

Stage Technicien

Generative Adversarial Networks

David Albert

INSA Rouen

Du 18 juin 2018 au 7 septembre 2018

1 Présentation générale

- Contexte du stage
- Thème et Outils

2 Projet Baidu

- Problème
- Modèle
- Résultats

3 Generative Adversarial Networks

- Principe des réseaux GAN
- Exemples d'applications
- Projet final

Contexte

University of Dundee

Contexte

University of Dundee

- 17 000 étudiants
- Top 200 - *Times Higher Education's World*
- Reconnu pour études en sciences de la vie



Figure: Queen Mother Building

Thématique

Deep learning

Problème général :

- $\{X_i, Y_i\}_{i=1..n}$
- Caractériser la relation entre les X et les Y
- Construire de f_θ
- $\theta = \min(\mathcal{L}(\theta))$
- Descente du gradient

Thématique

Deep learning

Problème général :

- $\{X_i, Y_i\}_{i=1..n}$
- Caractériser la relation entre les X et les Y
- Construire de f_θ
- $\theta = \min(\mathcal{L}(\theta))$
- Descente du gradient

Réseau de neurones artificiels

Façon de modéliser un système neuronal en informatique.

Problème général :

- $\{X_i, Y_i\}_{i=1..n}$
- Caractériser la relation entre les X et les Y
- Construire de f_θ
- $\theta = \min(\mathcal{L}(\theta))$
- Descente du gradient

Réseau de neurones artificiels

Façon de modéliser un système neuronal en informatique.

Apprentissage profond

Méthodes de machine learning utilisant des réseaux de neurones multicouches pour modéliser f_θ . Ce sera notre cas par la suite.

Les outils

Les outils

Liste non exhaustive :

- GPU
- Python - Matlab
- Tensorflow
- Colab



(a) Colaboratory

(b) Tensorflow

Figure: Tools

Projet Baidu

Présentation de la compétition



Projet Baidu

Présentation de la compétition

- L'entreprise

- Multinationale chinoise
- Fondée en 2000 par Robin Li
- Domaines d'activité principaux : Internet et IA
- Concurrent de Google



Projet Baidu

Présentation de la compétition

- L'entreprise
 - Multinationale chinoise
 - Fondée en 2000 par Robin Li
 - Domaines d'activité principaux : Internet et IA
 - Concurrent de Google
- Les compétitions Baidu
 - Compétition d'intelligence artificielle
 - Réservée aux étudiants
 - Récompense : 10 000 €
 - Finale à Pékin (11 équipes)



Projet Baidu

- L'objectif

Projet Baidu

- L'objectif
- Les données

Projet Baidu

- Les données
 - 3619 images



Projet Baidu

• Les données

- 3619 images
 - Fichier JSON de description des images

Projet Baidu

- Les données

- 3619 images
- Fichier JSON de description des images

$$unit = \begin{bmatrix} name \\ id \\ num \\ ignore_region \\ type \\ annotation \end{bmatrix} = \begin{bmatrix} "stage1/train/e6bc4d3e343cel6b01b947f7630c506.png" \\ 1851 \\ 5 \\ \square \\ "dot" \\ \left[\begin{array}{l} \{"y": 298, "x": 138\} \\ \{"y": 345, "x": 505\} \\ \{"y": 451, "x": 145\} \\ \{"y": 398, "x": 602\} \\ \{"y": 203, "x": 412\} \\ \{"y": 240, "x": 19\} \end{array} \right] \end{bmatrix}$$

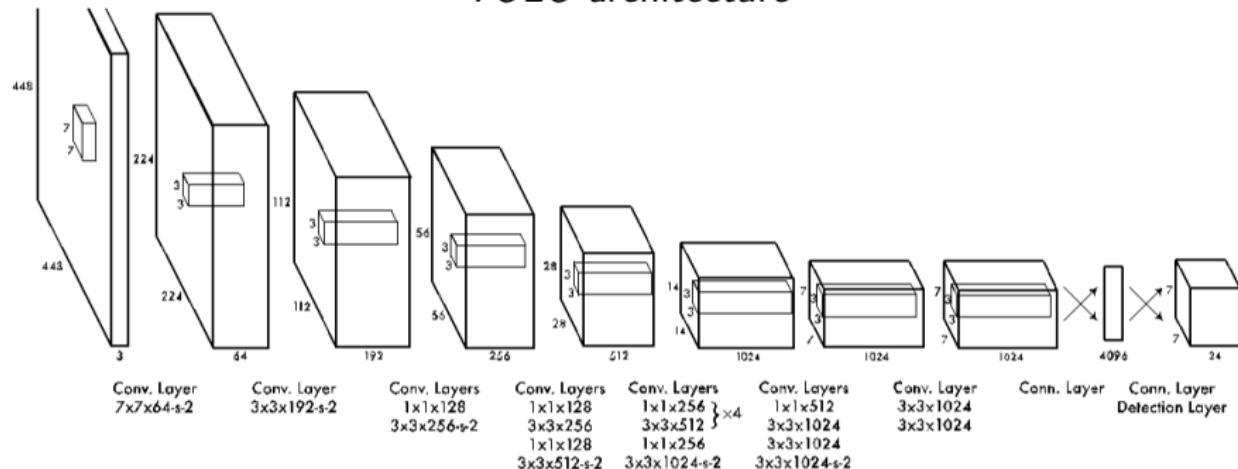
Projet Baidu

Choix du modèle

Projet Baidu

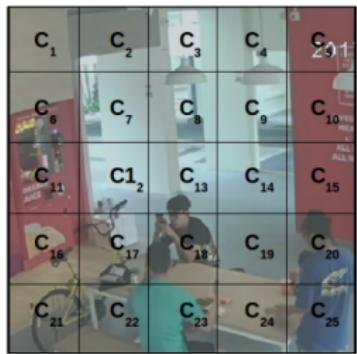
Choix du modèle

YOLO architecture

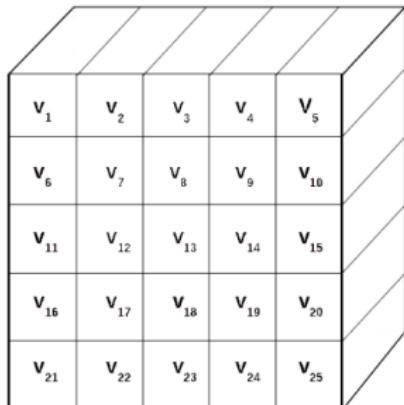


Choix du modèle

Modèle de prédiction



prediction →



$$v_i = (p_{i,1}, x_{i,1}, y_{i,1}, p_{i,2}, x_{i,2}, y_{i,2}, \dots, p_{i,5}, x_{i,5}, y_{i,5}), \forall i \in \{1, \dots, w \times h\}$$

Choix du modèle

Fonction de perte

$$\mathcal{L}_1 = \lambda_c \sum_{i=1}^{h \times w} \sum_{j=1}^B \mathbb{1}_{i,j}^{obj} [(x_{i,j} - \hat{x}_{i,j})^2 + (y_{i,j} - \hat{y}_{i,j})^2] + \lambda_{obj} \sum_{i=1}^{h \times w} \sum_{j=1}^B (p_{i,j} - \hat{p}_{i,j})^2$$

Entraînement du modèle

Entraînement du modèle

Paramètres d'apprentissage

- $\lambda_c = 0.5$ et $\lambda_{obj} = 1.0$

Entraînement du modèle

Paramètres d'apprentissage

- $\lambda_c = 0.5$ et $\lambda_{obj} = 1.0$

Temps d'apprentissage

- Apprentissage long ($\approx 1h/\text{époque}$)

Entraînement du modèle

Paramètres d'apprentissage

- $\lambda_c = 0.5$ et $\lambda_{obj} = 1.0$

Temps d'apprentissage

- Apprentissage long ($\approx 1h/\text{époque}$)

Problème rencontré

- Fort sur-apprentissage lié aux données

Entraînement du modèle

Paramètres d'apprentissage

- $\lambda_c = 0.5$ et $\lambda_{obj} = 1.0$

Temps d'apprentissage

- Apprentissage long ($\approx 1h/\text{époque}$)

Problème rencontré

- Fort sur-apprentissage lié aux données

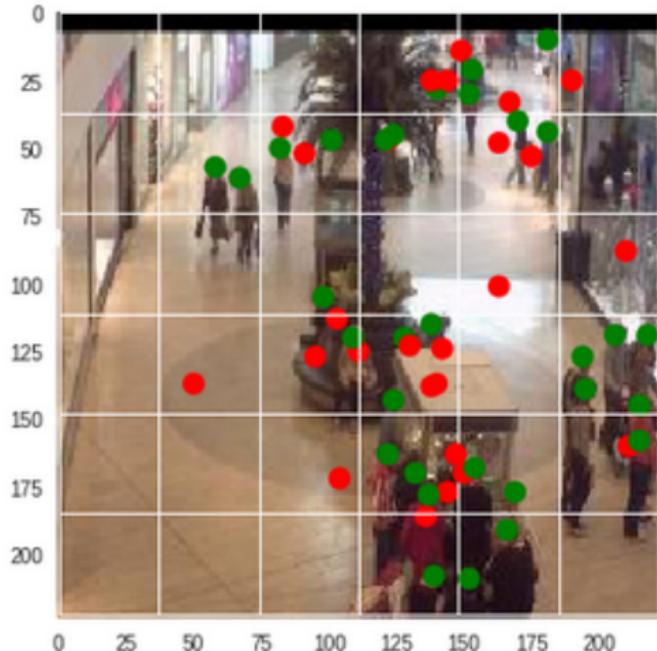
Solutions apportées

- Augmentation des données (72 500 images)
- Utilisation d'un modèle pré-entraîné
- Dropout

Quelques résultats qualitatif

Pred/True : 27 / 31

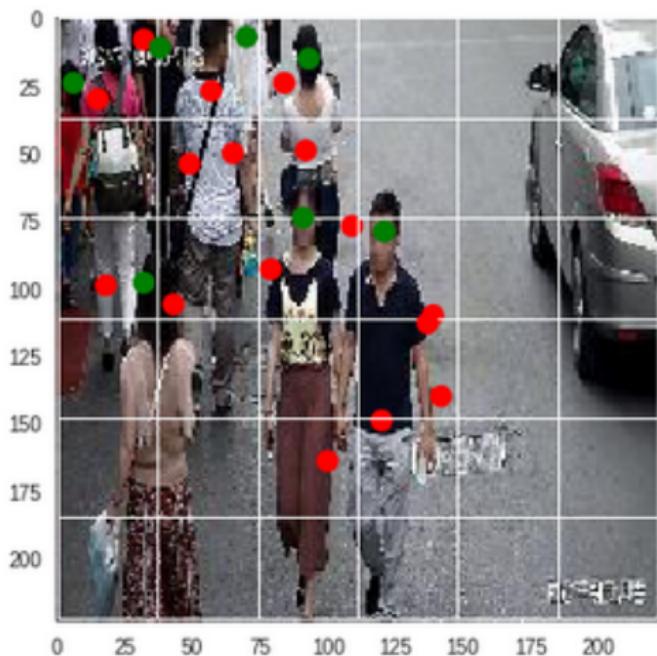
COUNT : 27



Quelques résultats qualitatif

Pred/True : 16 / 7

COUNT : 16



Résultats officiels

Critère d'évaluation du modèle

Minimiser la somme $\sum_{i=0}^n \frac{|P_i - G_i|}{G_i}$

où

- G_i est le nombre de personnes sur la photo
- P_i est la prédiction du nombre de personnes

Résultats officiels

Critère d'évaluation du modèle

Minimiser la somme $\sum_{i=0}^n \frac{|P_i - G_i|}{G_i}$

où

- G_i est le nombre de personnes sur la photo
- P_i est la prédiction du nombre de personnes

Résultats

David ≈ 0.33

Résultats officiels

Critère d'évaluation du modèle

Minimiser la somme $\sum_{i=0}^n \frac{|P_i - G_i|}{G_i}$

où

- G_i est le nombre de personnes sur la photo
- P_i est la prédiction du nombre de personnes

Résultats

David ≈ 0.33

George ≈ 0.24

Résultats officiels

Critère d'évaluation du modèle

Minimiser la somme $\sum_{i=0}^n \frac{|P_i - G_i|}{G_i}$

où

- G_i est le nombre de personnes sur la photo
- P_i est la prédiction du nombre de personnes

Résultats

David ≈ 0.33

George ≈ 0.24

1ère équipe ≈ 0.17

Résultats officiels

Critère d'évaluation du modèle

Minimiser la somme $\sum_{i=0}^n \frac{|P_i - G_i|}{G_i}$

où

- G_i est le nombre de personnes sur la photo
- P_i est la prédiction du nombre de personnes

Résultats

David ≈ 0.33

George ≈ 0.24

1ère équipe ≈ 0.17

Notre équipe ≈ 0.24

Résultats officiels

Critère d'évaluation du modèle

Minimiser la somme $\sum_{i=0}^n \frac{|P_i - G_i|}{G_i}$

où

- G_i est le nombre de personnes sur la photo
- P_i est la prédiction du nombre de personnes

Résultats

David ≈ 0.33

George ≈ 0.24

1ère équipe ≈ 0.17

Notre équipe ≈ 0.24 - 7 ème

Generative Adversarial Networks (GAN)

Generative Adversarial Networks (GAN)

Introduction

- Réseaux Antagonistes Génératifs
- Introduit par *Ian Goodfellow, 2015*
- Modèle génératif implicite

Definition

Un modèle génératif est dit **implicite** quand on ne peut pas déterminer explicitement l'expression de la loi de probabilité. On utilise dans ce cas une transformation d'une loi à une autre.

- Fait intervenir des réseaux de neurones
- Bons résultats pour la génération d'images

Generative Adversarial Networks (GAN)

Principe



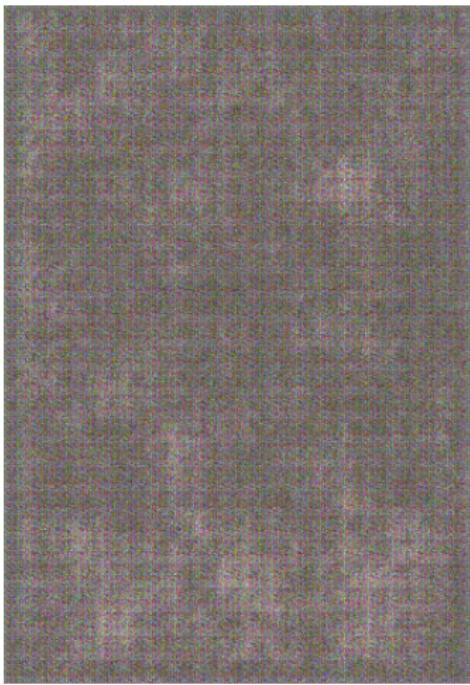
Dérivées du model GAN

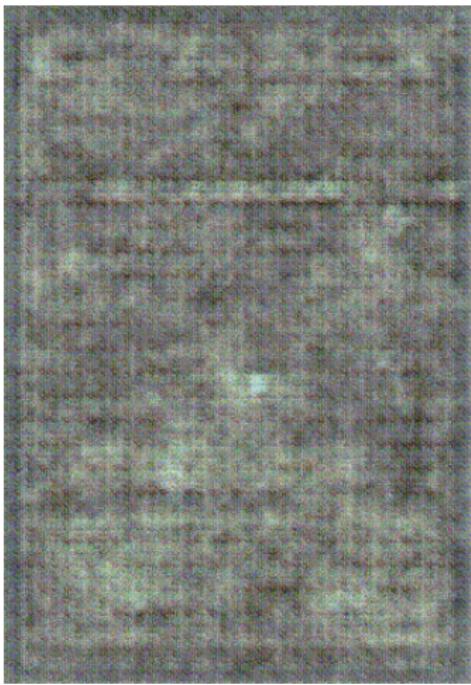
- Deep Convolutional GAN
- Conditional GAN
- Stack GAN
- pix2pix
- Cycle GAN
- 3D GAN

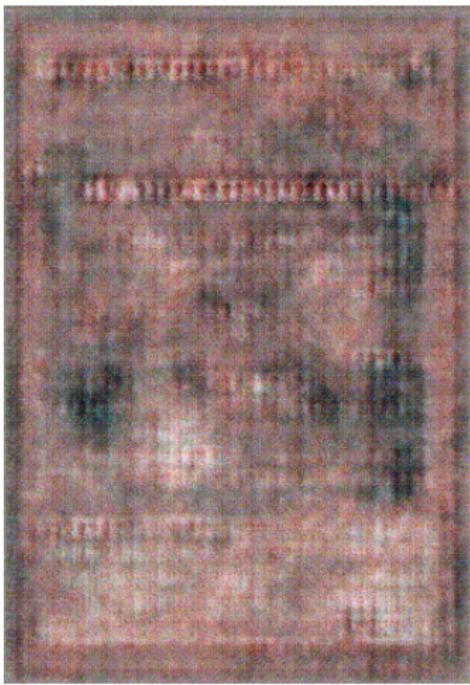
Dérivées du model GAN

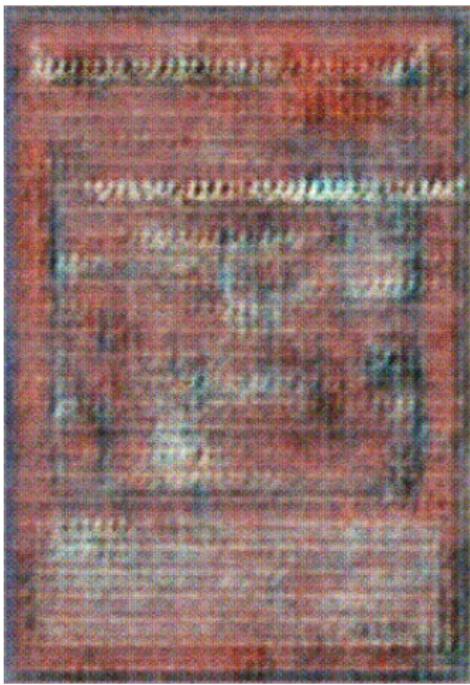
Deep Convolutional GAN









































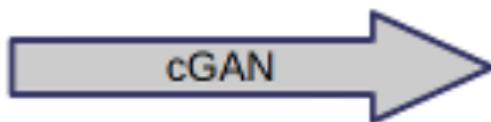




Dérivées du model GAN

Conditional GAN

0
0
0
0
0
0
0
0
0
1



Dérivées du model GAN

Stack GAN

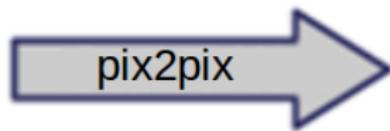
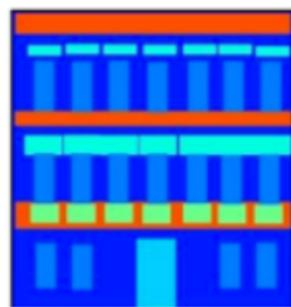
*This bird is blue
with white and has
a very short beak*

Stack GAN



Dérivées du model GAN

pix2pix



Dérivées du model GAN

Cycle GAN



zebra

Cycle GAN



horse

Dérivées du model GAN

3D GAN



zebra

Cycle GAN



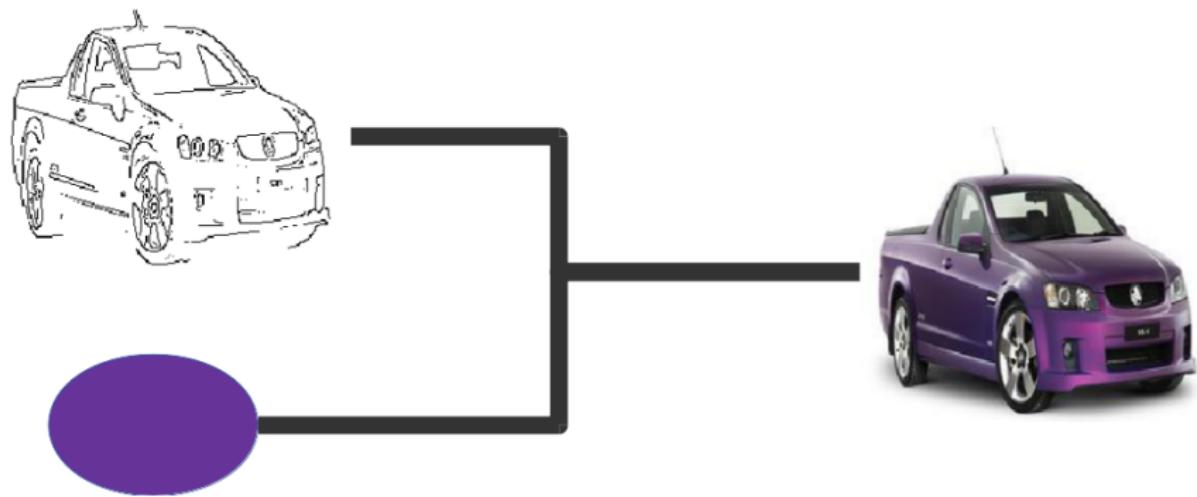
horse

Projet "*Edge to car*"

Objectif

Projet "Edge to car"

Objectif



Projet "Edge to car"

Pré-traitement des données

- Tâches réalisées

Projet "Edge to car"

Pré-traitement des données

- Tâches réalisées
 - Téléchargement de plus de 3 000 images

Projet "Edge to car"

Pré-traitement des données

- Tâches réalisées

- Téléchargement de plus de 3 000 images
- Prétraitement des données

Projet "Edge to car"

Pré-traitement des données

- Tâches réalisées

- Téléchargement de plus de 3 000 images
- Prétraitement des données
 - Récupération des contours



PIL



OpenCV



GradientAlgo

Projet "Edge to car"

Pré-traitement des données

- Tâches réalisées

- Téléchargement de plus de 3 000 images
- Prétraitement des données
 - Récupération des contours
 - Récupération de la couleur

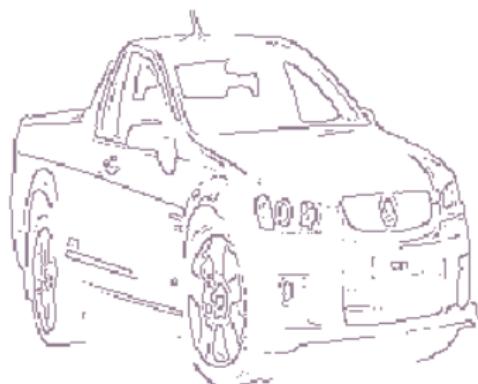


Projet "Edge to car"

Pré-traitement des données

- Tâches réalisées

- Téléchargement de plus de 3 000 images
- Prétraitement des données
 - Récupération des contours
 - Récupération de la couleur
 - Mise en commun



Projet "Edge to car"

Pré-traitement des données

- Tâches réalisées

- Téléchargement de plus de 3 000 images
- Prétraitement des données
 - Récupération des contours
 - Récupération de la couleur
 - Mise en commun
- Construction du modèle

Projet "Edge to car"

Pré-traitement des données

- Tâches réalisées

- Téléchargement de plus de 3 000 images
- Prétraitement des données
 - Récupération des contours
 - Récupération de la couleur
 - Mise en commun
- Construction du modèle
- Entrainement du modèle

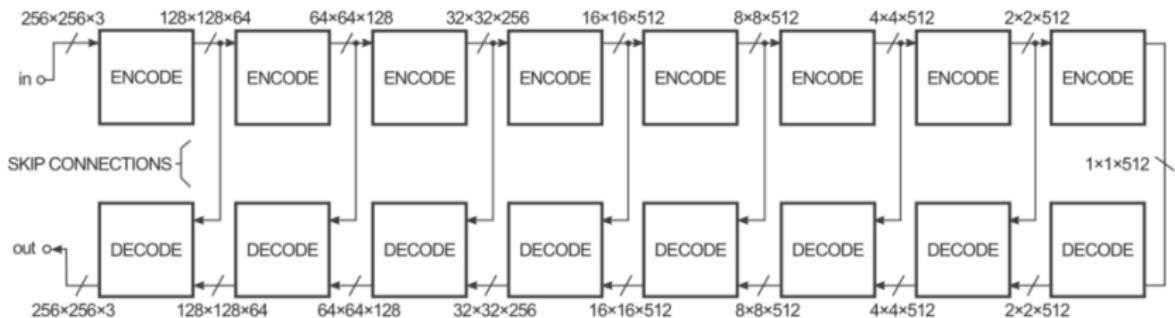
Construction du modèle

Architecture

Construction du modèle

Architecture

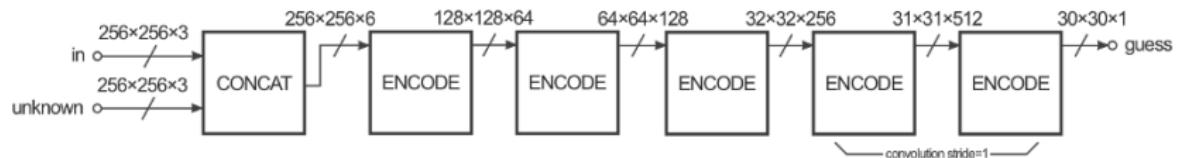
Architecture du générateur



Construction du modèle

Architecture

Architecture du réseau discriminant



Construction du modèle

Fonctions de perte

Discriminant:

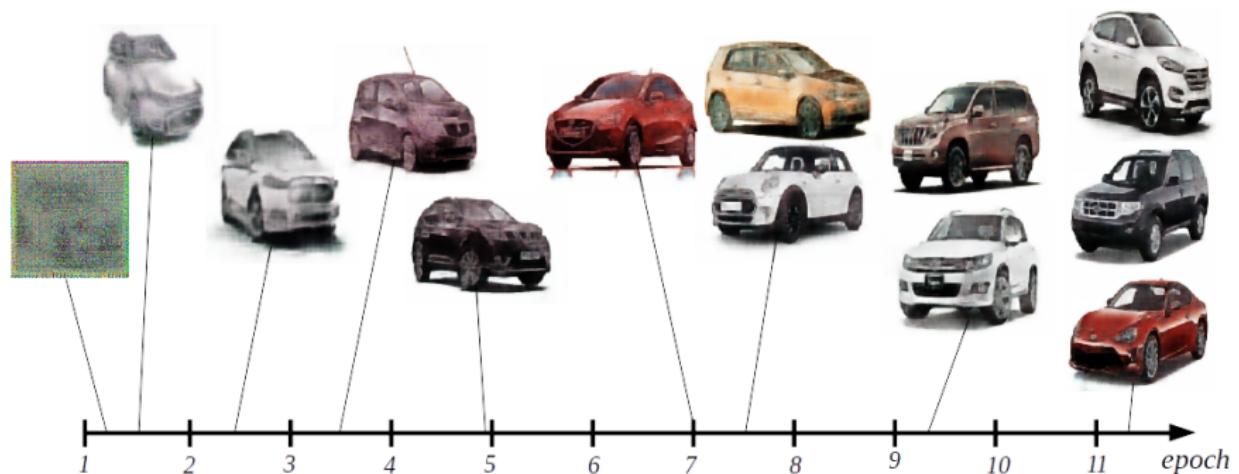
$$\mathcal{L}_D = - \sum_{i=1}^{k \times k} \log \left[D_{\theta_d}^{(i)}(y) \right] + \log \left[1 - D_{\theta_d}^{(i)}(G_{\theta_g}(x)) \right]$$

Générateur:

$$\mathcal{L}_G = -\frac{\lambda_{FD}}{k \times k} \sum_{i=1}^{k \times k} \log D_{\theta_d}^{(i)}(G_{\theta_g}(x)) + \frac{\lambda_{L1}}{n \times n \times 3} \sum_{i=1}^{n \times n \times 3} |y_i - G_{\theta_g}^{(i)}(x))|$$

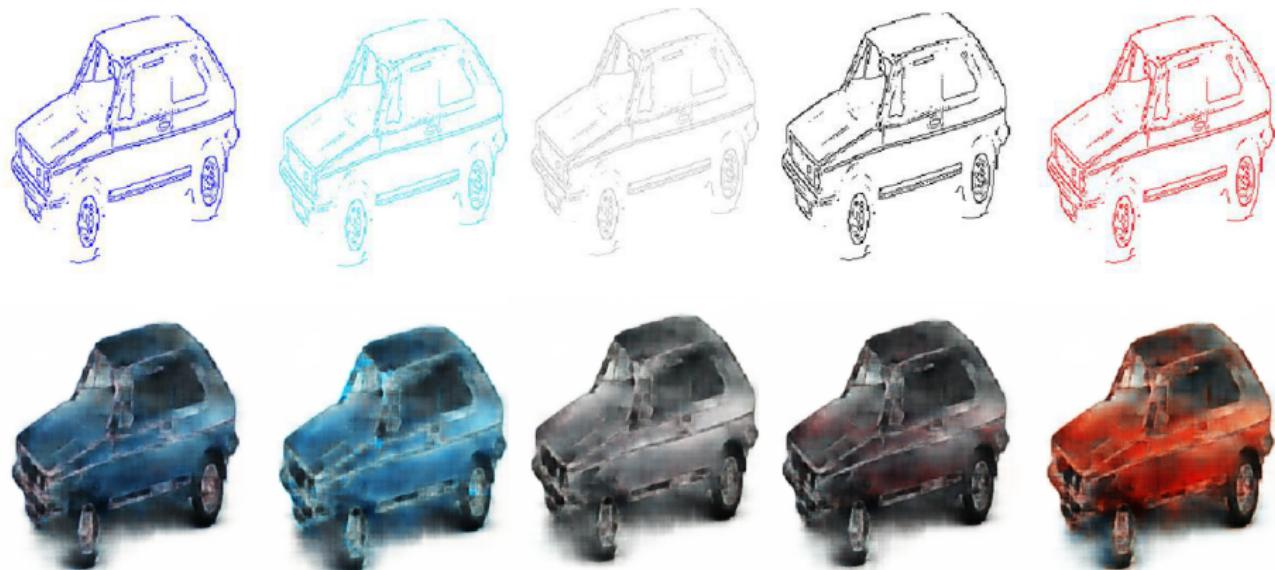
Entraînement du modèle

- $\lambda_{FD} = 1$ et $\lambda_{L1} = 100$
- Très rapide ($\approx 1h/5$ époque)



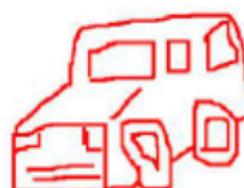
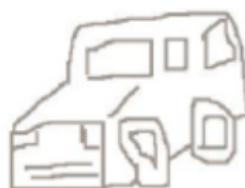
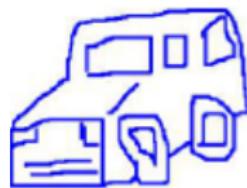
Quelques résultats

- Changement de la couleur d'une voiture

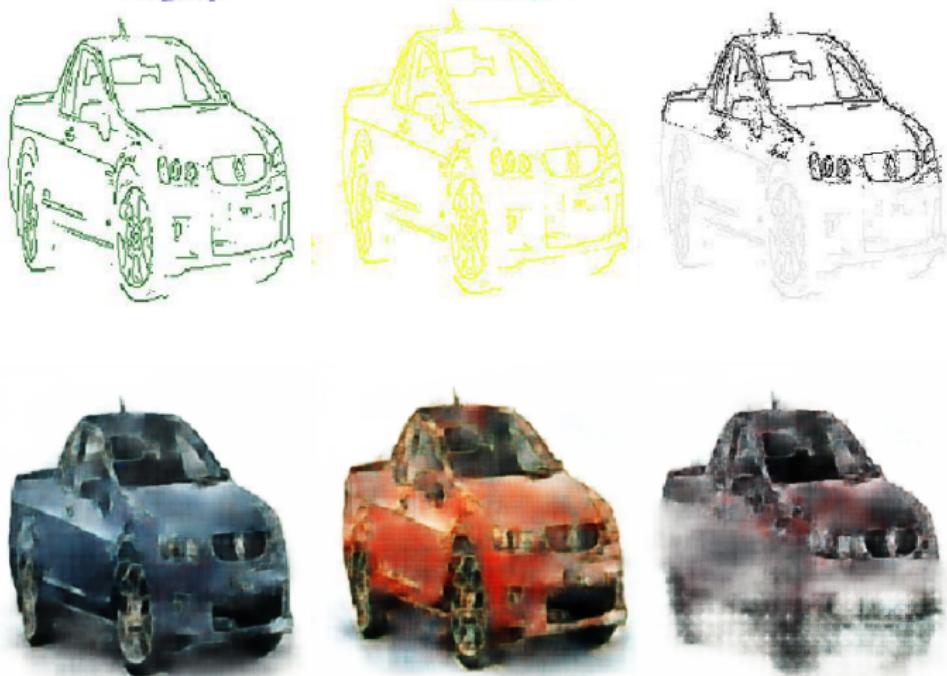


Quelques résultats

- Changement de la couleur d'une voiture
- Coloration d'un dessin de voiture



Autre résultats



Amélioration

Amélioration



Figure: Voiture bleu clair

Amélioration



Figure: Voiture bleu foncé

Amélioration



Figure: Voiture rouge

Amélioration



Figure: Voiture gris clair

Amélioration



Figure: Voiture jaune

Amélioration



Figure: Voiture "jaunâtre"

Amélioration



Figure: Voiture rose pâle

Amélioration



Figure: Voiture bicolore