

Rapport sur le Projet CycleGan

Paulin Brissonneau Thomas Kebaili Ilyas Moutawwakil
Valentin Laurent Lilian Lecomte

XX mai 2020

Table des matières

1	Le Multiperceptron	2
1.1	Le perceptron simple	2
2	Les réseaux à convolutions	3
2.1	Bases mathématiques	3
3	Les GAN (Réseaux Adverses Génératifs) (pas fini et pas relu hein)	4
3.1	Principe général des GAN	4
3.2	Le DCGAN (Deep Convolutionnal Adversarial Network)	4
3.3	Le W-GAN (GAN de Wasserstein)	4
3.4	Étude de la convergence des GAN	5
3.4.1	L'effondrement des modes	5
3.4.2	Perte de l'équilibre	6

Chapitre 1

Le Multiperceptron

1.1 Le perceptron simple

Le principe du perceptron est de bla bla bla

Chapitre 2

Les réseaux à convolutions

2.1 Bases mathématiques

Le principe du CNN est de bla bla bla

Chapitre 3

Les GAN (Réseaux Adverses Génératifs) (pas fini et pas relu hein

3.1 Principe général des GAN

Le principe général des GAN repose sur l'utilisation de deux réseaux, ayant des objectifs contraires, on dit qu'ils sont **adversaires**. Le premier réseau transforme du bruit en image, c'est le **générateur**. Le deuxième réseau prend en entrées des images et a pour but de les classer selon deux classes, c'est donc un classifieur binaire, il est appelé **discriminateur**. Le plus souvent, le discriminateur sera alimenté par des images de deux sortes : celles provenant de la base de donnée (images réelles), et celles générées par le générateur, on rôle sera donc de dire si une image est réelle ou générée.

Blabla explication Blabal invention Yann Goodfellow [ref]

3.2 Le DCGAN (Deep Convolutionnal Adversarial Network)

Le DCGAN est la première architecture de GAN qui a été proposée [ref GoodFellow]. Le générateur et le discriminateur sont tous les deux des **réseau à convolutions** [ref].

Blablabla

3.3 Le W-GAN (GAN de Wasserstein)

blablabla

3.4 Étude de la convergence des GAN

De par leur caractère adversaires, les GAN requièrent un équilibre fin entre la générateur et le discriminateur, ils sont donc par nature **instables**. L'étude de la convergence des GAN est un domaine encore très actif de la recherche. Deux allons discuter de deux phénomènes très communs qui peuvent gêner ou ruiner l'apprentissage des GAN : l'**effondrement des modes** (*mode collapse*), et la **non-convergence** due à la perte d'équilibre du système.

3.4.1 L'effondrement des modes

L'effondrement des modes survient quand le réseau générateur ne génère pas des images conformant à l'ensemble de la distribution des images réelles, mais seulement à une petite partie. L'effondrement des modes est très visible lorsque la distribution des images réelles forme des zones bien séparées, c'est à dire quand celle-ci comporte des classes bien définies. La manifestation de ce phénomène se traduit par des images générées qui se ressemblent toutes. La figure [ref figure] montre un exemple du phénomène sur la base de données MNIST et CelebA.



FIGURE 3.1 – legende

Pour mieux comprendre le phénomène, il est intéressant de regarder la distribution des images de MNIST dans son ensemble, cela est possible grâce à des algorithmes de réduction de dimension. Attention, la réduction de dimension se fait dans l'espace des pixels, et non pas dans un espace sémantique, la visualisation ne permet donc pas de séparer efficacement les différentes classes, elle permet seulement un aperçu de la distribution dans l'espace sémantique. La figure [ref figure] présente une visualisation de MNIST par transformation t-sne [ref tsne].

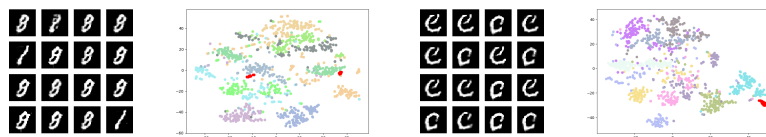


FIGURE 3.2 – legende

L'ensemble de points rouges correspond à un ensemble d'images générées par le réseau générateur lors de l'effondrement des modes. Sur les données MNIST,

on observe différents groupes de points (des *clusters*), ce sont les **modes** inhérents à la base de donnée MNIST : les chiffres de 1 à 9. Ce qu'il est intéressant de noter, c'est que les points générés sont rassemblés autour de un ou plusieurs pôles denses très localisés, qui ne sont pas répartis dans tout l'espace. Cela traduit l'effondrement des modes : les images générées ne couvrent qu'une petite partie de la distribution de la base de donnée d'entraînement.

Il n'y a pas de solution simple, directe et universelle pour lutter contre l'effondrement des modes, mais quelques solutions ont été proposées :

- La pénalisation de la similarité des images en sortie de générateur *mini-batch discrimination*. Cela consiste à ajouter un terme à la fonction de coût pour traduire la similarité (il peut s'agir de calculer une similarité pixel à pixel, ou d'estimer la similarité sémantique avec un autre réseau de neurones).
- Le *one-side label smoothing*. Cela consiste à changer l'objectif du discriminateur : son objectif ne sera plus de discriminer les fausses images avec une probabilité de 1, mais une probabilité plus faible, par exemple 0.9. Cela permet d'éviter la sur-confiance, et permet de laisser le générateur explorer tout l'espace des images réelles.
- Certaines architectures sont plus résistantes que d'autres à l'effondrement des modes. Par exemple, les GAN de Wasserstein ne présentent ce problème.

3.4.2 Perte de l'équilibre

Comme expliqué plus haut, l'apprentissage des GAN repose sur un équilibre fin entre le discriminateur et le générateur. Cet équilibre est parfois difficile à atteindre et est souvent instable, c'est pourquoi parfois le système s'effondre complètement. Cet effondrement vient souvent du fait que le discriminateur est devenu "trop fort" (sa fonction de perte tombe à zéro), et le générateur ne peut plus s'améliorer. Lorsque cela arrive, l'entraînement peut être arrêté : les images générées ne s'amélioreront plus. Un exemple de ce phénomène est illustré [ref figure], où l'on voit qu'à partir d'un cycle d'entraînement, la fonction de perte du discriminateur s'écroule et celle du générateur diverge.

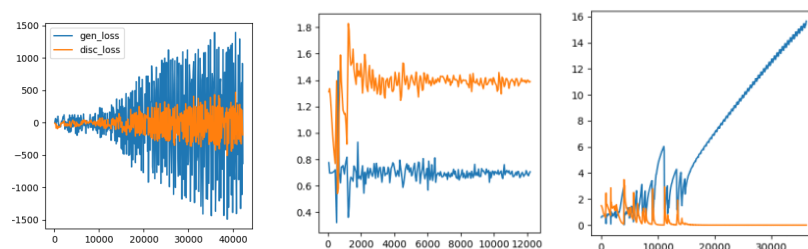


FIGURE 3.3 – légende

Il existe des solutions pour lutter contre ce problème, et cela consiste souvent à rééquilibrer les puissances ou les vitesses de convergence des différents réseaux. On peut par exemple diminuer la complexité du discriminateur, diminuer le taux d'apprentissage du discriminateur, ou mettre à jouer plus souvent le générateur que le discriminateur. Ajouter du bruit sur les images de la base de donnée

permet aussi se renforcer la stabilité de l'apprentissage. Par ailleurs, on peut noter que les GAN de Wassertein sont plus stables que les DCGAN, mais ne sont pas totalement immunisés aux problèmes de convergence.