ETH zürich

IKG
Institute of Cartography and Geoinformation

Stable Diffusion in Cartography – Opportunities and Challenges

Raimund Schnürer, Sidi Wu, Lorenz Hurni

ICC Pre-Conference Workshop on Cartography and AI (MapAI)

12. August 2023

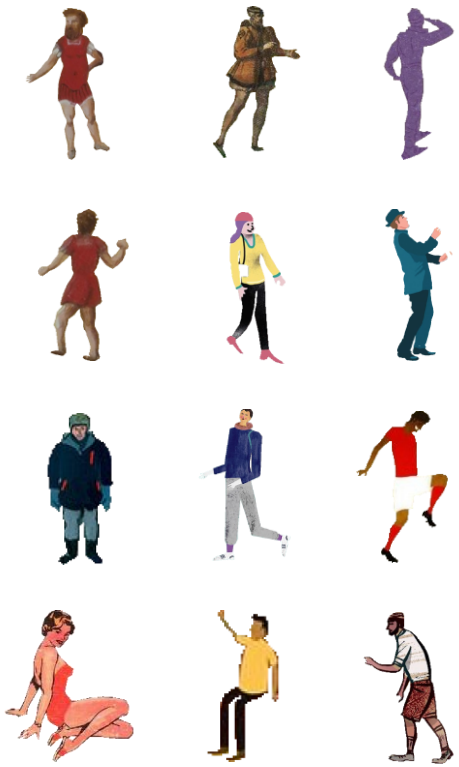
About the presenter

Raimund Schnürer

Doctoral student, Institute of Cartography and Geoinformation, ETH Zurich

Dissertation ”Storytelling with Animated Interactive Objects in Real-time 3D Maps”

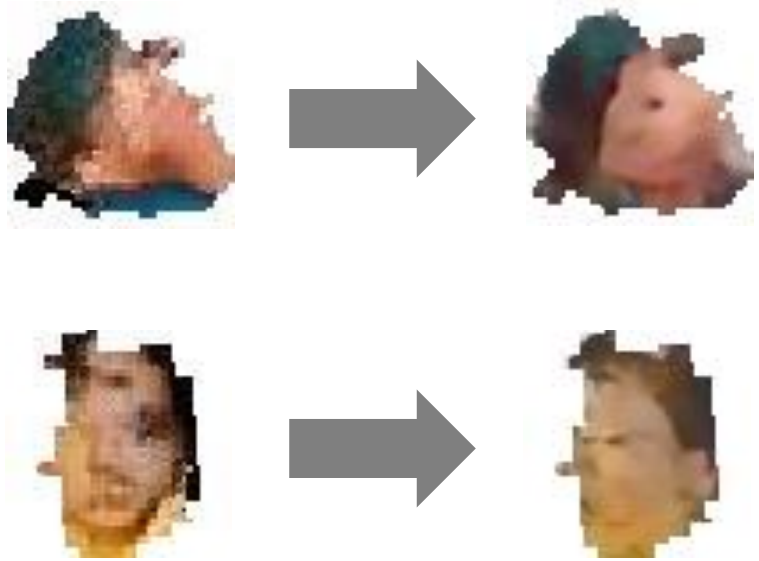
Recent work: Inferring Implicit 3D Representations from Human Figures on Pictorial Maps



http://narrat3d.ethz.ch/3d-humans-from-pictorial-maps/

Recent work: Inferring Implicit 3D Representations from Human Figures on Pictorial Maps

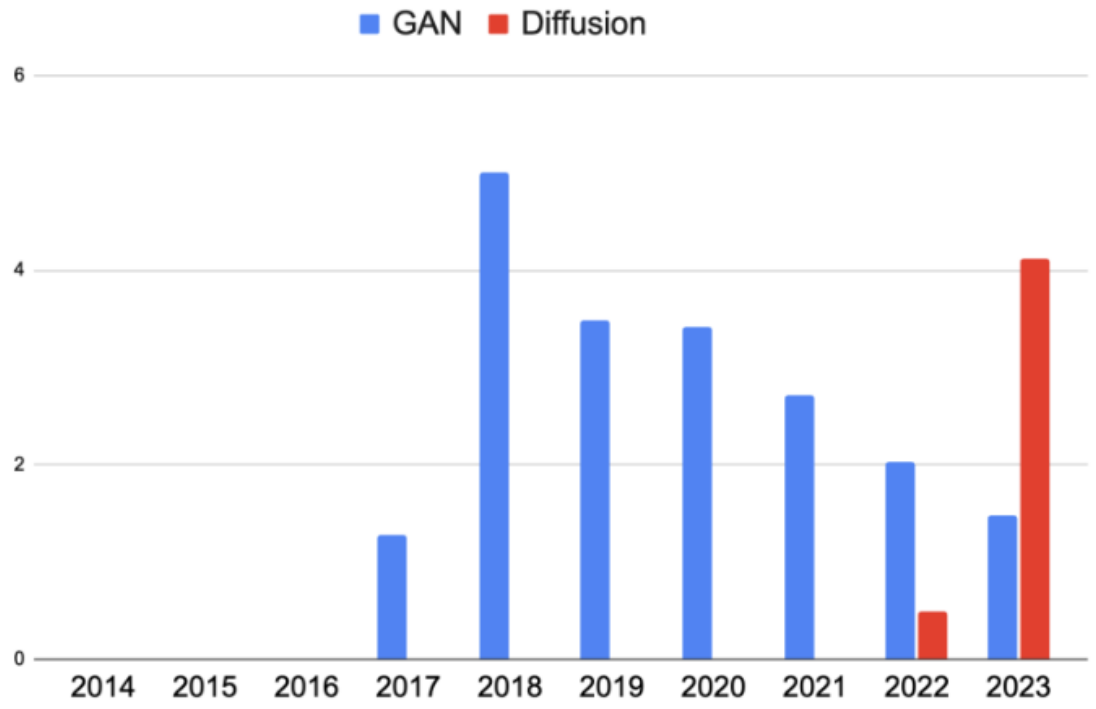
Face resynthesis



An example of a generative task

Diffusion: Motivation

Accepted Papers [%] at one of the most important conferences on computer vision (CVPR)



https://cvpr2023-tutorial-diffusion-models.github.io/

Diffusion: Motivation

Image generation

created in Midjourney with the prompt
“holy professor giving lecture to students in alpine mountains
and teaching them how to save the world”

Diffusion: Tools

https://labs.openai.com/

https://www.midjourney.com/

runway

https://app.runwayml.com/

DreamStudio

https://dreamstudio.ai/

Much more available → search for “AI image generators“

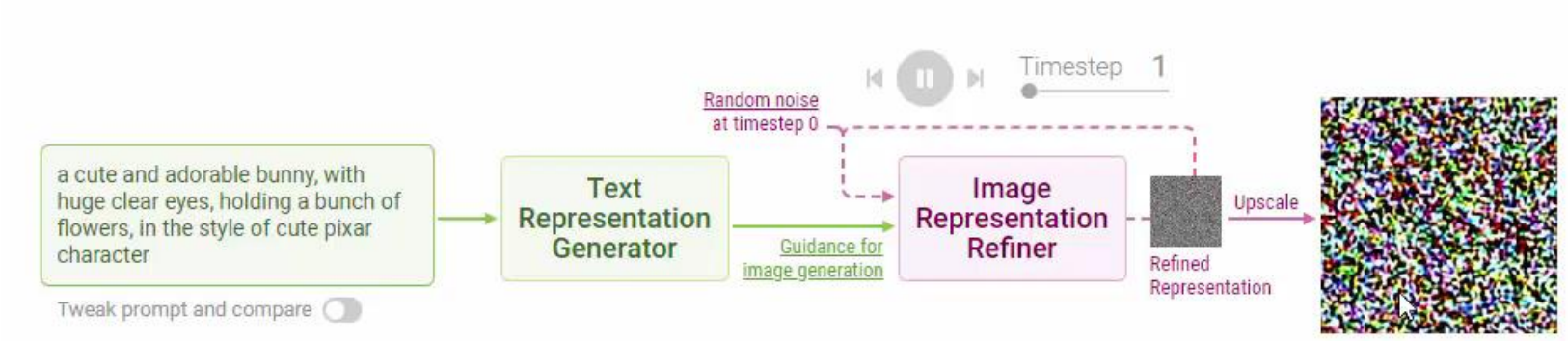
Diffusion: Tools – Open Source

Stable Diffusion web UI

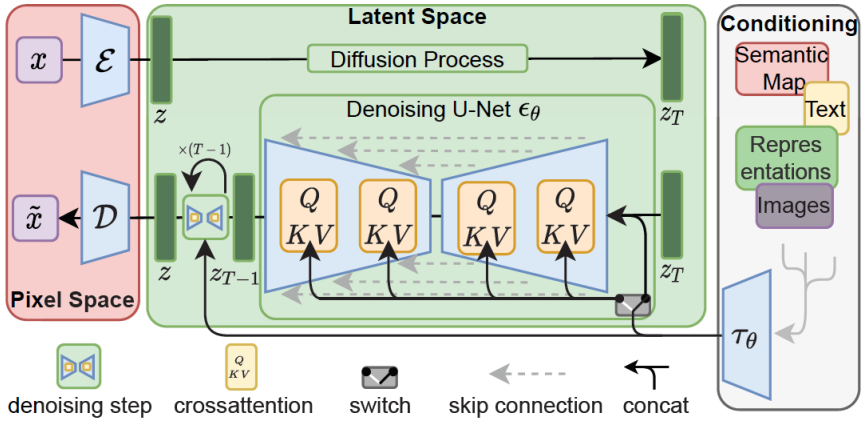
https://github.com/AUTOMATIC1111/stable-diffusion-webui

https://github.com/huggingface/diffusers

Stable Diffusion: Procedure



https://poloclub.github.io/diffusion-explainer/



Rombach et al. (2022) – High-Resolution Image Synthesis with Latent Diffusion Models

Stable Diffusion: Training data

Download

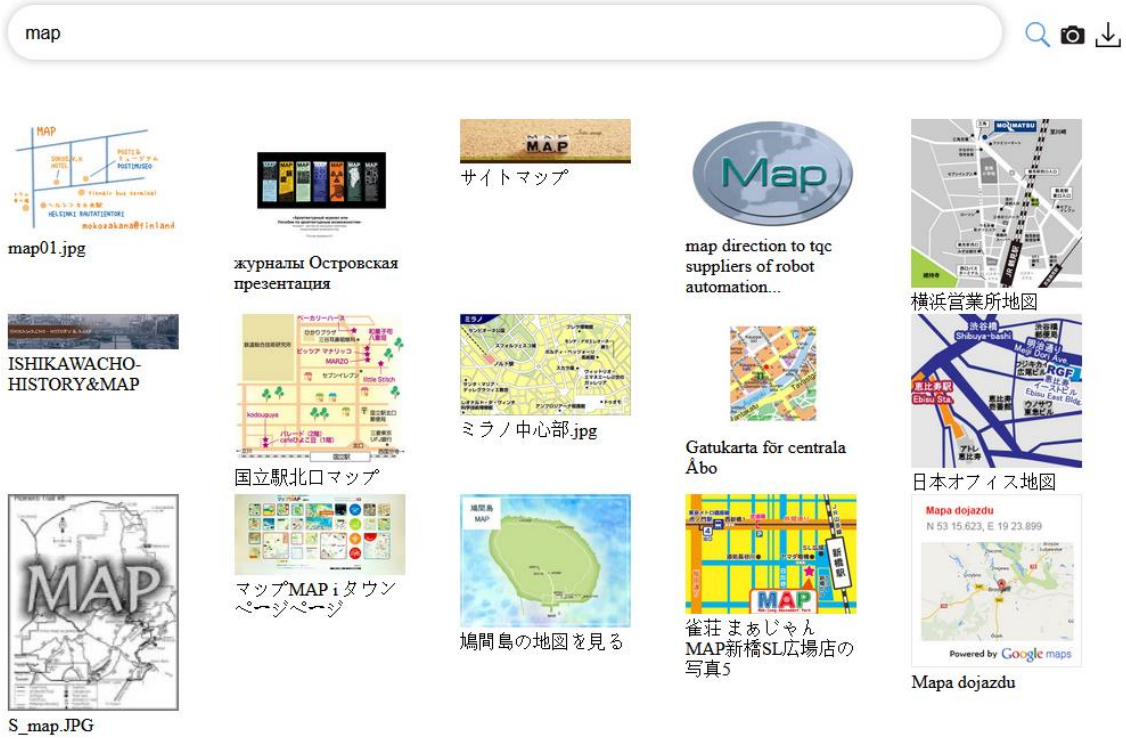
LAION-5B: A NEW ERA OF OPEN LARGE-SCALE MULTI-MODAL DATASETS

5.85 billion pairs of image URLs and the corresponding metadata (800GB)

- laion2B-en 2.32 billion of these contain texts in the English language
- laion2B-multi 2.26 billion contain texts from 100+ other languages
- laion1B-nolang 1.27 billion have texts where a particular language couldn't be clearly detected.

https://laion.ai/blog/laion-5b/

Explore



https://rom1504.github.io/clip-retrieval/

Time for some experiments

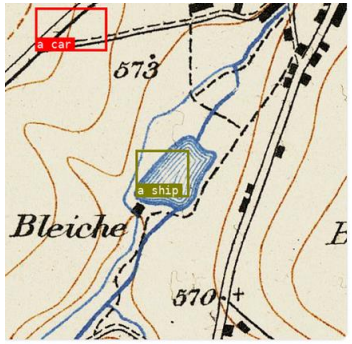
Preliminary tests: Pictorial maps and objects

Generation

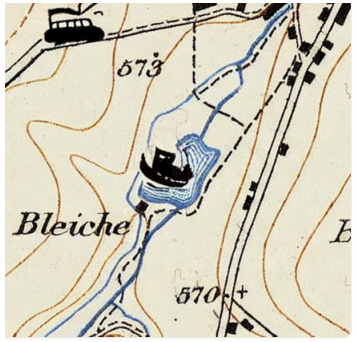
„a pictorial map
with forest and lakes“

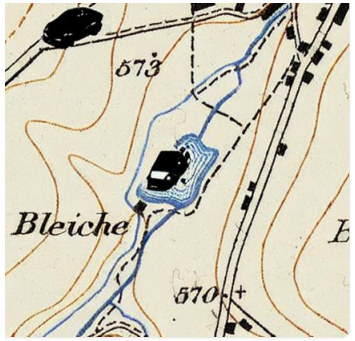
txt2img @ Stable Diffusion web UI

Inpainting



„a car and
a ship“





https://huggingface.co/spaces/gligen/demo

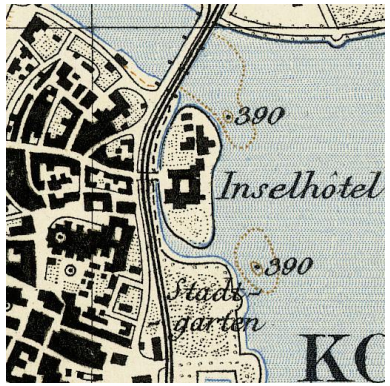
Abstraction

„comic style,
icon“

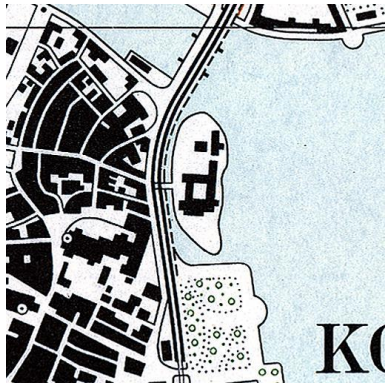
img2img @ Stable Diffusion web UI

Experiment I: Text-based map generation

Tiles of historic topographic maps of Switzerland



Siegfried map (*siegfr*)
around 1880



Old national map (*oldnat*)
around 1950

Framework: Stable Diffusion web UI

Method: Textual inversion

Pre-trained model: v1-5-pruned

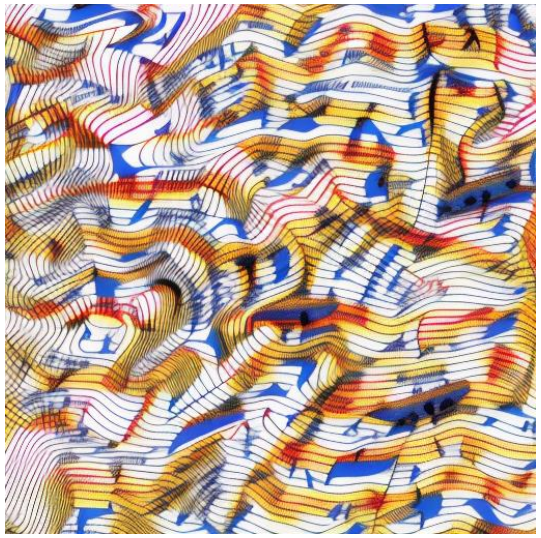
Prompt template: [*name*] map

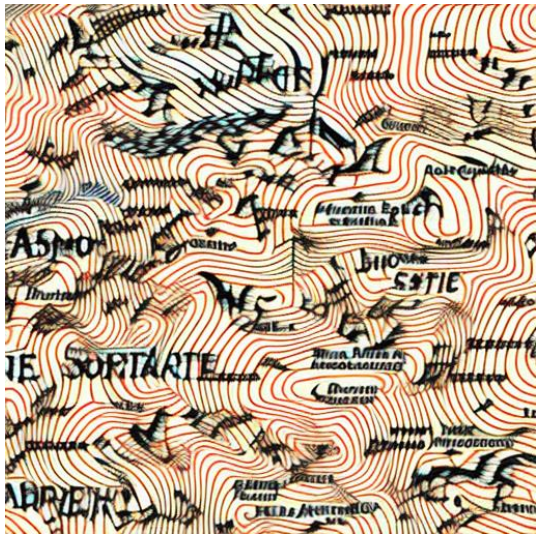
Hardware: NVIDIA RTX 3070 Ti (8GB)

https://arxiv.org/abs/2208.01618 https://youtu.be/2ityl_dNRNw

Experiment I: Text-based map generation

Intermediate results while training with 12 map tiles for 250 epochs



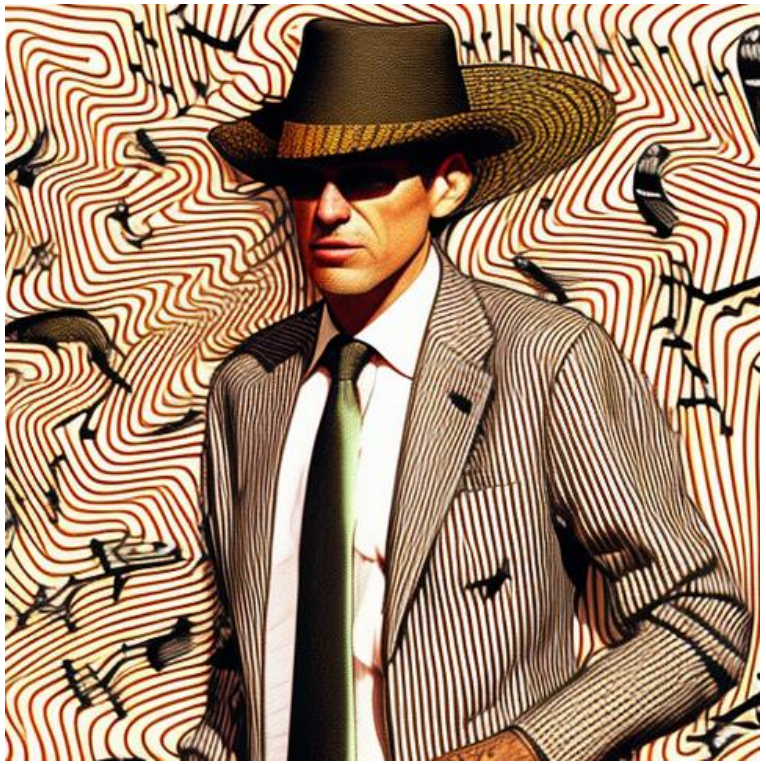



”*siegfr* map“

Experiment I: Text-based map generation

Results after training with 12 map tiles for 250 epochs

”a woman in the style of a *siegfr* map“

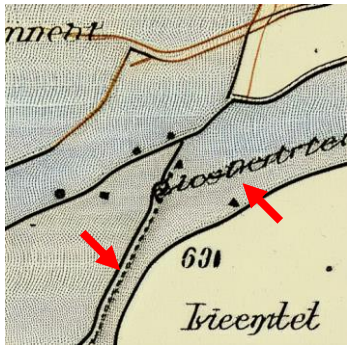


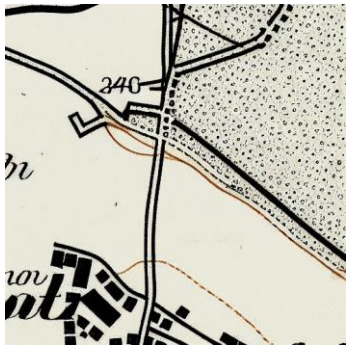
”a man in the style of a *siegfr* map“

Experiment I: Text-based map generation

Intermediate results while training with 800 map tiles for 50 epochs

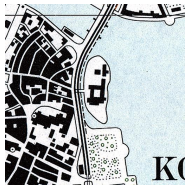
original

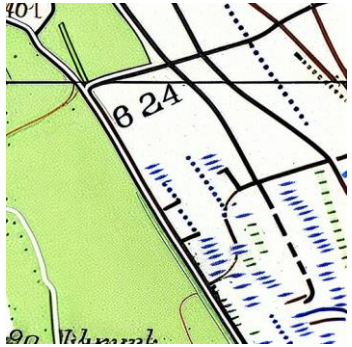




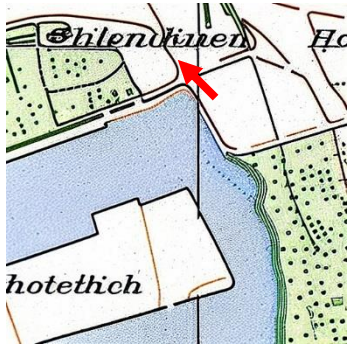
”*siegfr* map“

original





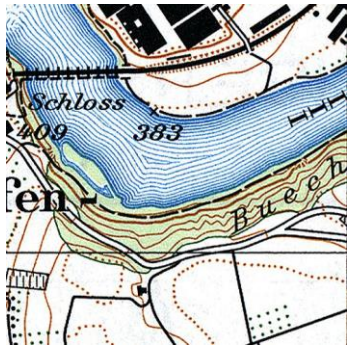
”*oldnat* map“



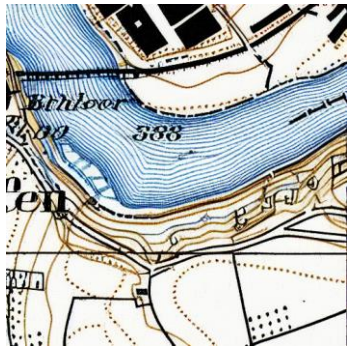
Problem: Overfitting

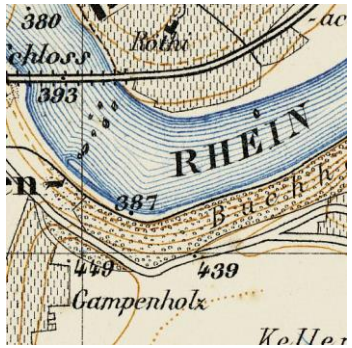
Experiment I: Text-based map generation

Results after training with 800 map tiles for 50 epochs



”*siegfr* map“



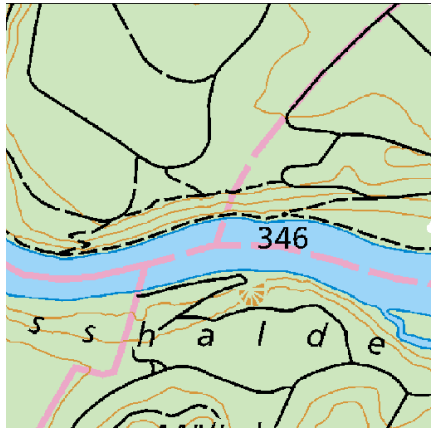


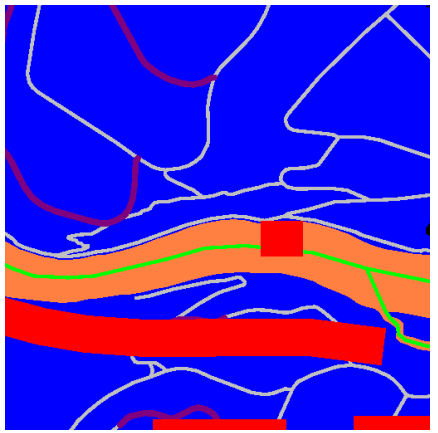
”*oldnat* map“

Style transfer attempt

Experiment II: Image-based map generation

1300 tiles and segmentation masks of the national map of Switzerland (2021)





Framework: Diffusers

Method: ControlNet

Pre-trained model: stable-diffusion-2-1-base

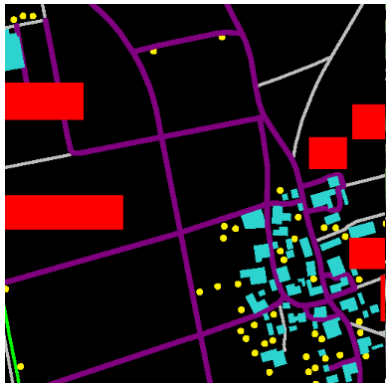
Hardware: NVIDIA RTX A4000 (16GB)

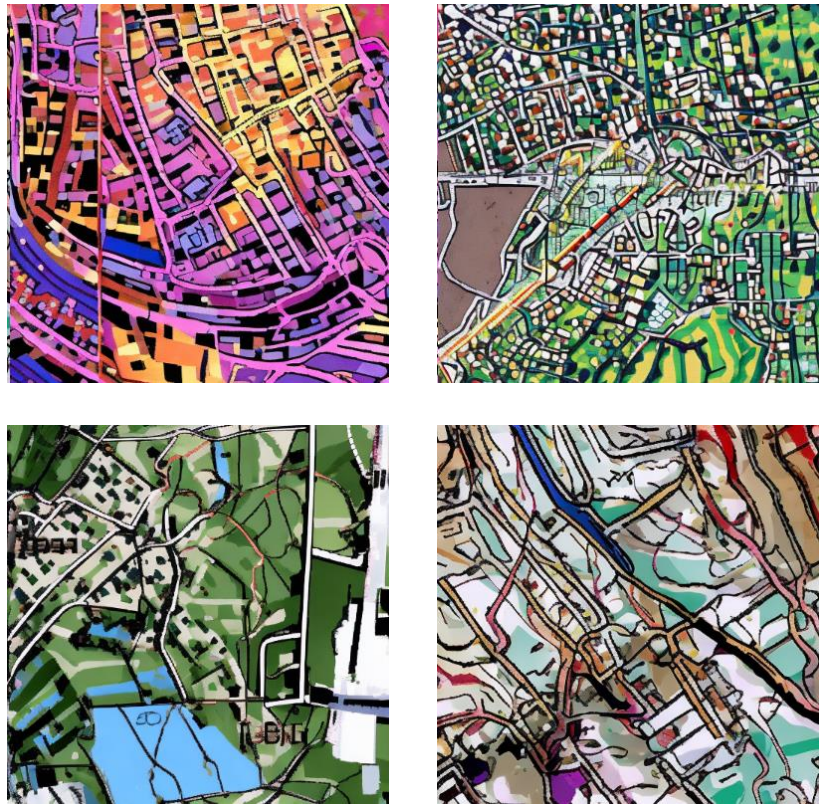
https://arxiv.org/abs/2302.05543

https://huggingface.co/blog/train-your-controlnet

Experiment II: Image-based map generation

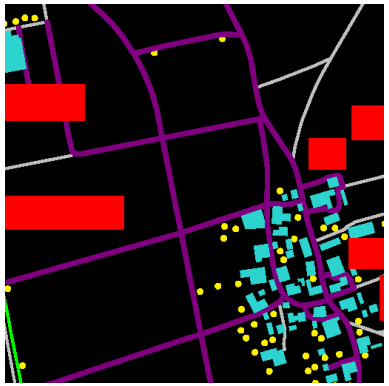
Intermediate results in the first training epochs



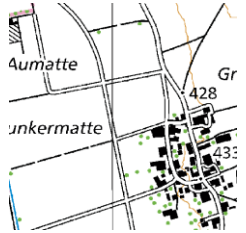


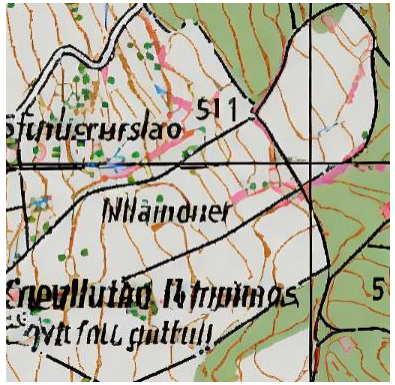
Experiment II: Image-based map generation

Results after training for 50 epochs

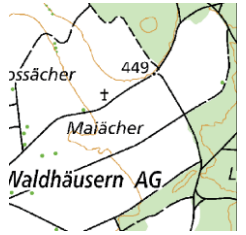


original





original



Summary

Opportunities	Challenges
Stable Diffusion can be applied to a variety of cartographic tasks.	Training a Stable Diffusion model is resource-intensive.
Existing map styles can be reproduced and artistic map styles can be produced.	The symbology, topology and labels of generated maps may not be plausible.
The generation process can be controlled by additional inputs.	Fine-grained control as known from GIS software may be difficult to achieve.

Outlook

3D

“a DSLR photo of a squirrel wearing a kimono riding a motorcycle“

https://dreamfusion3d.github.io/

Animation

“An animated painting of fluffy white clouds moving in sky”

https://research.nvidia.com/labs/toronto-ai/VideoLDM/

Thank you for your attention!

This presentation will be available at

https://blogs.ethz.ch/schnuerer

Let‘s discuss in the Group Q&A or via

schnuerer@ethz.ch

Don‘t miss the talk in the afternoon session

The Ethics of AI-Generated Maps: A Study of DALL·E 2 and Implications for Cartography

Yuhao Kang, Qianheng Zhang and Robert Roth, University of Wisconsin Madison