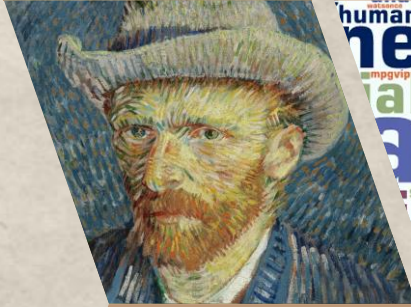# VanGoghGAN

Group project for Data Mining Master's
Course 2020-2021 at Sapienza

Felice Massotti 1699768
Daniela Moretti 1708855
Redemptor Jr Laceda Taloma 1704899

1 Description of the task

2 Related works

3 Implementation

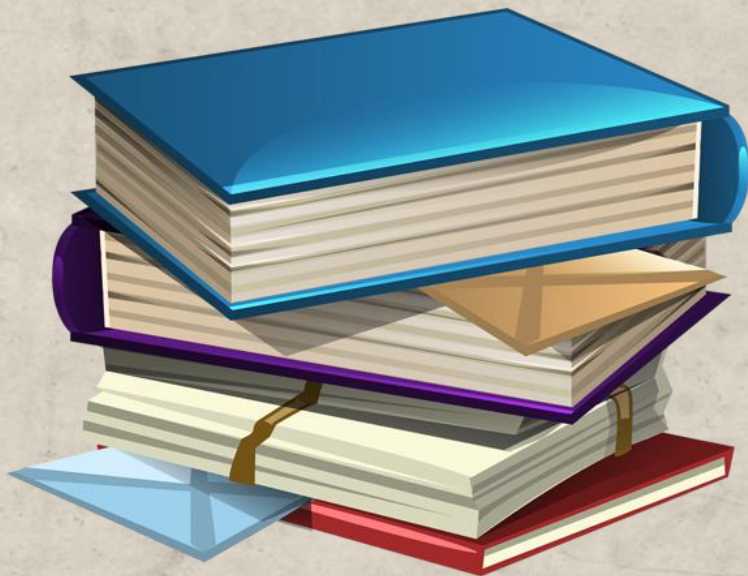4 Results and Conclusions

5 Future works

# Description of the task

- Painting the unfinished drawings of landscapes left by Van Gogh with his style.

- Building an appropriate dataset with Van Gogh's paintings is hard.

- Relaxing the task: learn from 20+ impressionists.

## 2

# Related works

- Pix2Pix

- Auto-Painter

- CycleGAN

- SketchyGAN (State of the art)

# *Related works*

———————————— • ————————————

## Pix2Pix

A conditional GAN for general purpose image-to-image translation trained on paired data, optimized under adversarial loss and reconstruction loss (e.g. L1). It is a typical baseline model.
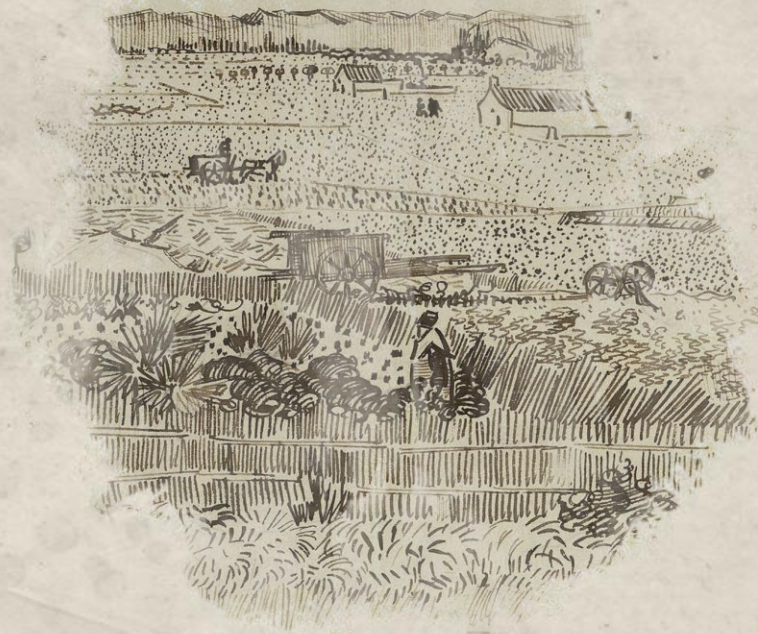
## Auto-Painter

An improved Pix2Pix specific for cartoon sketch-to-image translation, optimized also under perceptual loss and total variance loss.

## CycleGAN

A general purpose architecture for image-to-image translation made of 2 GANs. It learns a mapping between an input and output image and vice-versa in the absence of paired training examples.

# SketchyGAN

A modified **Label-Supervised Pix2Pix** that synthesizes colorful images from human drawn sketches from 50 categories. It is optimized under some specific losses to enhance the quality of the results.

It uses **Masked Residual Units** (MRUs) to repeatedly condition the the layers of both generator and discriminator on the input image, thus improving the information flow.

# ③ Implementation

## Dataset:

*Train set: 2840 paintings and synthetic sketches by 24 artists*
*Test set: 134 original drawings by Van Gogh and 47 by realists*

Boudin: 145
Caillebotte: 58
Cézanne: 110
Chase: 62
Corot: 219
Gauguin: 153
Gorbatov: 43
Guillaumin: 40
Hassam: 135
Korovin: 22
Loiseau: 158
Martin: 112

Maufra: 81
Metcalf: 83
Monet: 483
Pisarro: 123
Renoir: 39
Robinson: 22
Rose: 65
Shishkin: 110
Sisley: 168
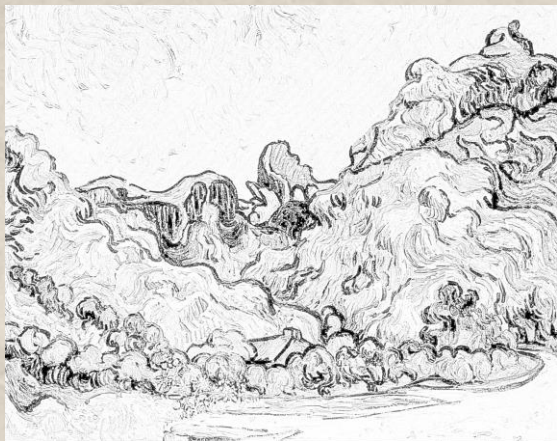Thaulow: 15
Twachtman: 52
Van Gogh: 334
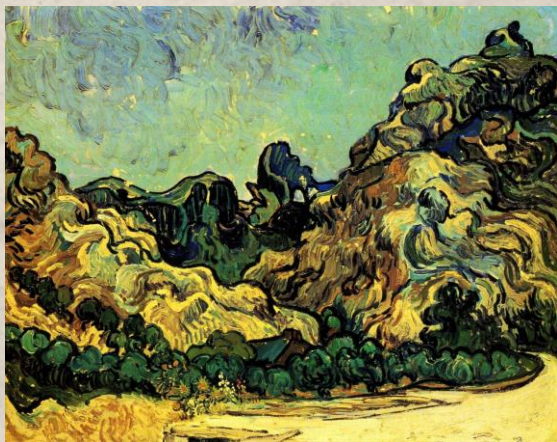
# Preprocessing

1. OpenCV used to convert paintings into (some sort of) grayscale pencil sketches.

2. Synthesized sketches retain edges and shades.

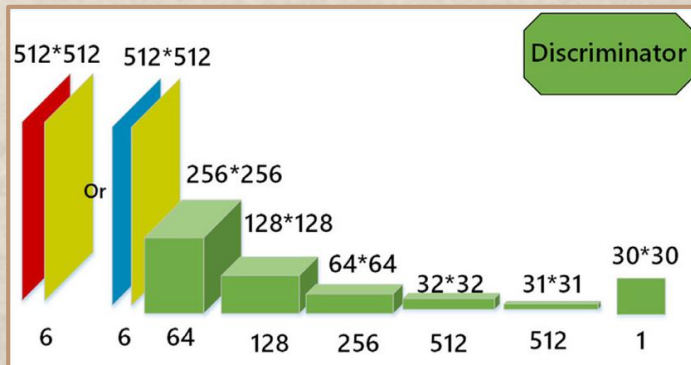3. Paintings and sketches resized to 256x256 and cropped at the center.

# Pix2Pix

## Discriminator

The 70 × 70 discriminator architecture is C64-C128-C256-C512.
It is a PatchGAN classifying 70×70 patches of the input image as real or fake, producing a probabilistic map.
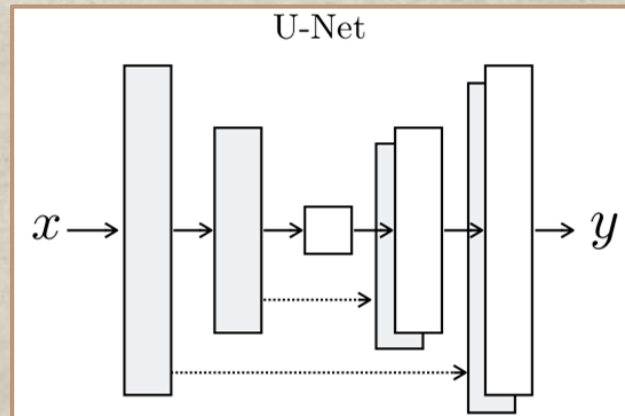


## Generator

Encoder:
C64-C128-C256-C512-C512-C512-C512-C512
Decoder: CD512-CD1024-CD1024-C1024-C1024-C512-C256-C128

A U-net encoder decoder with skip connections leveraging the low-level information between input and target.

# Pix2Pix

—————————————— • ——————————————

$$G^* = \arg\min_G \max_D L_{cGAN}(G, D) + \lambda L_{L1}(G) + L_{percep}(G)$$

**Adversarial loss**   $L_{cGAN}(G, D) = E_{x,y}[log(D(x, y))] + E_{x,z}[log(1 - D(x, G(x, z)))]$

**Reconstruction loss**   $L_{L1}(G) = E_{x,y,z}[||y - G(x, z)||_1]$

**Perceptual loss**   $L_{percep}(G) = E_{x,y,z}[||\phi(y) - \phi(G(x, z))||_1]$

## Training

➔ Train on Van Gogh's paintings but it is insufficient
➔ Then, repaint drawings with general impressionist style
➔ Train on 2511 paintings and synthetic sketches by 22 artists
➔ 20 epochs under Adam Optimizer

Van Gogh Drawings

Van Gogh Drawings

Van Gogh Drawings

Constable

Matisse

Pissarro

Rousseau

Sargent

Seurat

**Pro:** colors slightly faded, as expected by impressionist paintings.

**Con**: trained on heterogeneous dataset. Pix2Pix general style does not resemble any artist.
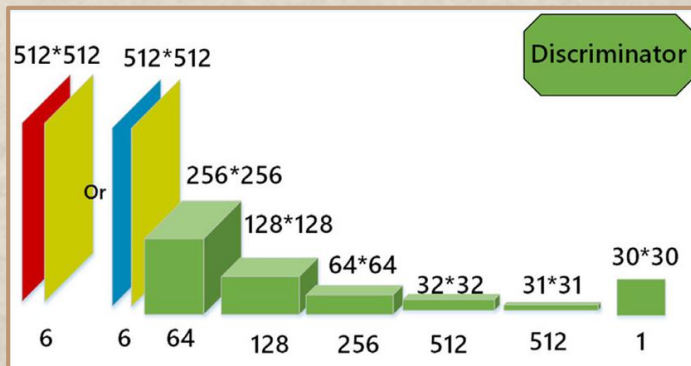
**Con**: outputs depend a lot on the quality of sketches. Worse results on original Van Gogh's drawings as many are deteriorated.

# StylePix2Pix

## Discriminator

70 × 70 PatchGAN as in Pix2Pix with the integration of Auxiliary Classifier (so called in ACGAN from which we have just taken the idea) for predicting the class of an artwork.



## Generator

U-net generator as in Pix2Pix taking also the class label, which is transformed into a channel by embedding and fully connected layers, then concatenated with the input image.
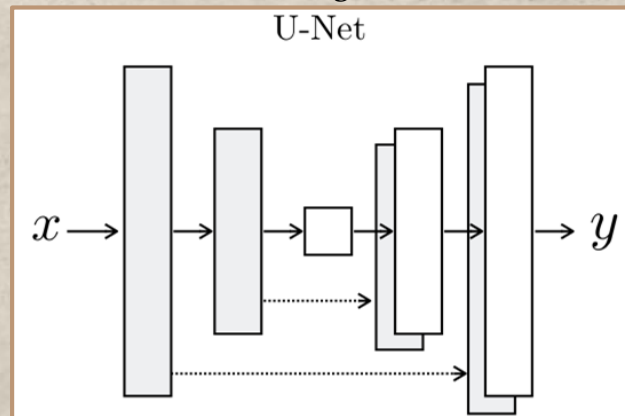
# StylePix2Pix

───────────── • ─────────────

$$G^* = \arg\min_G \max_D L_{cGAN}(G, D) + \lambda L_{L1}(G) + L_{ac}(D)$$

**Adversarial loss**
$$L_{cGAN}(G, D) = E_{x,y}[log(D(x, y))] + E_{x,z}[log(1 - D(x, G(x, z)))]$$

**Reconstruction loss**
$$L_{L1}(G) = E_{x,y,z}[||y - G(x, z)||_1]$$

**Auxiliary Classification loss**
$$L_{ac}(D) = E[log(P(C = c|y))]$$

## Training

➜ Repaint drawings with a given style
➜ Train on 1156 paintings and synthetic sketches by 3 classes
   ◆ Impressionism: 0
   ◆ Post-Impressionism: 1
   ◆ Realism: 2
➜ 50 epochs under Adam Optimizer

Van Gogh Drawings

Van Gogh Drawings

Van Gogh Drawings

Rousseau - Impressionism

Rousseau - Post-Impressionism

Rousseau - Realism

Seurat - Impressionism

Seurat - Post-Impressionism

Seurat - Realism

**Pro**: StylePix2Pix applies a different style for each class, partly solving the generic style problem of Pix2Pix.

**Con**: unable to generate paintings exactly with the true style. Colors a little bit faded like in Pix2Pix as different classes tend to influence each other.

**Con**: a painter with few artworks is not suitable for a class. Subsets should be sufficiently large and balanced.
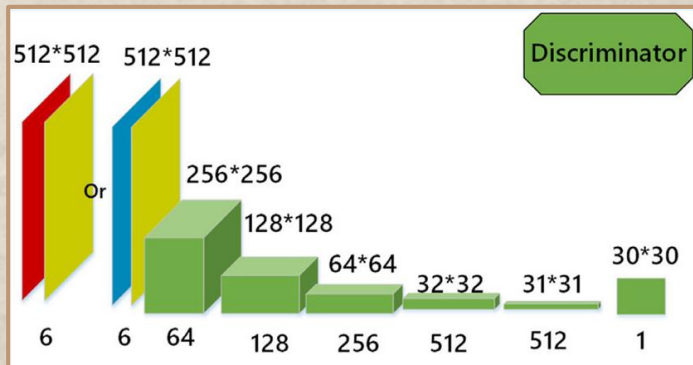
# CycleGAN

## Discriminator
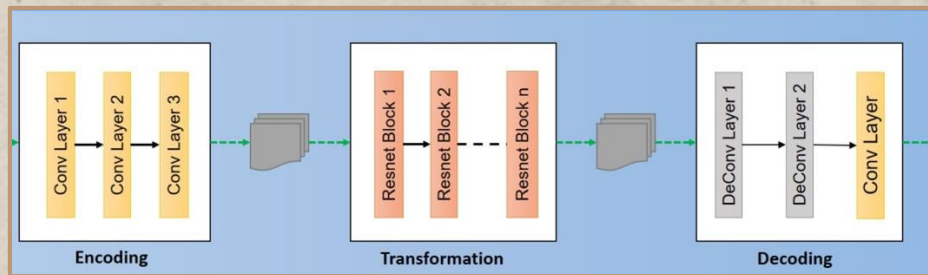
70 × 70 PatchGAN as in Pix2Pix.

This model classifies 70×70 patches of the input image as real or fake, producing a probabilistic map.



## Generator

A modified ResNet generator with encoder, 9 ResNet blocks with additive skip connections and decoder.
c7s1-64, d128, d256, R256, R256, R256, R256, R256, R256, R256, R256, R256, u128, u64, c7s1-3

# CycleGAN

$$G^*, F^* = \arg \min_{G,F} \max_{D_X, D_Y} L(G, F, D_X, D_Y)$$

**Objective function**
$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, Y, X) + \lambda L_{cyc}(G, F) + 0.5\lambda L_{identity}(G, F)$$

**Adversarial loss**
$$L_{GAN}(G, D, X, Y) = E_x \left[ (D(G(x)) - 1)^2 \right] + E_y \left[ (D(y) - 1)^2 \right] + E_x \left[ D(G(x))^2 \right]$$

**Cycle consistency loss**
$$L_{cyc}(G, F) = E_x[\| F(G(x)) - x \|_1] + E_y[\| G(F(y)) - y \|_1]$$

**Identity loss**
$$L_{identity}(G, F) = E_y[\| G(y) - y \|_1] + E_x[\| F(x) - x \|_1]$$

## Training

➤ Repaint existing paintings with Van Gogh's style
➤ Train on 344 Van Gogh's paintings and 2511 synthetic sketches by 22 artists
➤ Discriminators updated sampling from a pool of 50 previously generated images to reduce the model oscillation
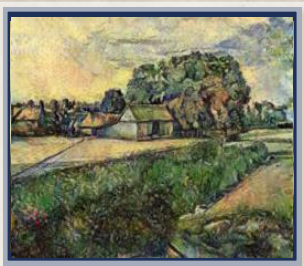➤ 120 epochs under Adam Optimizer

Van Gogh Drawings

Van Gogh Drawings

Van Gogh Drawings

Constable

Matisse

Pissarro

Rousseau

Sargent

Seurat

**Pro**: brighter colors and a bit vaguely closer to Van Gogh's style. Brushstrokes marked across the painting.

**Pro**: no need for paired training data. Results less affected by how much similar train and test sketches are.

**Con**: more expensive and unstable training.

# Results and Conclusions

- Painting Van Gogh's drawings as he would have done is hard as his sketches are deteriorated and visually different from the synthetic ones extracted from training paintings.
- A model, e.g. a GAN, could be developed to rebuild original human drawings similarly to training synthetic sketches.



Van Gogh - Pix2Pix

Van Gogh - StylePix2Pix

Van Gogh - CycleGAN

Van Gogh - Pix2Pix

Van Gogh - StylePix2Pix

Van Gogh - CycleGAN

Van Gogh - Pix2Pix

Van Gogh - StylePix2Pix

Van Gogh - CycleGAN

# Future works

- **StyleCycleGAN**: a Label-Supervised CycleGAN can potentially merge the advantage of training on unpaired data and the ability of conditioning the result on a class label.

- **SketchyGAN**: rebuild the discriminator and generator with MRU blocks for injecting the input image at multiple scales, thus improving the information flow.

Thank you!