriminateur s'entraîne à distinguer les images de la base d'exemple de celles générées. Ces réseaux sont donc particulièrement adaptés pour générer des images ressemblant à des cartes (Courtial et al., 2021b).

Comme notre but n'est pas d'étudier l'intérêt de ces réseaux pour la génération de cartes, mais plutôt de démontrer qu'une meilleure organisation de l'information initiale permet un meilleur apprentissage des tâches relatives à la cartographie nous avons uniquement testé l'entraînement du réseau GAN pix2pix (Isola et al., 2017) dans son implémentation en Pytorch avec leur paramètre et construction par défaut.

### 3.3 Générateur et combinaison

Le générateur du GAN a pour but d'encoder l'information initiale puis de la décoder en une prédiction, nous nous intéressons donc à la façon dont il est capable de combiner l'information géographique. Le générateur de pix2pix comme pour de nombreux GAN le générateur est en forme de U-Net Ronneberger et al. (2015). Il s'agit d'un réseau d'apprentissage qui encode et décode l'information dans une image via une série de convolutions ascendantes et descendantes, chaque convolution permet de synthétiser la valeur d'un pixel en fonction de sa valeur et celle de ses voisins. Durant la phase descendante, le nombre d'informations observées diminue, mais la quantité de valeurs pour chaque nœud augmente, durant la phase ascendante, c'est le contraire, on reconstruit une image de la taille initiale grâce à une pile d'information toujours plus petite. En connectant chaque étape de décodage au niveau d'encodage de même taille d'image, il permet de saisir différents niveaux d'information dans l'image, notamment la structure générale et le détail. Mais cette méthode d'encodage ne permet pas réellement la combinaison de différents niveaux d'information.

Au contraire le FuseNet (Hazırbas et al., 2016) est un réseau d'apprentissage conçu pour apprendre à combiner une image d'information principale avec une information additionnelle qui sert d'indication. Le cas d'étude pour lequel ce réseau a été conçu est la segmentation des images d'intérieurs à partir de photographies et d'images de profondeurs en niveau de gris qui servent d'informations additionnelles. Ce réseau permet de mieux distinguer les objets sur différents plans lors de la segmentation. L'architecture de ce réseau est faite de deux chemins d'encodage parallèles, un pour les images en niveau de gris et l'autre pour les images couleurs, à chaque étape des informations de profondeurs sont réintégréées et fusionnées à l'image en couleurs encodée, ensuite l'image en couleurs est décodée en une segmentation. Il peut être adapté à la combinaison d'information géographique où la forme et la localisation des entités géographiques sont l'information principale, mais non suffisante à la génération de cartes.

On propose donc de comparer une architecture de GAN dont le générateur est un U-Net qui encode directement une pile d'images comme pour Pix2pix (Isola et al., 2017) et une architec-

## Représentation et combinaison de l'information géographique pour l'apprentissage profond

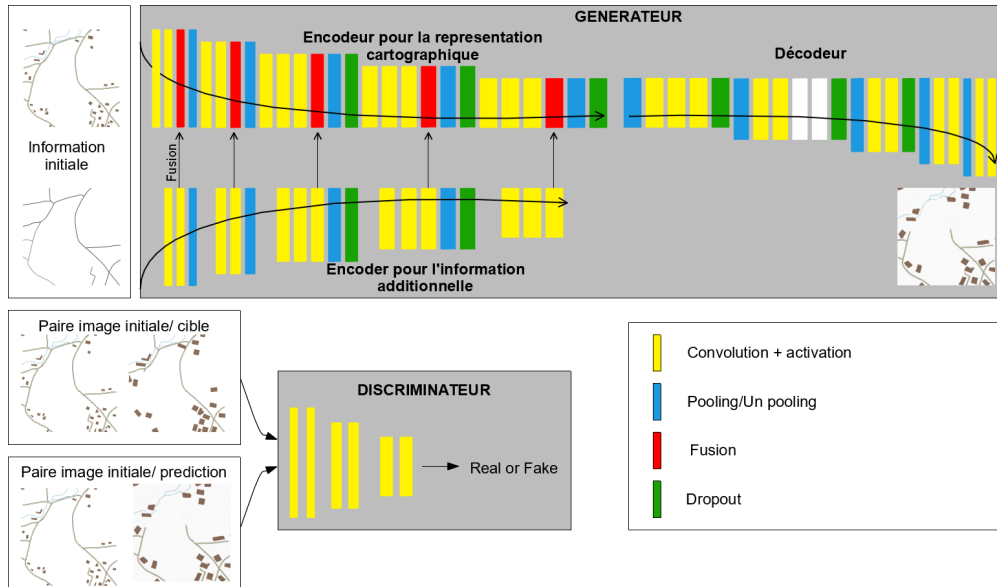


FIG. 1 – Architecture du GAN proposé opposant un FuseNet à un discriminateur.

ture de GAN dont le générateur est un FuseNet qui fusionne progressivement les informations additionnelles avec l'image de carte initiale puis reconstruit une carte généralisée. Cette architecture est décrite dans la Figure 1.

#### 4 Représentation de données vectorielles avec un tenseur

Nous proposons de comparer la représentation des données initiales sous forme d'extrait de cartes avec une autre représentation initiale de l'information qui limiterait les problèmes de coalescence entre objets de types différents. (Dans tous les cas, l'image cible est un extrait de la carte cible.) Cette représentation consiste à utiliser une pile de masques pour représenter la base de données initiale, chaque masque de la pile représentant un type d'objet. Dans sa version simple où on représente uniquement l'emprise des objets sans informations supplémentaires la pile d'images binaire route, hydrographie, bâtiments, serait alors suffisante pour représenter une ville comme dans l'exemple Figure 2. L'apprentissage des relations entre objets de types différents (Par exemple : la rivière longe une route.) serait facilité, car les deux éléments sont toujours représentés même en cas de colocalisation, et les relations entre objets ou groupes d'objets du même type seraient également plus lisibles, car non parasités par d'autres symboles en superposition. Pour aller plus loin, on peut sur le même principe proposer des images en niveau de gris pour représenter des attributs ou informations additionnelles (gradation ou catégorie) dans chaque image de la pile (et donc thème de la carte).

La Figure 3 présente un extrait de notre jeu test et les prédictions associées à chacune des approches.

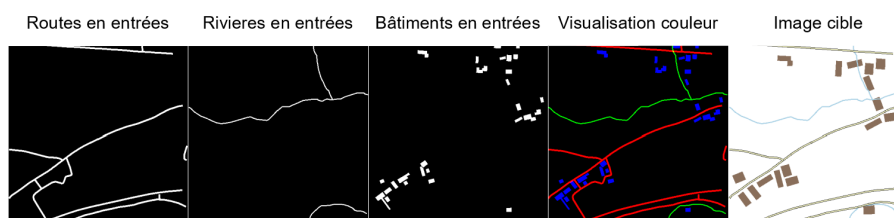


FIG. 2 – Extrait des images par thème de notre jeu de donnée.

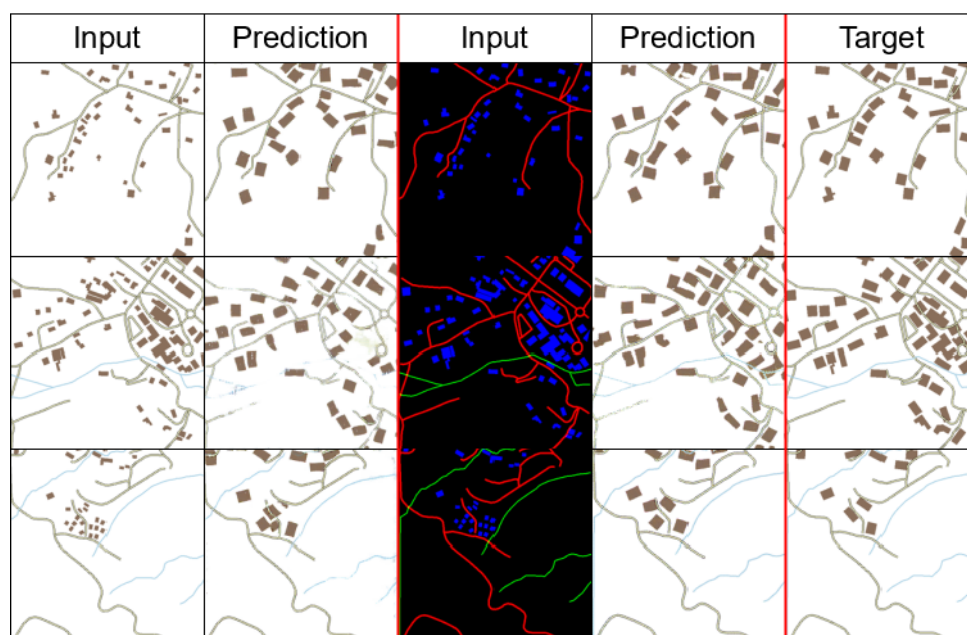


FIG. 3 – Comparaison des images de cartes générées par le GAN en fonction de la représentation initiale de l'information.



## Représentation et combinaison de l'information géographique pour l'apprentissage profond

On observe que pour la plupart des situations, comme la première ligne de la Figure 3, le mode de représentation n'impacte pas la qualité de la prédiction. Mais pour les cas où l'information est dense et la superposition entre symboles est importante dans l'image initiale, comme dans le cas décrit par la deuxième ligne de la Figure 3 la représentation par thème limite les effacements d'informations liés à ces superpositions. La superposition des symboles dans la prédiction n'est cependant pas totalement évitée. Enfin, dans le cas où l'information serait dense, mais où il n'y a pas de superposition dans l'image initiale, la prédiction basée sur la représentation par thème comporte moins de superpositions liées au grossissement des bâtiments, mais elle n'est pas totalement évitée.

## 5 Ajouter le contexte autour de l'imagerie

On se concentre ici sur les informations additionnelles pouvant être représentées sous forme d'un tenseur de même dimension que la représentation symbolisée de la donnée géographique, ainsi le lien avec les données initiales est aisé : le pixel de coordonnée  $(i,j)$  de l'image symbolisée représente et décrit la même portion de terrain que le pixel de même coordonnée  $(i,j)$  du tenseur additionnel. Ces informations additionnelles peuvent être de 3 types principaux décrits et illustrés dans la table 1.

Type	Description	Exemple
<b>Attributaire</b>	Renseigne sur les valeurs d'un attribut pour un type d'objet	La fonction d'un bâtiment.
<b>Contextuel</b>	Renseigne sur le contexte d'un objet géographique, permet d'avoir une information sur ce qu'il se passe en dehors de la tuile	Densité dans un îlot urbain.
<b>Relationnel</b>	Renseigne sur les relations spatiales implicites entre les objets dans l'espace représenté	Alignement de bâtiments.

TAB. 1 – Type et exemple d'information additionnelle.

Chacune des informations dans la Table 1 est représentée par un tenseur couvrant la même emprise que la représentation initiale en Figure 4. Ainsi, les pixels des images d'informations additionnelles peuvent prendre la forme d'une image en niveau de gris, une segmentation en  $n$  classes, ou d'une image binaire, mais on pourrait imaginer des cas plus complexes où l'information serait sous forme d'une image en couleur.

Pour nos expérimentations nous nous sommes concentrés sur le cas d'étude d'une information sur les routes en encodant la probabilité de chaque portion de route d'être gardée après généralisation. L'information est issue d'un premier modèle de prédiction sur les arcs du réseau routier avec un réseau d'apprentissage par convolution sur les graphes (Courtial et al., 2021a). Cette information est représentée sous forme d'une image en niveau de gris, les pixels de route ont pour valeur la probabilité de ne pas être conservé de la route à laquelle ils appartiennent, les autres pixels valent 1. Nous espérons que la combinaison de l'information additionnelle choisie améliore l'apprentissage de l'abstraction de ce thème. C'est-à-dire que les arcs peu importants



FIG. 4 – Représentation des exemples d'information additionnelle. a) Fonction du bâtiment, segmentation en  $n$  classes. b) Densité des îlots, niveau de gris. c) Axe d'alignement, segmentation binaire.

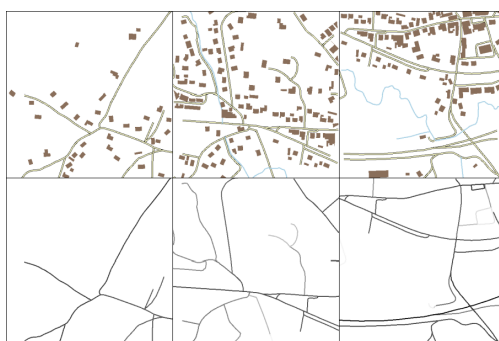


FIG. 5 – Extrait des images symbolisées et additionnelles de notre jeu de données.

et/ou causant des problèmes de lisibilité ne soient pas représentés, sans que cela altère la forme et la structure générale du réseau routier. La Figure 5 est un exemple de cette information pour quelques tuiles de notre zone d'étude.

Nous proposons de comparer l'empilement dans un unique tenseur de cette information avec l'image symbolisée de l'information initiale, avec l'utilisation de mécanisme de fusion de cette information additionnelle. Ces deux stratégies se traduisent par l'utilisation de deux générateurs différents dans une architecture GAN 3.3. D'une part, un générateur en forme de U-Net qui prend en entrée le tenseur d'empilement, et d'autre part, un générateur de type FuseNet qui fusionne deux tenseurs. Dans les deux cas, le générateur est opposé à un réseau de convolutions (CNN) qui sert de discriminateur. Cela permet de générer des images qui ressemblent à des cartes généralisées. La Figure 6 présente les images générées par chacun de ces GAN pour un extrait de notre jeu test.

On observe dans un premier temps que les images sont bruitées et les formes de bâtiments manquent de régularité. On peut expliquer cette diminution de qualité par l'augmentation du niveau d'abstraction entre image initiale et finale qui complexifie l'apprentissage par rapport à celle testée dans (Courtial et al., 2021b). On note également que les deux réseaux apprennent à effacer des routes pour éviter la superposition entre symboles de routes et de bâtiments. Mais le réseau procédant par empilement ne semble pas tirer profit de la carte de probabilités, en effet, il efface des portions d'arcs et non des arcs complets créant des déconnexions et les arcs sélectionnés ne permettent pas de préserver la structure générale du réseau routier. De plus, cette méthode semble aussi dégrader la représentation des rivières. Au contraire, le réseau

Représentation et combinaison de l'information géographique pour l'apprentissage profond

Probabilité	Input	Prediction fusion	Prediction empilement	Cible symbolisée

FIG. 6 – Comparaison des images de cartes générées par un GAN à partir de l'empilement ou de la fusion d'une image de carte et de l'information additionnelle de probabilité de sélection.

ayant un générateur en forme de FuseNet ne sur-sélectionne pas le réseau routier et préserve davantage la structure générale. Par exemple, pour la première situation générée la structure principale en Y du réseau routier est préservée par fusion (colonne 3) mais pas par empilement (colonne 4)

## 6 Conclusion

En conclusion, la représentation et la combinaison de ces informations figurent parmi les défis majeurs de la génération de cartes. Nos expérimentations permettent de décrire et de comparer des méthodes pour ce problème. Nous avons notamment montré qu'une représentation initiale par thème plutôt que symbolisée permet une meilleure description des relations spatiales de l'espace à cartographier. De plus, l'utilisation d'informations additionnelles est essentielle à l'amélioration des réseaux de générations de cartes, et cela, d'autant plus qu'ils demandent un niveau d'abstraction important, et que la fusion d'information est dans notre cas d'étude plus efficace que l'empilement pour intégrer de l'information géographique additionnelle. À l'issue de ce travail, nous notons qu'en plus de la recherche en cours dans la littérature sur des architectures mieux adaptées à l'information géographique, il est utile d'étudier l'enrichissement de l'information initiale fournit aux réseaux de neurones et d'élaborer des architectures de fusion d'informations plus performantes (GAN de fusion avec données non appariées, ou semi-appariées, permettant la fusion de plus d'informations additionnelles,

...). Enfin, nous pensons que la possibilité d'utiliser une information additionnelle sous une autre forme (image de taille différente, graphe, table de valeur, ...) figure parmi les perspectives importantes de ce travail, puisque la représentation sous forme d'image peut limiter la transmission de l'information géographique nécessaire à la généralisation cartographique.

## Références

- Chen, X., S. Chen, T. Xu, B. Yin, J. Peng, X. Mei, et H. Li (2020). SMAPGAN : Generative Adversarial Network-Based Semisupervised Styled Map Tile Generation Method. *IEEE Transactions on Geoscience and Remote Sensing*, 1–19. Conference Name : IEEE Transactions on Geoscience and Remote Sensing.
- Chen, X., B. Yin, S. Chen, H. Li, et T. Xu (2021). Generating Multi-scale Maps from Remote Sensing Images via Series Generative Adversarial Networks. *arXiv :2103.16909 [cs, eess]*.
- Courtial, A., A. El Ayedi, G. Touya, et X. Zhang (2020). Exploring the Potential of Deep Learning Segmentation for Mountain Roads Generalisation. *ISPRS International Journal of Geo-Information* 9(5), 338. Number : 5 Publisher : Multidisciplinary Digital Publishing Institute.
- Courtial, A., G. Touya, et X. Zhang (2021a). Can Graph Convolution Networks Learn Spatial Relations? *Abstracts of the ICA 3*, 1–2. DOI : 10.5194/ica-abs-3-60-2021.
- Courtial, A., G. Touya, et X. Zhang (2021b). Generative adversarial networks to generalise urban areas in topographic maps. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XLIII-B4-2021, pp. 15–22. Copernicus GmbH. ISSN : 1682-1750.
- Du, J., F. Wu, R. Xing, X. Gong, et L. Yu (2021). Segmentation and sampling method for complex polyline generalization based on a generative adversarial network. *Geocarto International* 0(0), 1–23. Publisher : Taylor & Francis \_eprint : <https://doi.org/10.1080/10106049.2021.1878288>.
- Feng, Y., F. Thiemann, et M. Sester (2019). Learning cartographic building generalization with deep convolutional neural networks. *International Journal of Geo-Information*.
- Fu, H., M. Gong, C. Wang, K. Batmanghelich, K. Zhang, et D. Tao (2019). Geometry-Consistent Generative Adversarial Networks for One-Sided Unsupervised Domain Mapping. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 2422–2431. IEEE.
- Ganguli, S., P. Garzon, et N. Glaser (2019). GeoGAN : A Conditional GAN with Reconstruction and Style Loss to Generate Standard Layer of Maps from Satellite Images. *arXiv :1902.05611 [cs]*.
- Guo, X., R. Nie, J. Cao, D. Zhou, L. Mei, et K. He (2019). FuseGAN : Learning to Fuse Multi-Focus Image via Conditional Generative Adversarial Network. *IEEE Transactions on Multimedia* 21(8), 1982–1996. Conference Name : IEEE Transactions on Multimedia.
- Hazirbas, C., L. Ma, C. Domokos, et D. Cremers (2016). FuseNet : Incorporating Depth into Semantic Segmentation via Fusion-Based CNN Architecture.

Représentation et combinaison de l'information géographique pour l'apprentissage profond

- Isola, P., J.-Y. Zhu, T. Zhou, et A. A. Efros (2017). Image-to-image Translation with Conditional Adversarial Networks.
- Jenny, B., M. Heitzler, D. Singh, M. Farmakis-Serebryakova, J. C. Liu, et L. Hurni (2020). Cartographic Relief Shading with Neural Networks. *arXiv :2010.01256 [cs]*.
- Kang, Y., S. Gao, et R. Roth (2019). Transferring Multiscale Map Style Using Generative Adversarial Network.
- Li, J., Z. Chen, X. Zhao, et L. Shao (2020). MapGAN : An Intelligent Generation Model for Network Tile Maps. *Sensors* 20(11), 3119. Number : 11 Publisher : Multidisciplinary Digital Publishing Institute.
- Ronneberger, O., P. Fischer, et T. Brox (2015). U-Net : Convolutional Networks for Biomedical Image Segmentation. Volume abs/1505.04597, pp. 234–241.
- Sester, M. (2005). Optimization approaches for generalization and data abstraction. *International Journal of Geographical Information Science* 19(8), 871–897.
- Touya, G., X. Zhang, et I. Lokhat (2019). Is deep learning the new agent for map generalization? *International Journal of Cartography* 5(2-3), 142–157. Publisher : Taylor & Francis \_eprint : <https://doi.org/10.1080/23729333.2019.1613071>.
- Weibel, R., S. Keller, et T. Reichenbacher (1995). Overcoming the knowledge acquisition bottleneck in map generalization : The role of interactive systems and computational intelligence. In A. U. Frank et W. Kuhn (Eds.), *Spatial Information Theory A Theoretical Basis for GIS*, Volume 988 of *Lecture Notes in Computer Science*, pp. 139–156. Berlin Heidelberg : Springer.
- Zhang, X. (2021). Deep Learning-based Multi-focus Image Fusion : A Survey and A Comparative Study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1. Conference Name : IEEE Transactions on Pattern Analysis and Machine Intelligence.
- Zhang, Y., Y. Yin, R. Zimmermann, G. Wang, J. Varadarajan, et S.-K. Ng (2020). An Enhanced GAN Model for Automatic Satellite-to-Map Image Conversion. *IEEE Access* 8, 176704–176716. Conference Name : IEEE Access.
- Zhu, J.-Y., T. Park, P. Isola, et A. A. Efros (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, pp. 2242–2251. IEEE.

## Summary

Deep learning allows the generation of maps from images. However, images only convey a limited view of the vectorgeographic information used to create a map. In this article, we explore the methods to combine and represent cartographic information as tensors to improve map generation, and especially generalised map generation using deep learning. First, we propose a new information representation that separates geographic themes (roads, rivers or buildings,...) to avoid some limitations of usual image representation. Then, we study how to combine additional information with our input image using a FuseNet architecture. Our first results permit us to show the interest of information combination compared to simple map images.