Can Data Scientists make Comic Books?

An Implementation of Diffusion Models using LoRA and Dream Booth

Arianna Michelangelo | Maëlys Boudier | Natalia Beltrán

Project Goal:

Given limited data (approx. 5 images) as input, train a model to generate variations of a cartoon character based on text prompts to aid comic book illustrators.

Use VAEs and Diffusion Models focusing on using LoRA and/or Dream Booth.

Project Importance:

We aim to provide a tool for artists in the industry to save time and resources in first mock-up designs

From a environmental perspective the training of neural networks is becoming a major concern. The expected results from our project is to achieve a certain level of visual quality with a lower computational cost.

Main Use Cases

1. Generate Images for Mockups

Save time on making mockups by generating the character in different positions based on the artist's initial drawings

Early Stages

Late Stages

2. Generate Images to aid Artists Taking Over

Use existing comic books to create a dataset and aid new artists joining the process to pick up on the style quicker

Comics Research in Data Science

Comic Art Generation using GANs

Uses GANs (generative adversarial networks) to generate images for comic book strips. Our thesis is a good extension of his work given new tech advancements (Click Link Here, 2022)

Survey of Comics Research in Computer Science

Enumerates available research on Content Generation including colorization, character generation, and media conversion. We will focus on character generation (Click Link Here, 2018)

Automatic Anime Creation with GANs

Explore the training of GAN models specialized on an anime facial image dataset specifically implementing DRAGAN, but is limited to face generation for anime (Click Link Here, 2017)

Neural Style Transfer on Photos

Generating anime images / cartoons by applying a neural style transfer on photos: AnimeGAN, but these images still look very far from proper cartoons for comic or anime strips (Click Link Here, 2020)

Comics Research in Data Science

Photo   CartoonGAN   ComixGAN   AnimeGAN
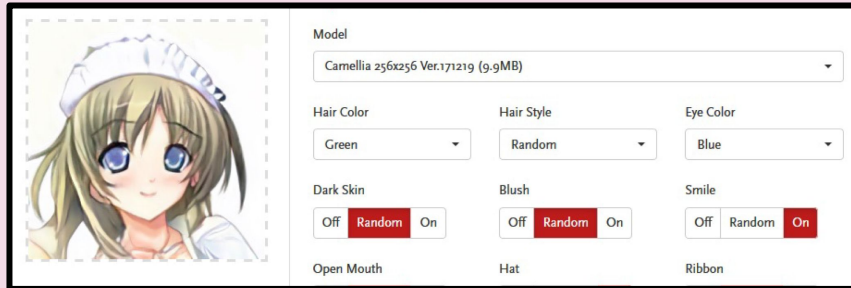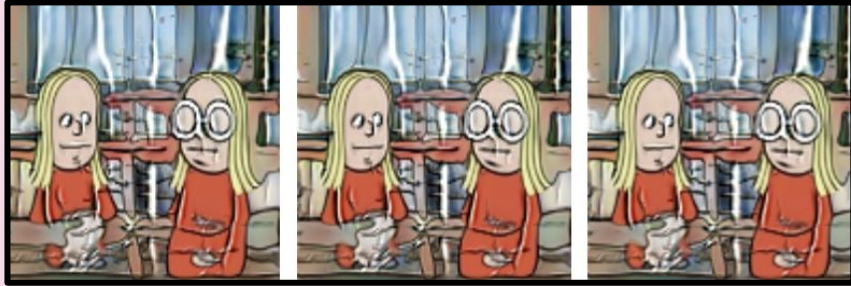
Data Scarcity

Tackle data scarcity by implementing models with 5-10 different images of a specific character

Property of Jordane Meignaud

[imagine digital, colored cartoon images]

Available in May

Online Data

Pros: easy to find and download
Cons: copyright & concerns if base model trained on them
Use: to run initial models before we obtain original data

Original Sketch Data

Pros: faster to obtain than digital
Cons: lower quality of image
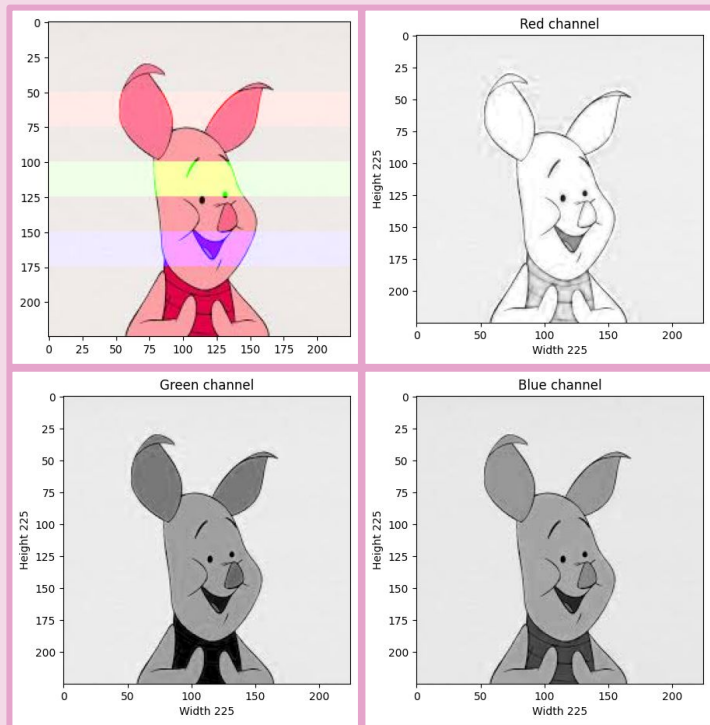**Use**: original sketches from artists we will be working with but will not be used for modeling

Original Digital Data

Pros: closer to real-life conditions
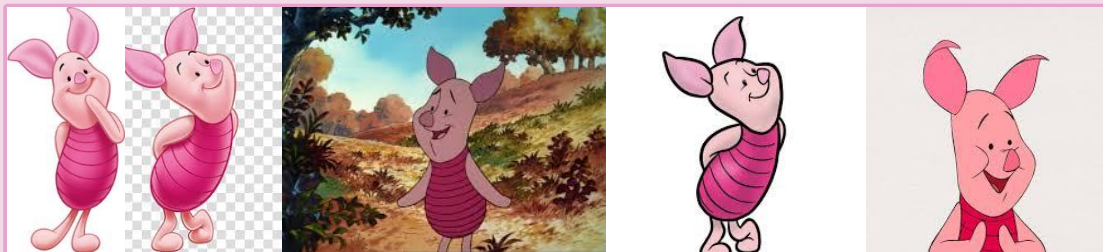**Cons**: hardest data to obtain
Use: we have 1 - 2 artists willing to provide a character for our thesis, these will be ready by May

Data Overview: Example of Piglet Dataset

General Statistics:

	width	height	aspect ratio	file size (bytes)	mean red	mean green	mean blue
min	144	195	0.48	5106	127	106	87
mean	210	239	0.92	7393	214	136	189
max	258	300	1.32	12203	240	221	226

All images are property of their respective owners

Models

1. Diffusion Models

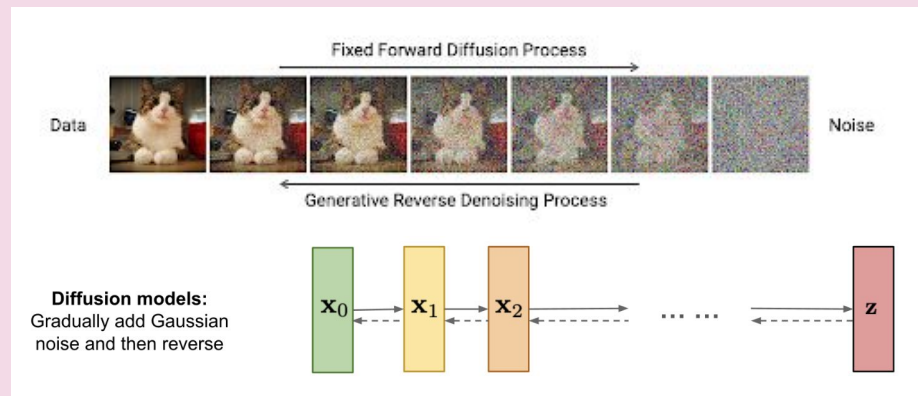2. LoRA

3. Dream Booth

Diffusion Models

Applications:

- **Image and Audio Generation**
- Control and reinforcement learning
- Life-science applications
- Black box optimizations

Process:

Across Models:

Models:	Quality	Diversity	Speed
VAEs	✗	✓	✓
GANs	✓	✗	✓
Diffusions	✓	✓	✗

Fixed Forward Diffusion Process

Data — Noise

Generative Reverse Denoising Process

Diffusion models: Gradually add Gaussian noise and then reverse
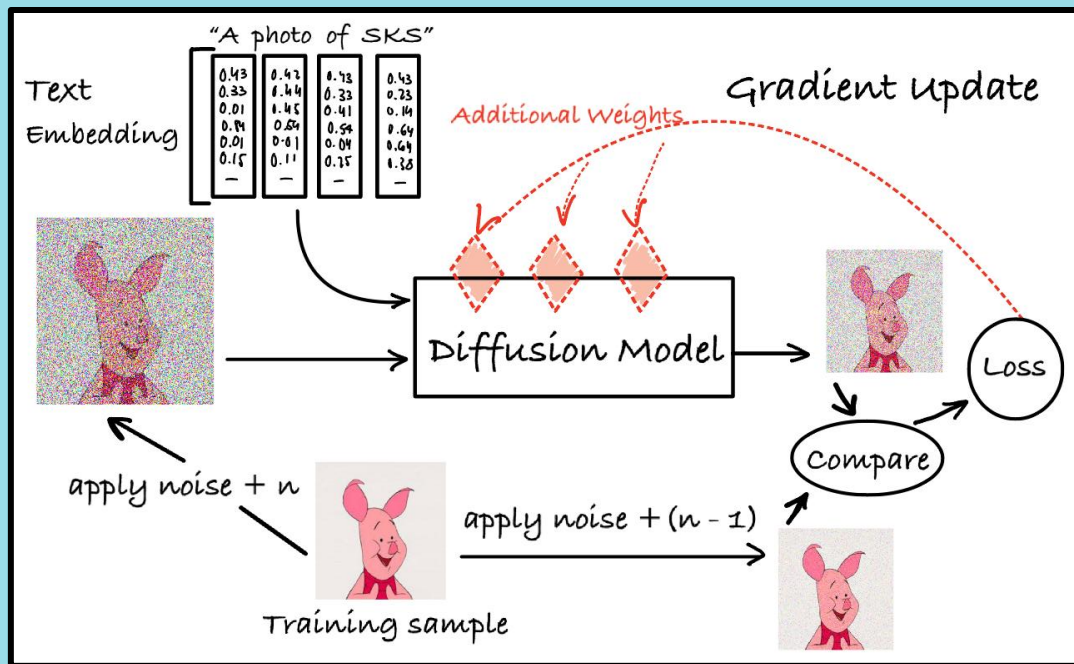
x_0 — x_1 — x_2 — ... ... — z

How much noise is added and the quality of the final picture will depend on the numbers of steps used.

LoRA (Low Rank Adaptation)

Objective: Teach the model a new concept without creating a whole new model. Add a number of layers to the diffusion model and then train the model. The weights (of the new layers) are modified as the model understands the concept.

Dream Booth

Objective: Teach the model to associate a unique identifier to a specific concept. Fine-tunes the diffusion model itself until it understands the new concepts.
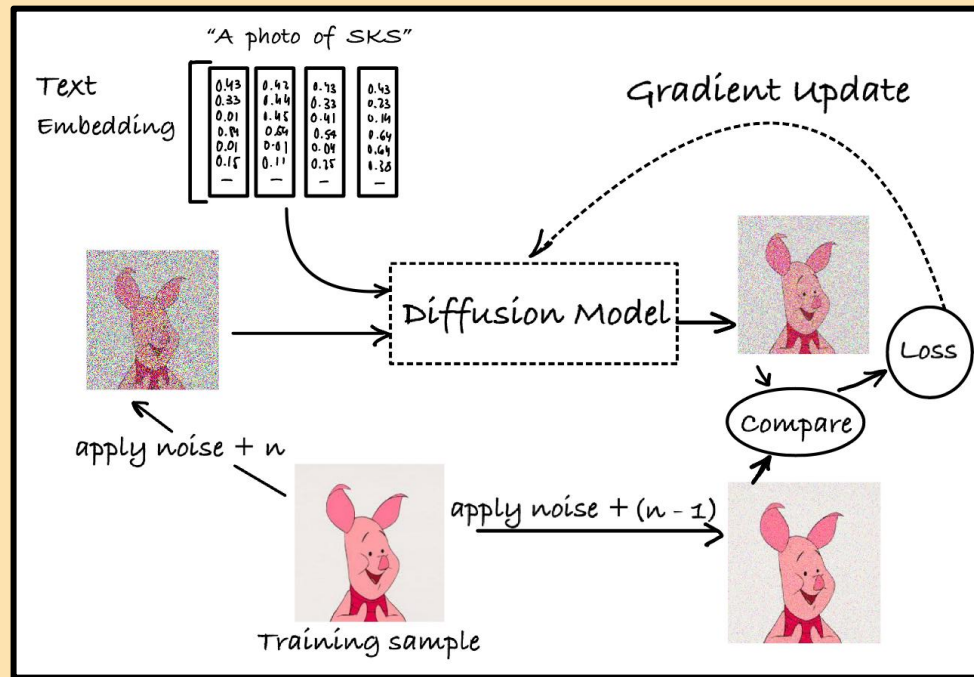
Methodology: Model Tuning

Goal: *demonstrate the usability of LoRA and/or Dream Booth to generate cartoon characters and how some parameters may affect results*

Compare Across Models

# of Steps

Learning Rate

Dataset Size

Implement 1 model successfully, and if time & resources permit, implement another to compare

Balance underfitting and computational resources

Determine optimal learning rate for convergence

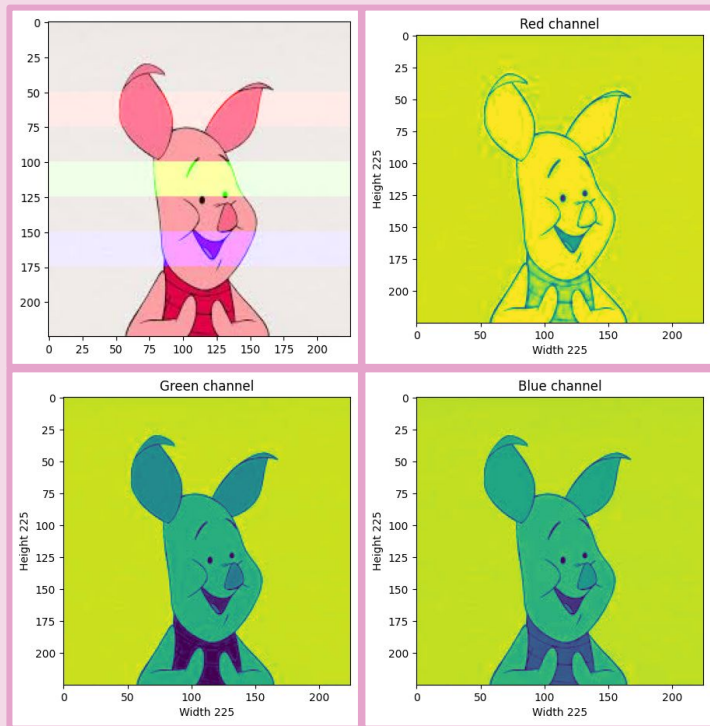Determine whether models can converge with less data and how this affects learning time

The End.

Thank you for listening.

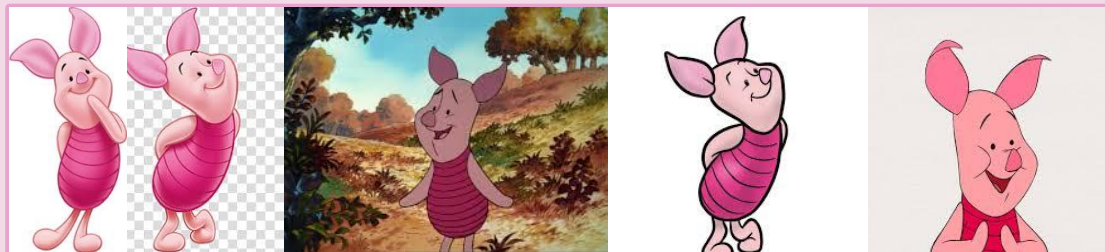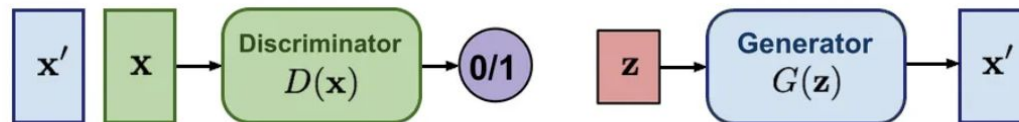Any Questions?

Data Overview: Example of Piglet Dataset

General Statistics:

	width	height	aspect ratio	file size (bytes)	mean red	mean green	mean blue
min	144	195	0.48	5106	127	106	87
mean	210	239	0.92	7393	214	136	189
max	258	300	1.32	12203	240	221	226

All images are property of their respective owners

GAN: Adversarial training

$\mathbf{x}'$ | $\mathbf{x}$ → Discriminator $D(\mathbf{x})$ → 0/1 $\mathbf{z}$ → Generator $G(\mathbf{z})$ → $\mathbf{x}'$

VAE: maximize variational lower bound

$\mathbf{x}$ → Encoder $q_\phi(\mathbf{z}|\mathbf{x})$ → $\mathbf{z}$ → Decoder $p_\theta(\mathbf{x}|\mathbf{z})$ → $\mathbf{x}'$

Diffusion models: Gradually add Gaussian noise and then reverse

$\mathbf{x}_0$ → $\mathbf{x}_1$ → $\mathbf{x}_2$ → ... ... → $\mathbf{z}$
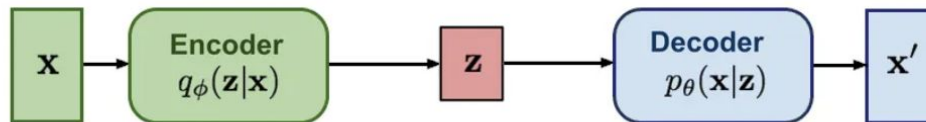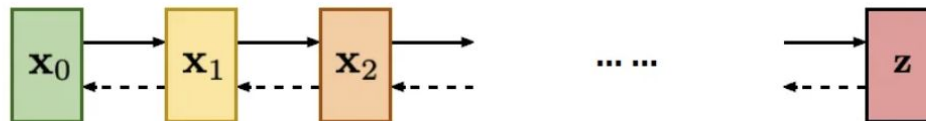
Challenge for Comic Books
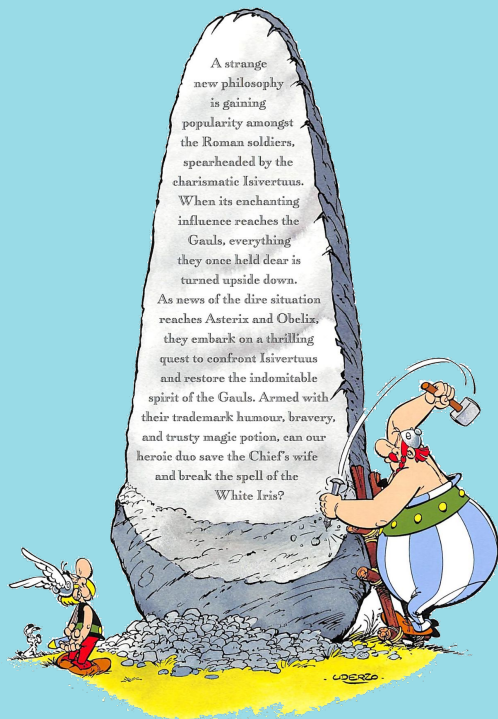
Asterix and Obelix was illustrated by Uderzo until 2009. Frédéric Mébarki took over, but **quit because of the pressure to follow in Uderzo's footsteps**.

Ultimately, Didier Conrad took over the drawing process (releasing numbers 35 - 39).

Dataset for Existing Franco-Belgian Style Comic Books:
46 pages x 8-12 panels/page x # volumes

But, there are also challenges at **the start of new projects** in creating mock-ups and exploring new possibilities

Literature Review

Comic Art Generation using GANs

Student Master's Thesis (Marnix Verduyn) from Germany using GANs written in 2021-2022. Our thesis feels like a good extension of his work with new advancements ([Click Link Here](#))

Step-by-Step LoRA Implementation

Guide on Data Camp which covers LoRA and Dream Booth Fine Tuning with examples, our goal is to start by replicating this with for a cartoon character ([Click Link Here](#))

Stable Diffusion using LoRA code

An alternative code on GitHub which also covers LoRA and Dream Booth Fine Tuning, well documented and can also be used as a baseline to implement ([Click Link Here](#))

Dream Booth Example on Keras

Documentation for Dream Booth on Keras Website ([Click Link Here](#))