

Veille GAN

Sources :

- <https://www.lebigdata.fr/gan-definition-tout-savoir>
- https://fr.wikipedia.org/wiki/R%C3%A9seaux_antagonistes_g%C3%A9n%C3%A9ratifs
- <https://datascientest.com/gan-machine-learning>
- <https://theconversation.com/competition-entre-reseaux-de-neurones-artificiels-pour-creer-des-images-realistes-110005>
- <https://blent.ai/gan-deep-learning/>
- <http://www.xavierdupre.fr/app/papierstat/helpsphinx/lectures/others/gan.html>
- <https://www.tensorflow.org/tutorials/generative/dcgan>
- <https://github.com/nashory/gans-awesome-applications#super-resolution-1>
- https://developers.google.com/machine-learning/gan/gan_structure
- <https://neptune.ai/blog/6-gan-architectures>

Définition :

Un GAN ou Generative Adversarial Network (réseau antagoniste génératif en français) est une technique de Machine Learning. Elle repose sur la mise en compétition de deux réseaux au sein d'un framework.

- Apprentissage non supervisé
- réseaux adverses génératifs
- permettent de générer des images avec un fort degré de réalisme
- 2 réseaux de neurones qui s'entraînent en simultanément l'un contre l'autre pour créer des fausses images
 - Le premier est le générateur (G), il génère des fausses images de visage à partir d'un vecteur de bruit aléatoire (z). Cela permet à l'algorithme de générer à chaque nouvelle itération un visage différent.
 - Le discriminateur (D) quant à lui apprend à distinguer les images fausses créées par le générateur d'images de la vie réelle. En pratique c'est donc un classificateur.

Définition Wikipédia :

Un GAN est un modèle génératif où deux réseaux sont placés en compétition dans un scénario de théorie des jeux. Le premier réseau est le générateur, il génère un échantillon (ex. une image), tandis que son adversaire, le discriminateur essaie de détecter si un échantillon est réel ou bien s'il est le résultat du générateur. Ainsi, le générateur est entraîné avec comme but de tromper le discriminateur. L'apprentissage peut être modélisé comme un jeu à somme nulle.

L'apprentissage de ces réseaux est difficile en pratique, avec des problèmes importants de non convergence.

Fonctionnement :

Ces deux réseaux sont appelés » générateur » et » discriminateur « . Le générateur est un type de réseau neuronal convolutif dont le rôle est de créer de nouvelles instances d'un objet. De son côté, le discriminateur est un réseau neuronal » déconvolutif » qui détermine l'authenticité de l'objet ou s'il fait ou non partie d'un ensemble de données.

Pendant le processus d'entraînement, ces deux entités sont en compétition et c'est ce qui leur permet d'améliorer leurs comportements respectifs. C'est ce que l'on appelle la rétropropagation.

L'objectif du générateur est de produire des outputs sans que l'on puisse déterminer s'ils sont faux, tandis que l'objectif du discriminateur est d'identifier les faux. Ainsi, au fil du processus, le générateur produit des outputs de meilleure qualité tandis que le discriminateur détecte de mieux en mieux les faux. De fait, l'illusion est de plus en plus convaincante au fil du temps.

Dans un premier temps, il convient de déterminer ce que l'on souhaite que le GAN produise (l'output). La seconde étape consiste à composer un ensemble de données basé sur ces paramètres. Ces données sont ensuite entrées dans le générateur jusqu'à ce qu'il commence à produire des outputs convaincants.

Les images générées sont ensuite transmises au discriminateur, aux côtés des véritables points de données du dataset. Le discriminateur filtre les informations, et établit une probabilité comprise entre 0 et 1 pour déterminer l'authenticité de l'image. Le chiffre 1 signifie que l'image est réelle, le 0 indique qu'il s'agit d'un faux. Ces valeurs sont ensuite vérifiées manuellement, et le processus est répété jusqu'à ce que le résultat souhaité soit atteint.

Cas d'usages

Les GAN peuvent avoir de multiples cas d'usage. Il est possible de s'en servir pour imiter du contenu multimédia, des textes ou encore des discours. De manière générale, la principale utilité d'un GAN est de créer des données à partir de rien.

Parmi les applications populaires de cette technologie, on peut citer la modification ou la création d'images, la colorisation d'images en noir et blanc, ou encore l'amélioration de la définition d'une image. Il est par exemple possible de créer de faux visages grâce à cette technologie.

Il est aussi possible de s'en servir pour créer des animations de comportement ou de mouvement humains pour des vidéos.

De même, les GAN peuvent servir à recréer une peinture ou une photo populaire. C'est ainsi que Mona Lisa a pu prendre vie grâce à l'IA de Samsung. On peut aussi s'en servir pour entraîner une IA à écrire des articles pour un blog ou un site web en se basant sur du contenu existant. Il est aussi possible de générer une image à partir de texte, ou encore de produire des représentations photoréalistes de prototypes de produits.

GAN et Data Augmentation

Parmi les avancées réalisées grâce à l'utilisation du Deep Learning dans le domaine de la vision par ordinateur, on compte une technique appelée » data augmentation « ou augmentation des données.

Cette technique permet l'élaboration de modèles plus performants, avec un taux d'erreur fortement réduit. Elle repose sur la création de nouveaux exemples artificiels mais plausibles sur le domaine pour lequel le modèle est entraîné.

Dans le cas des données d'images, la data augmentation peut se présenter sous forme de modifications très simples : rognage, zoom, retournement... et autres transformations des images de l'ensemble de données d'entraînement.

Toutefois, les GAN offrent une alternative et une approche plus spécifique pour la data augmentation. En réalité, la data augmentation elle-même peut être considérée comme une version simplifiée des modèles génératifs.

Tout particulièrement, dans le cas des domaines les plus complexes, où le volume de données est limité, les modèles génératifs permettent un meilleur entraînement des modèles. On utilise notamment les GAN pour le Deep Learning par renforcement.

Ces réseaux génératifs antagonistes peuvent compenser le manque de données, et offrir des résultats multiples. La définition des images peut être améliorée, des images artistiques peuvent être entièrement créées, et il est même possible de modifier les images en changeant par exemple la saison ou en passant du jour à la nuit.

Limites

- Il est difficile de générer des images de haute résolution pour le GAN. En effet, le nombre de paramètres à apprendre dès le début est particulièrement élevé.
- Les images générées manquaient de diversité. Les GANs sont assez sensibles à ce qu'on appelle le 'mode collapse'. C'est-à-dire que le générateur génère toujours les mêmes images qui arrivent à tromper le discriminateur.
- Enfin, on ne peut pas contrôler les caractéristiques de sortie des images générées.

Le progressive growing ou croissance progressive permet de résoudre les limitations de résolution. Concrètement, l'adaptive growing consiste à entraîner le modèle sur des images de résolution croissante. Au fur et à mesure de l'apprentissage, on ajoute des couches au générateur et discriminateur pour travailler sur des images de meilleures résolutions tout en tirant parti de l'apprentissage déjà effectué.

De plus, le progressive growing améliore aussi la stabilité de l'algorithme et la diversité des images créées.

Applications

- Deepfake : génération de fausses images.

- Augmentation de la résolution d'une image
- Supprimer des éléments d'un visage
- Transformation d'objets (ex : pomme en orange...)
- Pix to pix : transforme dessin en image réelle

Architecture

- Pix to pix
- CycleGAN
- StyleGAN
- pixelRNN
- text-2-image
- DiscoGAN
- IsGAN