

Nvidia's GauGAN AI turns rough sketches into photorealistic images in real-time

Semantic Image Synthesis with Spatially-Adaptive Normalization

R.A.J. Wacanno

03-04-2019

Informatica in het Nieuws
Academische Vaardigheden Informatica 2 - Groep A

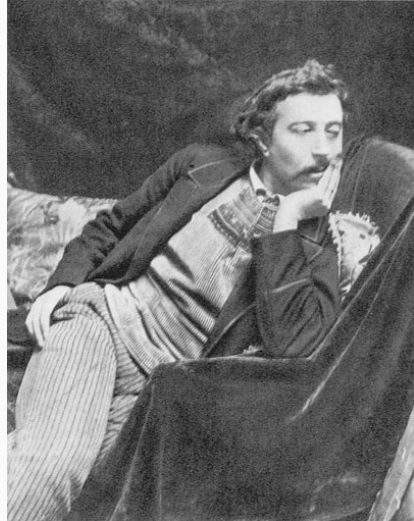
Table of contents

1. Introduction
2. The Article
3. The Paper
4. Article vs. Paper
5. Explanation of SPADE
6. Impact of these Developments
7. Discussion
8. Questions?

Introduction

Introduction

- Published March 19th 2019
- Techspot
- Nvidia, UC Berkely and MIT
- GauGAN
- Generating photo-realistic images
- SPADE



The Article

Nvidia's GauGAN AI turns rough sketches into photorealistic images in real-time

Anyone can be an artist with this MS Paint for the AI era

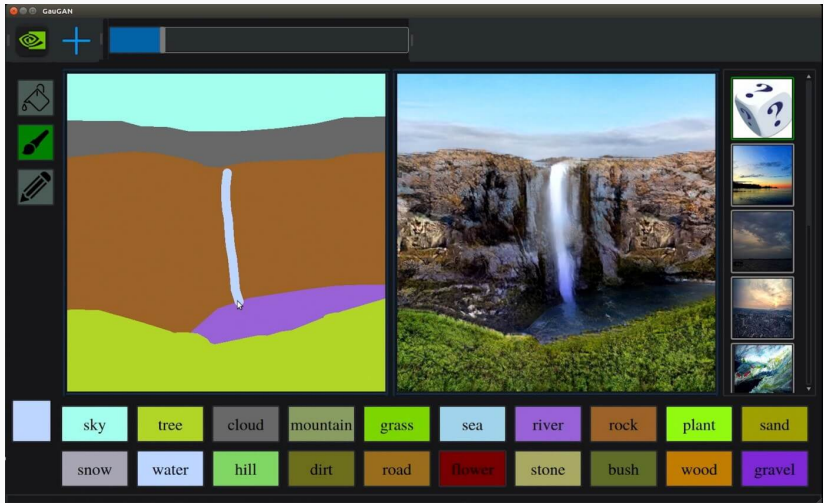
By Cal Jeffrey on March 19, 2019, 12:41 PM



A novice painter might set brush to canvas aiming to create a stunning sunset landscape — craggy, snow-covered peaks reflected in a glassy lake — only to end up with something that looks more like a multi-colored inkblot.” Nvidia has created a tool that can turn the simplest drawings into photo-realistic images.

- Introducing GauGAN
- Brief technical details
- Describing usage

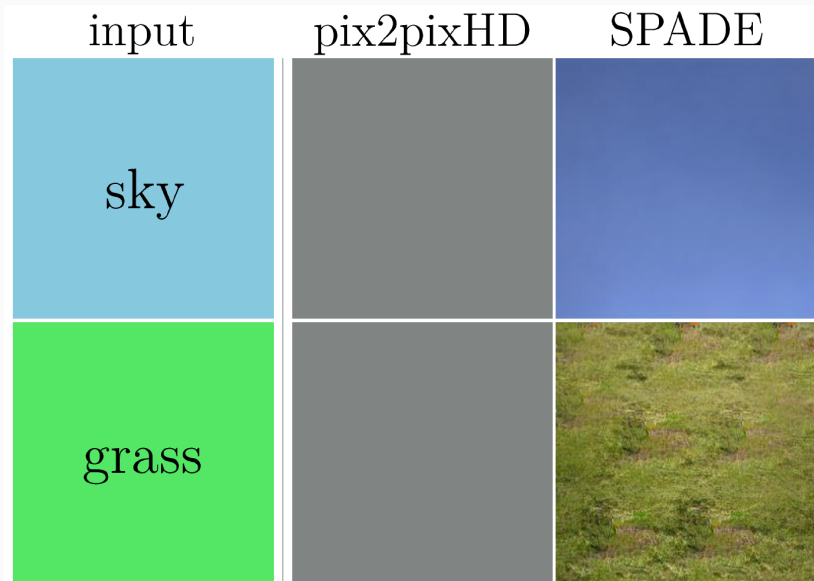
The Interface



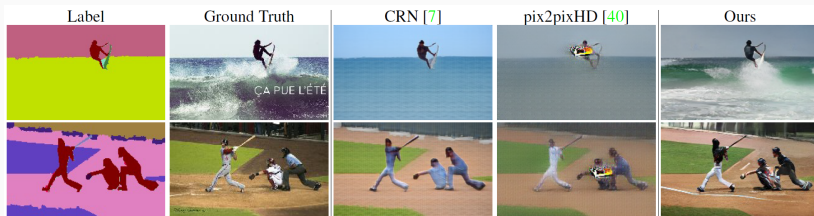
The Paper

- Introduce problem/solution
- Outline earlier work and used research
 - Earlier methods (eg. Pix2PixHD)
- Describing the SPADE algorithm
- Used training data sets
- Testing and evaluating

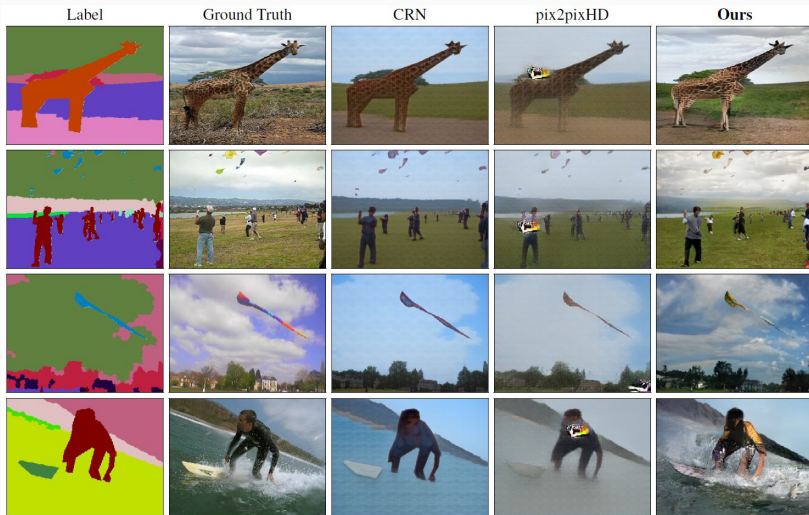
Comparison



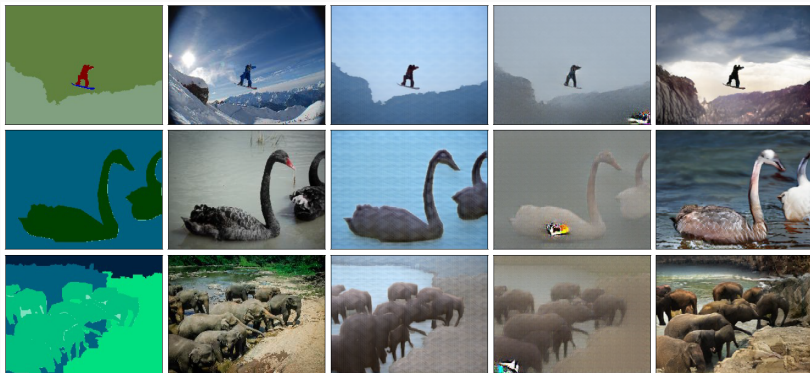
Testing the Network



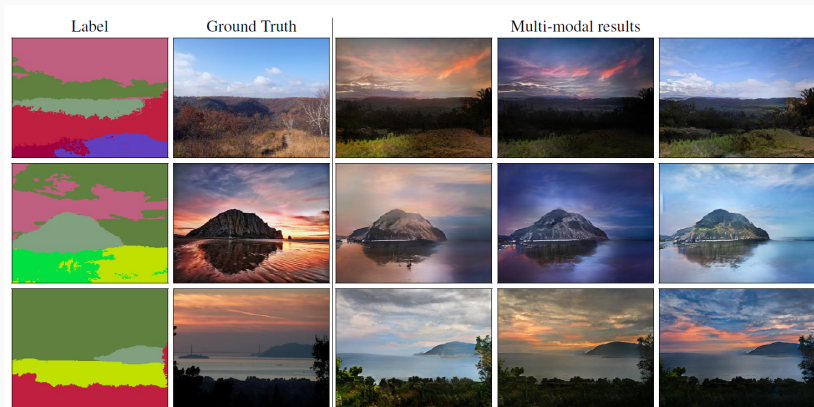
Testing the Network



Testing the Network



Multiple Styles



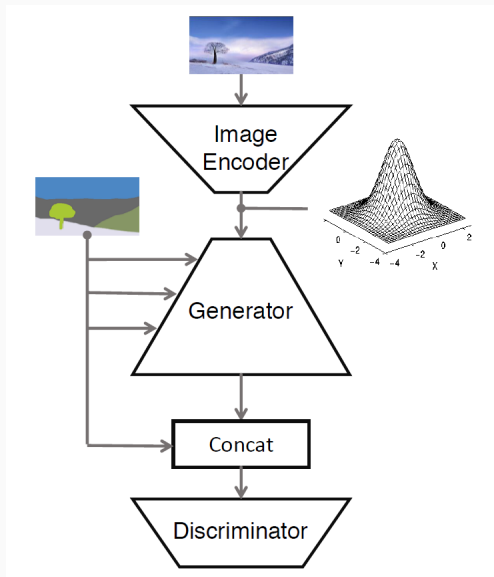
Article vs. Paper

Article vs. Paper

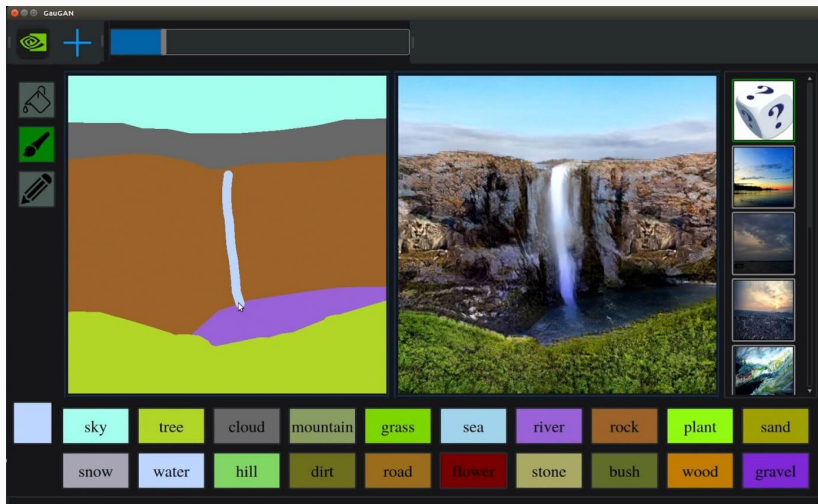
- Article:
 - All important details
 - Skims over uninteresting details
 - No exaggeration
- Paper:
 - More in depth
- Usage vs. implementation

Explanation of SPADE

Generative Adversary Networks



Semantic Segmentation Mask



Semantic Segmentation Mask

$$\mathbf{m} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}$$

$$x_{nm} \in \mathbb{L}$$

for example:

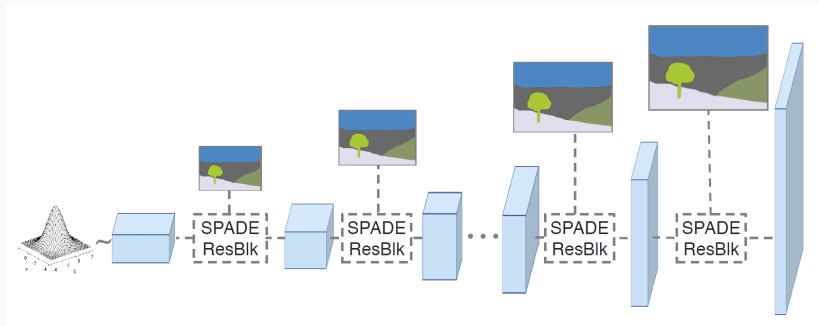
$$\mathbb{L} = \{0 \text{ (sky)}, 1 \text{ (grass)}, \dots\}$$

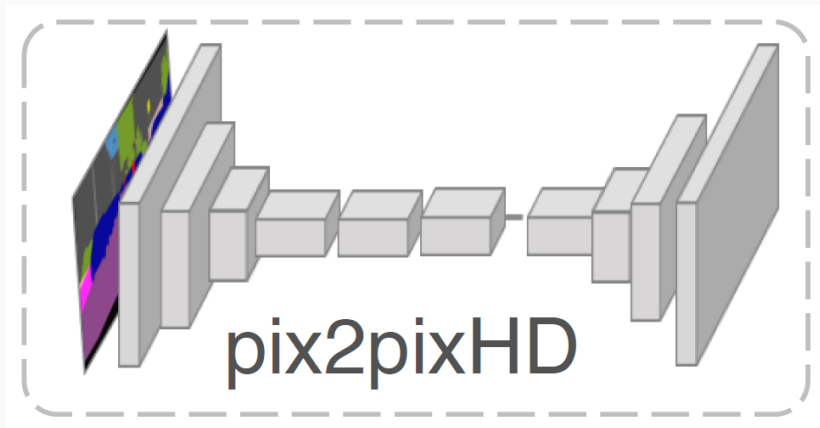
$$\gamma^i \frac{h^i - \mu^i}{\sigma^i} + \beta^i$$

With SPADE:

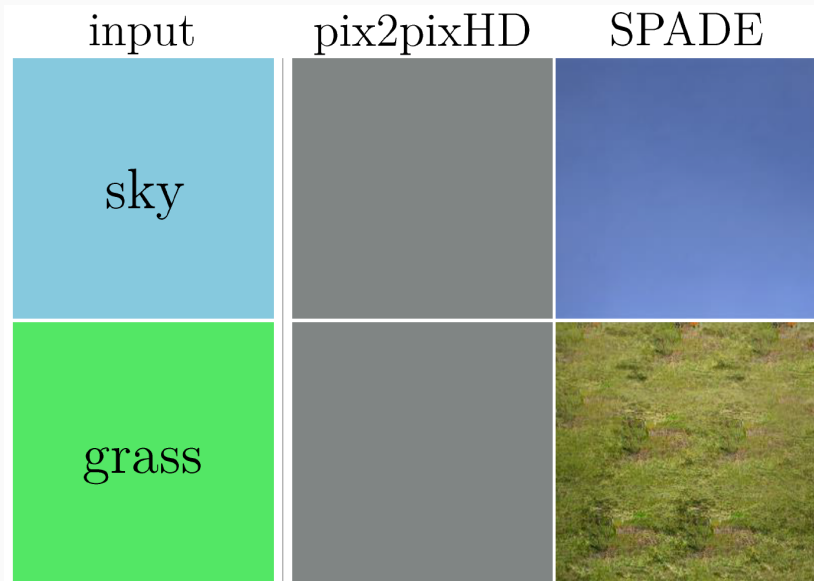
$$\gamma_{xy}^i(m) \frac{h^i - \mu^i}{\sigma^i} + \beta_{xy}^i(m)$$

Generator

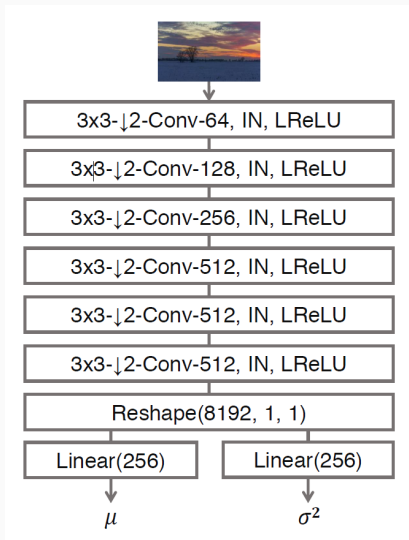




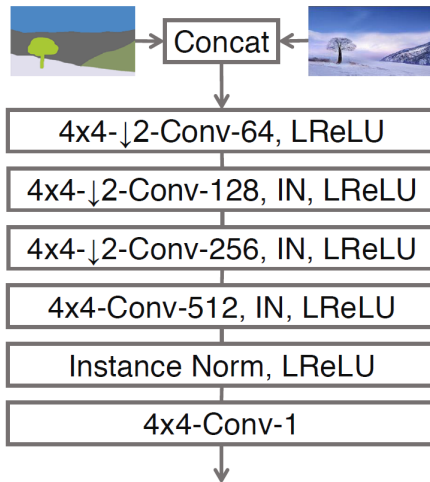
Comparison after Normalisation



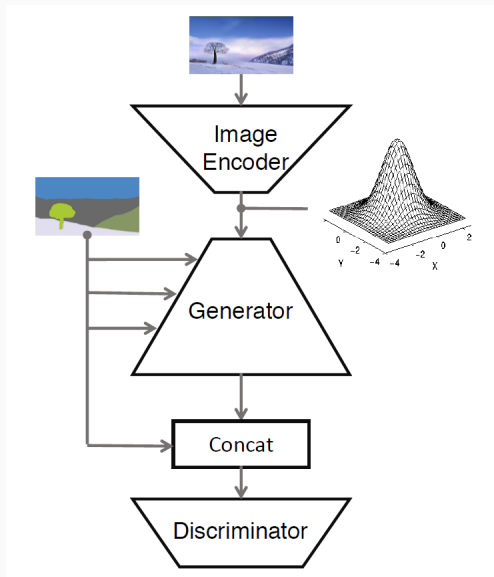
Encoder



Discriminator



Full Algorithm



Impact of these Developments

- Use cases:
 - Architecture
 - Urban planning
 - Landscape design
 - Game development
 - Data sets for other NN applications
- Further research *spatial adaptive* GAN's

Discussion

Someone using tools like these is not a real artist.

Questions?
