

Génération d'images de documents synthétiques à l'aide de réseaux de types GAN

Laboratoire :

LaBRI, CNRS UMR 58000, Université Bordeaux, Bordeaux-INP – Talence, France

Encadrants :

Boris Mansencal (boris.mansencal@labri.fr)

Nicholas Journet (journet@labri.fr)

Antoine Pirrone (antoine.pirrone@labri.fr)

Contexte :

Ce travail de recherche est en lien direct avec la thématique de recherche en analyse d'images de documents. Afin de pouvoir tester leurs algorithmes de traitement ou d'analyse d'images (par exemple pour un logiciel de reconnaissance de caractères) les chercheurs ont besoin de données labellisées. Afin d'éviter une annotation manuelle de très gros volumes de données, une solution consiste à créer de manière synthétique de fausses images de documents dont on contrôle la génération. L'image de la figure 1 montre un exemple de ce que nous savons faire actuellement. L'image d'origine (à gauche) est analysée, les caractères sont extraits par un utilisateur puis un algorithme génère une image synthétique similaire, avec un texte différent. Si les résultats sont visuellement bons, cette génération nécessite encore une intervention manuelle trop importante.

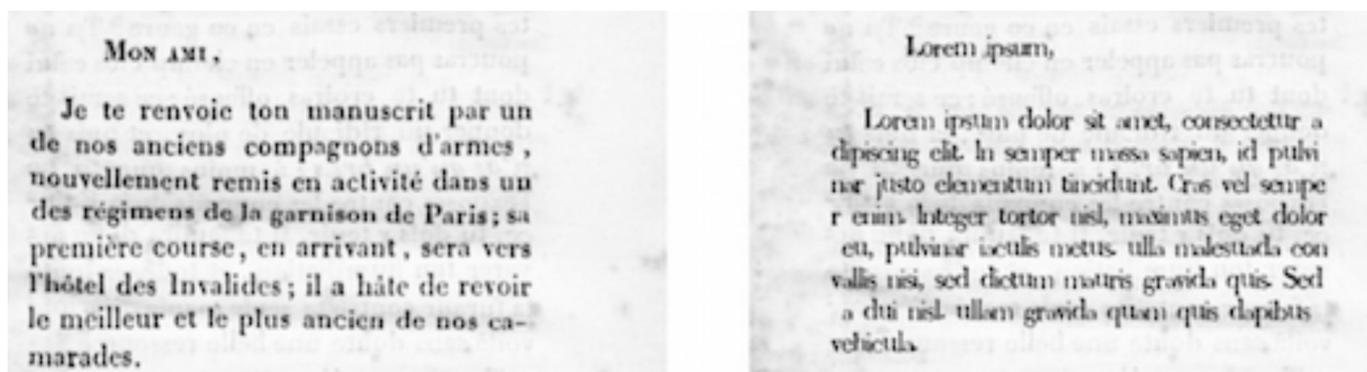


Figure 1 : Génération d'images synthétiques

D'autres exemples sur <http://doc-creator.labri.fr/>

Objectifs du Projet:

Un nouveau type de réseaux de neurones, les Generative Adversarial Networks (GAN), permet de générer des images synthétiques. L'idée consiste à entraîner simultanément deux réseaux de neurones : le premier génère des images synthétiques et le second, qui reçoit ces images synthétiques générés ainsi que de vraies images, doit déterminer quelles images sont vraies et quelles sont synthétiques. Au fur et à

mesure de l'entraînement des deux réseaux, le premier réseau apprend à améliorer sa génération de données synthétiques et le second est de plus en plus difficile à tromper.

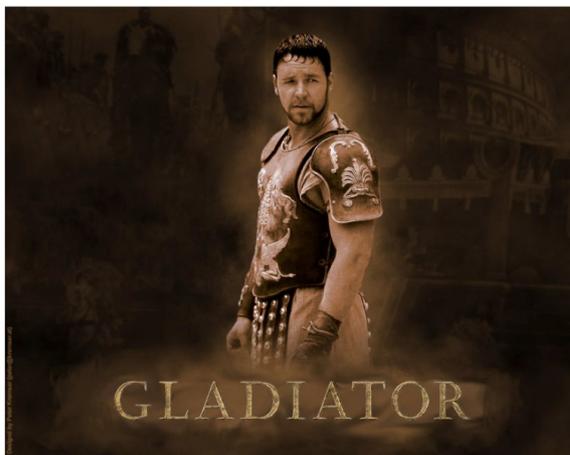
La vidéo suivante, proposée par Nvidia, montre les résultats obtenus sur la génération synthétique de visages humains :

<https://www.youtube.com/watch?v=G06dEcZ-QTg&feature=youtu.be>

Des expériences similaires ont été menées sur de la génération synthétique de textes :

- <https://bair.berkeley.edu/blog/2018/03/13/mcgan/>
- <https://kaonashi-tyc.github.io/2017/04/06/zi2zi.html>
- <https://multi threaded.stitchfix.com/blog/2016/02/02/a-fantastic-voyage/>
- <https://github.com/nashory/gans-awesome-applications>

Le réseau développé pour le projet de Berkeley permet notamment d'obtenir ce type de résultats :



L'image de gauche a été produite uniquement à partir de l'image de droite. Les lettres manquantes ont été créées par le réseau.

Le but de ce projet est d'étudier les travaux existants autour des GAN dans un contexte d'analyse d'images de documents.

Lors de ce projet vous aurez en particulier à :

- Faire un état de l'art des méthodes existantes.
- Recenser les codes sources fonctionnels permettant de générer du texte synthétique
- Tester ces codes sur un corpus d'images de documents anciens
- Fournir une analyse des résultats obtenus

Profil du candidat:

- Des connaissances en traitement et analyse d'images.
- Des notions de machine learning / deep learning

Outils

- Frameworks deep learning en python (Keras, Pytorch, Tensorflow...)