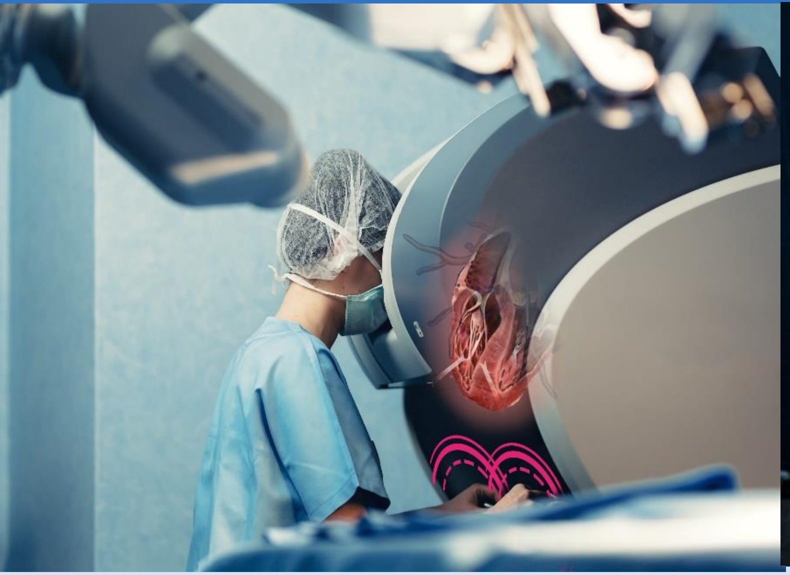
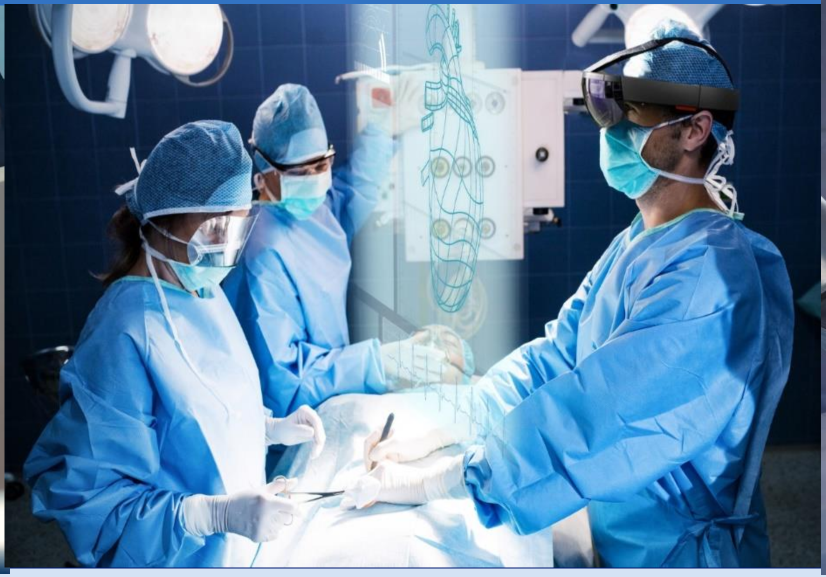
Synthesizing realistic surgical images with sim2real GANs


Sebastian Bodenstedt
Translational Surgical Oncology

NCT
NATIONALES CENTRUM
FÜR TUMORERKRANKUNGEN
PARTNERSTANDORT DRESDEN
UNIVERSITÄTS KREBSCENTRUM UCC

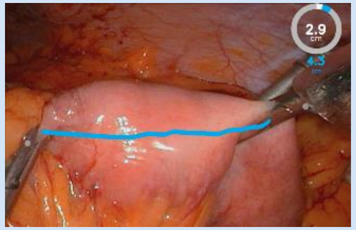
getragen von:
Deutsches Krebsforschungszentrum
Universitätsklinikum Carl Gustav Carus Dresden
Medizinische Fakultät Carl Gustav Carus, TU Dresden
Helmholtz-Zentrum Dresden-Rossendorf

Motivation: Example applications of ML in surgery

Robotics

Surgical Training

Data Analysis



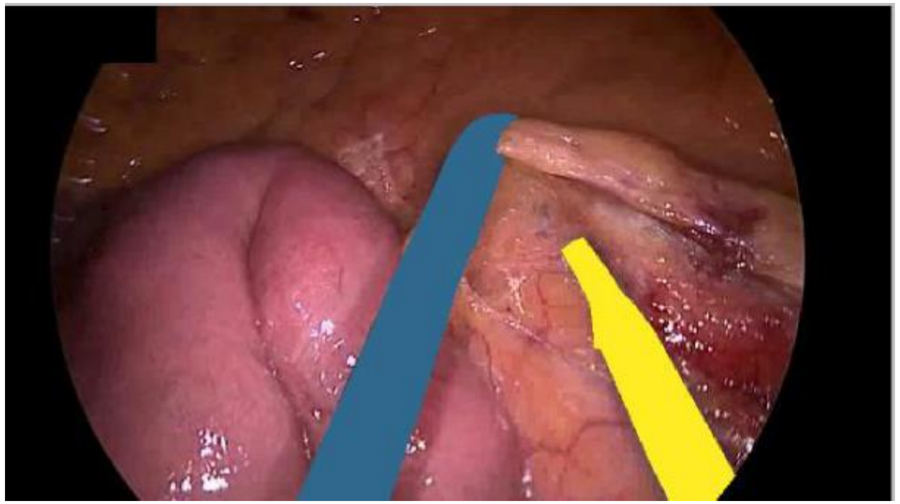
Human-Machine Interaction

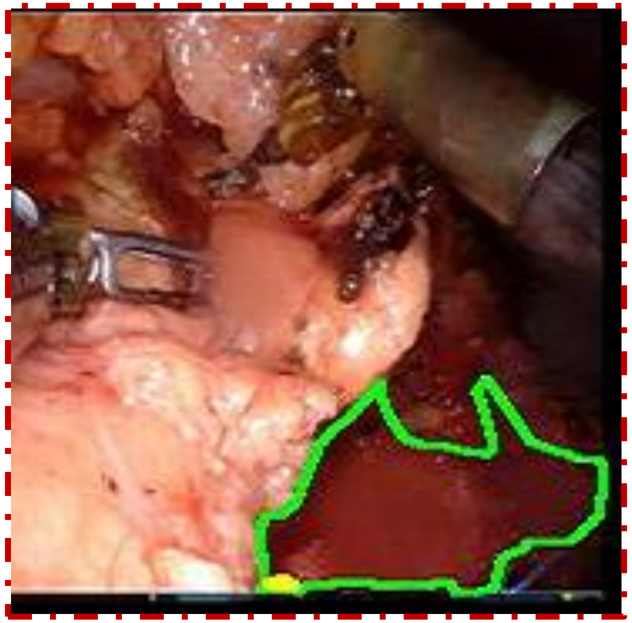
Navigation

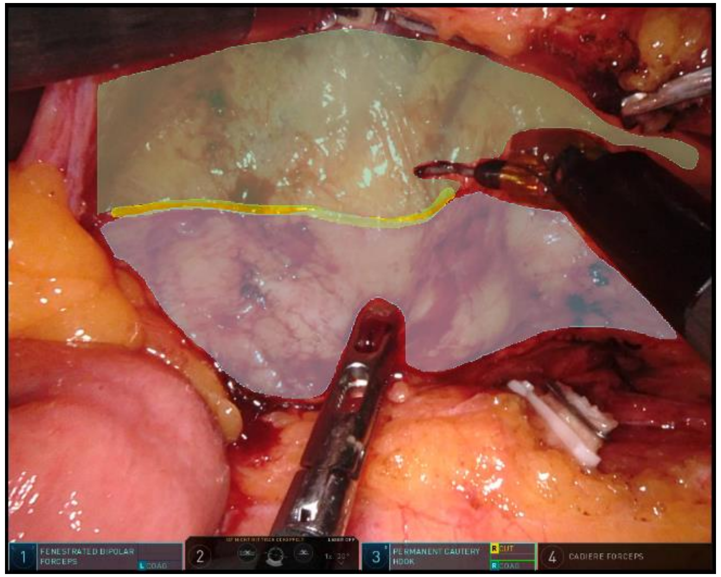
NCT

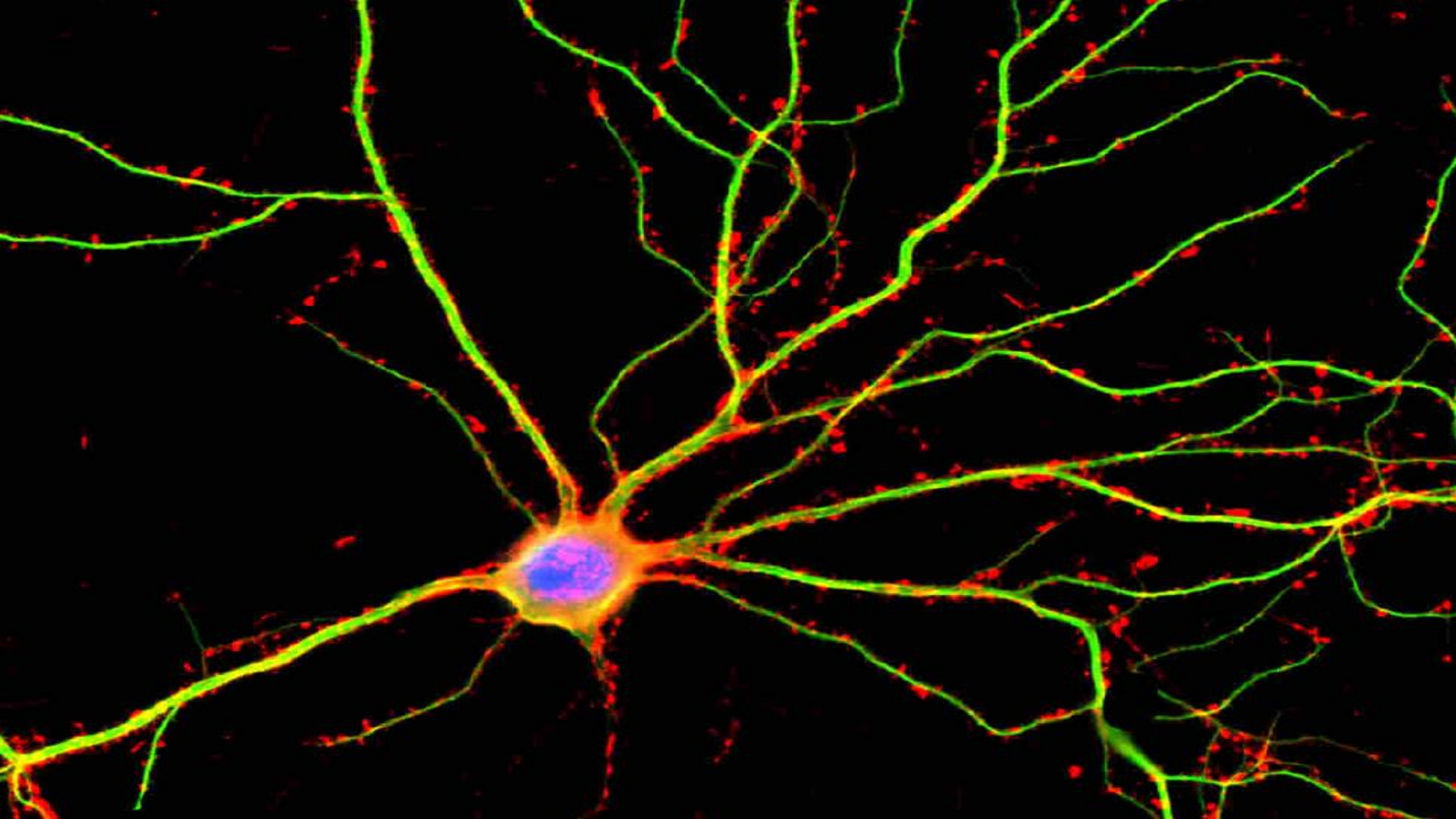
Motivation: Annotation of data

- Machine learning relies on large amounts of annotated training examples
- Annotation often requires expert knowledge is time consuming



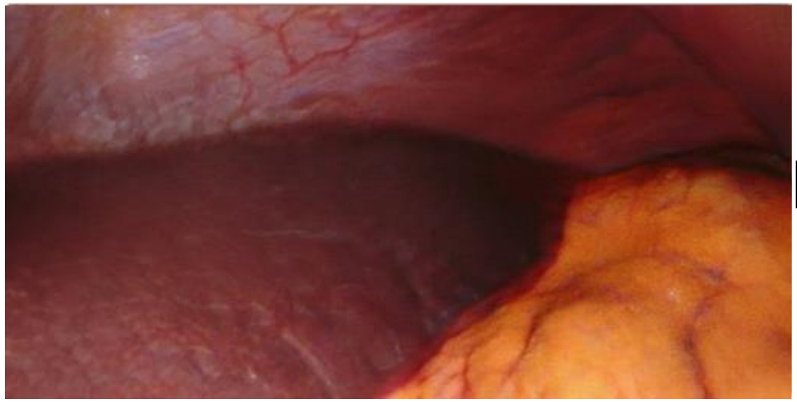


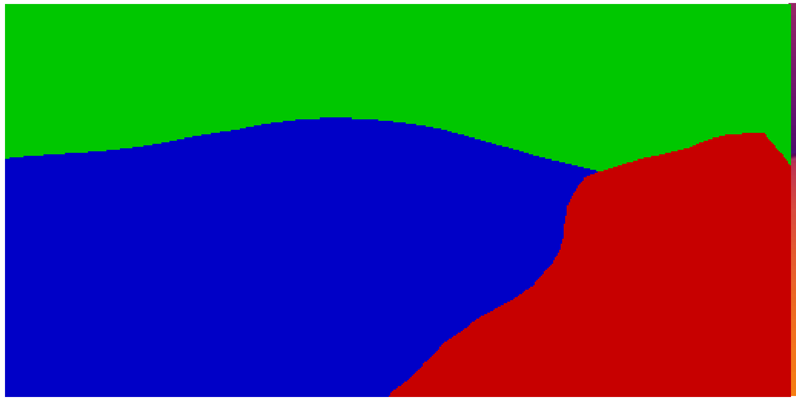




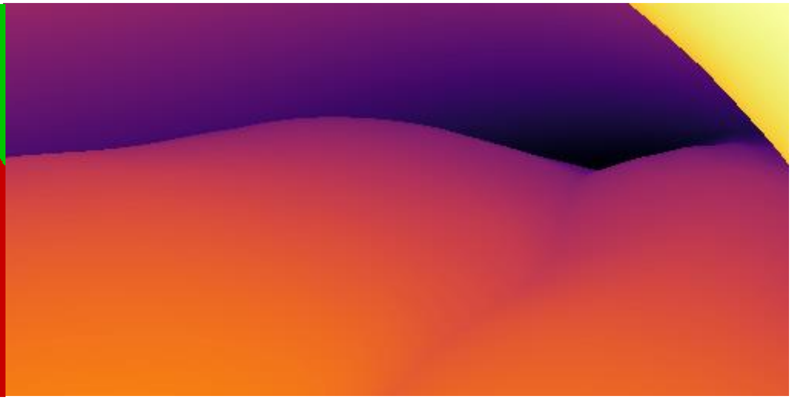
Motivation: Synthetic Data Generation

- Generating ground truth data for synthetic scenes is trivial
- Apply realistic textures to synthetic scenes





Segmentation



Depth

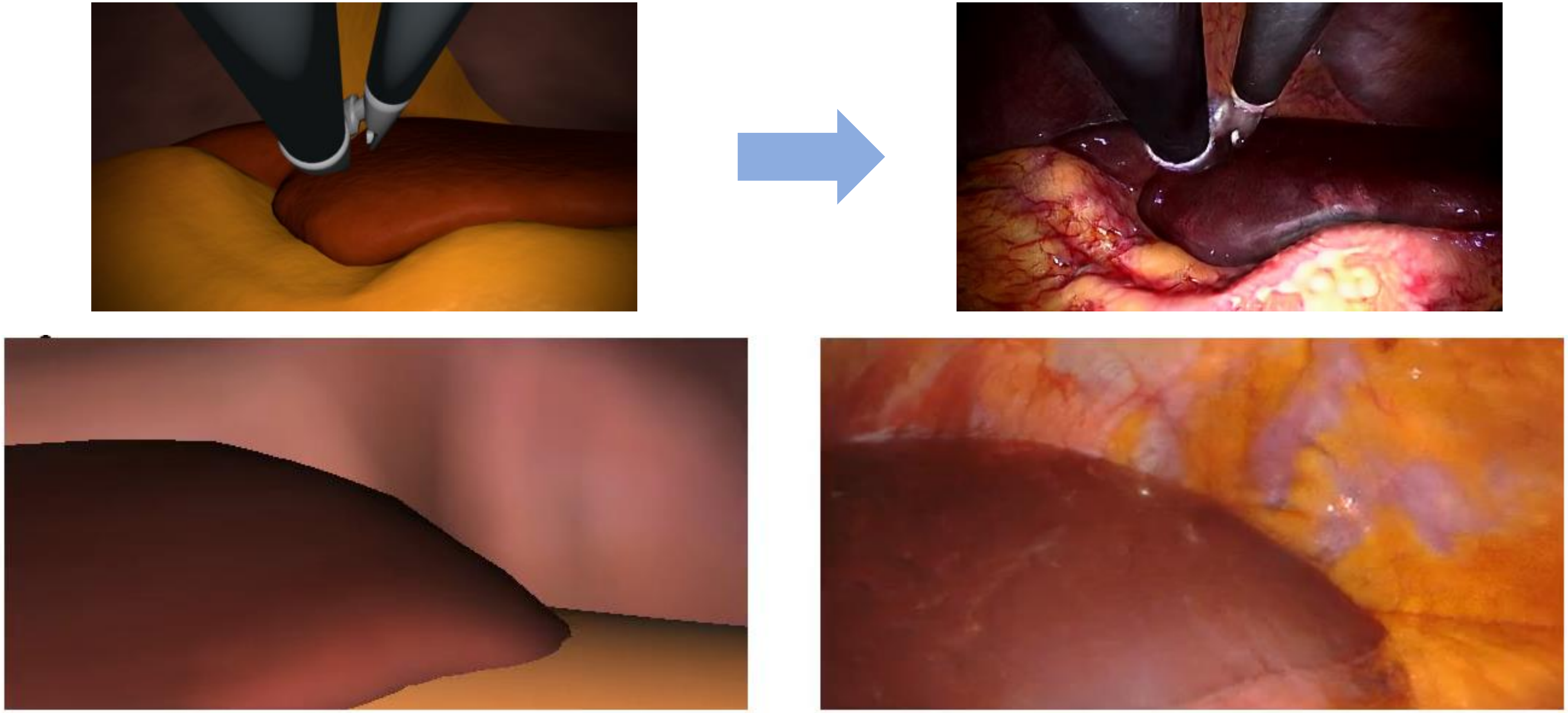
Optical Flow

Pfeiffer,..,Weitz, Speidel MICCAI 2019: „Generating large labeled data sets for laparoscopic image processing tasks using unpaired image-to-image translation"
Rivoir,..,Weitz, Speidel ICCV 2021: „Long-Term Temporally Consistent Unpaired Video Translation from Simulated Surgical 3D Data"

NCT

Motivation: Synthetic Data Generation

- Apply realistic textures to synthetic scenes



Pfeiffer,..,Weitz, Speidel MICCAI 2019: „Generating large labeled data sets for laparoscopic image processing tasks using unpaired image-to-image translation"
Rivoir,..,Weitz, Speidel ICCV 2021: „Long-Term Temporally Consistent Unpaired Video Translation from Simulated Surgical 3D Data"

Example Deep Fakes

Image: https://github.com/iperov/DeepFaceLive

Image: Chris Ume

Image: https://this-person-does-not-exist.com

Image: https://thenextweb.com

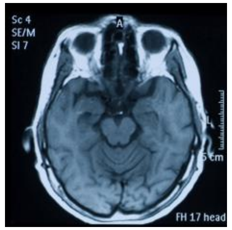
Overview lecture

- Introduction to Generative Adversarial Networks (GANs)
  - What are generative models
  - Different modes of learning
  - Components of GANs
- Example usages of GANs
  - Image2Image translation

Supervised vs unsupervised learning

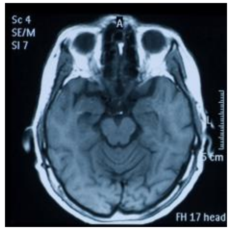
Supervised learning (concept learning)

- Learning by labeled example, i.e. we “tell” the algorithm what to learn
- Two forms of output
  - **Symbolic (classification)**
    - Output is a discrete value/category, e.g. “tumor”/“no tumor”



“no tumor”

  - **Subsymbolic/regression**
    - Output is continuous value, e.g. age or temperature

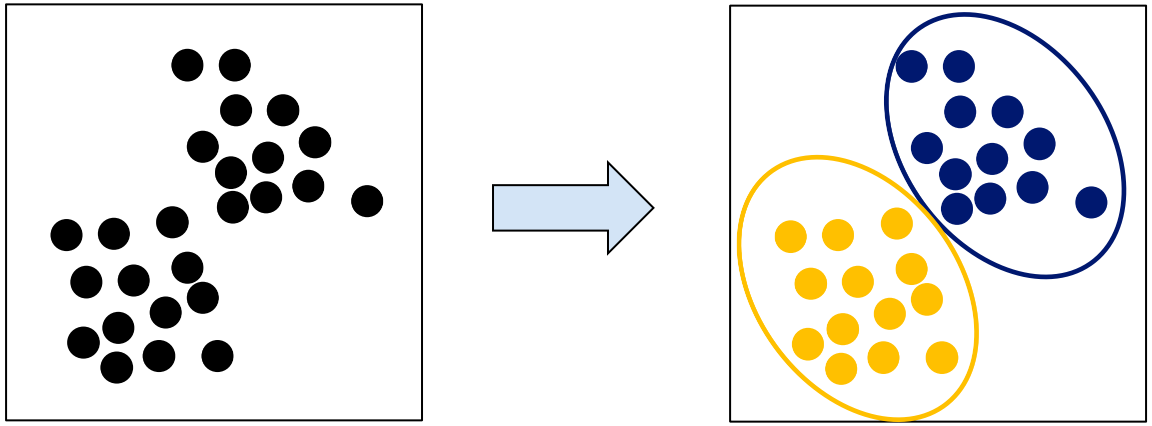


“42.0 years old”

Supervised vs unsupervised learning

Unsupervised learning

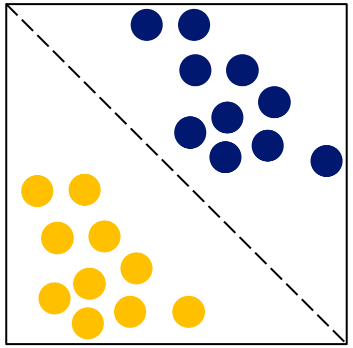
- Detect previously unknown patterns in unlabeled data
  - No desired output!
- Example applications
  - Find anomalies in data, e.g. credit card usage
  - Find clusters, e.g. grouping pictures



Discriminative vs generative learning

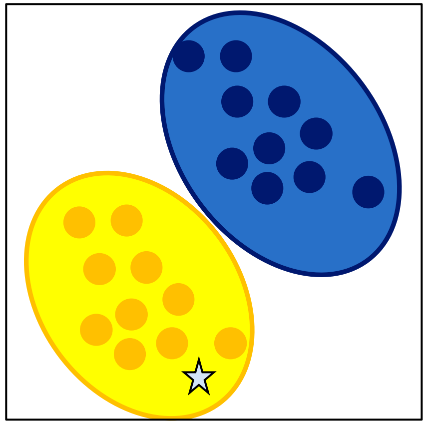
Discriminative learning a.k.a classification

- Supervised learning problem
- A class label is assigned to an input
- => can be used to **discriminate** between given examples



Generative Learning

- Unsupervised learning problem
- A model learns the distribution of input data
- => can be used to **generate** new examples of that distribution



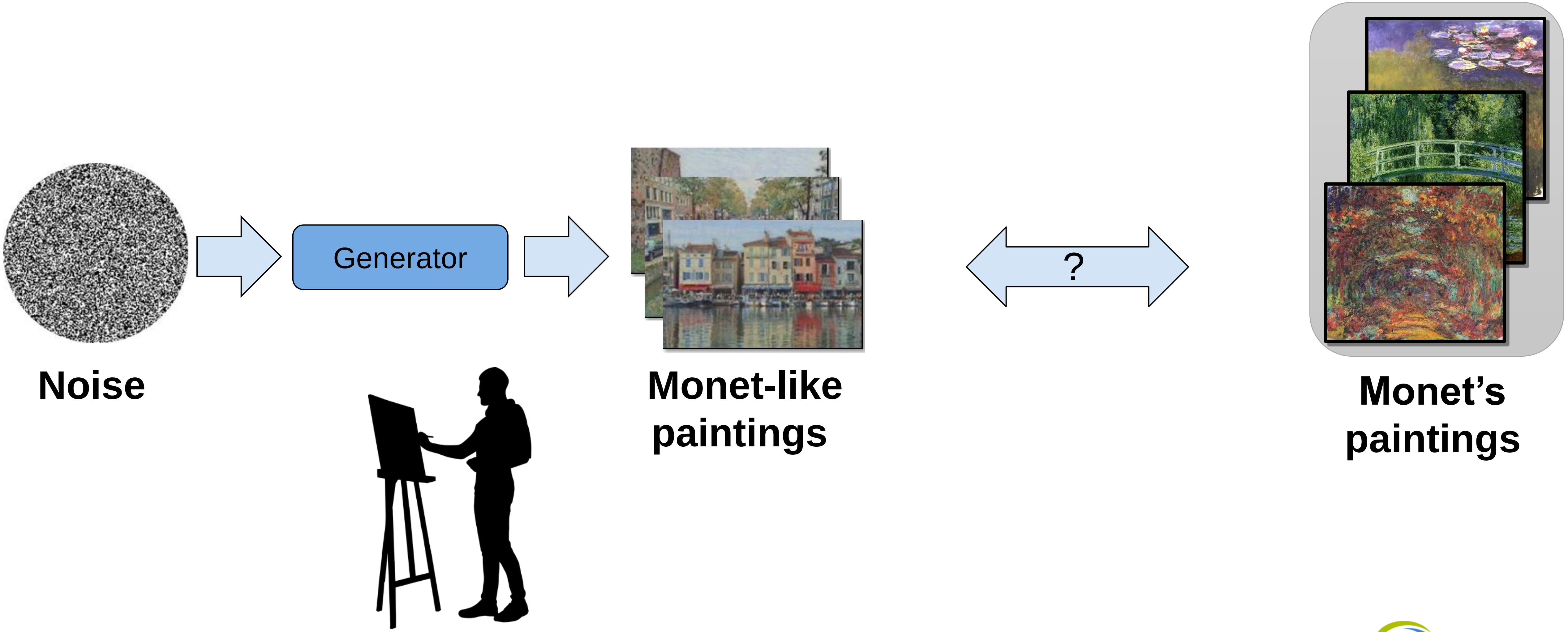
Both play a role in Generative Adversarial Networks!

Generative Adversarial Networks (GANs)

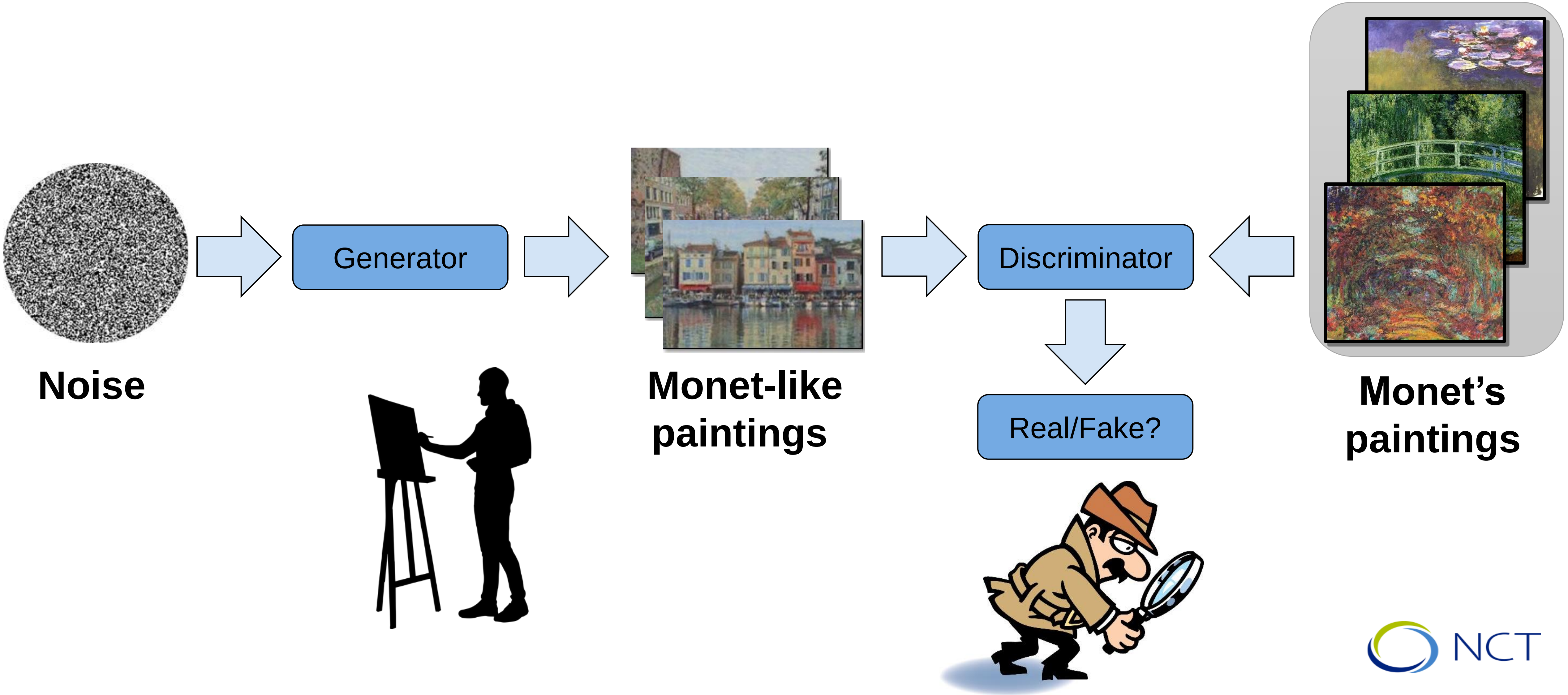
Generative models based on Deep Neural Networks

- An approach for training generative models
  - Introduced by Goodfellow et al. *Generative Adversarial Networks*
- Generally consists of two separate models
  - **Generator**: generates new samples from a target domain
  - **Discriminator:** determines whether a sample is real (from domain) or generated
- Training a GAN is comparable to an arms race between the separate models
  - Example: art forger vs. detective

Example: Teaching an art style

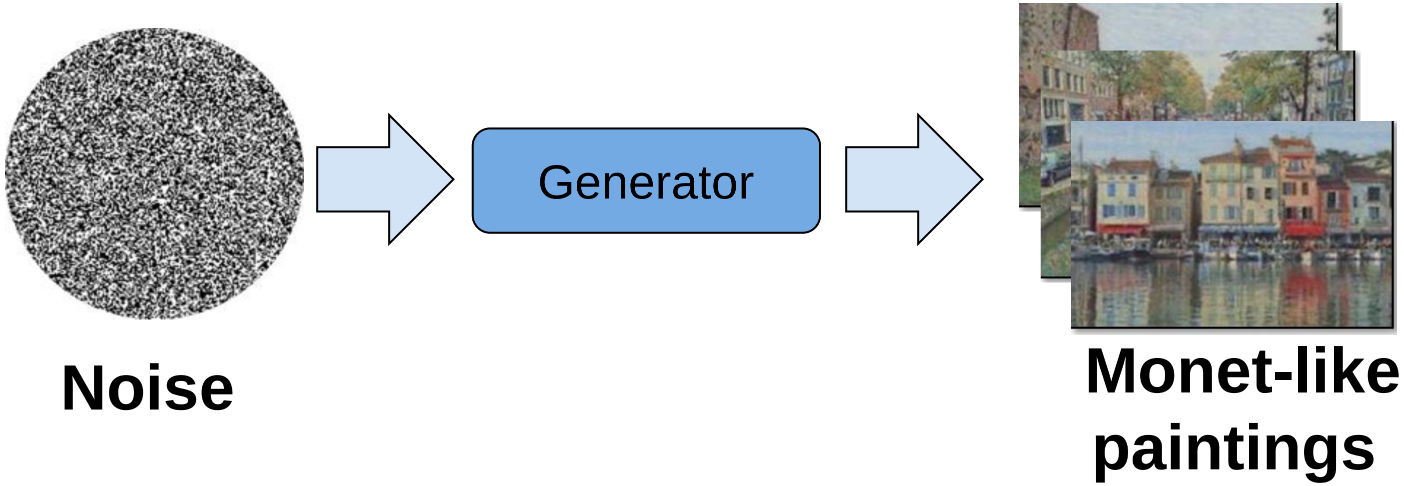


Example: Teaching an art style



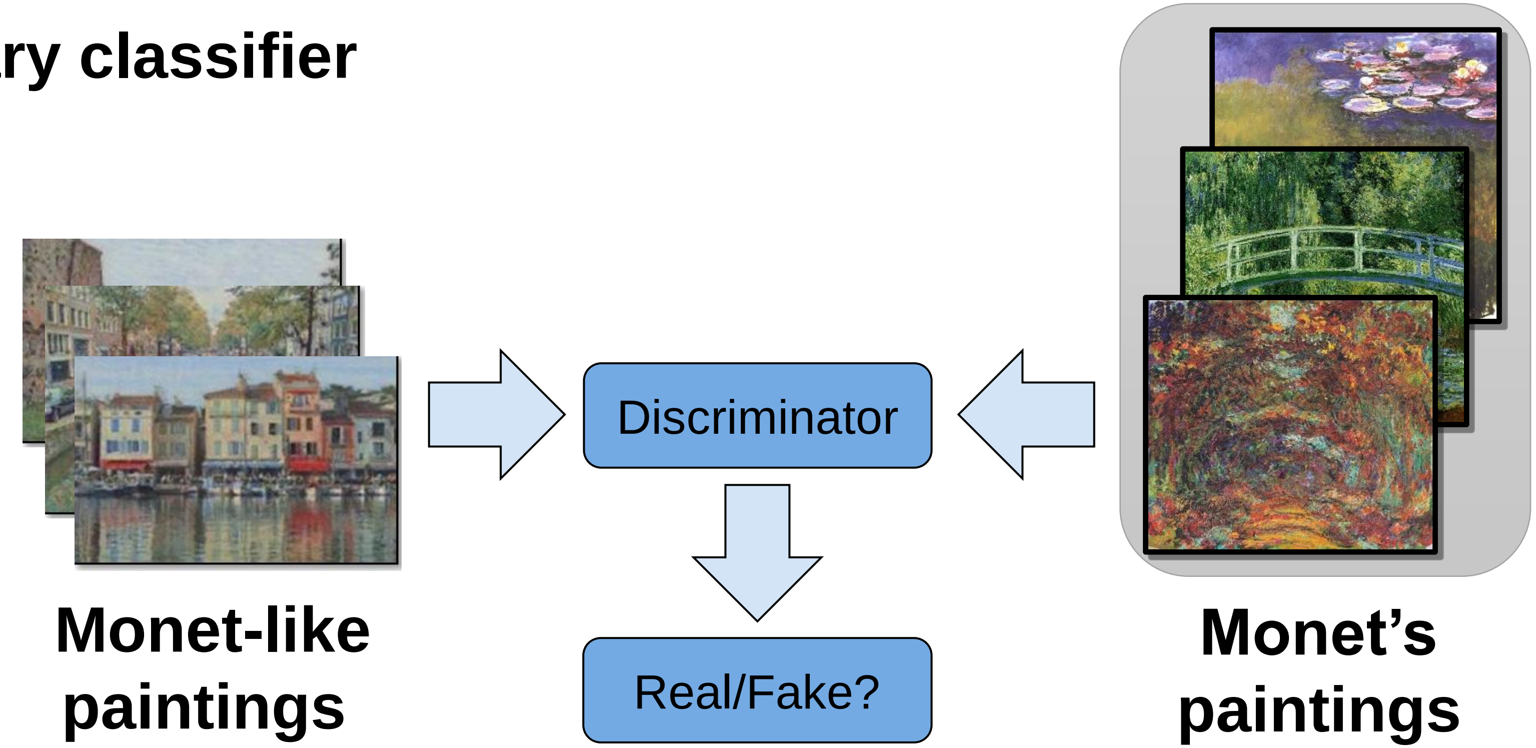
Generator

- Learns **distribution** of **a target domain**
  - Given a random input vector, generates a sample
    - Points in vector (latent) space correspond to points in target domain
  - **Latent space** => compression of the target domain
  - Often a Convolutional Neural Network (CNN)



Discriminator

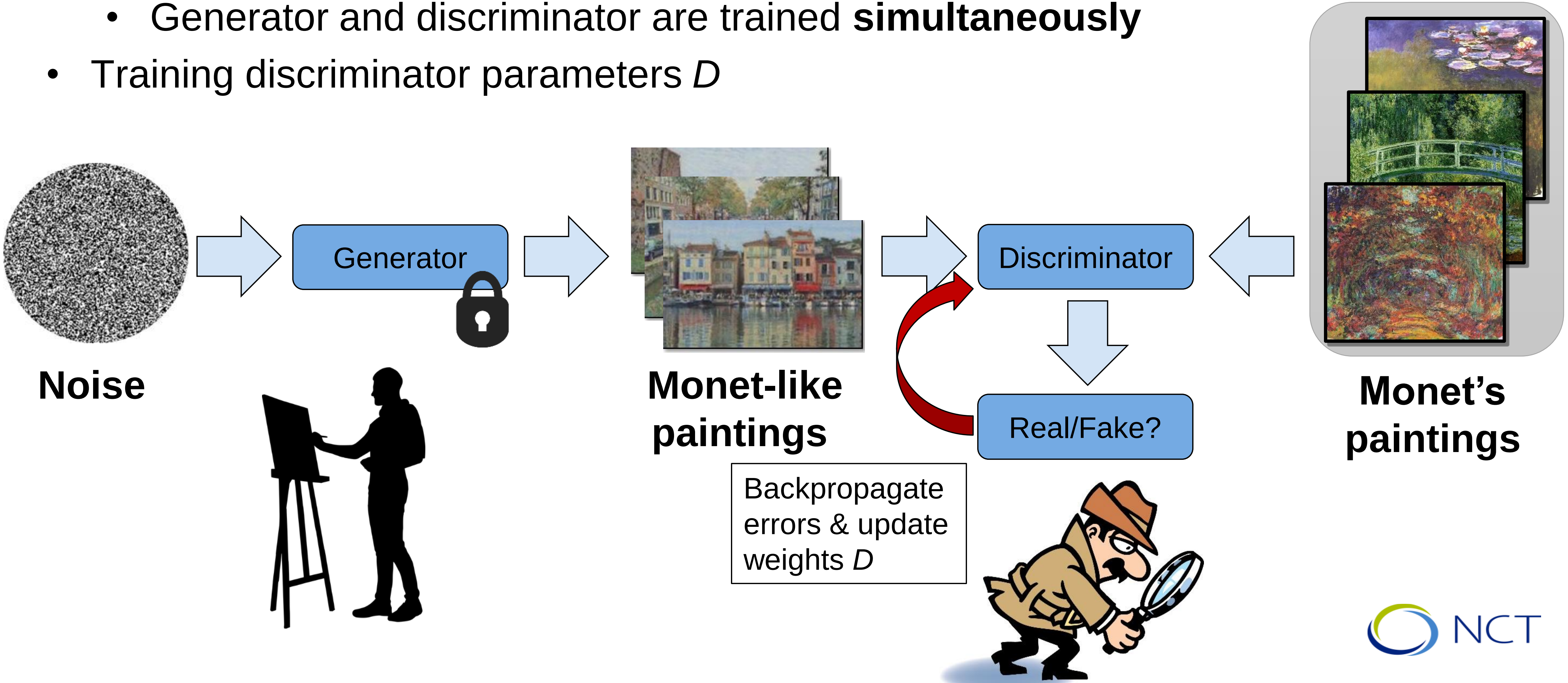
- Distinguishes real from fake (generated) examples
  - Real examples come from training set
  - Generated examples come from generator
  - Indirectly **connects** generator to target domain
  - Often a Convolutional Neural Network (CNN)
    - **Binary classifier**



NCT

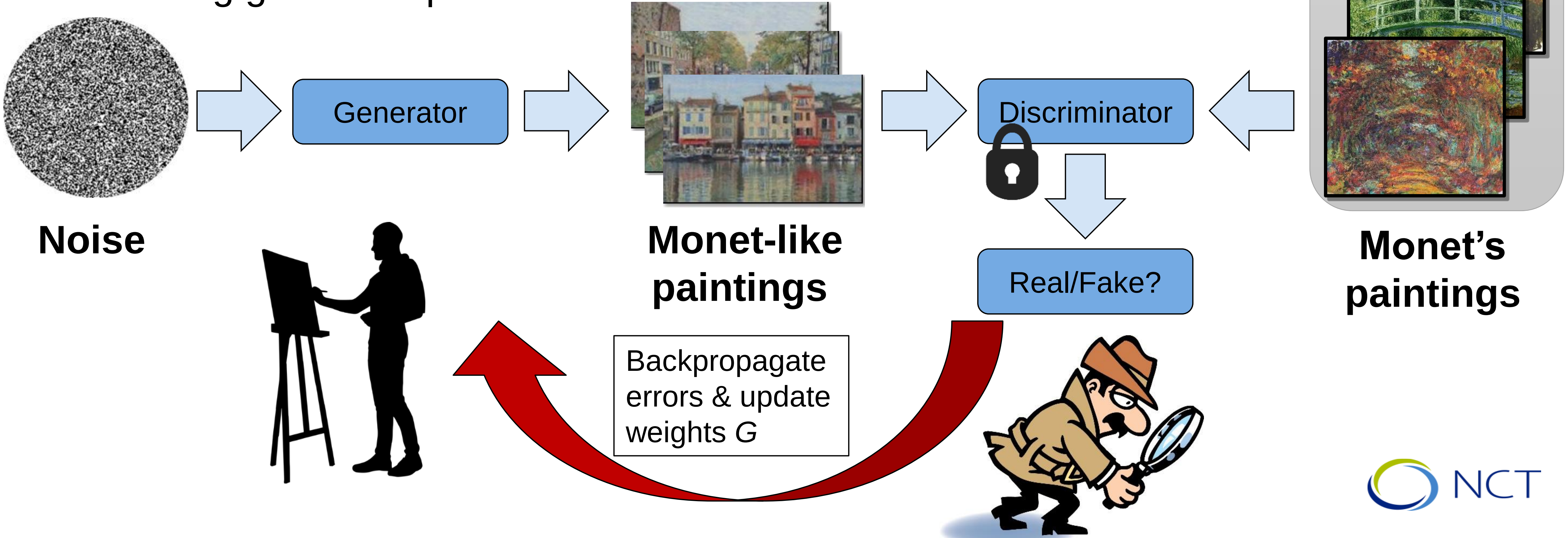
Training a GAN

- Two step process
  - Generator and discriminator are trained **simultaneously**
- Training discriminator parameters *D*



Training a GAN

- Two step process
  - Generator and discriminator are trained simultaneously
- Training discriminator parameters *D*
- Training generator parameters *G*

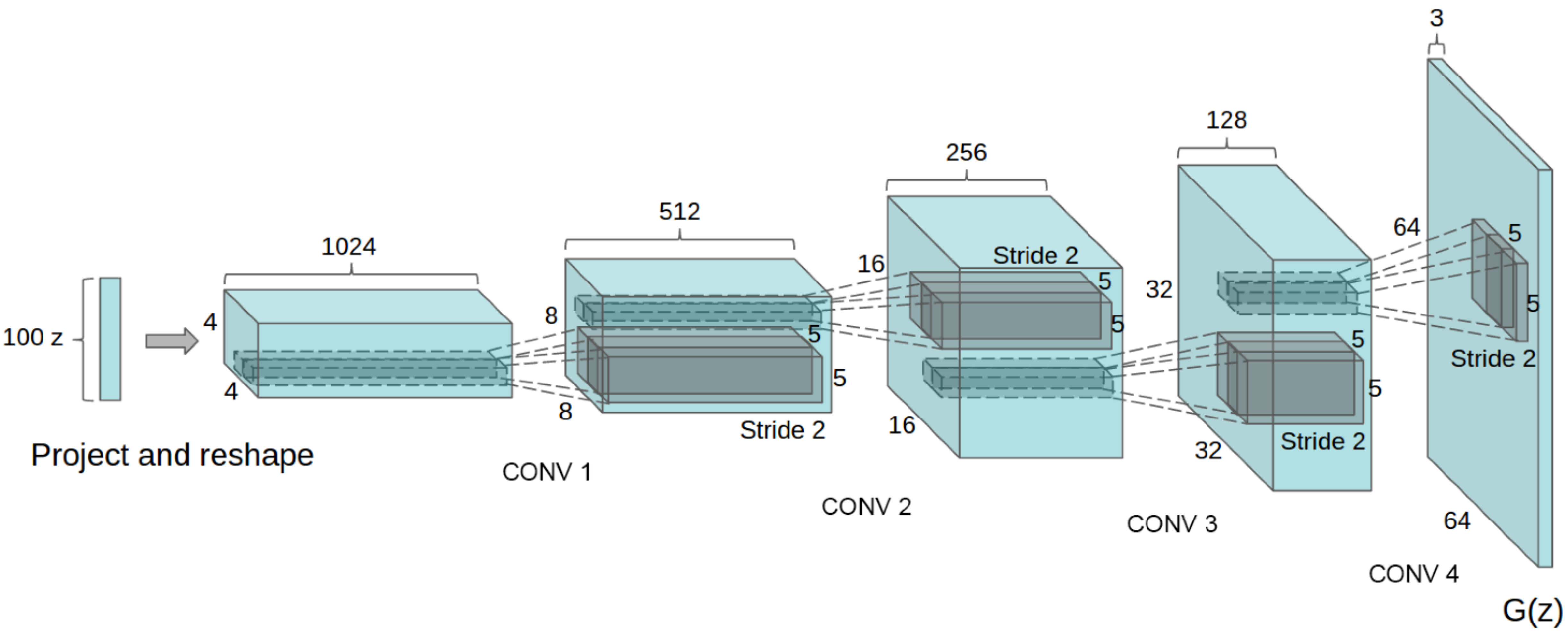


NCT

Training a GAN

- Two step process
  - Generator and discriminator are trained simultaneously
- Training discriminator parameters *D*
- Training generator parameters *G*
- Formulated as a **minimax** game
  - The discriminator aims to maximize its reward, i.e. its fake/real accuracy
  - The generator aims to minimize the reward of the discriminator

Some examples

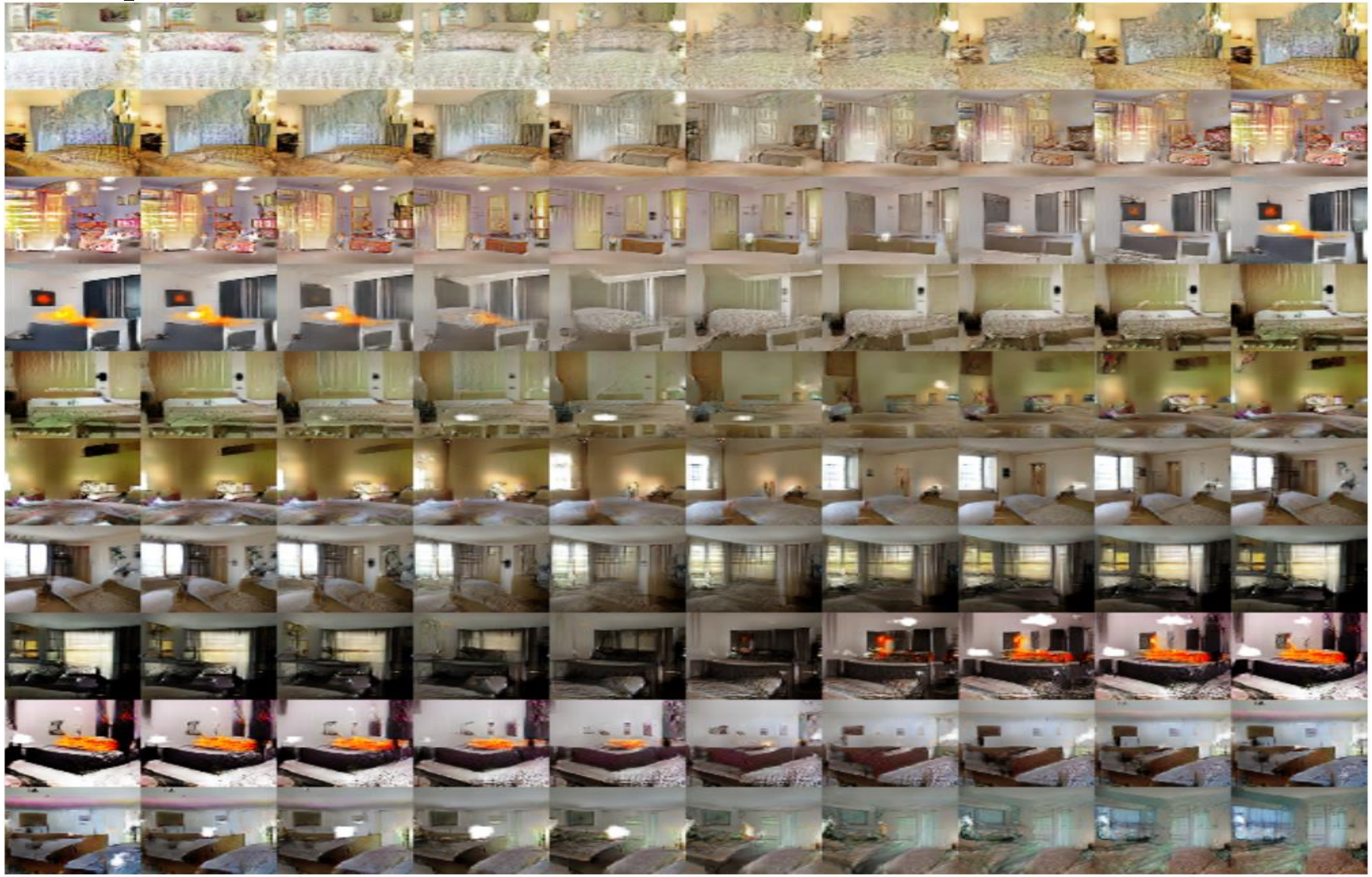


Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

Some examples

Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

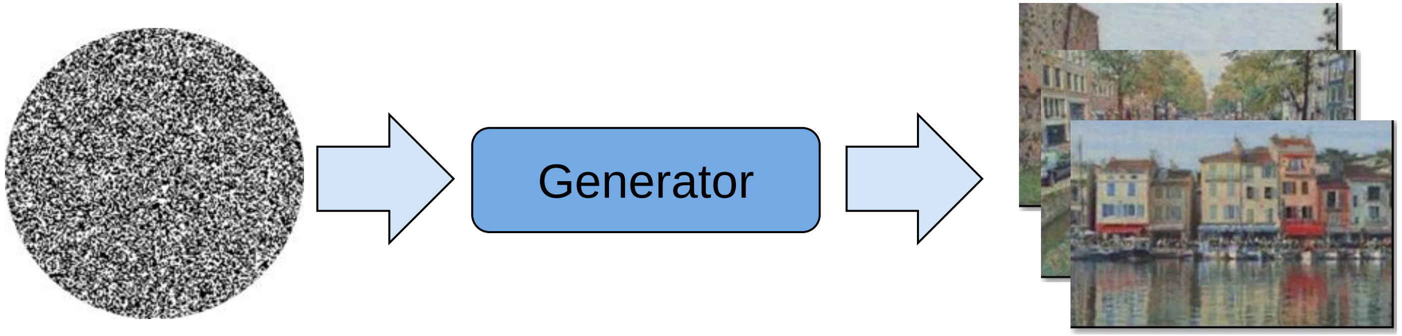
Some examples



Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

So far

- Image generation from noise



- Next: Image to Image translation

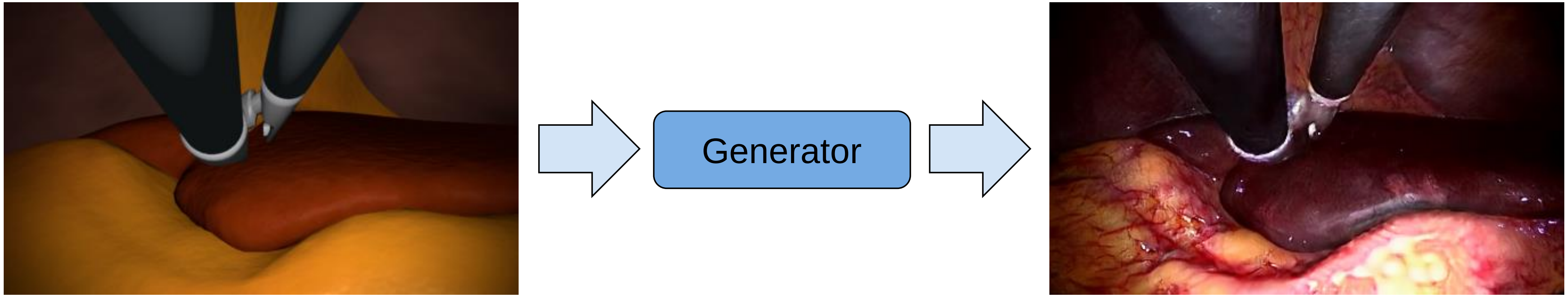
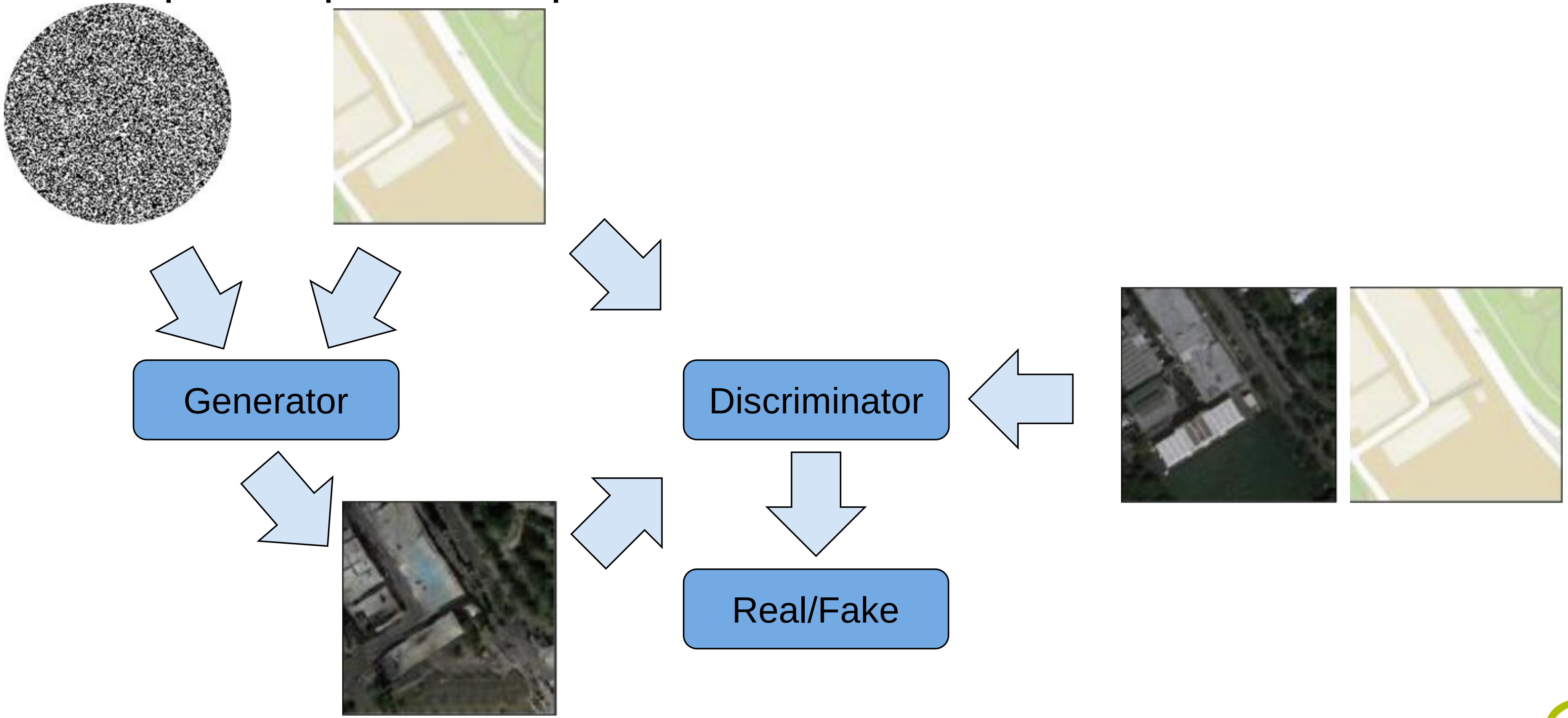


Image to image translation (with paired data)

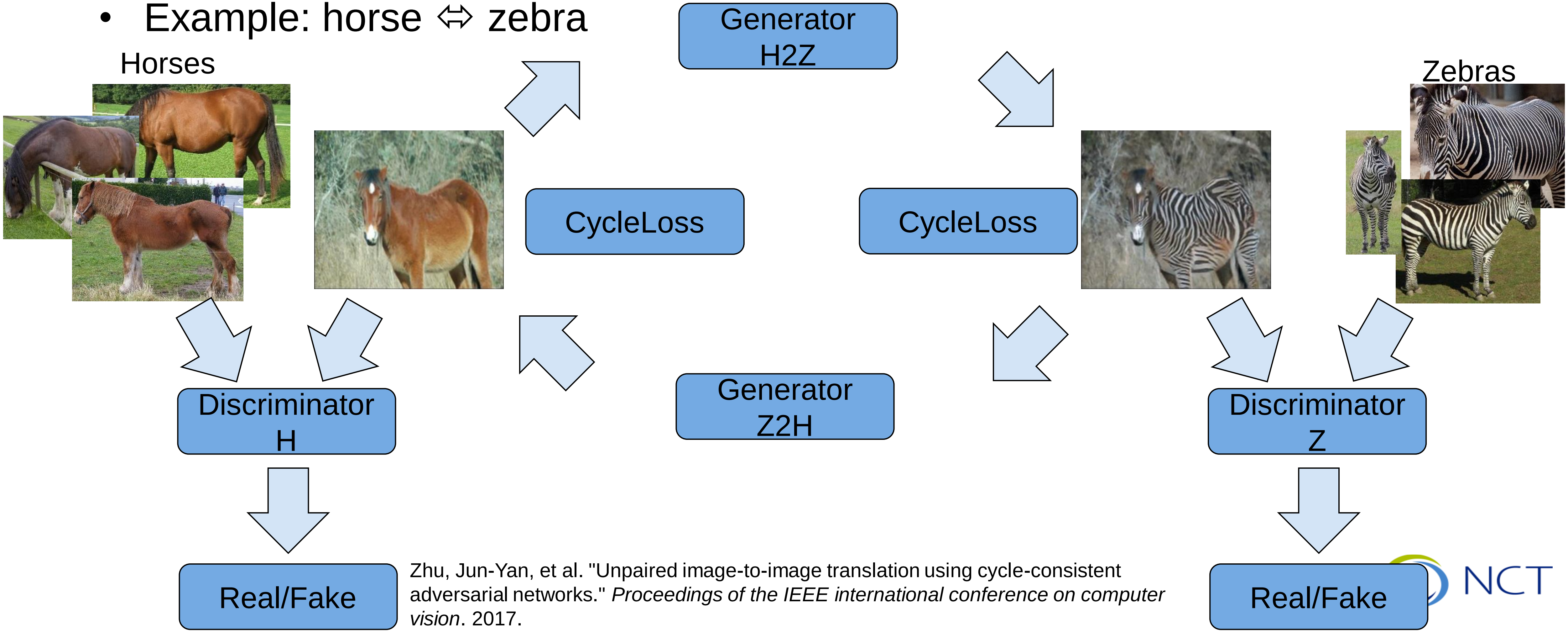
- Generally based on conditional GANs
  - Example: Map to aerial photo



Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. “Image-to-image translation with conditional adversarial networks”. arXiv preprint arXiv:1611.07004. (2016)

Image to image translation (with unpaired data)

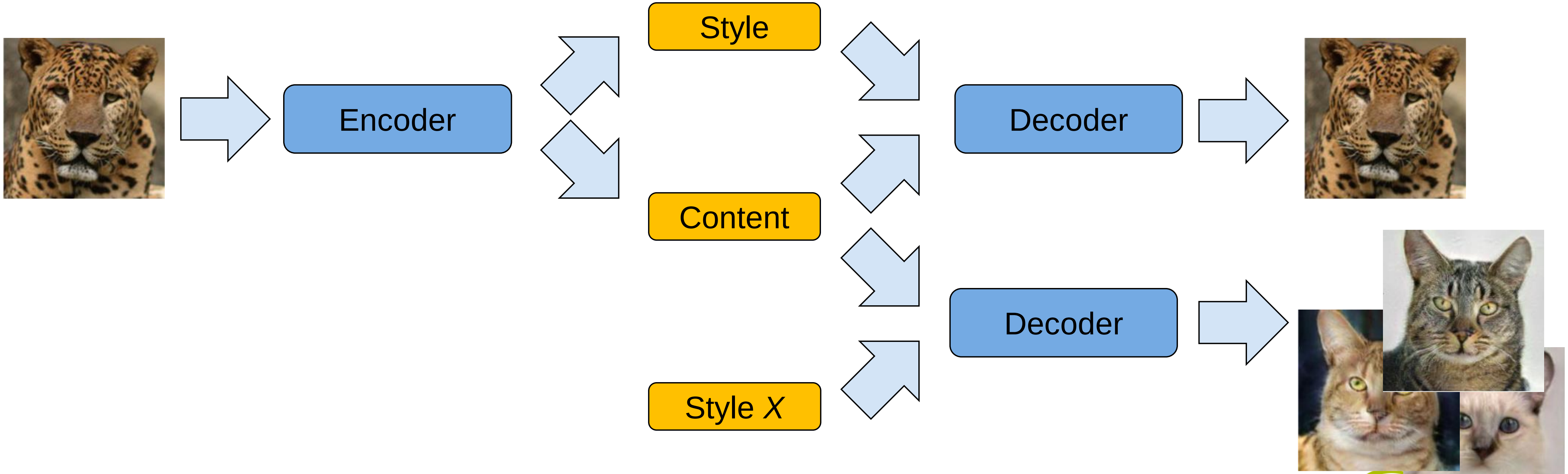
- Dataset with no corresponding classes
- Generally based on Cycle GANs
  - Example: horse ⬄ zebra



Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *Proceedings of the IEEE international conference on computer vision*. 2017.

Multimodal image to image translation

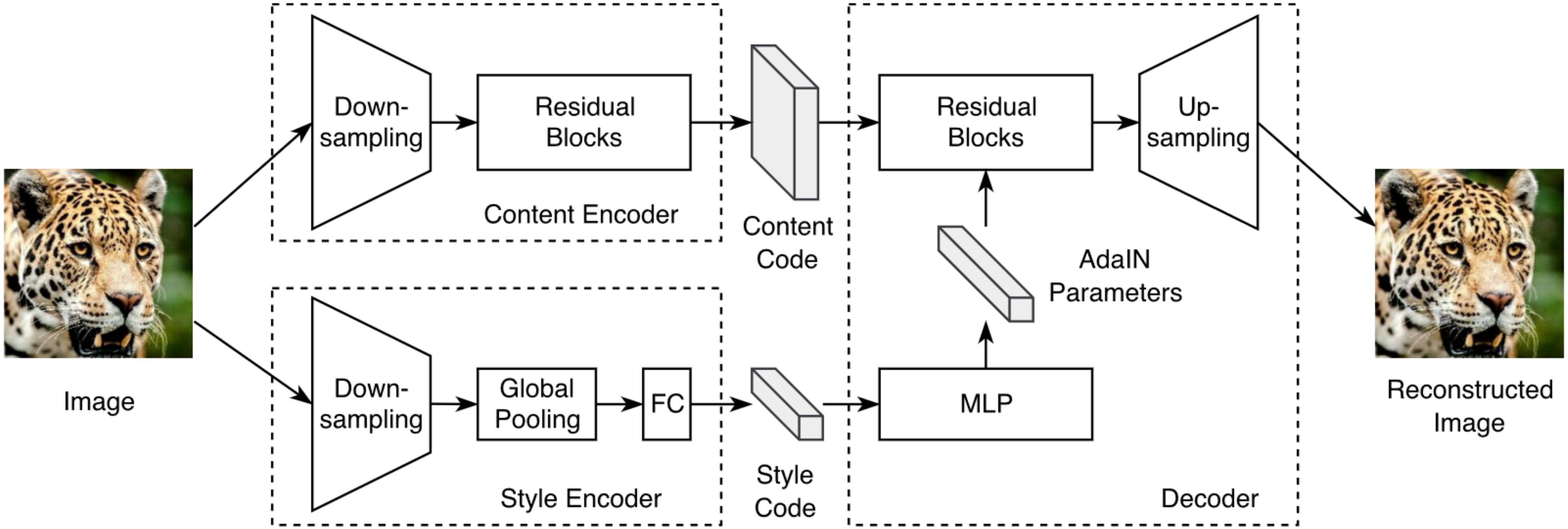
- Assumption:
  - An image can be decomposed into a **content code** and a **style code**
  - Image translation through combining a content code with a different style code



Huang, Xun, et al. "Multimodal unsupervised image-to-image translation." *Proceedings of the European conference on computer vision (ECCV)*. 2018.

Multimodal image to image translation

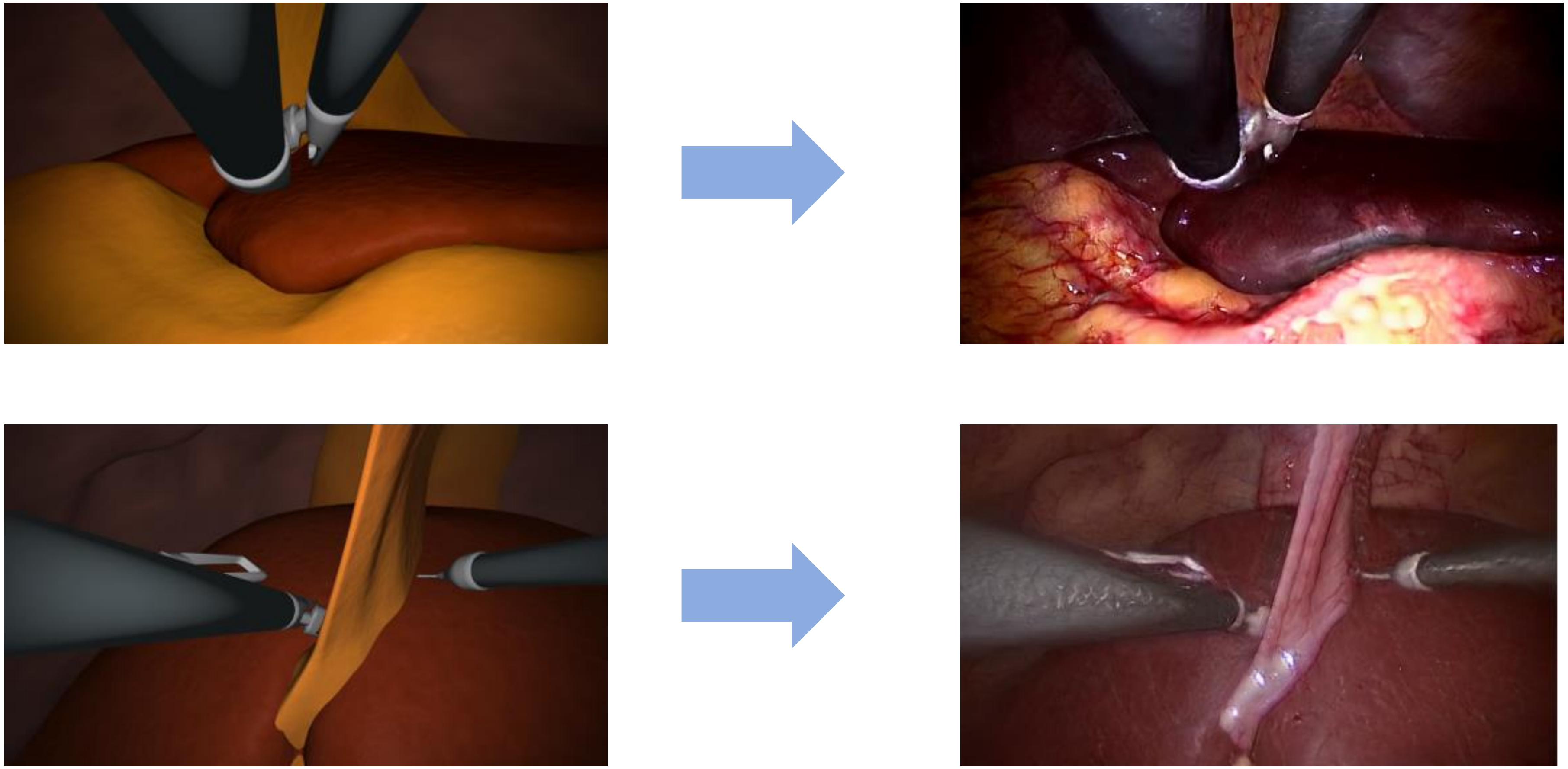
- Trained in similar manner as unpaired image to image
- Additional loss terms for:
  - Reconstruction of images with original codes
  - Reconstruction of codes by decomposing image with know codes



Huang, Xun, et al. "Multimodal unsupervised image-to-image translation." *Proceedings of the European conference on computer vision (ECCV)*. 2018.

Image to image translation for surgical images

- Unpaired image-to-image translation for generating realistic images

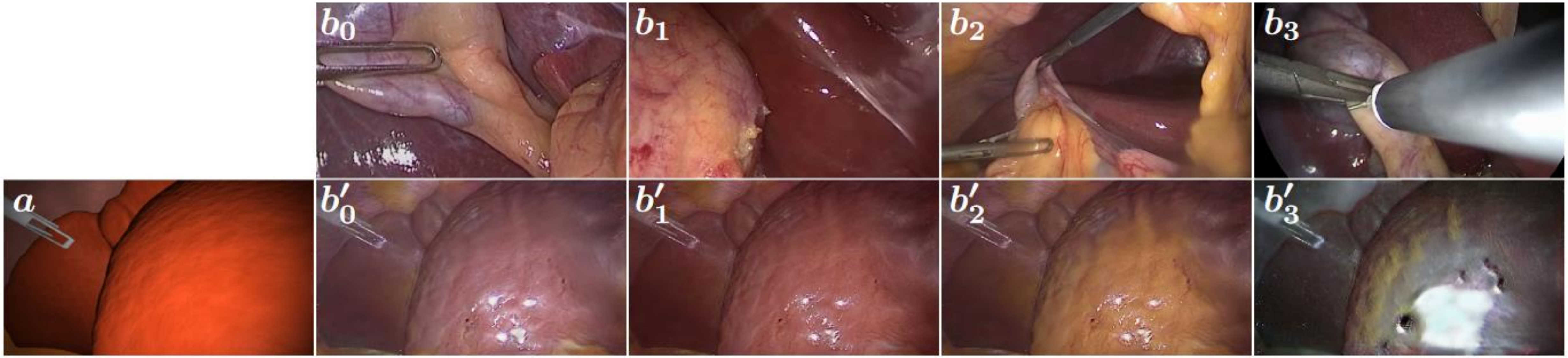


Pfeiffer,..,Weitz, Speidel MICCAI 2019: „Generating large labeled data sets for laparoscopic image processing tasks using unpaired image-to-image translation"

NCT

Image to image translation for surgical images

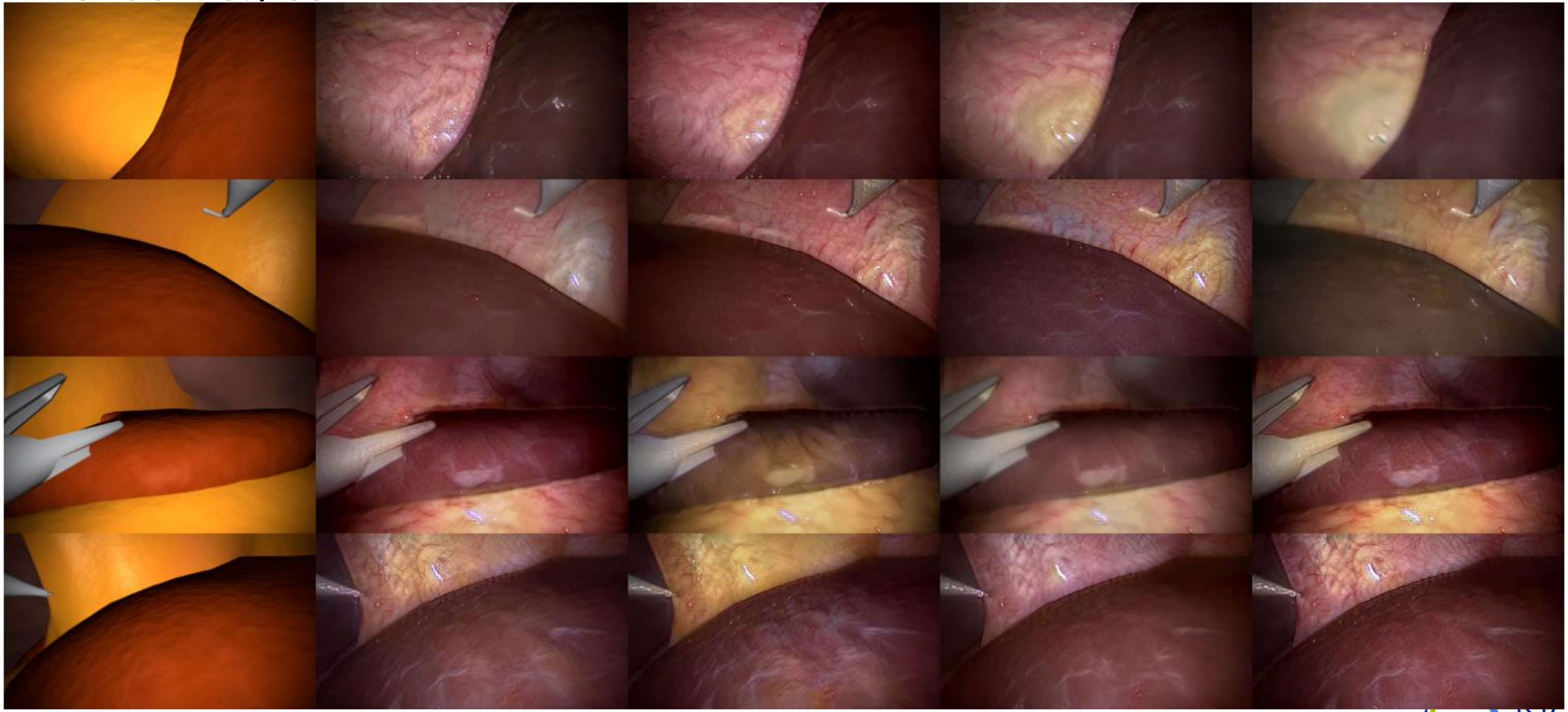
- Different styles extracted from data



Pfeiffer,..,Weitz, Speidel MICCAI 2019: „Generating large labeled data sets for laparoscopic image processing tasks using unpaired image-to-image translation"

Image to image translation for surgical images

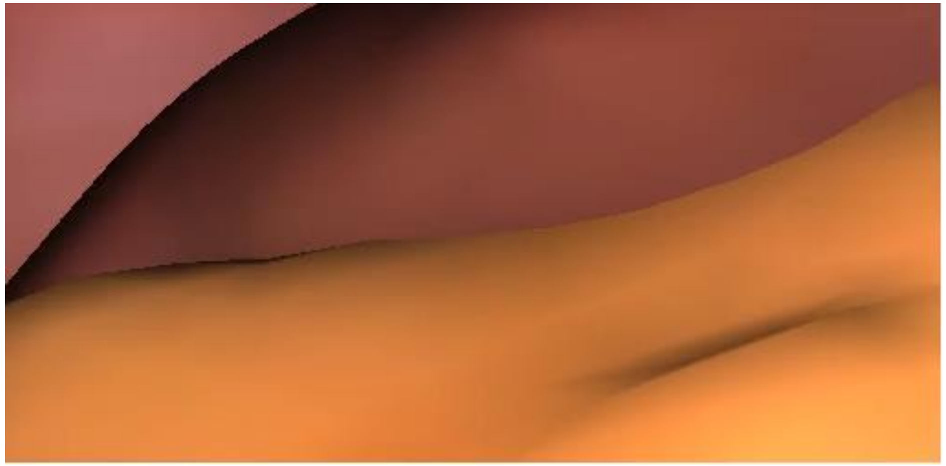
- Random styles

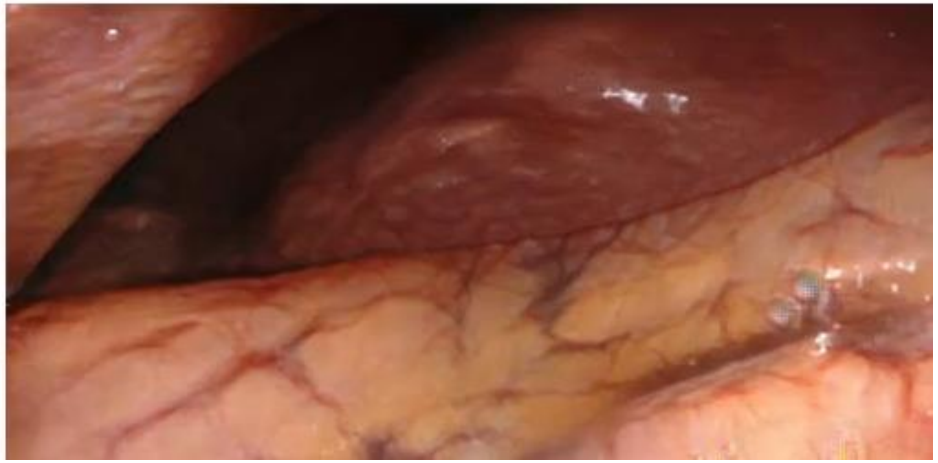


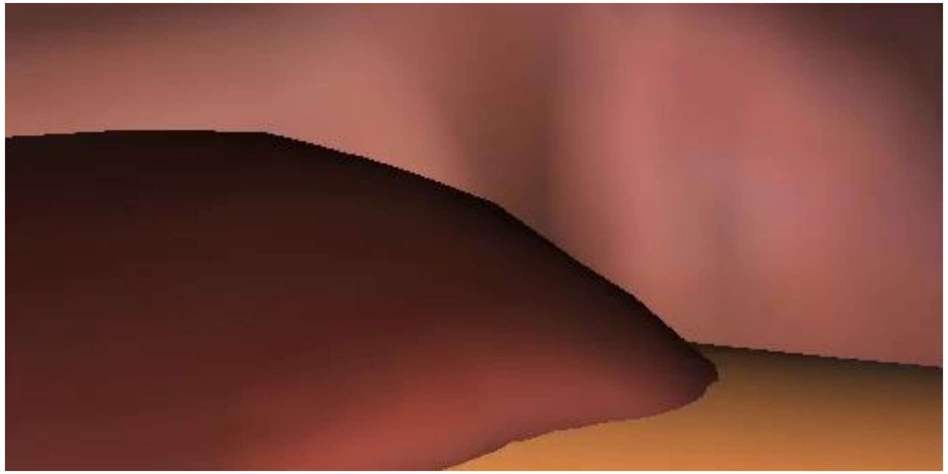
Pfeiffer,..,Weitz, Speidel MICCAI 2019: „Generating large labeled data sets for laparoscopic image processing tasks using unpaired image-to-image translation"

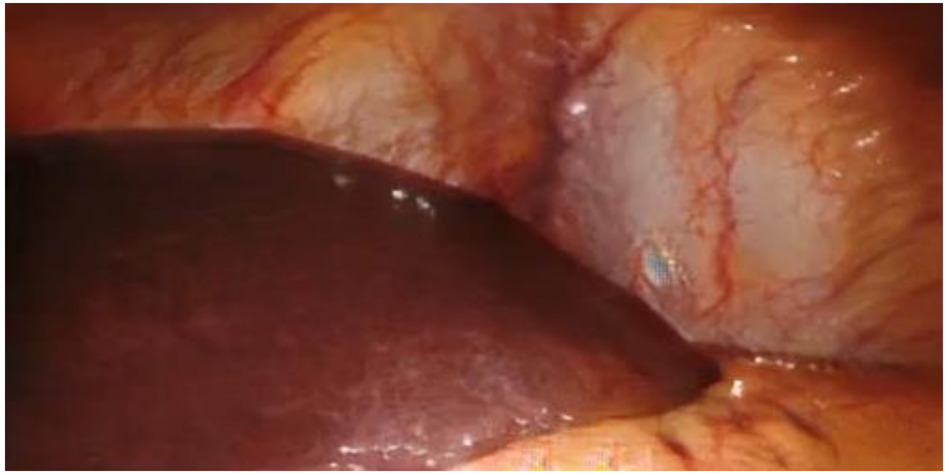
Simulation to video translation for surgical images

- Applying image to image translation to sequences can lead to temporal inconsistencies







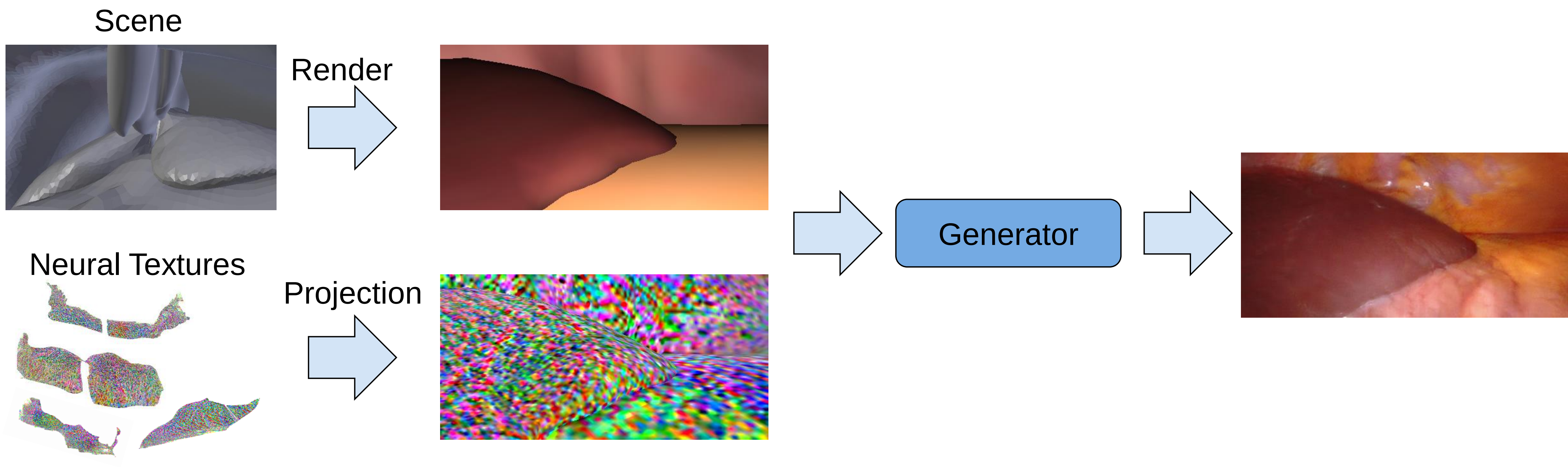


Rivoir,..,Weitz, Speidel ICCV 2021: „Long-Term Temporally Consistent Unpaired Video Translation from Simulated Surgical 3D Data"

NCT

Simulation to video translation for surgical images

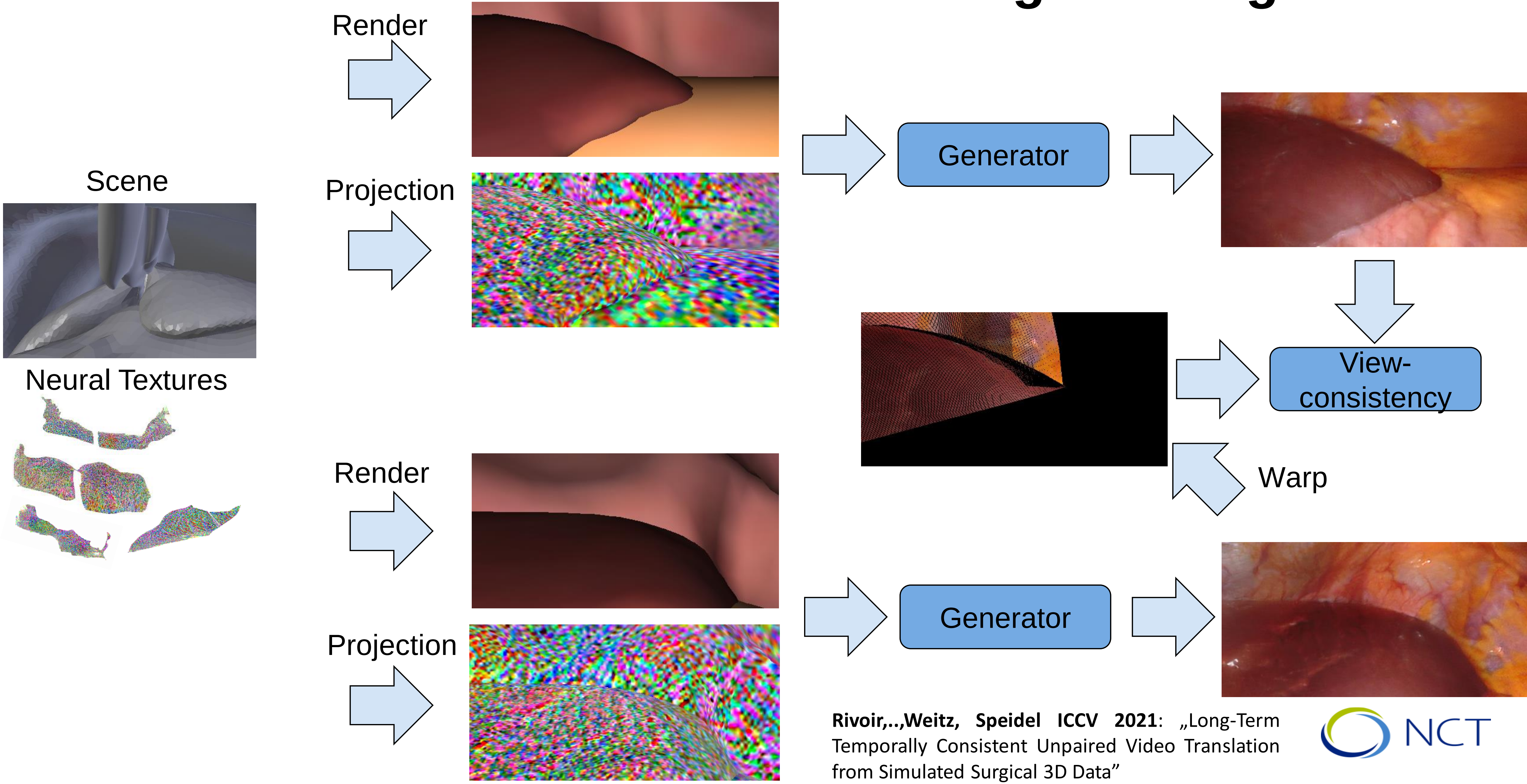
- Applying image to image translation to sequences can lead to temporal inconsistencies
- Enforce consistency via neural textures



Rivoir,..,Weitz, Speidel ICCV 2021: „Long-Term Temporally Consistent Unpaired Video Translation from Simulated Surgical 3D Data"

NCT

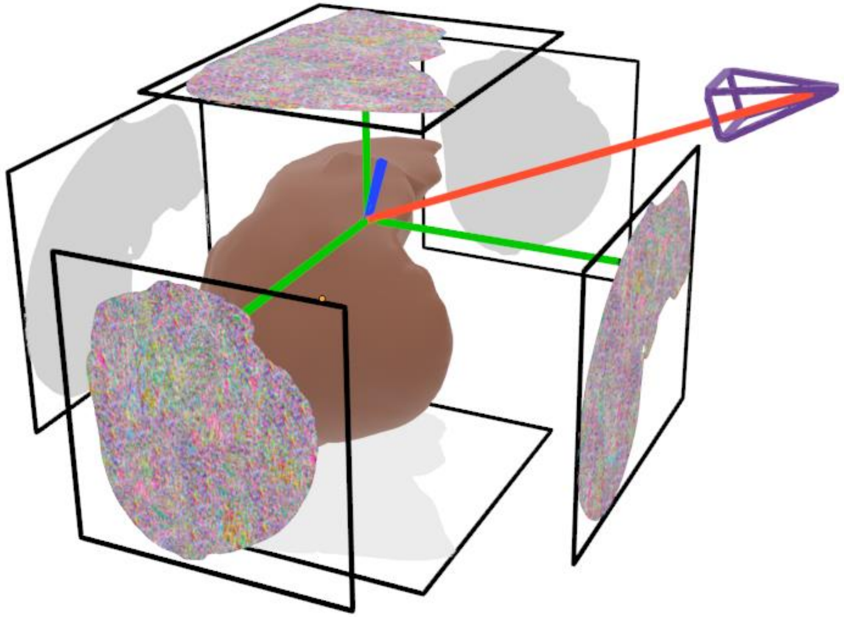
Simulation to video translation for surgical images

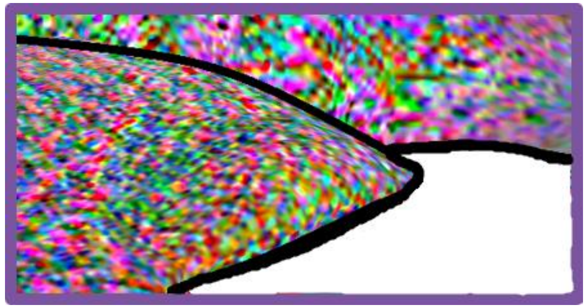


Rivoir,..,Weitz, Speidel ICCV 2021: „Long-Term Temporally Consistent Unpaired Video Translation from Simulated Surgical 3D Data"

NCT

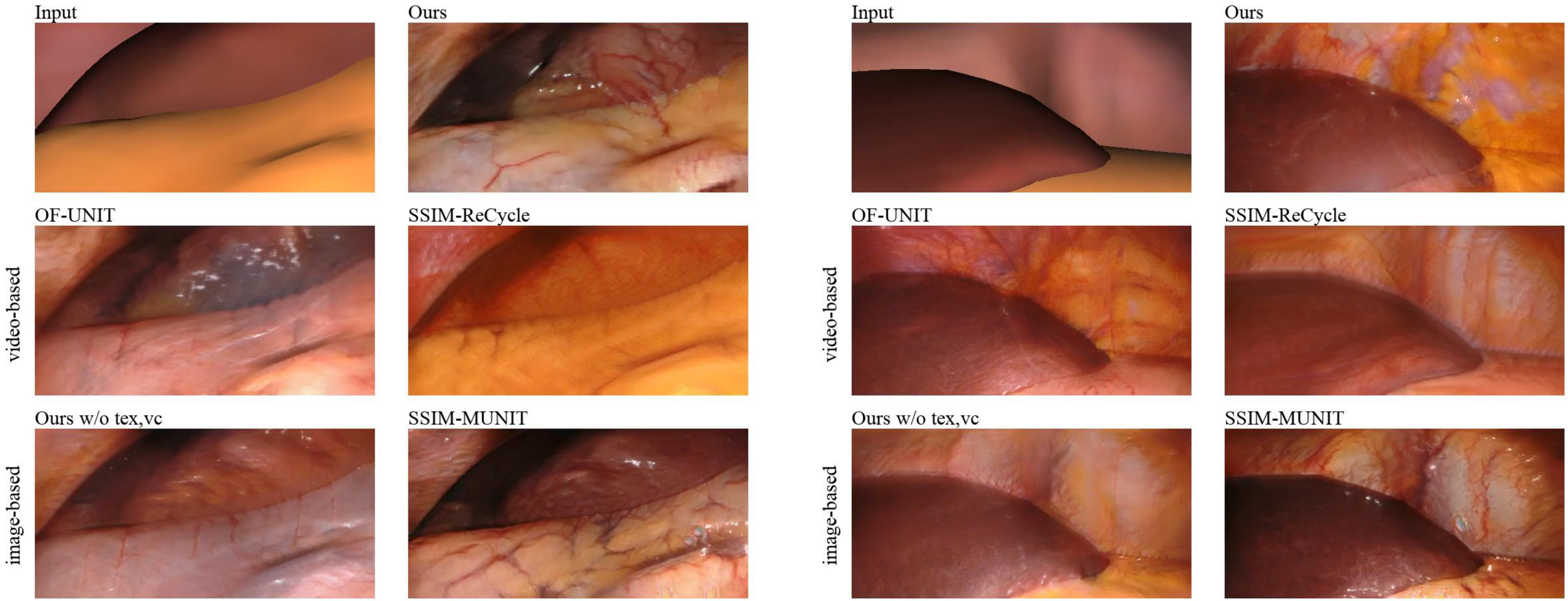
Simulation to video translation for surgical images





Rivoir,..,Weitz, Speidel ICCV 2021: „Long-Term Temporally Consistent Unpaired Video Translation from Simulated Surgical 3D Data”

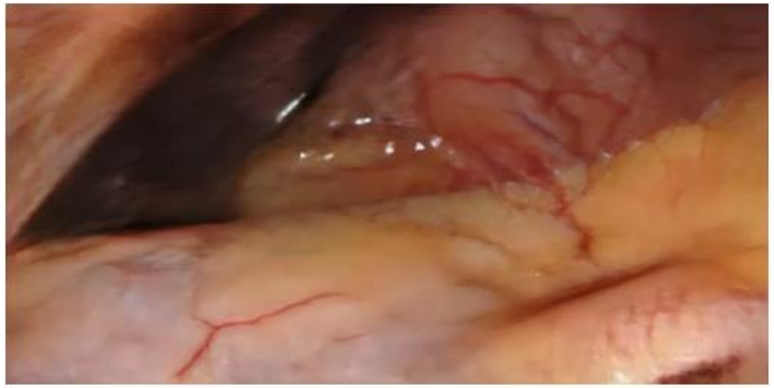
Simulation to video translation for surgical images

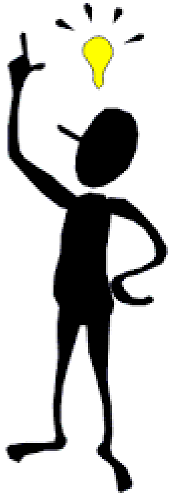


Rivoir,..,Weitz, Speidel ICCV 2021: „Long-Term Temporally Consistent Unpaired Video Translation from Simulated Surgical 3D Data"

Summary

- Introduction into the basics of GANs
  - How to train a GAN to generate images
- GANs for image to image transition
  - Paired
  - Unpaired
- GANs for simulation to video transition
  - Using neural textures





NCT
NATIONAL CENTER FOR TUMOR DISEASES
PARTNER SITE DRESDEN
UNIVERSITY CANCER CENTER UCC

Supported by:
German Cancer Research Center
University Hospital Carl Gustav Carus Dresden
Carl Gustav Carus Faculty of Medicine, TU Dresden
Helmholtz-Zentrum Dresden-Rossendorf

dkfz. DEUTSCHES KREBSFORSCHUNGSZENTRUM IN DER HELMHOLTZ-GEMEINSCHAFT

NEARDATA

CloudSkin

UniversitätsKlinikum Heidelberg

Contact: sebastian.bodenstedt@nct-dresden.de

Universitätsklinikum Carl Gustav Carus
DIE DRESDNER.

CeTI Centre for Tactile Internet with Human-in-the-Loop

Else Kröner Fresenius Stiftung

DFG Deutsche Forschungsgemeinschaft

Federal Ministry of Education and Research

Federal Ministry for Economic Affairs and Energy

Federal Ministry of Health

ESF
Chancen fördern
EUROPÄISCHER SOZIALFONDS IN BADEN-WÜRTTEMBERG

NCT