

Generative Adversarial Networks pour le design génératif

Mots clés : modèles génératifs, apprentissage profond, données structurées.

1 Contexte

Les réseaux antagonistes génératifs (*generative adversarial networks*, GANs), introduits en 2014 par [5] sont des modèles génératifs qui cherchent à modéliser une distribution de probabilités à partir de données. Un GAN est basé sur la compétition entre deux composantes (le plus souvent des réseaux de neurones profonds), un « générateur » et un « discriminateur ». Le générateur produit des données candidates (par ex. des images) et le discriminateur cherche à déterminer si les données présentées sont issues de la « vraie » distribution ou non. Les données « authentiques » de départ servent à entraîner le discriminateur avant le début de la compétition. Suit une étape d'apprentissage commune (à la fois pour le générateur et pour le discriminateur) durant laquelle le générateur cherche à tromper le discriminateur en produisant des candidats qui ressemblent de plus en plus aux données authentiques, alors que le discriminateur s'améliore aussi en essayant de détecter comme non authentiques tous les candidats produits par le générateur. Cette étape d'apprentissage étant connue pour être difficile et souvent instable, de nombreux travaux ont visé à améliorer la qualité de cet apprentissage, voir par ex. l'introduction des Wasserstein GAN [1, 6].

2 Enjeux et objectifs

Les GANs sont principalement employés pour générer des images réalistes (visages, animaux, scènes, etc.), sans tenir compte de la structure interne de ces images, c'est à dire des objets présents et des relations entre ces objets. Quelques applications au design d'objets (voir par ex. [3]) emploient la même méthodologie pour la génération d'images d'objets. Utilisés dans [7] pour la génération d'objets 3D, les GAN produisent des objets qui sont des volumes composés de voxels (éléments de volume) mais suivant une approche similaire, qui ignore la structure interne de ces objets. Peu de travaux s'intéressent à la génération d'objets **composés** et emploient pour cela une architecture hiérarchique, voir par ex. [2].

Le premier objectif du stage proposé, correspondant à une première étape du travail, est d'évaluer des versions récentes des GAN, comme les Wasserstein GAN [1, 6], sur des images d'objets composés (constitués de composantes qui présentent des relations particulières entre elles), sans tenir compte de cette structure. Les résultats de cette phase doivent servir de référence aux développements ultérieurs.

Le second objectif du stage est de chercher à modéliser par apprentissage la structure des objets composés et à la prendre en compte lors du processus génératif. Les résultats issus de cette seconde phase du travail seront comparés à la référence obtenue lors de la première phase.

Suivant le temps disponible, il est envisageable de s'intéresser aussi à la séparation entre « style » et structure, suivant les idées de [4] ou [8].

3 Profil recherché

Le (la) candidat(e) doit avoir une bonne maîtrise du domaine de l'apprentissage statistique et de l'apprentissage profond, avec à la base de bonnes connaissances mathématiques. De bonnes capacités à programmer en Python sont également nécessaires, l'emploi de modèles d'apprentissage profond nécessitant l'utilisation de bibliothèques comme TensorFlow, PyTorch, etc. avec portage du code sur GPU.

Envoyez vos candidatures (avec CV, lettre de motivation, notes obtenues) à Michel Crucianu, Marin Ferecatu et Nicolas Thome (courriels : prénom.nom@cnam.fr).



FIGURE 1 – Exemples d’objets générés dans [7]

4 Organisation

Le stage débutera au printemps 2019 et durera 6 mois, sous la co-direction de Michel Crucianu, Marin Ferecatu (équipe Vertigo) et Nicolas Thome (équipe MSDMA).

Le stage se déroulera au CNAM Paris (<http://www.cnam.fr/>) dans l’équipe de recherche Vertigo (<http://cedric.cnam.fr/vertigo/>) du laboratoire CEDRIC (<http://cedric.cnam.fr/>). Plusieurs enseignants-chercheurs, doctorants et post-doctorants du CEDRIC travaillent sur l’apprentissage profond et certains sur les GANs.

Références

- [1] Martín Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *Proceedings of the 34th International Conference on Machine Learning, ICML 2017, Sydney, NSW, Australia, 6-11 August 2017*, pages 214–223, 2017.
- [2] Wei Chen, Ashwin Jeyaseelan, and Mark Fuge. Synthesizing designs with inter-part dependencies using hierarchical generative adversarial networks. In *ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Quebec City, Canada, Aug 2018. ASME.
- [3] Jaime Deverall. Using generative adversarial networks to design shoes : The preliminary steps. In *New Product Design with Popular Fashion Style Discovery Using Machine Learning : Proceedings of the Artificial Intelligence on Fashion and Textiles (AIFT) Conference 2018, Hong Kong, July 3–6, 2018*, 01 2018.
- [4] L. A. Gatys, A. S. Ecker, and M. Bethge. Image style transfer using convolutional neural networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2414–2423, June 2016.
- [5] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, NIPS’14*, pages 2672–2680, Cambridge, MA, USA, 2014. MIT Press.

- [6] Ishaan Gulrajani, Faruk Ahmed, Martín Arjovsky, Vincent Dumoulin, and Aaron C. Courville. Improved training of wasserstein gans. *CoRR*, abs/1704.00028, 2017.
- [7] Jiajun Wu, Chengkai Zhang, Tianfan Xue, William T Freeman, and Joshua B Tenenbaum. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. In *Advances in Neural Information Processing Systems*, pages 82–90, 2016.
- [8] Rui Zhang, Sheng Tang, Yu Li, Junbo Guo, Yongdong Zhang, Jintao Li, and Shuicheng Yan. Style separation and synthesis via generative adversarial networks. In *Proceedings of the 26th ACM International Conference on Multimedia*, MM '18, pages 183–191, New York, NY, USA, 2018. ACM.