

# Cartoonisation d'image

Samuel Helye, Sylvain Leclerc

30 Octobre 2022

## 1 Contextualisation

### 1.1 Cartoonisation à partir d'image

Un article [1] est revenu à de nombreuses reprises dans les recherches, il traite d'une méthode de transformation d'image réelle dans un style cartoon avec l'utilisation d'un Generative Adversarial Network (GAN). Le processus peut se découper en trois étapes :

- **Surface representation** : consiste à extraire les surfaces continue d'une image, cette étape est inspiré de méthode de compositions utilisée par des artistes.
- **Structure representation** : contrairement à la méthode précédente qui conserve bien les textures, ici, on va s'attarder sur une découpe en aplats pour découper en bloc de couleur (le papier indique une combinaison de Superpixel et Selective search pour cette étape).
- **Texture representation** : cette étape va réduire le nombre de composantes de 3 à 1, l'objectif est de supprimer la luminance et a couleur de l'image pour garder l'intensité.

Ces étapes sont dans le but de générer une image cartoon à partir d'une image en entrée. Ensuite va intervenir un discriminant (qui est la seconde étape du GAN) pour classifier le résultat obtenu, ainsi ce travail en deux étapes va se compléter pour se corriger l'un, l'autre.

### 1.2 Génération d'image cartoon

En préparation de la présentation de nos pistes, il est intéressant de discuter à propos d'un mouvement plutôt populaire en ce moment : **la génération d'image**. Deux I.A on récemment fait beaucoup de bruit à ce propos, DALL-E 2 et Midjourney. Les résultats de ces deux programmes sont très impressionnants et surtout est une preuve que la génération peut avoir un impact important sur l'avenir, et cela semble intéresser les grosses entreprises comme TikTok qui offre un espace pour la génération d'image également. Dans le cas de Dall-e la donnée d'entrée est un texte, alors dans un premier une décomposition sémantique de ce texte est faite, par la suite un encodeur d'image va pouvoir lier cette phrase à des images connues (cette étape est le CLIP Training). Après avoir cela, que l'on peut grossièrement, à une base de donnée, Dall-E a recourt à un Diffusion Model, de manière simplifier l'IA va créer un bruit initial qui va être inversé pour obtenir une image qui correspond aux patterns attendus.

## 2 Nos pistes

### 2.1 Analyse sémantique et génération

Cette popularité de la génération d'image offre de nouvelles perspectives dans beaucoup de domaine et notre cas n'y échappe pas. En gardant l'utilisation d'une photo en donnée d'entrée, il est envisageable de décomposer sémantiquement la composition de cette photo pour ensuite générer une image équivalente en spécifiant un style. Bien que le processus soit encore très hypothétique à ce stade, on peut imaginer une photo simple tel que :



(a) Photo arbitrairement choisie

Ici on peut imaginer que l'analyse sort une liste d'éléments tel que :

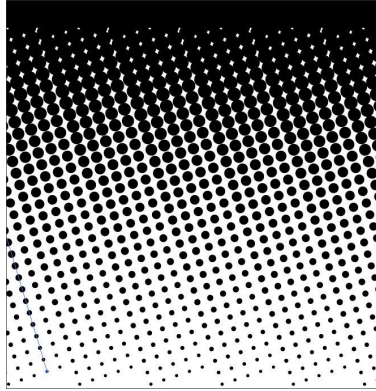
- Trottoir
- Route
- Banc
- Panneau de signalisation
- Arbre
- etc.

Ainsi, on pourrait utiliser un outil (pré-fait ou de notre travail) pour générer chacun de ses composants, et après une analyse composition, replacer les différents éléments générés, ce qui composera notre image finale.

## 2.2 Découpage de forme

Ici l'idée pourrait être appliquée avec des méthodes classiques et améliorée avec l'utilisation de CNN. Nous pourrions utiliser une méthode classique pour la détection de contours, comme le filtre de canny[2] et d'un procédé de propagation de couleur pour obtenir des aplats. Ainsi, en jouant avec des paramètres tels que le seuil, on pourrait dans un premier temps déterminer la forme globale d'une surface, puis obtenir les zones d'ombres et y appliquer des trames. On s'approche ainsi d'un procédé classique d'une composition graphique par un artiste.

Une approche possible pour le remplissage des zones, pourrait être de détecter la direction principale du gradient, puis de générer un dégradé dans la même direction. Mais en appliquant une texture caractéristique de la BD.



Ce procédé est basique et il existe de nombreuses manières de le perfectionner avec l'utilisation de méthode (pour la détection de contours, pour la propagation de couleur, pour déterminer la couleur de l'aplat, etc.) plus perfectionnée.

### 2.3 Utilisation d'une GAN

Pour ne pas plus se répéter, il nous est possible de choisir d'approfondir les recherches sur les méthodes qui sont abordées dans l'article présenté plus tôt.

## Références

- [1] Rudresh Joshi, Shreya Asoba. CARTOONIZE IMAGE FORMAT USING GENERATIVE ADVERSARIAL NETWORK (January 2021).
- [2] [Canny 1986] (en) J. Canny, « A Computational Approach To Edge Detection », IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, 1986, p. 679-698.