

Espace latent compact et sémantique pour l'édition désenchevêtrée d'images

Alasdair NEWSON, ISIR - Paris

Yann GOUSSEAU, LTCI - Palaiseau

Gwilherm LESNÉ, LTCI - Palaiseau

Saïd LADJAL, LTCI - Palaiseau

Les progrès récents dans le domaine des modèles génératifs, en particulier des réseaux antagonistes génératifs (GAN) et les modèles de diffusion, ont permis des avancées significatives pour l'édition contrôlée d'images. Cependant, malgré leur capacité à appliquer des modifications réalistes à une image, ces méthodes peinent souvent à permettre l'édition d'attributs indépendamment les uns des autres, autrement dit à *désenchevêtrer* ces attributs. Dans ce travail, nous proposons un réseau de neurones sous la forme d'un auto-encodeur latent afin d'améliorer cette caractéristique tout en permettant un contrôle des attributs. Plus précisément, nous proposons de structurer l'espace de représentation des images d'une manière compacte, de sorte à mettre en évidence des directions d'édition pour chaque attribut souhaité. Nous testons notre approche sur plusieurs modèles, dont la classique architecture StyleGAN et montrons sa capacité à éditer des attributs de visages ainsi qu'à désenchevêtrer ces derniers. De plus, comme conséquence directe de notre méthode, nous montrons la capacité de notre réseau à résoudre des problèmes inverses contrôlés ou effectuer de la génération conditionnelle.