

Segmentation des images satellitaires en utilisant GAN

AYADI Hajar, ZEFFATE Najoua

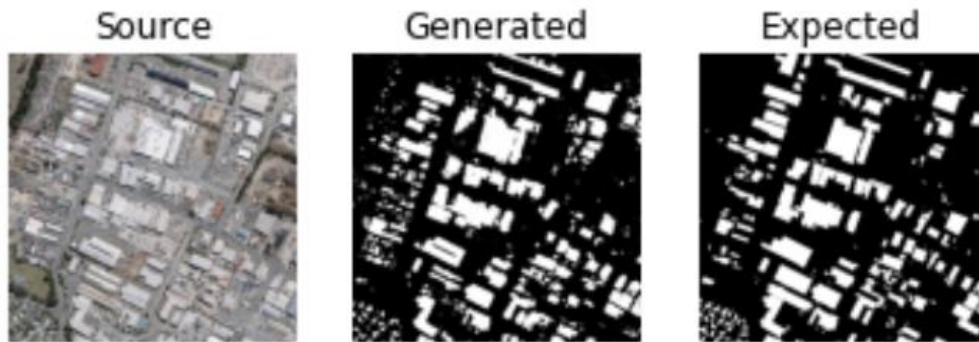
¹National Institute of Statistics and applied
Economies, Rabat, Morocco

nzeffate@insea.ac.ma, hayadi@insea.ac.ma

Introduction :

Durant ce projet nous avons essayé d'effectuer la segmentation des images satellitaires en utilisant les réseaux antagonistes génératifs (GAN).

L'objectif de ce projet c'est d'aboutir au résultat illustré dans la figure suivante :



Pour obtenir ce résultat, c'est toute une procédure derrière.

Dans un premier temps il faut chercher les données sur lesquelles on cherche à travailler, dans notre cas il s'agit des images satellitaires.

Ensuite il faut préparer cette data pour entraîner notre modèle ainsi pour le tester. Pour ce projet, nous avons utilisé le modèle GAN et plus précisément pix2pix GAN model car nous avons effectué du « image to image translation ».

Alors à travers ce rapport nous tentons répondre aux interrogations suivantes :

C'est quoi la segmentation des images ?

Quels sont les types de segmentation des images ?

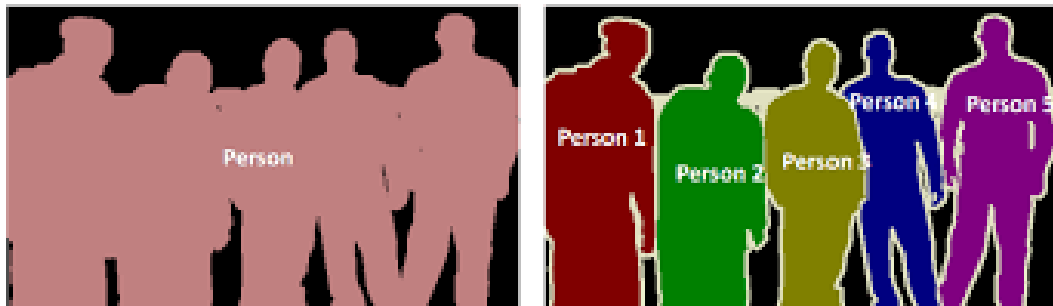
Comment segmenter une image via les GAN ?

1. Segmentation des images :

La segmentation d'images est une technique de computer vision qui consiste à découper de façon automatique une image en zones de pixels appartenant à une même classe d'objets. La segmentation d'images a de nombreuses applications, notamment en imagerie médicale.

1.2. Types de segmentation :

La segmentation d'images se divise en deux types, la segmentation sémantique (semantic segmentation) et la segmentation par instance (instance segmentation). Dans le cadre de la segmentation sémantique, on cherche à classifier les pixels de l'image comme appartenants ou non à une certaine catégorie. La segmentation par instance en revanche, permet d'avoir plus d'informations sur l'image, en divisant les différentes instances d'un même objet.

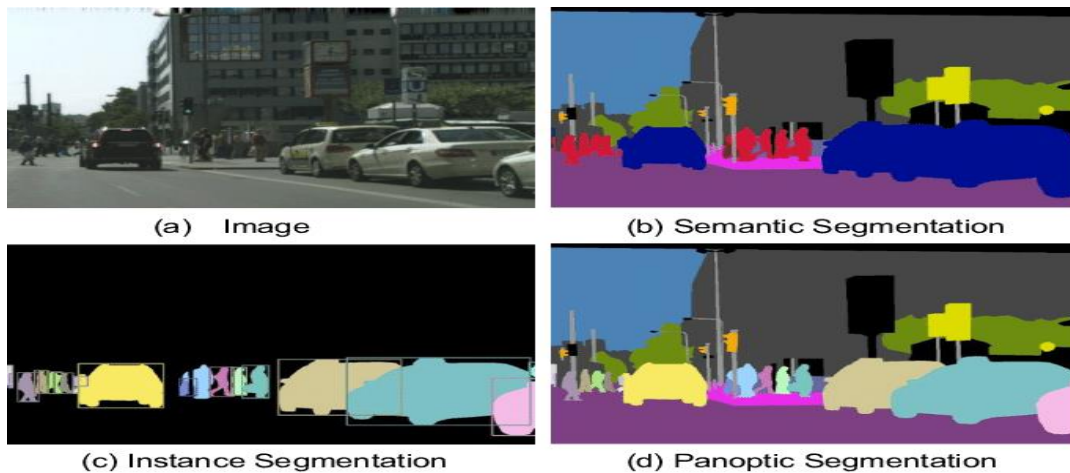


Semantic Segmentation

Instance Segmentation

Illustration de la différence entre segmentation sémantique et segmentation par instance

Lorsque l'on souhaite classifier toutes les zones de l'images, en divisant les instances, on parle de segmentation panoptique (panoptic segmentation). Elle consiste en une combinaison des deux types de segmentation.



(a) Image

(b) Semantic Segmentation

(c) Instance Segmentation

(d) Panoptic Segmentation

1.3. Image satellitaire :

L'imagerie spatiale est une technique d'observation à distance qui repose sur la prise d'images dans le domaine optique depuis l'espace par des équipements installés à bord de satellites artificiels.

1.4. Applications de la segmentation d'images en Analyse d'images satellitaires :

La segmentation d'images est aussi très utilisée pour l'analyse d'images satellitaires. Des modèles de segmentation permettent de détecter les routes, les bâtiments ou les champs. Récemment une polémique a éclaté concernant l'utilisation des images satellitaires pour détecter les citoyens qui ne déclarent pas leurs piscines. La limitation n'est pas technique ici, un modèle de détection de ce type serait facile à construire, elle est d'ordre juridique et éthique.

2. Réseaux antagonistes génératifs (Generative Adversarial Networks (GAN)):

Dans le domaine de l'intelligence artificielle, les réseaux antagonistes génératifs (ou GAN, generative adversarial networks) appartiennent à une catégorie de technique d'apprentissage automatique dans laquelle deux réseaux sont placés en compétition dans un scénario de jeu à somme nulle.

Généralement, les GAN sont non supervisés et apprennent par eux-mêmes à imiter n'importe quelle répartition de données.

Les deux réseaux neuronaux qui composent un GAN sont appelés générateur et discriminateur. Le générateur est un type de réseau neuronal convolutif qui crée de nouvelles instances d'un objet. Le discriminateur est un type de réseau neuronal déconvolutif qui détermine l'authenticité de cet objet ou son appartenance à un jeu de données.

Ces deux entités sont en compétition pendant la phase d'apprentissage où les pertes se confrontent les unes aux autres afin d'améliorer les comportements, ce mécanisme étant appelé rétropropagation.

L'objectif du générateur est de produire une sortie passable sans être pris en faute tandis que celui du discriminateur est d'identifier les contrefaçons. A mesure que la double boucle de rétroaction se déroule, le générateur produit une sortie de meilleure qualité et le discriminateur identifie mieux les contrefaçons.

Les GAN sont de plus en plus connus comme une forme évoluée d'apprentissage automatique. Des chercheurs et des développeurs ont expérimenté l'utilisation de GAN pour produire des copies, même imparfaites, d'œuvres célèbres telles que la Joconde et des portraits de personnes qui n'existent pas.

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

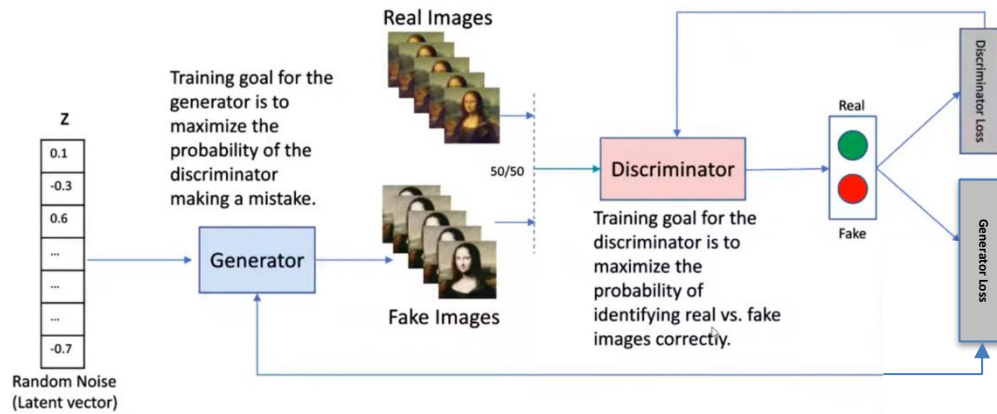
2 .1. Fonctionnement d'un GAN :

Pour créer un GAN, nous devons commencer par déterminer la sortie finale souhaitée et compiler un jeu initial de données d'apprentissage fondé sur ces paramètres. Ensuite, ces données sont envoyées de manière aléatoire dans le générateur jusqu'à ce qu'il obtienne une précision minimum dans la production des sorties.

Puis les images générées sont introduites dans le discriminateur accompagnées des points de données réelles provenant de la conception d'origine. Le discriminateur filtre les informations et renvoie une probabilité entre 0 et 1 pour représenter l'authenticité de chaque image (1 pour une image réelle et 0 pour une image contrefaite).

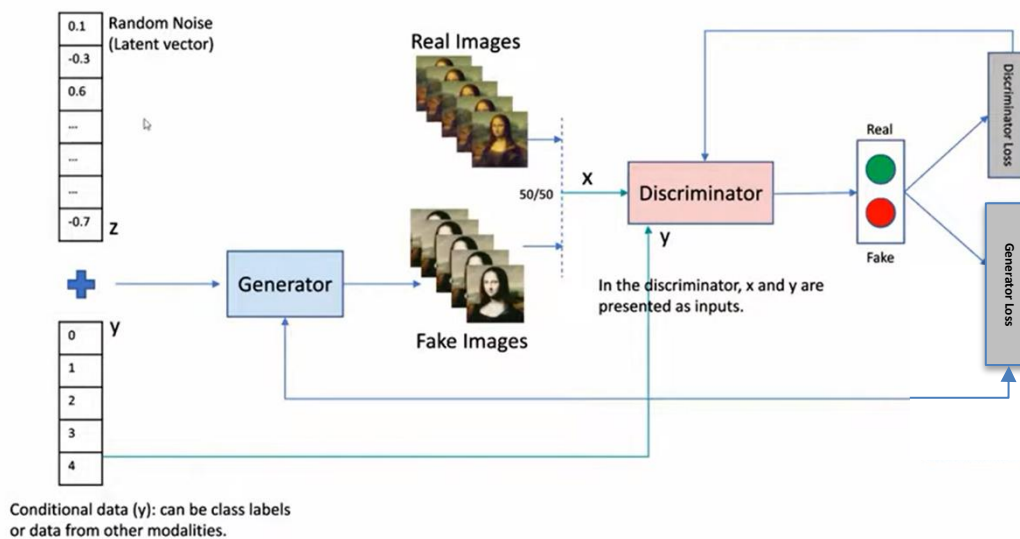
Enfin, ces valeurs sont contrôlées manuellement et répétées jusqu'à obtention du résultat souhaité. Les deux réseaux essaient d'optimiser des fonctions différentes, opposées, dans un jeu à somme nulle.

2.2. Architecture de GAN :



2.3. Architecture de CGAN :

Conditional Generative Adversarial Network



2.4. Utilisations des GAN

Les GAN peuvent apprendre à imiter la vie réelle afin de générer des contenus tels que du multimédia, du texte ou un discours. La principale fonction d'un GAN est de créer des données à partir de zéro. Les applications les plus courantes sont, par exemple :

- La conversion d'une image en une autre image ou la création d'images complètes à partir d'une ébauche.
- L'augmentation de la résolution d'une image.

- La modélisation de comportements et de mouvements humains récurrents pour des vidéos.
- L'entraînement d'un auteur artificiel chargé de rédiger un blog pour un site Web en s'inspirant d'un contenu précédent.
- La recreation d'une photo ou d'une peinture très connue.
- La prévision d'images vidéo successives.
- La génération d'une image à partir d'un texte.
- La création de vidéos truquées (deepfakes) ou l'imitation de comportements de célébrités.
- La production de descriptions photoréalistes de prototypes de produits.

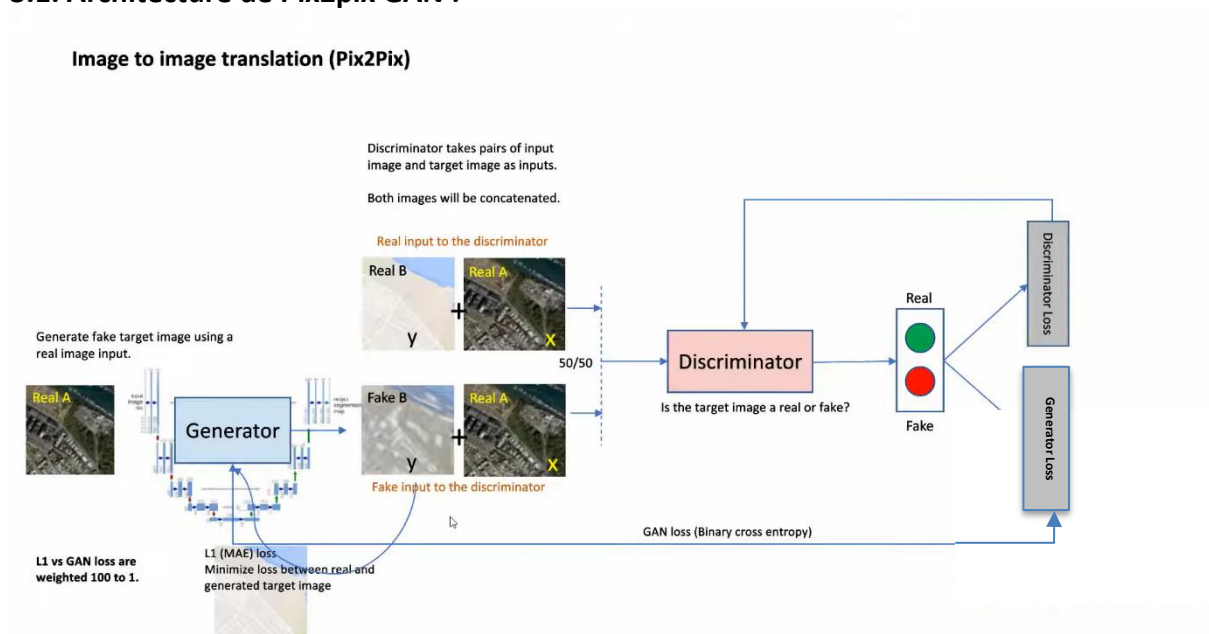
NB : Ce qui nous intéresse dans ce projet c'est l'application de GAN dans l'image to image translation, pour ce faire nous allons introduire le model pix2pix GAN.

3. Pix2Pix GAN :

Pix2pix est l'une des architectures GAN les plus célèbres et les plus utilisées pour toute tâche de traduction d'image à image.

Pix2pix sont des réseaux contradictoires conditionnels comme solution polyvalente aux problèmes de traduction d'image à image. Ces réseaux apprennent non seulement le mappage de l'image d'entrée à l'image de sortie, mais apprennent également une fonction de perte pour former ce mappage. Cela permet d'appliquer la même approche générique à des problèmes qui nécessiteraient traditionnellement des formulations de pertes très différentes. Ici, cette approche est efficace au niveau de la segmentation et nous n'avons plus à concevoir manuellement la fonction de cartographie.

3.1. Architecture de Pix2pix GAN :

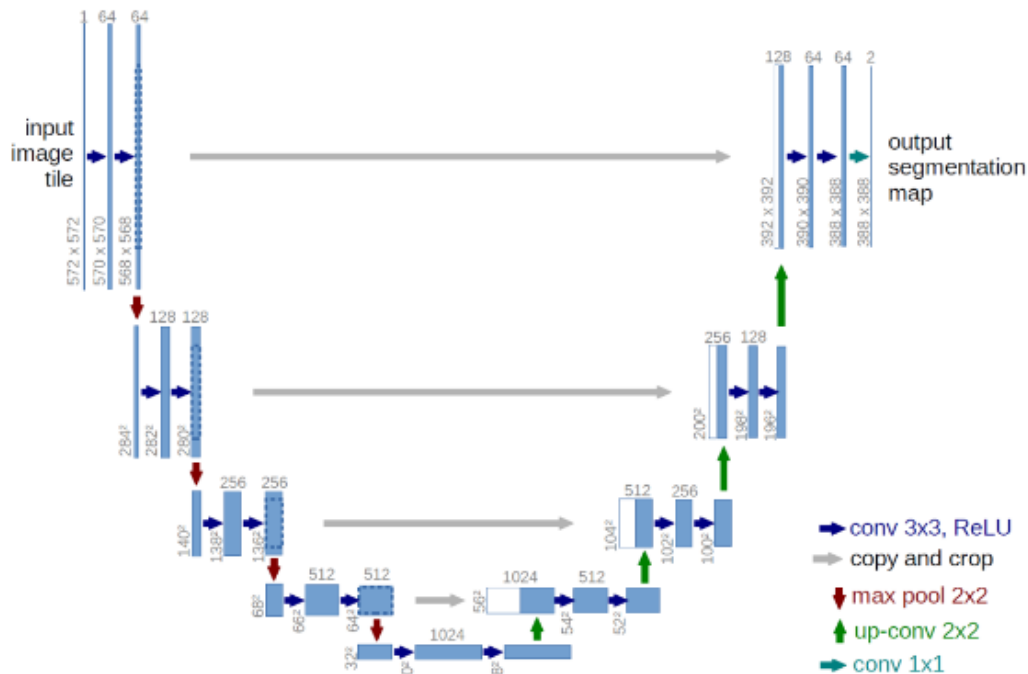


Pour le générateur il s'agit d'un model U-Net qui est un Modèle de Réseau de Neurones Entièrement Convolutif. Ce modèle fut initialement développé par Olaf Ronneberger, Phillip Fischer, et Thomas Brox en 2015 pour la segmentation d'images médicales. Un Modèle de Réseau de Neurones Entièrement Convolutif. Ce modèle fut initialement développé par Olaf Ronneberger, Phillip Fischer, et Thomas Brox en 2015 pour la segmentation d'images médicales.

L'architecture de U-NET est composée de deux « chemins ». Le premier est le chemin de contraction, aussi appelé encodeur. Il est utilisé pour capturer le contexte d'une image.

Il s'agit en fait d'un assemblage de couches de convolution et de couches de « max pooling » permettant de créer une carte de caractéristiques d'une image et de réduire sa taille pour diminuer le nombre de paramètres du réseau.

Le second chemin est celui de l'expansion symétrique, aussi appelé décodeur. Il permet aussi une localisation précise grâce à la convolution transposée.



Pour le discriminateur il s'agit d'un model apache GAN.

4. Partie application :

Pour réaliser ce projet nous avons utilisé la base de donnée "(Resized) aerial-imagery-for-roof-segmentation" présentée dans Kaggle "<https://www.kaggle.com/atilol/resized-aerialimageryforroofsegmentation>".

Composer des images pour les training process (repository image and repository target) et des images pour le testing process (repository image and repository target).

Pour la partie prétraitement de la data nous avons tous d'abord éliminé les images vides présentent dans les label quel que soit pour les images de training process ou bien les images de testing process.

Ensuite nous faire un resize dee images de la taille [1024,1024] à la taille [512,512].

Après ce pré-process nous avons stocké notre data dans le drive.

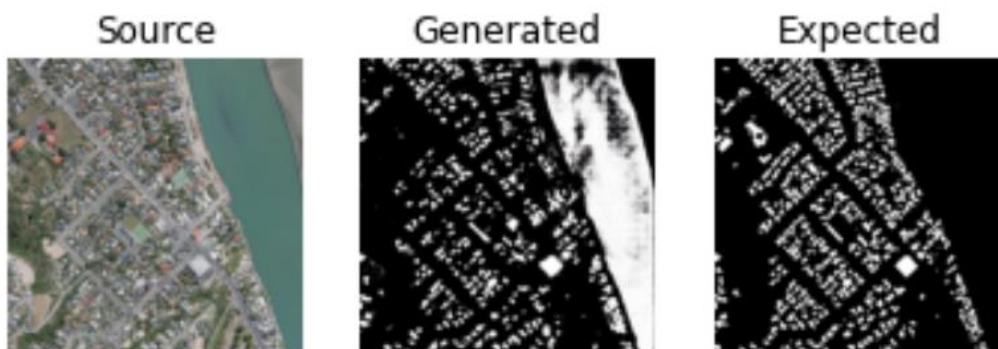
Maintenant notre data est prête pour l'entraînement, il suffit juste d'implémenter le code définissant le model pix2pix GAN.

Et au final l'applé pour le training process.

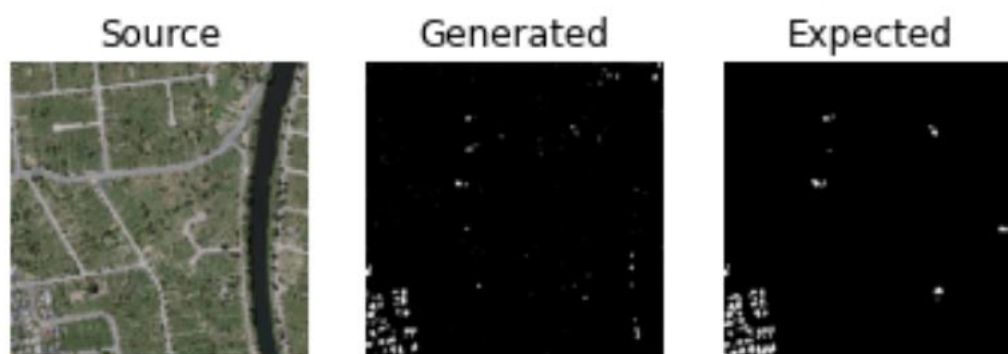
Après avoir trainer notre modèle nous avons importer les images de test pour verifier la performance de notre modèle. Qui a donné de bonne résultats.

Voici quelque screens résultante du testing process:

1/1 [=====] - 0s 24ms/step



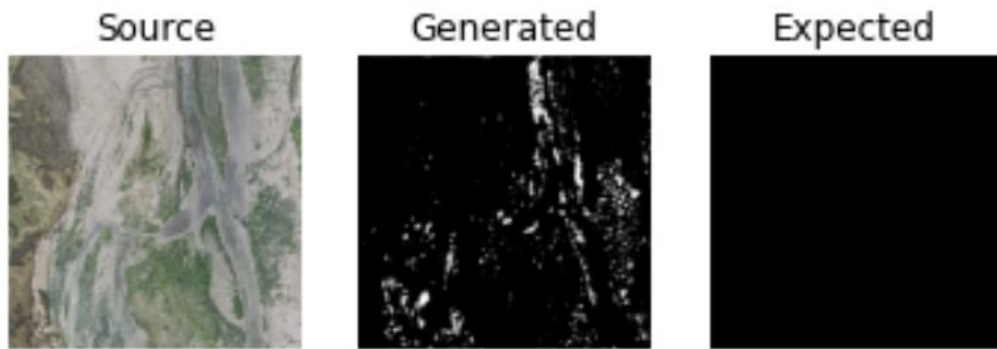
1/1 [=====] - 0s 29ms/step



1/1 [=====] - 0s 20ms/step



1/1 [=====] - 0s 21ms/step



Implémentation:

Vous trouverez ci-joint le lien vers notre notebook colab dont vous trouverez la partie implémentation bien détaillée:

<https://colab.research.google.com/drive/1L1150S1vCpHTrCjlsjH7ifZz--821WUK?usp=sharing>

Conclusion:

Durant ce projet nous avons senti la puissance des reseaux GAN dans le domaine des image-to-image translation et particulièrement dans le domaine de segmentation des images satellitaire.

Pour conclure, nous pouvons dire que les GAN offrent de très belles perspectives, notamment en analyse d'images ou en montage vidéos. Néanmoins, comme souvent lorsque les technologies sont aussi puissantes, il peut y avoir de gros inconvénients.

Evidemment ces inconvénients et les dangers qui en découlent ne doivent pas être une barrière au progrès technologique. Mais ils demandent une meilleure régulation, notamment grâce à des outils de traçabilité des données et des algorithmes utilisés.

Référence :

1. <https://medium.com/intel-student-ambassadors/segmentation-using-generative-adversarial-networks-80a161cf33c0>
2. https://github.com/bnsreenu/python_for_microscopists/tree/master/251_satellite_image_to_maps_translation
3. <https://datascientest.com/u-net>
4. <https://www.lemagit.fr/definition/Reseau-antagoniste-generatif-GAN>
5. <https://www.kaggle.com/atilol/resized-aerialimageryforroofsegmentation>
6. <https://dept-info.labri.fr/~vialard/Image3D/cours/cours-segmentation.pdf>