

AI for the Media

Week 7, CycleGAN + SuperRes



Today we will learn about a model capable of:



Overview

(pre-recorded lecture)

Domain to domain models, CycleGANs and Super resolution:

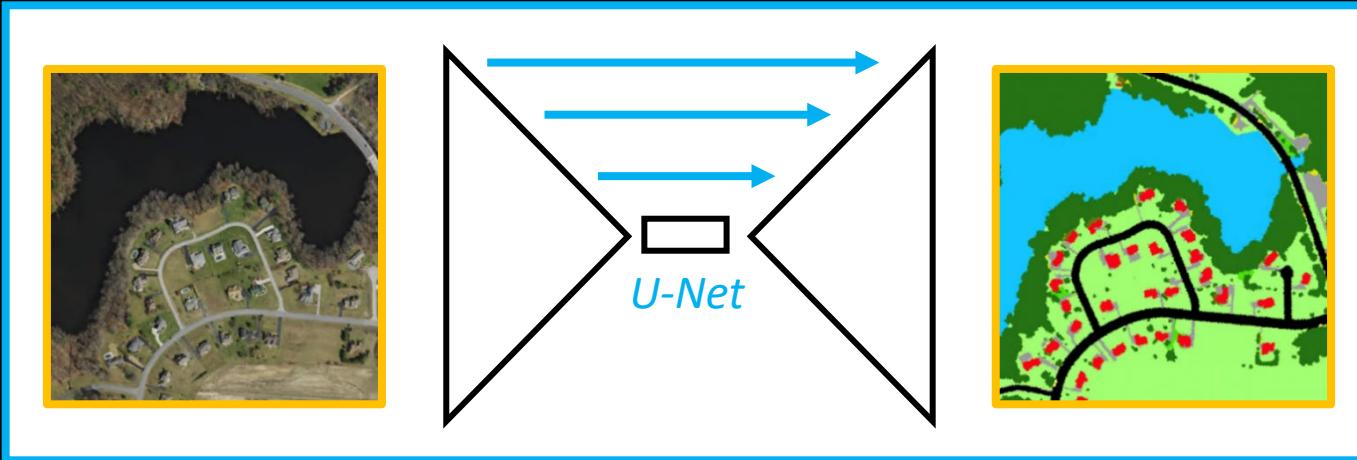
- **CycleGAN** as a simple extension to pix2pix
- **SuperResolution**
 - As a tool to increase quality of generated images with some cost involved
- **Limitations and follow-up readings**

Practical session (*during the live session*):

- **Code:** Trying a CycleGAN model
- **Demo:** Easy SuperResolution tools to use

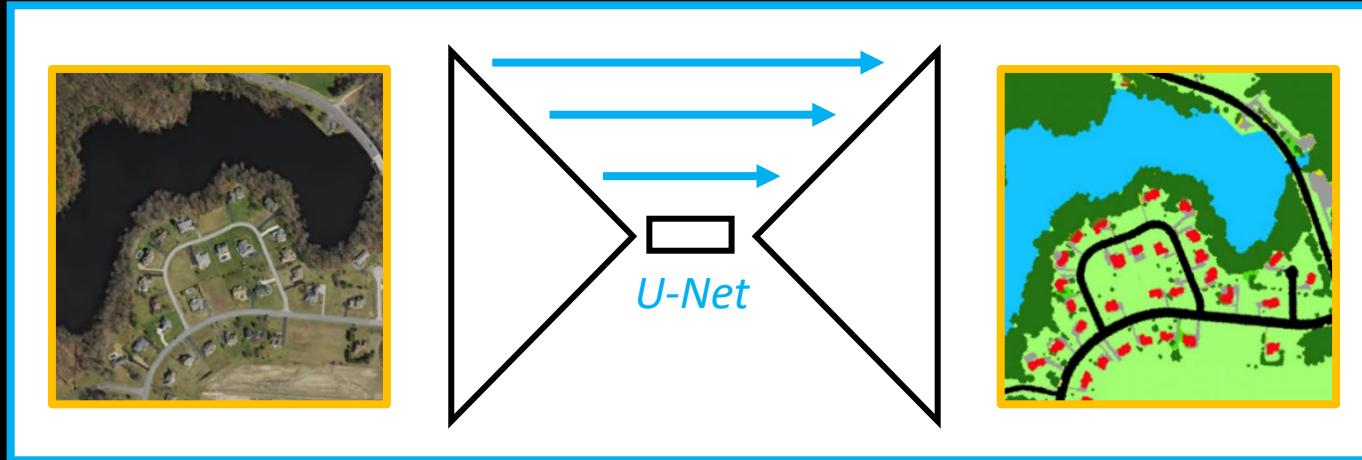
Recap

U-Net



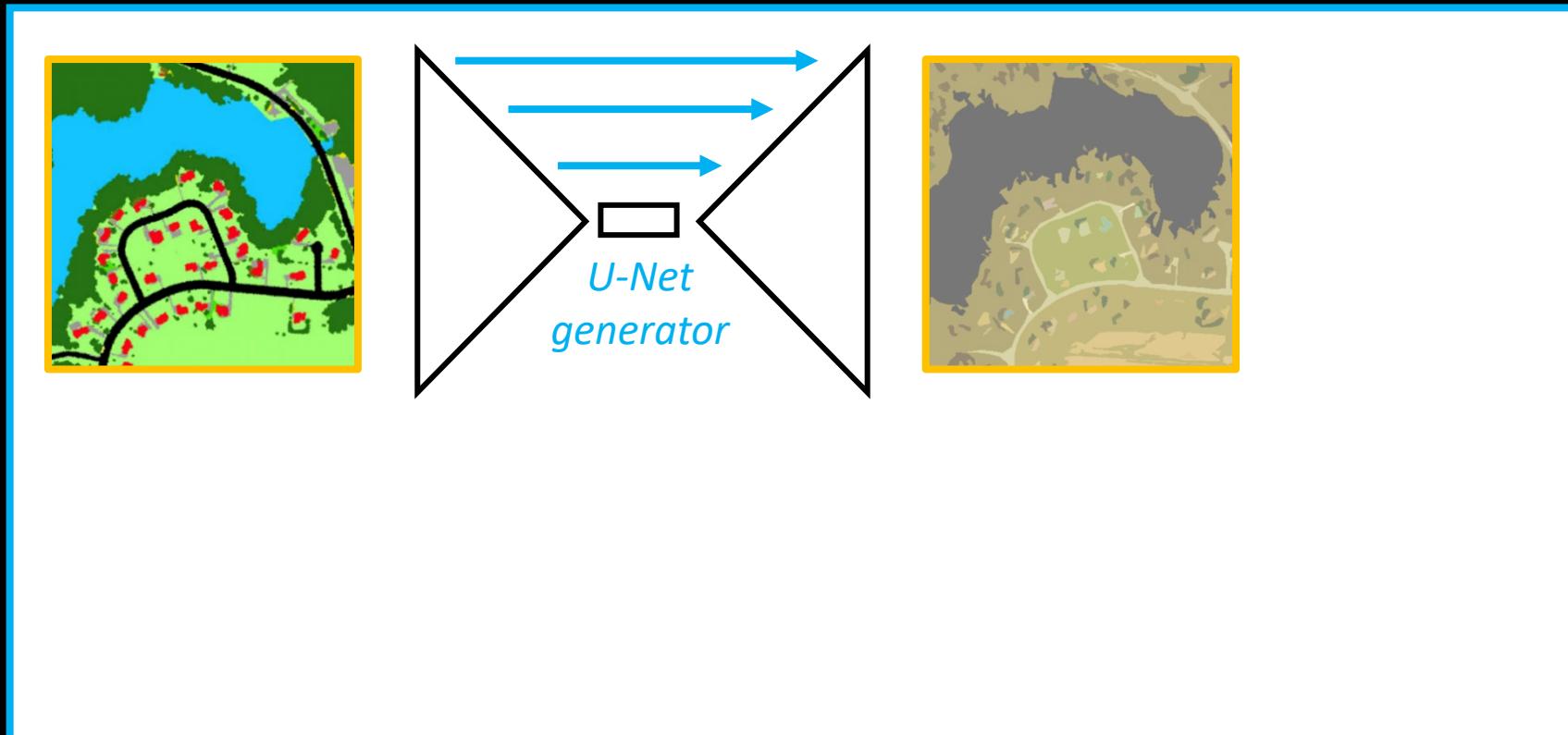
- Typical model for **semantic segmentation**

U-Net

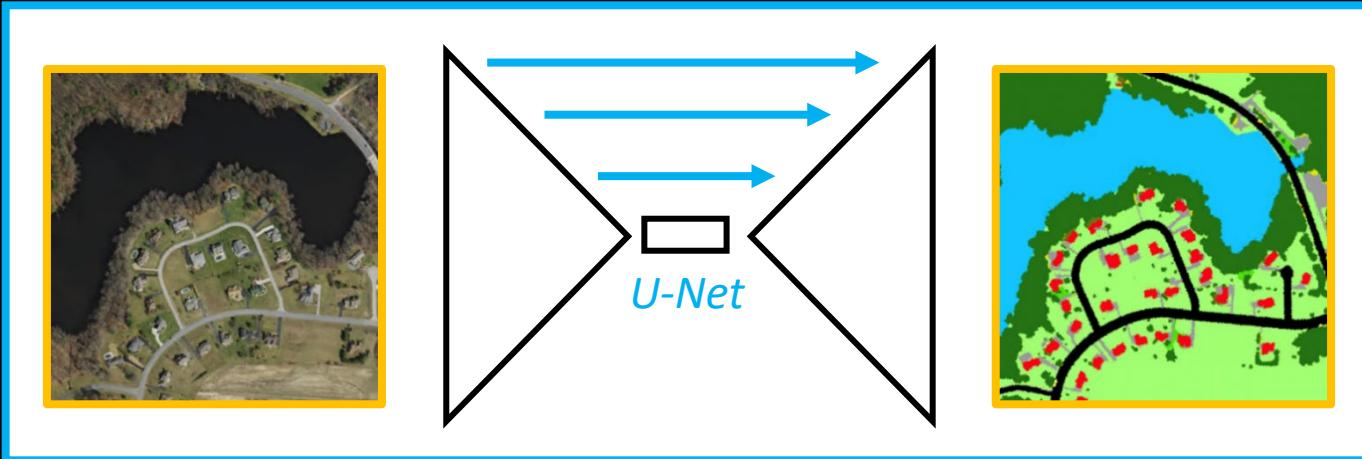


- Typical model for **semantic segmentation**

pix2pix

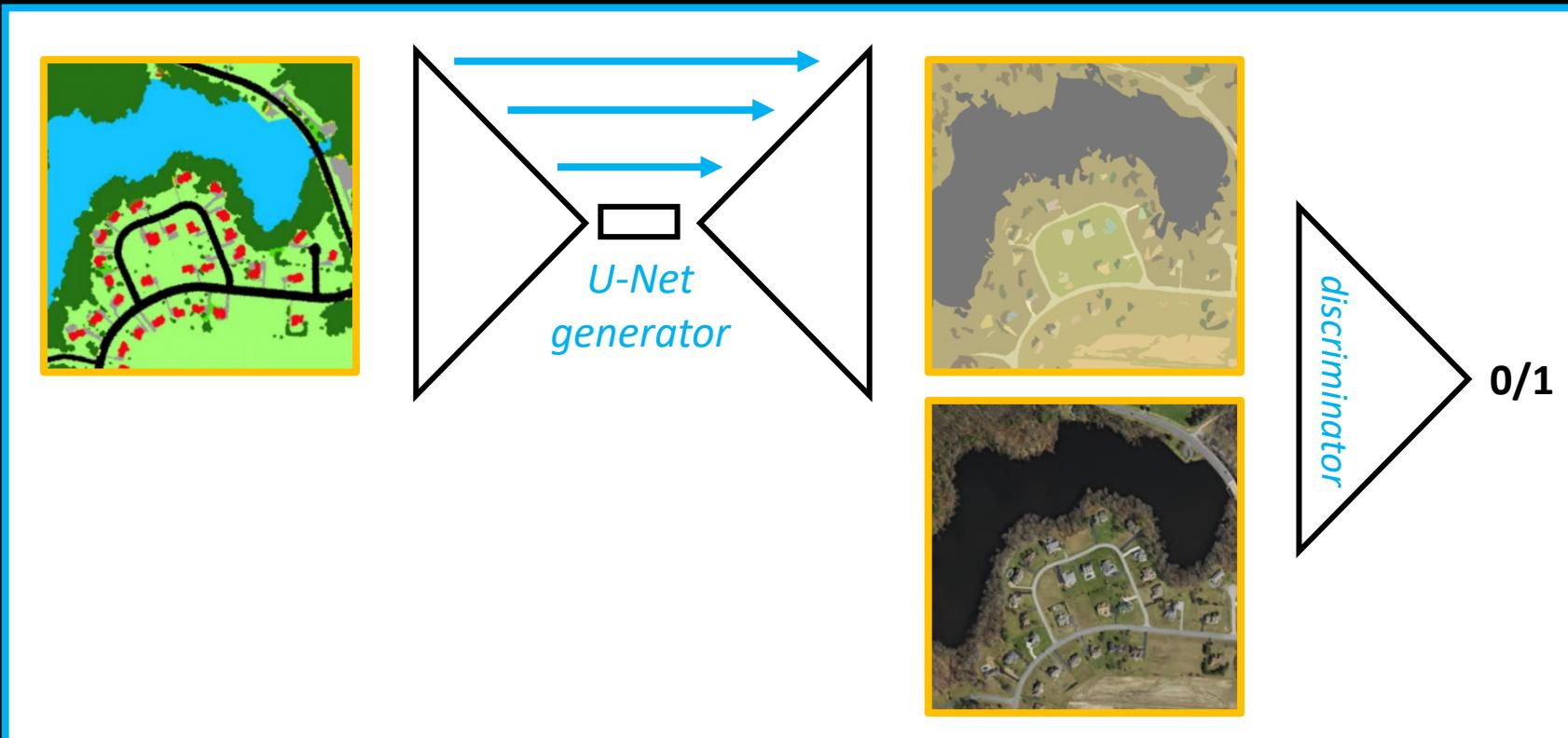


U-Net



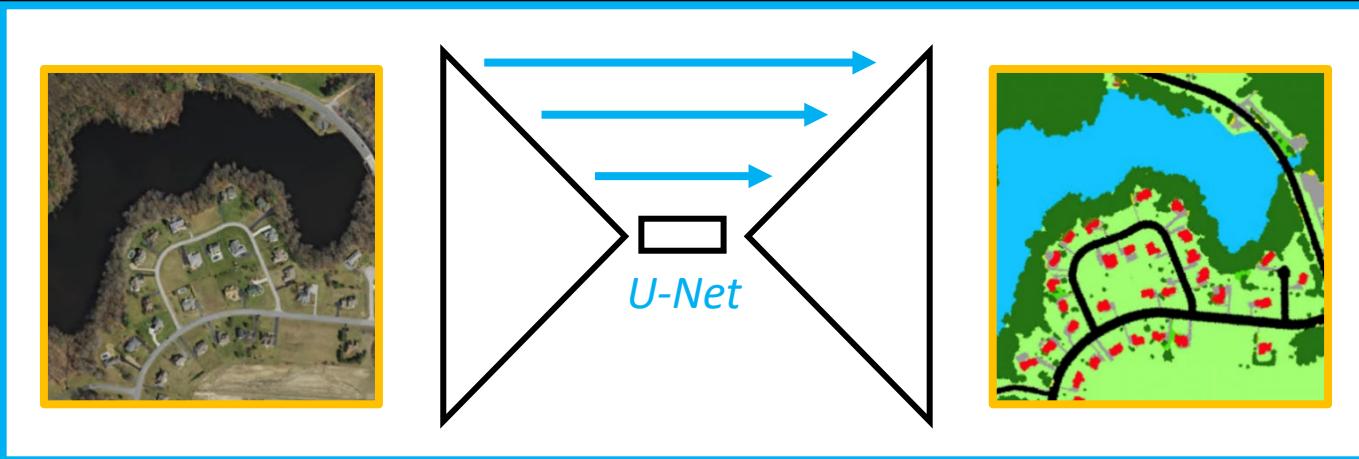
- Typical model for **semantic segmentation**

pix2pix



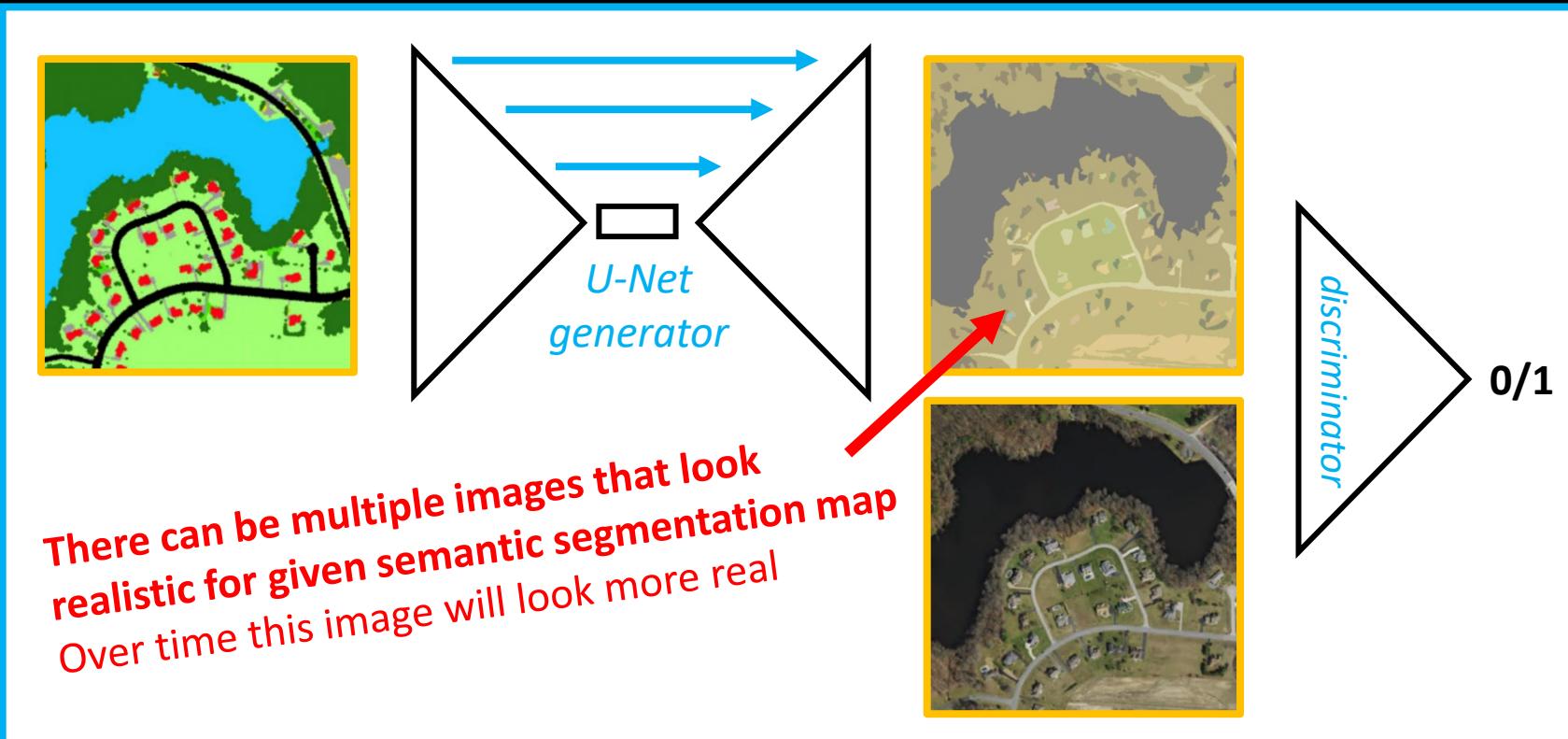
- Typical model when the output has **multiple possible realities**

U-Net



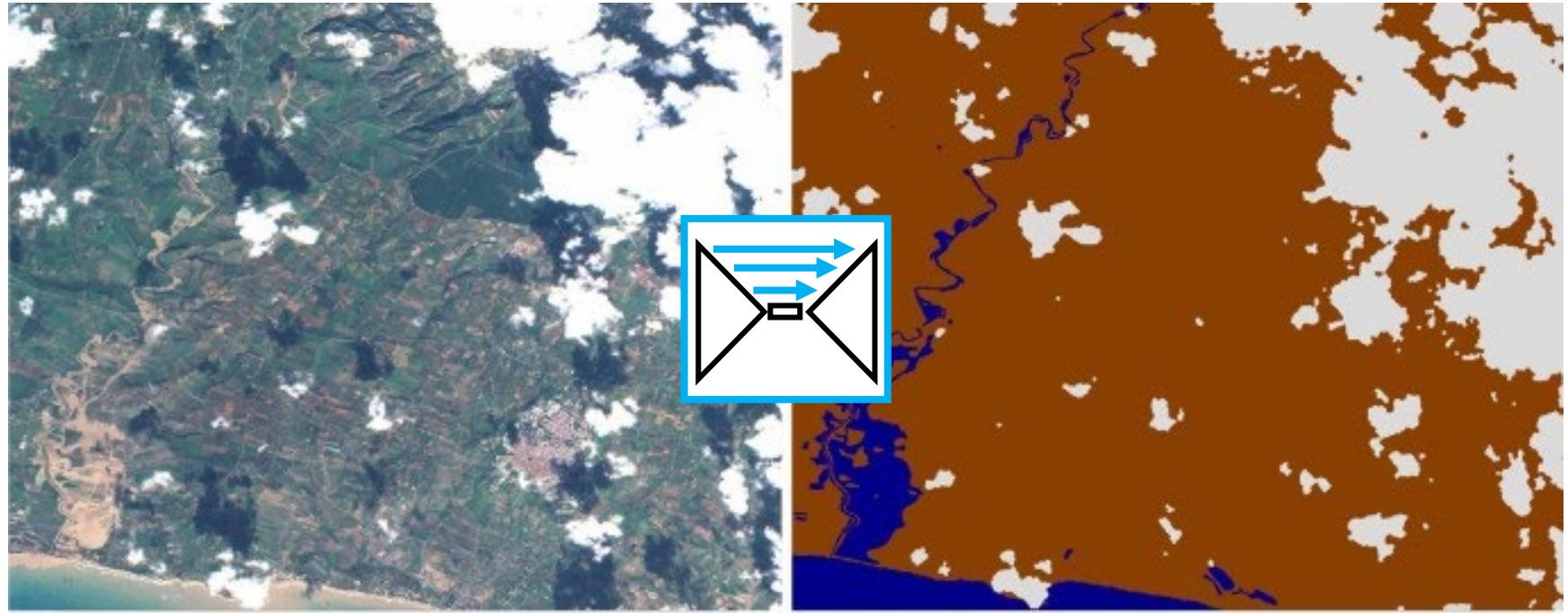
- Typical model for **semantic segmentation**

pix2pix



- Typical model when the output has **multiple possible realities**

U-Net: Real-world example

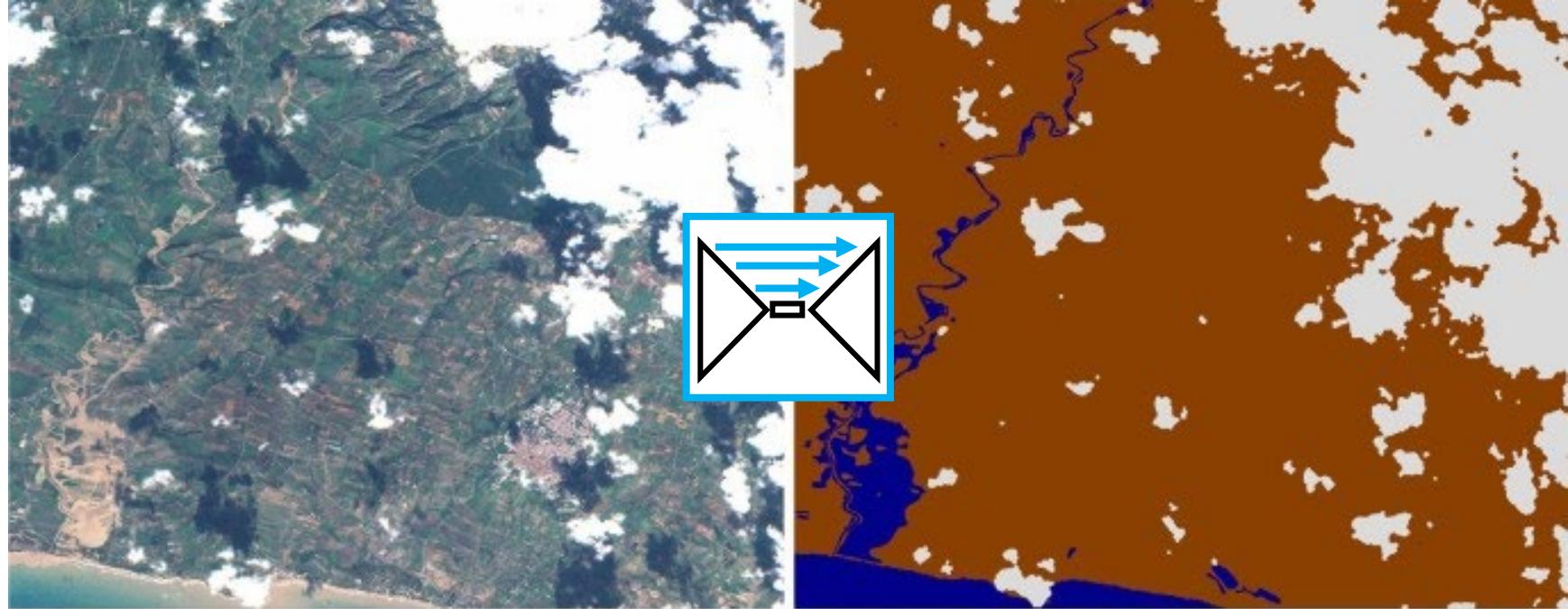


*Nature paper: Towards global **flood mapping** onboard **low cost satellites** with ML*

U-Net: Real-world example



radiation resistant HW



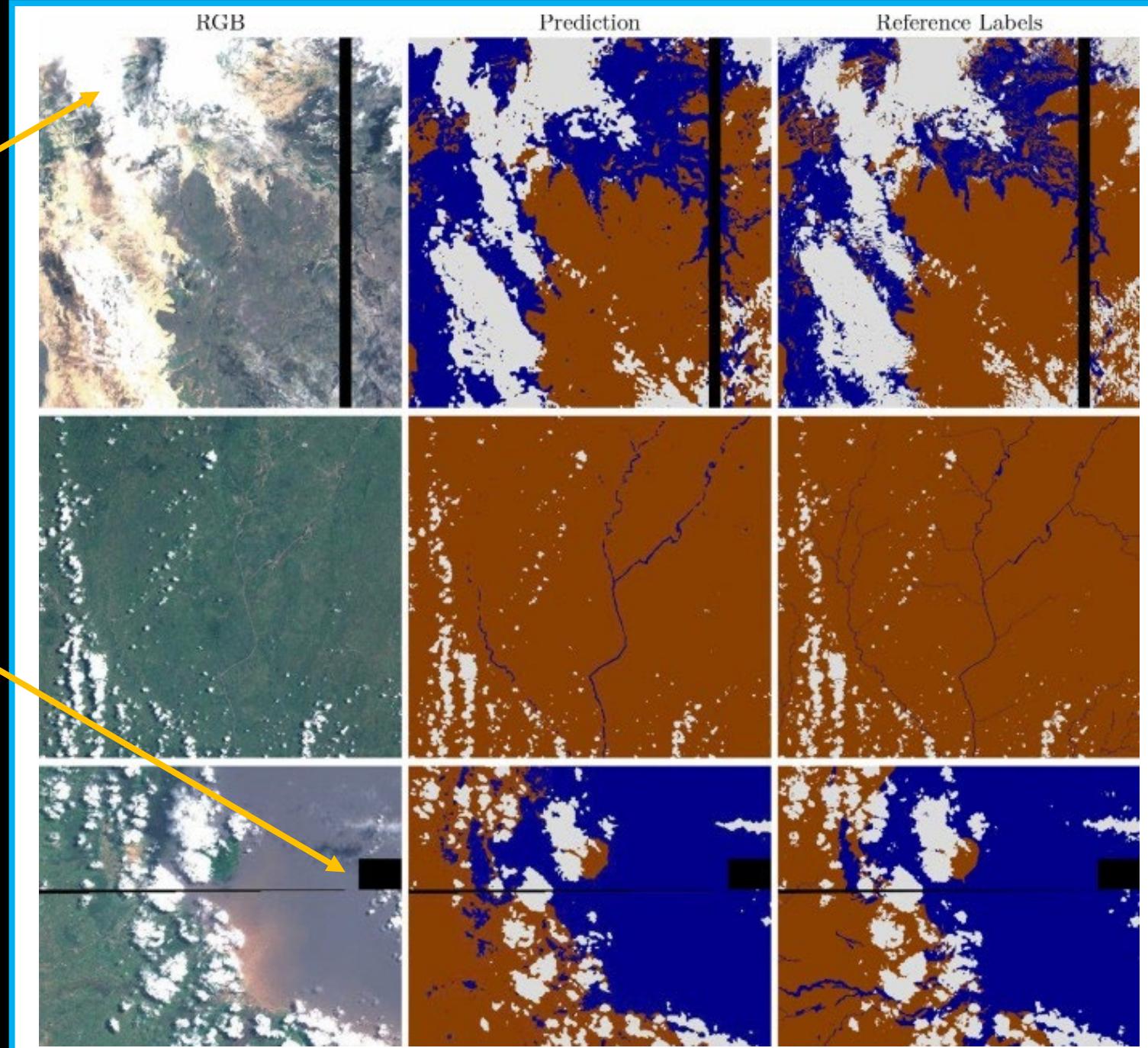
*Nature paper: Towards global **flood mapping**
onboard low cost satellites with ML*

Real-world is messy!

- Clouds everywhere
- Erroneous data, Noisy



- Limited viewing capability



Topic I: CycleGANs

Introduction

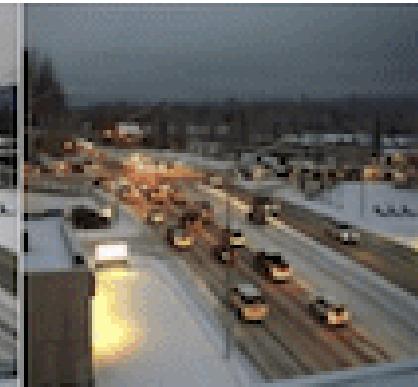
Motivation

- We can't always get paired data!

- Hard to collect pairs
 - horse2zebra



- We want it to work with new unseen samples
 - day2night
 - (*also not exactly pairs*)



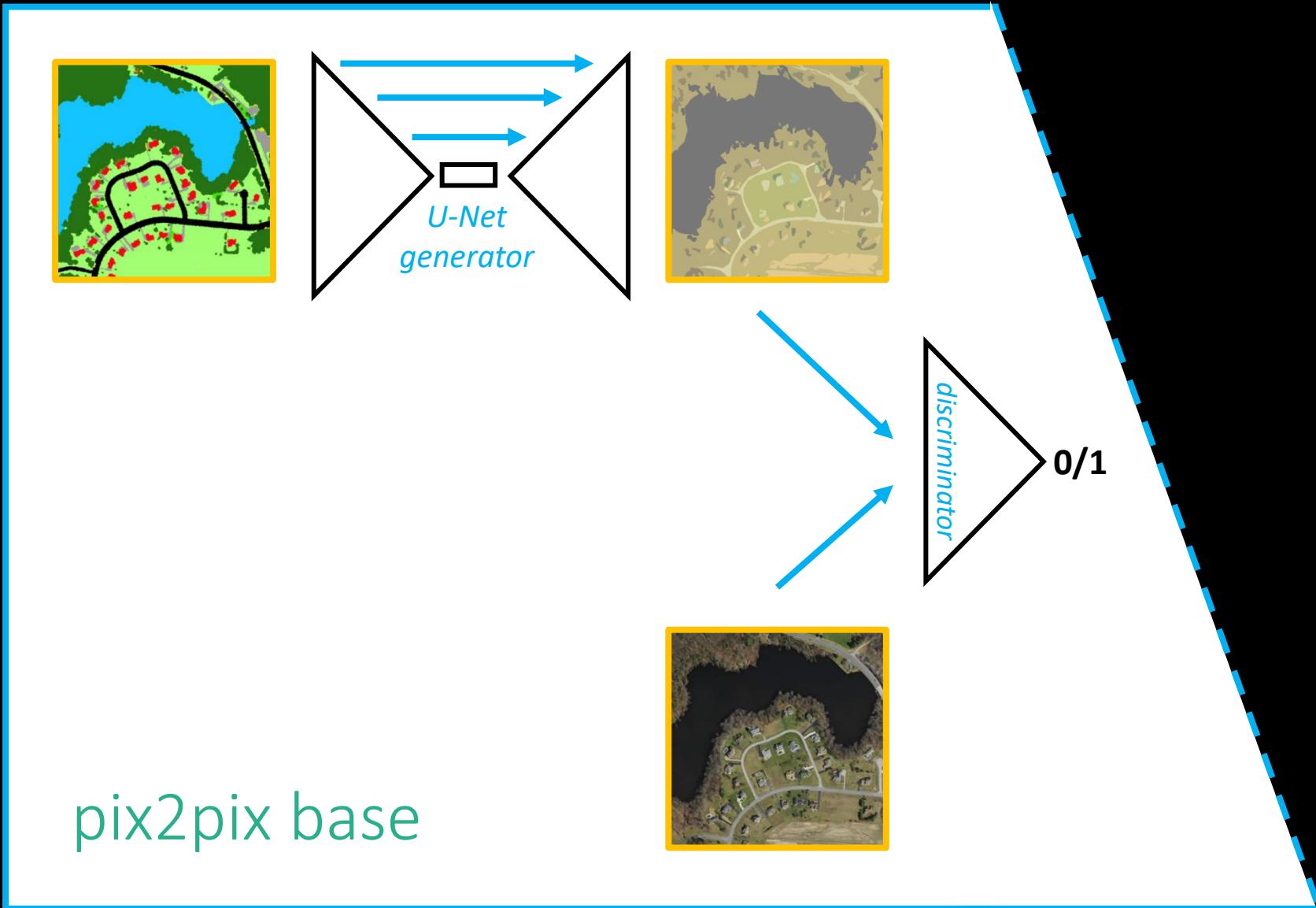
- When accessing archive data, we may have incompatible recordings
 - Optical2sar
 - (*same location, but different sensor type*)



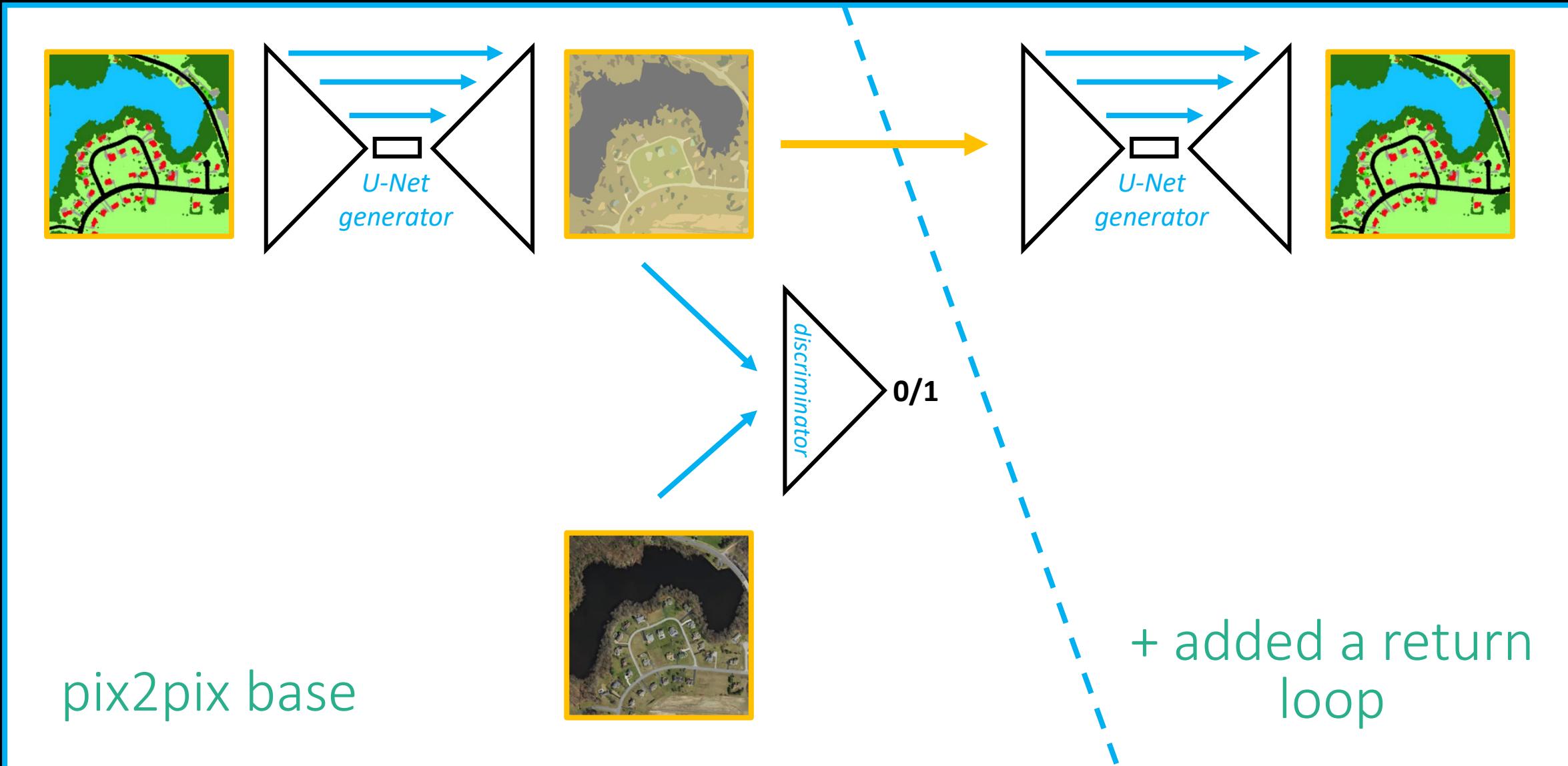
13 channels

8 channels

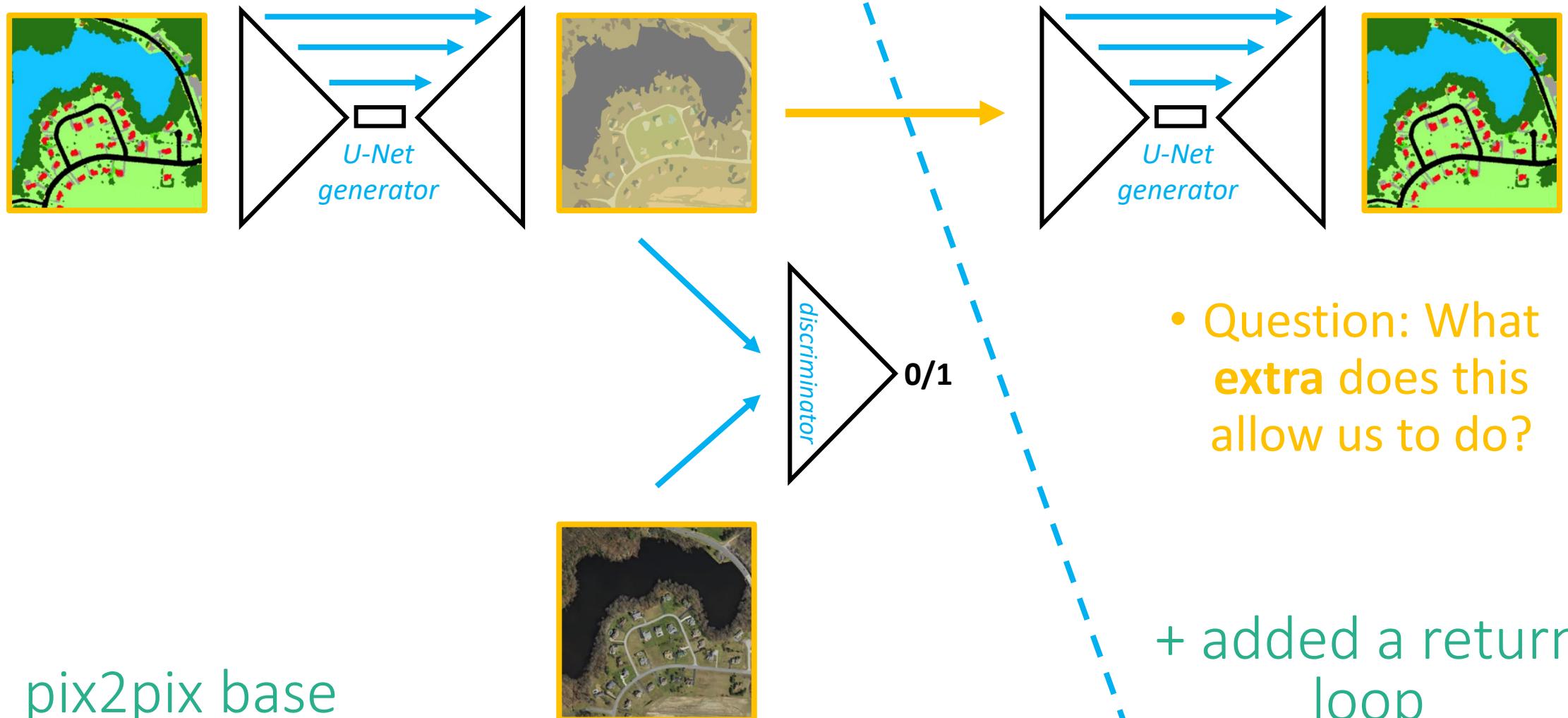
CycleGAN



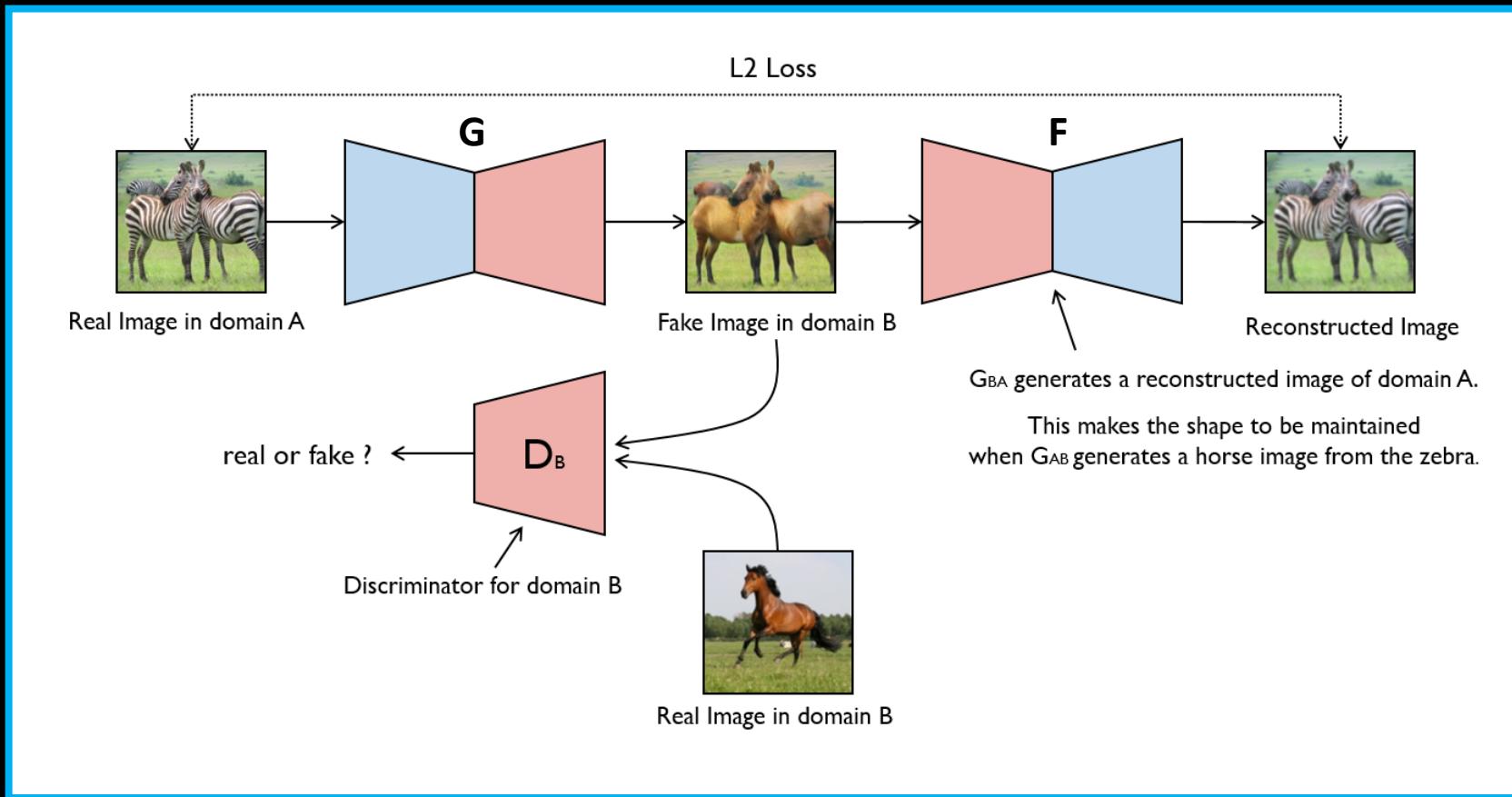
CycleGAN



CycleGAN

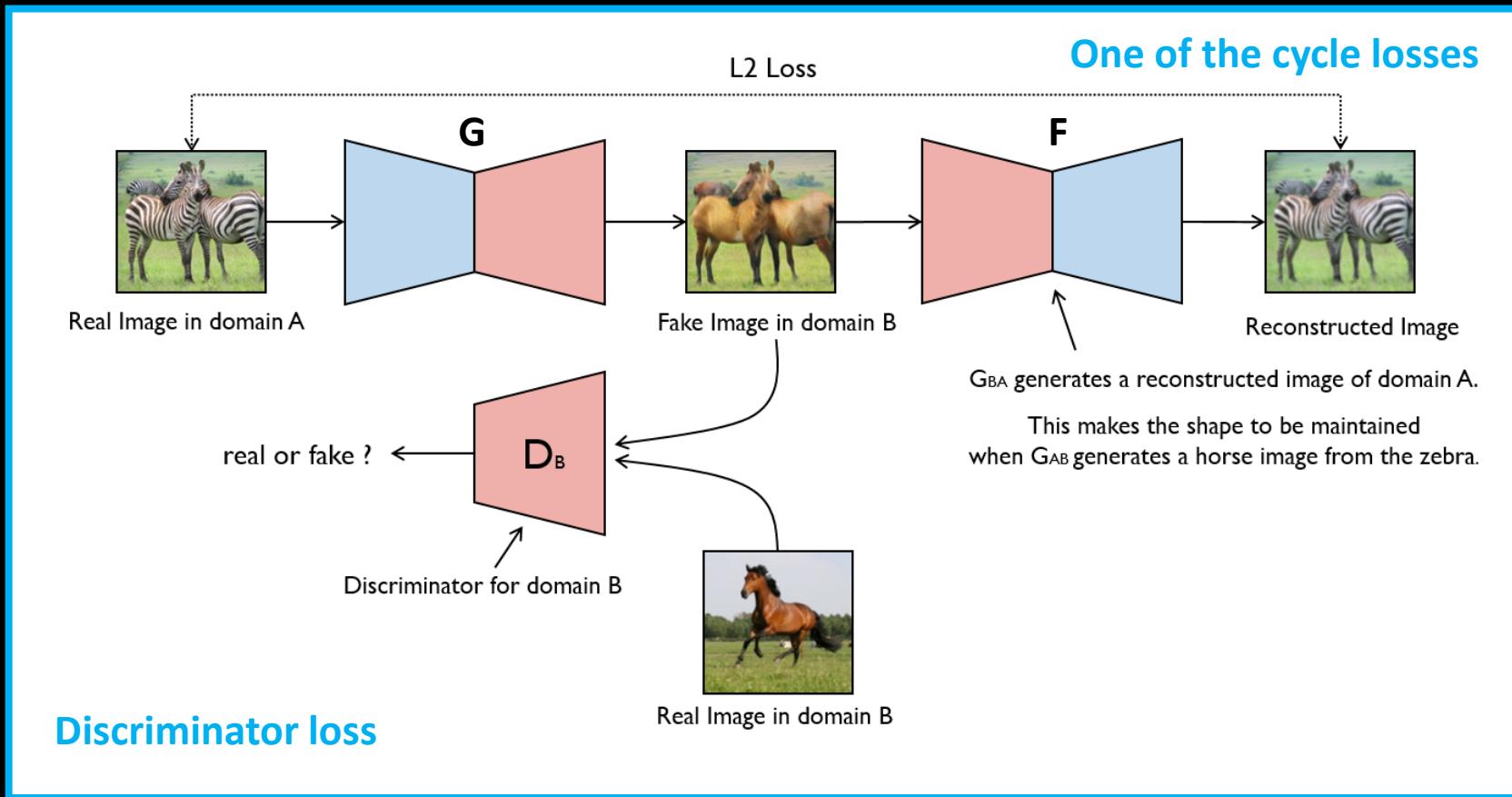


CycleGAN, two generators



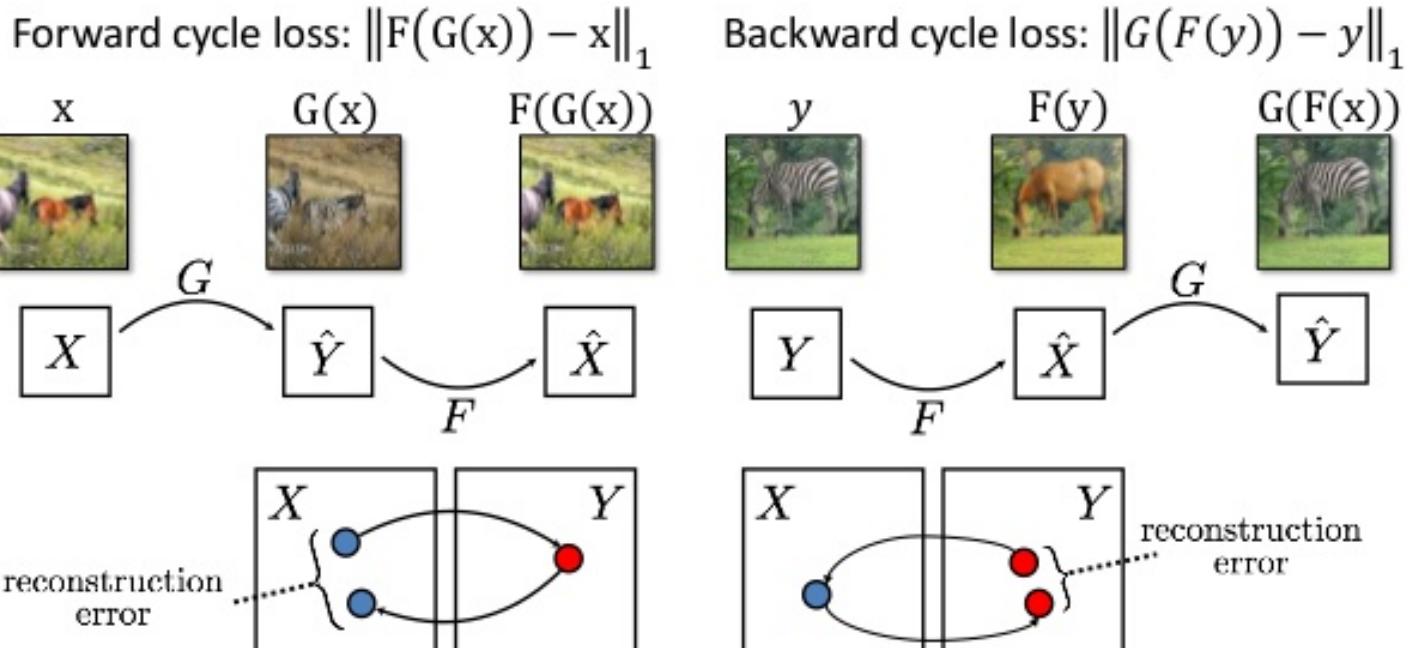
- Setup with two generators going between two domains (**G** and **F**) and a special cyclic loss function. This model **doesn't require paired data** in the training datasets.

CycleGAN, two generators



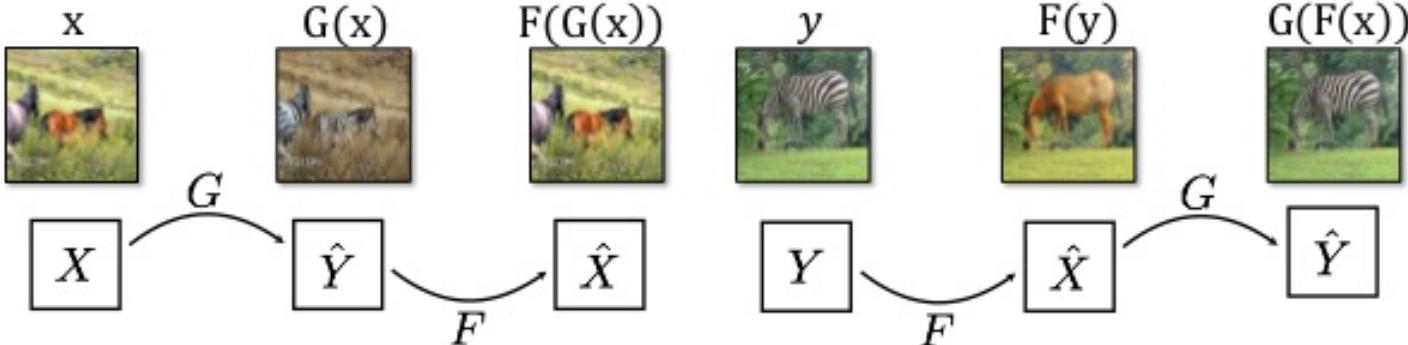
- Setup with two generators going between two domains (**G** and **F**) and a special cyclic loss function. This model **doesn't require paired data** in the training datasets.

CycleGAN, loss functions



CycleGAN, loss functions

Forward cycle loss: $\|F(G(x)) - x\|_1$ Backward cycle loss: $\|G(F(y)) - y\|_1$



reconstruction
error

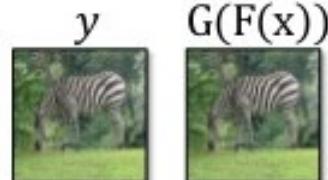
reconstruction
error

- We call these **cyclic loss functions**

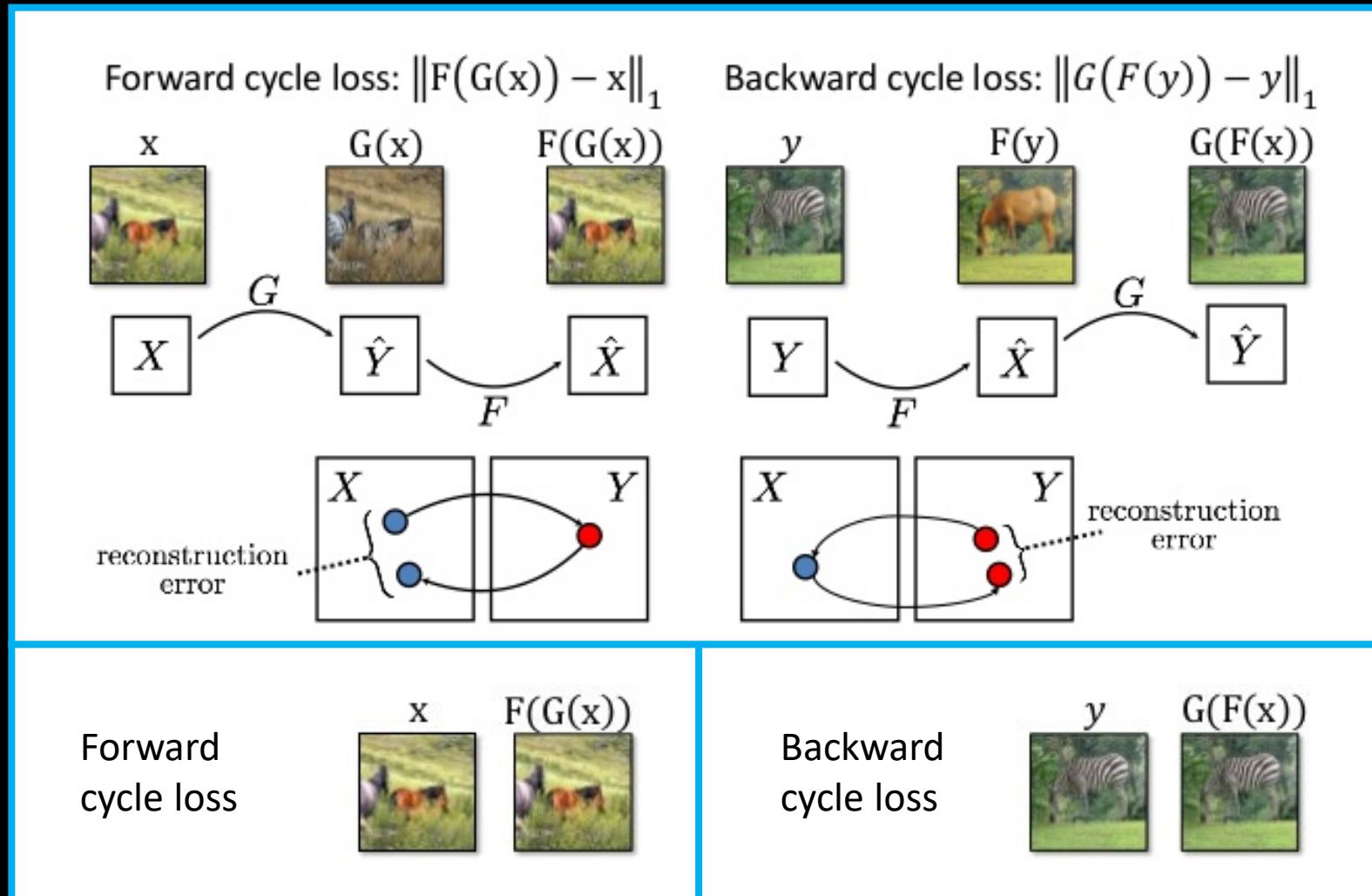
Forward
cycle loss



Backward
cycle loss



CycleGAN, loss functions



- We call these **cyclic loss functions**

- Alongside the original **discriminator loss functions** (to prevent just learning identity)

Topic I: CycleGANs

Examples

CycleGAN: Real-world example



- Models on autonomous cars usually trained on clean data
- When they encounter night images, the accuracy goes down!



CycleGAN: Real-world example



- Models on autonomous cars usually trained on clean data
- When they encounter night images, the accuracy goes down!

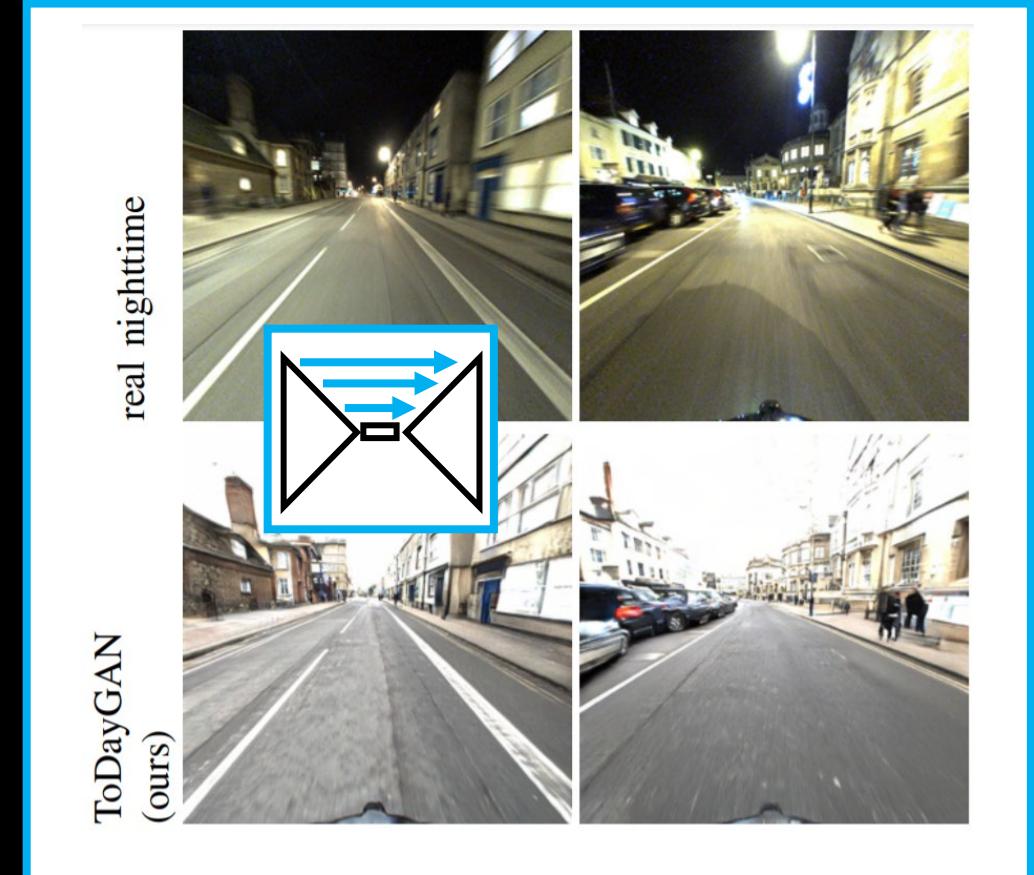


- We want something like a **night2day model**
- We **can't collect** photos from everywhere from as {day, night} **pairs**

CycleGAN: Real-world example



- Models on autonomous cars usually trained on clean data
- When they encounter night images, the accuracy goes down!



Night 2 Day, followed by more tasks

CycleGAN: Real-world example

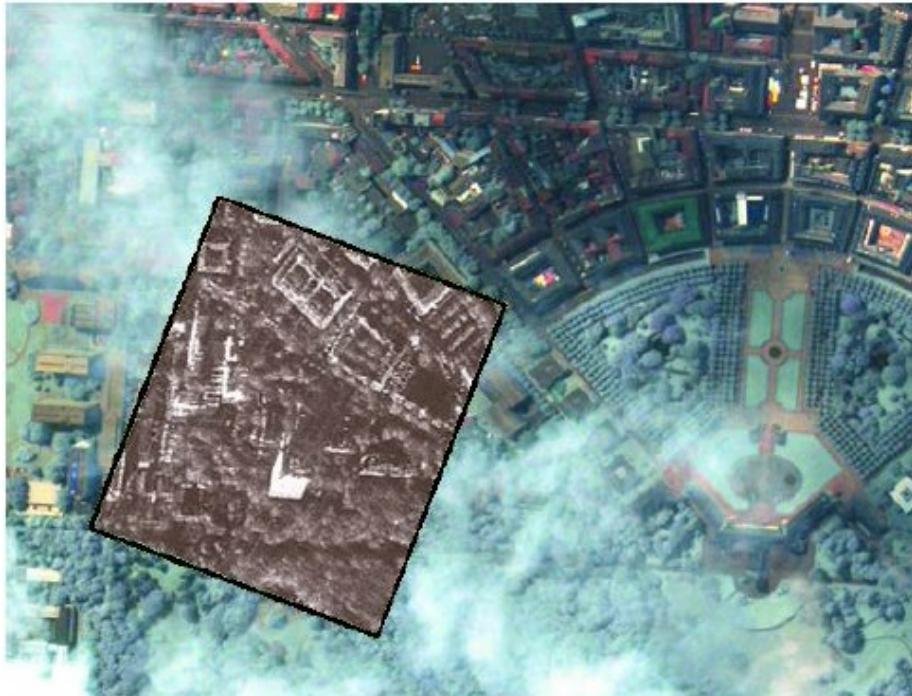
- **Change detection** with remote sensing images:



- But what if these two images come from different types of satellites?

CycleGAN: Real-world example

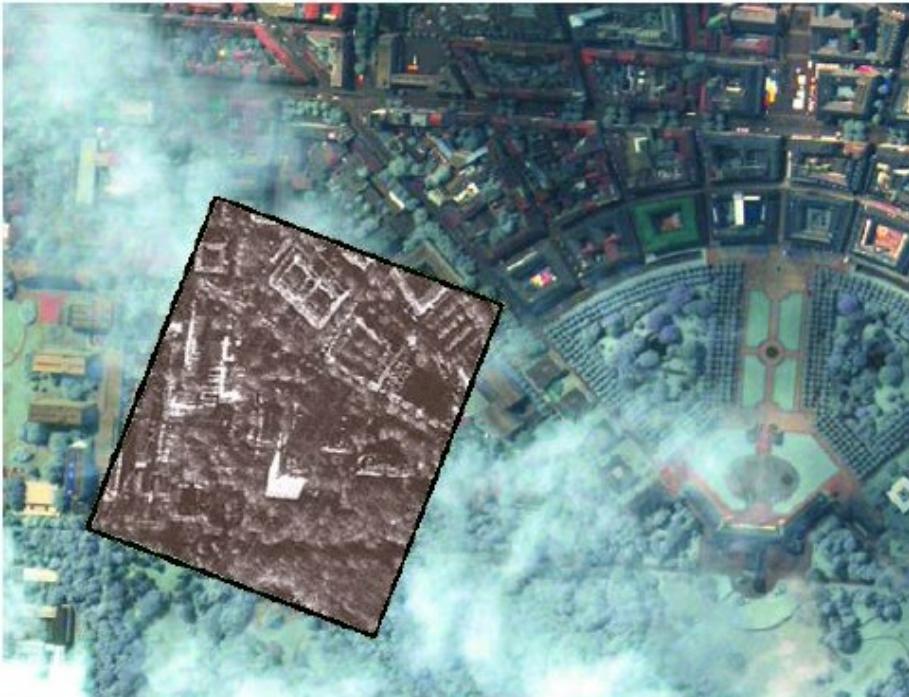
- If the location is covered by clouds, only so-called SAR satellites can see through:



Q: Can we still compare with these?

CycleGAN: Real-world example

- If the location is covered by clouds, only so-called SAR satellites can see through:



Q: Can we still compare with these?

Hard to compare
between these



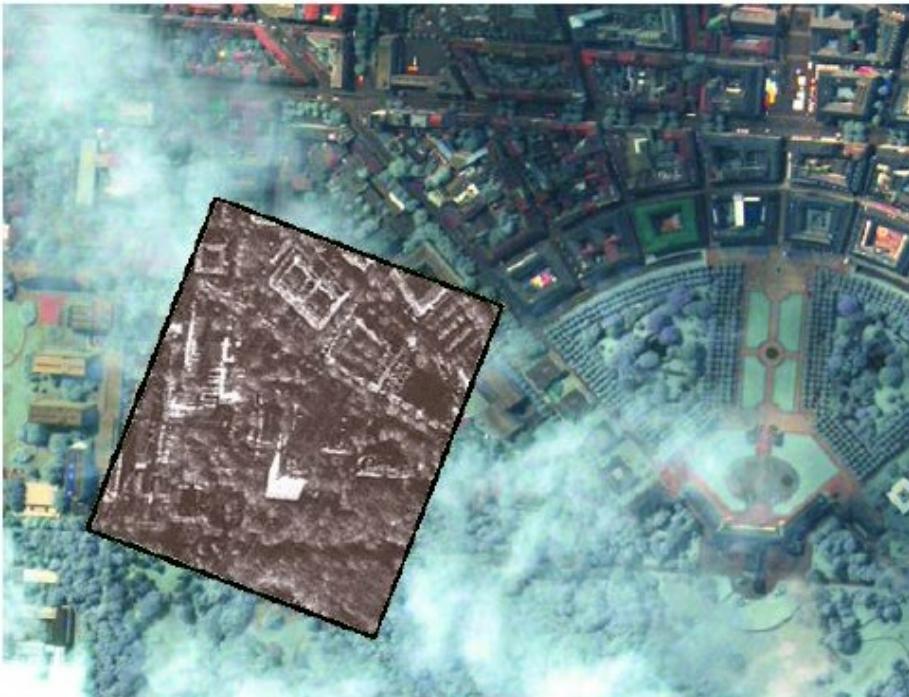
Image at T1



Image at T2

CycleGAN: Real-world example

- If the location is covered by clouds, only so-called SAR satellites can see through:



Q: Can we still compare with these?

Image at T1



Then we can compare them:

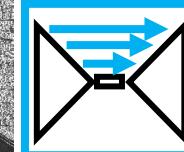
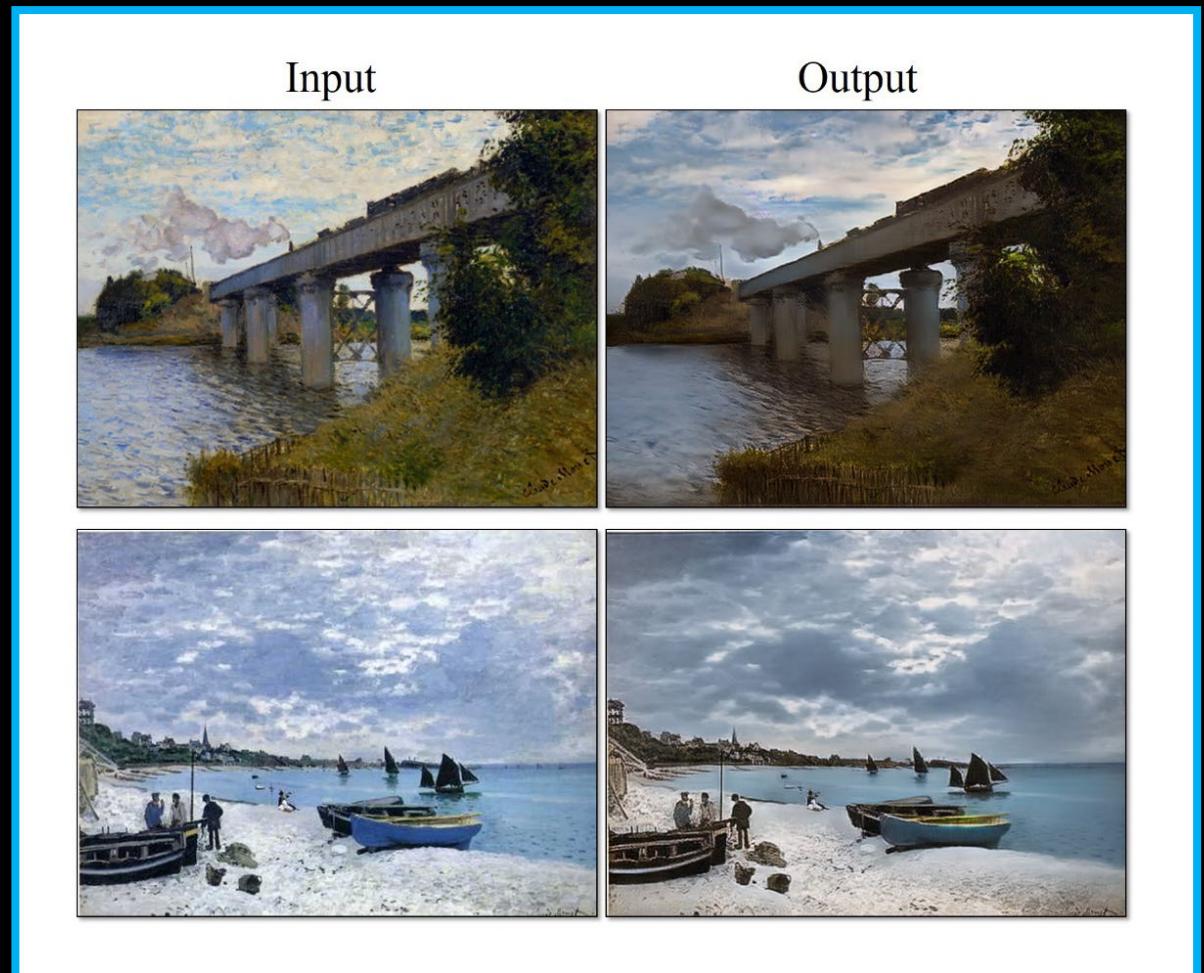


Image at T2

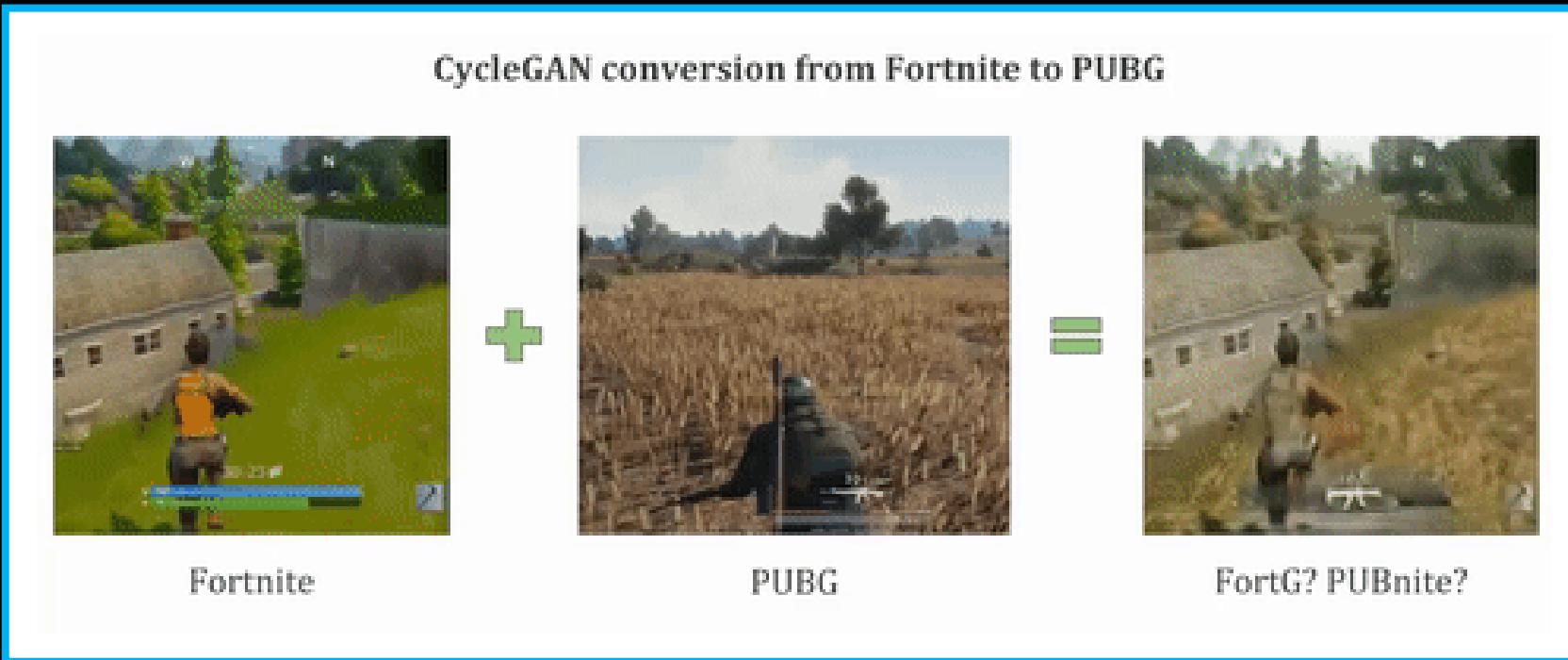
Image at T2
CycleGANed

CycleGAN: Art examples

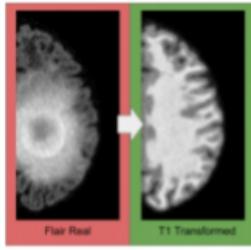
Monet 2 Real,
and more examples at
junyanz.github.io/CycleGAN/



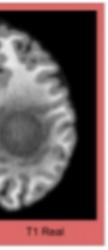
CycleGAN: Art examples



Game 2 Game, Fortnite Into PUBG
youtu.be/xkLtgwWxrec



(a) A translation removing tumors



(b) A translation adding tumors

to a sky to make it look like it was painted by Van Gogh, it may add tumors in medical images where none exist, or remove those that do. More information on dangers like this can be found in **Cohen et al.**

Resurrecting Ancient Cities



Jack Clark used our code to convert ancient maps of **Babylon**, **Jerusalem** and **London** into modern Google Maps and satellite views.

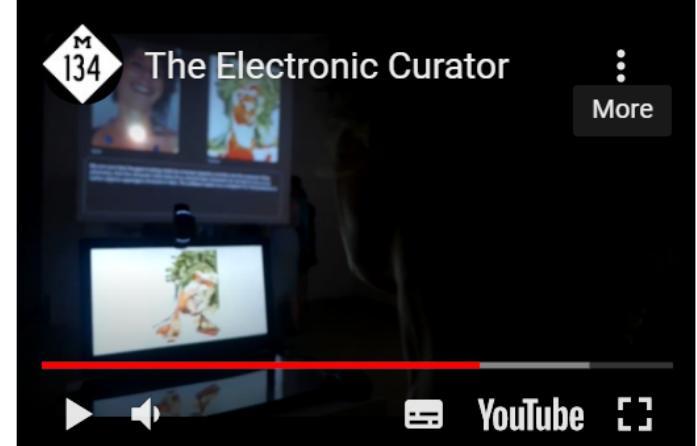
How to interpret CycleGAN results: CycleGAN, as well as any GAN-based method, is fundamentally hallucinating part of the content it creates. Its outputs are predictions of "what might it look like if ..." and the predictions, thought plausible, may largely differ from the ground truth. **CycleGAN should only be used with great care and calibration in domains where critical decisions are to be taken based on its output.** This is especially true in medical applications, such as translating MRI to CT data. Just as CycleGAN may add fanciful clouds

Face ↔ Ramen



Takuya Kato performed a magical and hilarious Face ↔ Ramen translation with CycleGAN. Check out more results [here](#)

The Electronic Curator



Eran Hadas and Eyal Gruss used CycleGAN to convert human faces into vegetable portraits. They built a real-time art demo which allows users to interact with the model with their own faces.

Topic I: CycleGANs

Limitations

Less control

... on which data gets transferred
between the two domains >



Failure Cases



Our model does not work well when a test image looks unusual compared to training images, as shown in the left figure. See more typical failure cases [\[here\]](#). On translation tasks that involve color and texture changes, as many of those reported above, the method often succeeds. We have also explored tasks that require geometric changes, with little success. For example, on the task of dog \leftrightarrow cat transfiguration, the learned translation degenerates into making minimal changes to the input. Handling more varied and extreme transformations, especially geometric changes, is an important problem for future work. We also observe a lingering gap between the results achievable with paired training data and those achieved by our unpaired method. In some cases, this gap may be very hard -- or even impossible -- to close: for example, our method sometimes permutes the labels for tree and building in the output of the cityscapes photos \rightarrow labels task. Resolving this ambiguity may require some form of weak semantic supervision. Integrating weak or semi-supervised data may lead to substantially more powerful translators, still at a fraction of the annotation cost of the fully-supervised systems.

weak semantic supervision. Integrating weak or semi-supervised data may lead to substantially more powerful translators, still at a fraction of the annotation cost of the fully-supervised systems.

Steganography



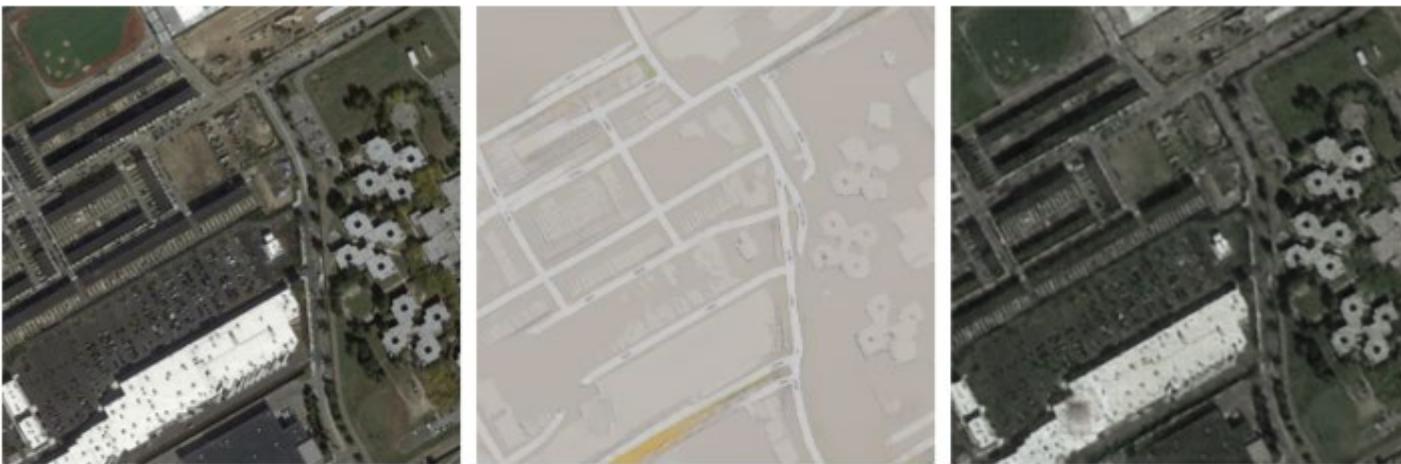
No message



Eagle has landed!

Steganography = encoding secret information into an image

CycleGAN steganography

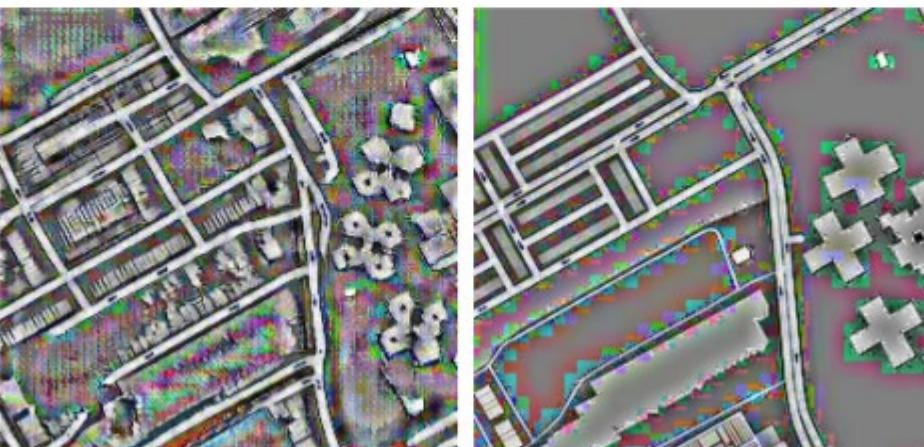


(a) Aerial photograph: x .

(b) Generated map: Fx .

(c) Aerial reconstruction: GFx .

Figure 1: Details in x are reconstructed in GFx , despite not appearing in the intermediate map Fx .



(a) Generated map.

(b) Training map, for comparison.

Figure 2: Maps with details amplified by adaptive histogram equalization. Information is present in the generated map even in regions that appear empty to the naked eye.

- Beware of what we are actually learning with the two generators ...

- **Steganography** = encoding secret information into an image (*such as the recipe of what to reconstruct*)

Topic II: SuperResolution

Introduction

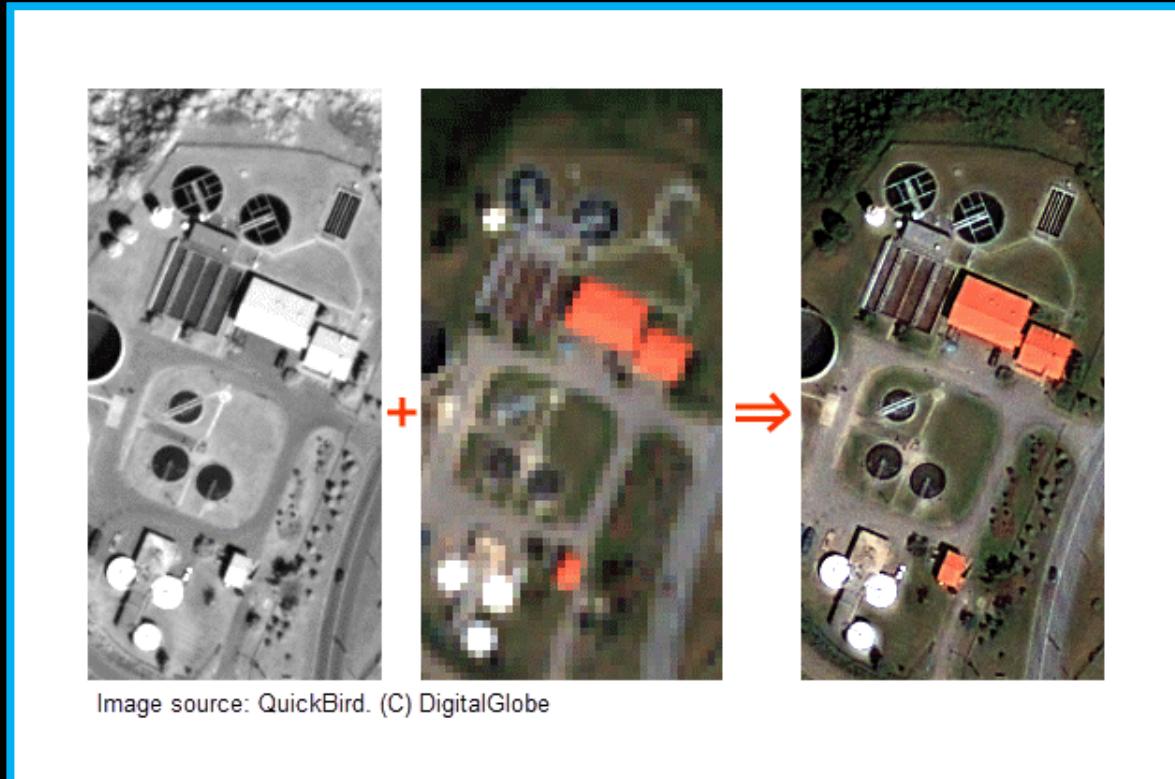
SuperResolution

- **Task motivation**
 - We saw that we can transfer from one domain to another domain, even if the data isn't paired ...
 - **Intuition:** Could we consider *low resolution images* and *high resolution images* as two domains?
 - Would we benefit from having them paired (converting HQ image into LQ is easy!)

(inspiration from the past)

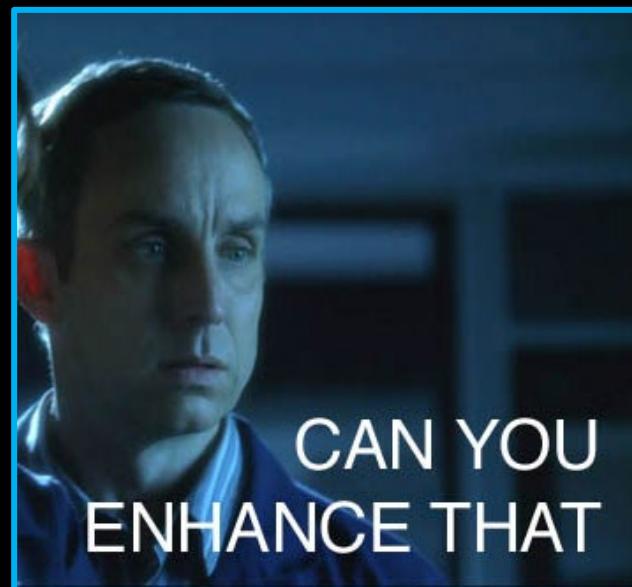
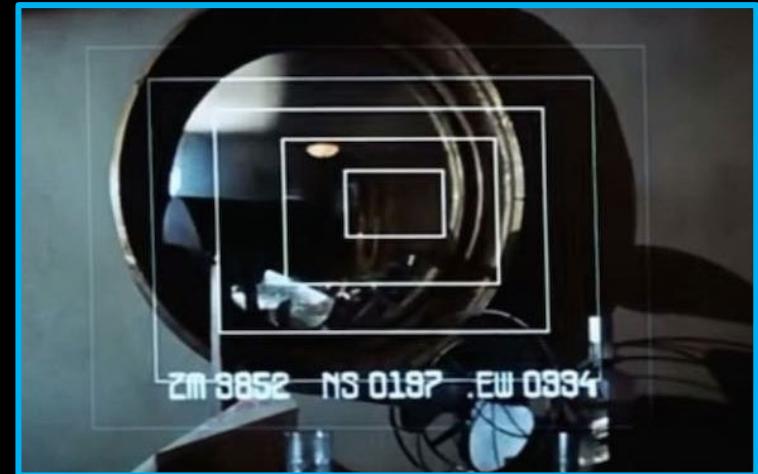
Pan sharpening

- Similar situation occurs when we have data of mixed resolutions ...



- High resolution black and white photo + low resolution colour photo
- Pan sharpening is a technique where we combine the detail of the BW photo with additional information of the colour image
- **Old technique: fully manual!**

Zoom and enhance!



SuperResolution CNN (SRCNN)

paper 2015

LR

Conv Net

HR

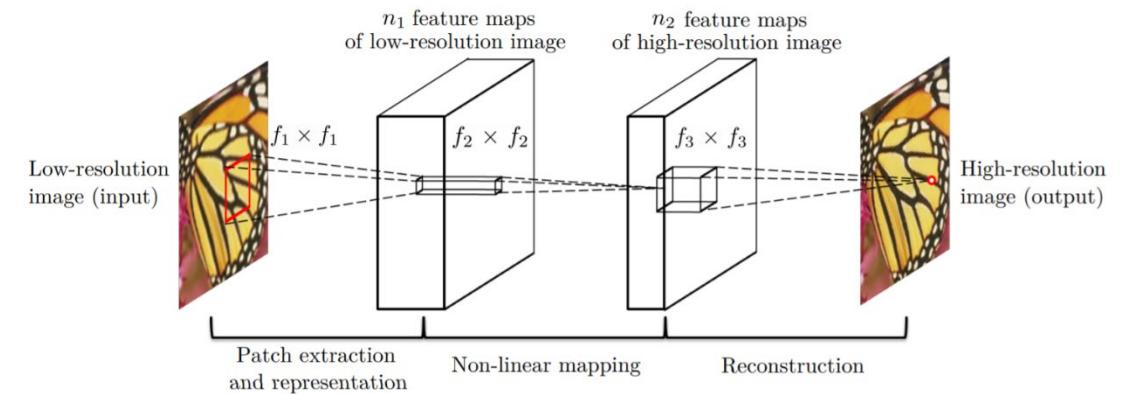
SuperResolution CNN (SRCNN)

paper 2015

LR

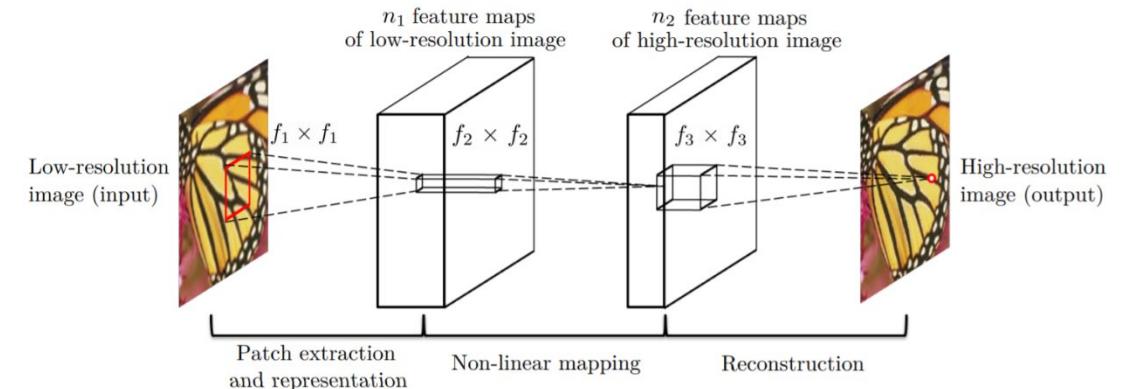
Conv Net

HR



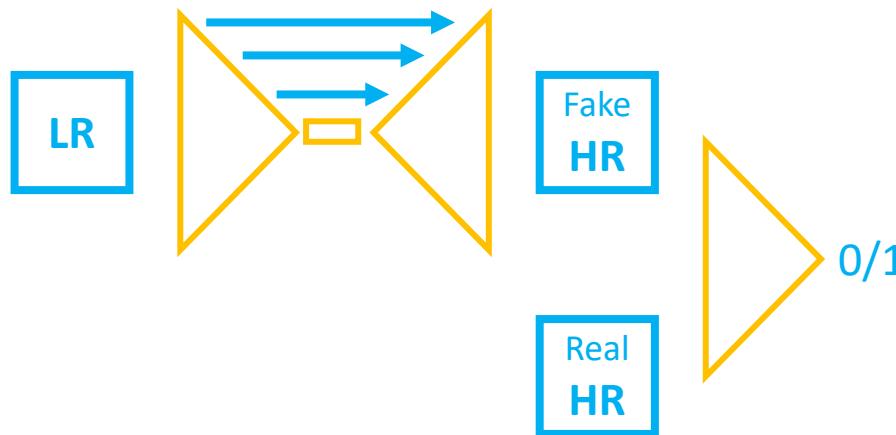
SuperResolution CNN (SRCNN)

[paper 2015](#)



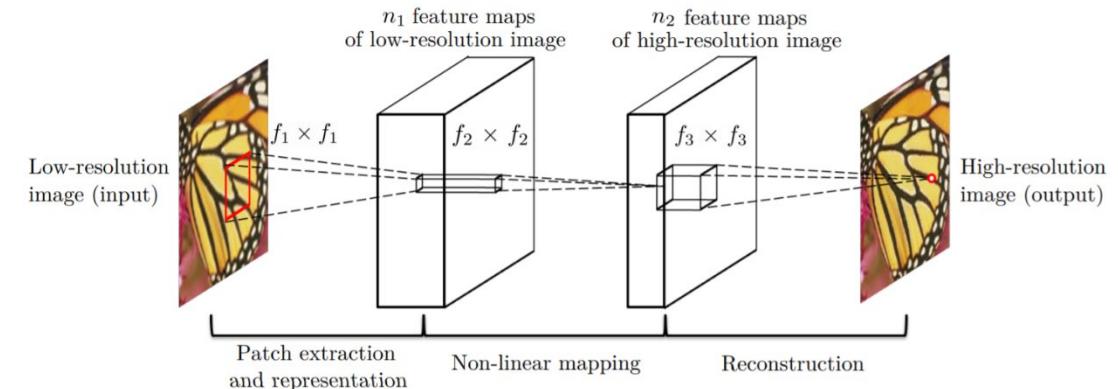
SuperResolution GAN (SRGAN)

[paper 2017](#)



SuperResolution CNN (SRCNN)

paper 2015



SuperResolution GAN (SRGAN)

paper 2017

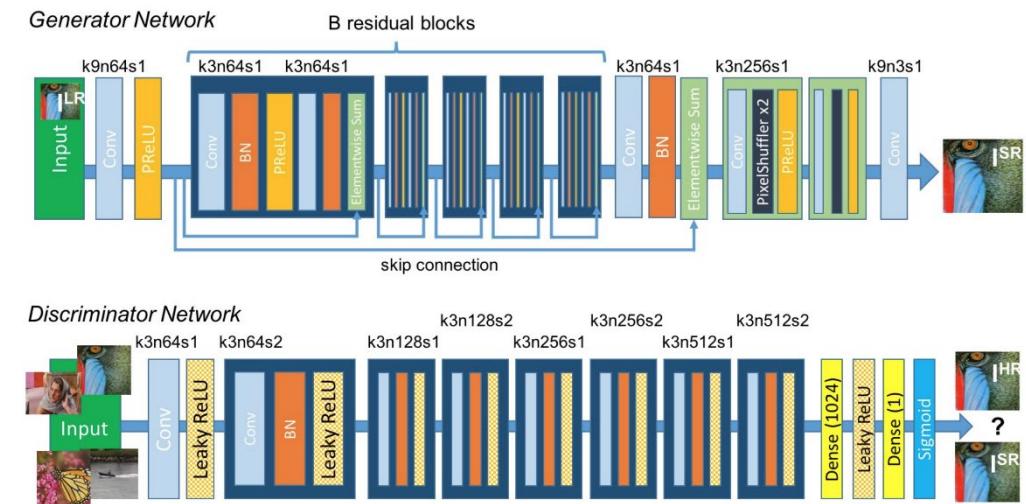
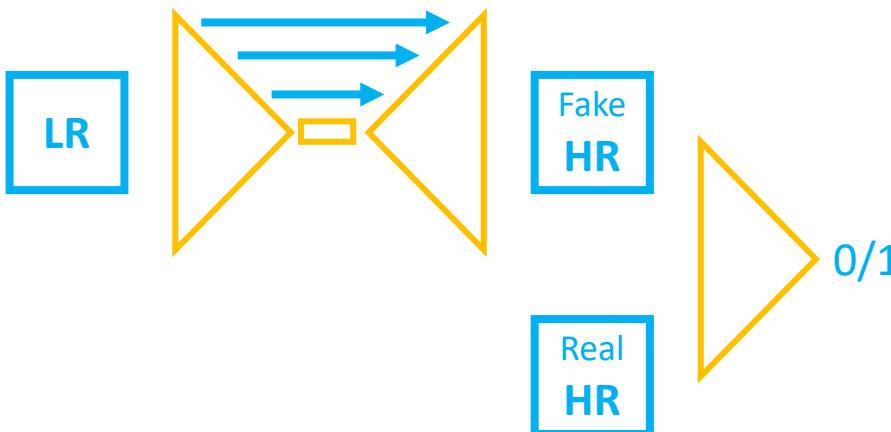
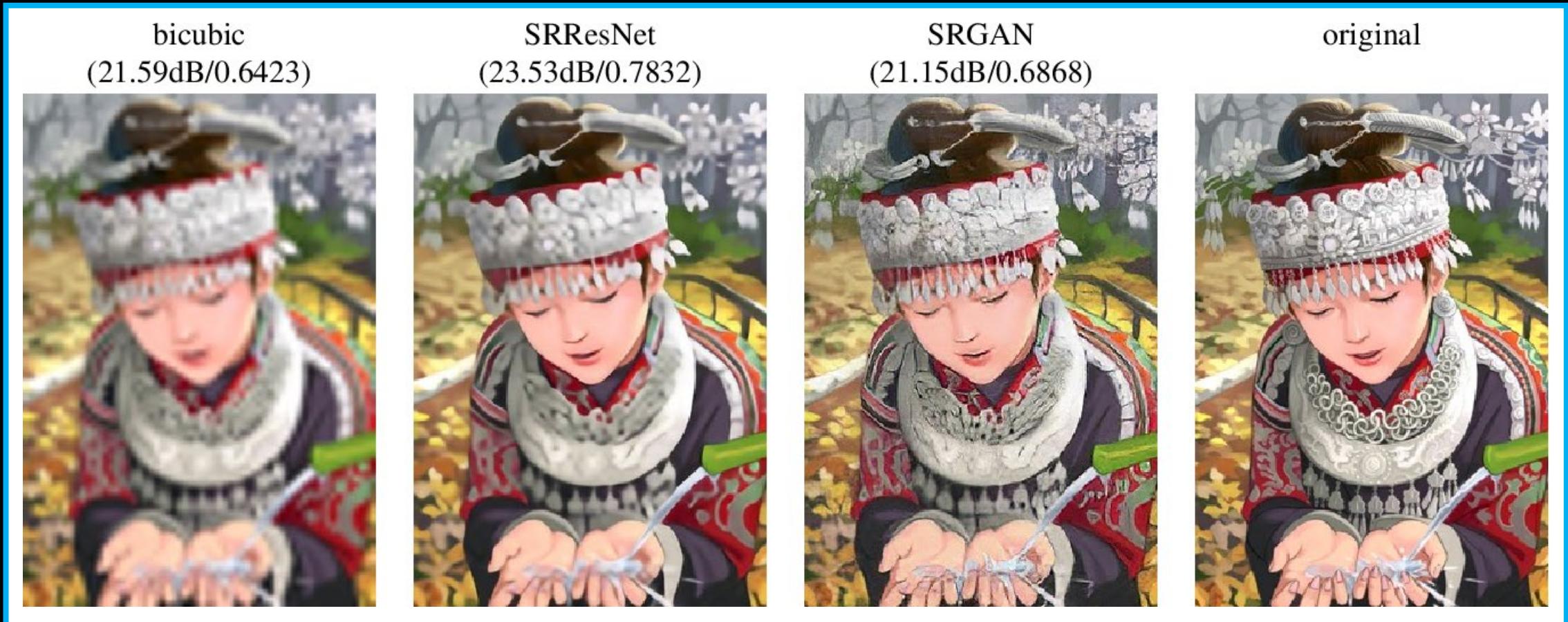


Figure 4: Architecture of Generator and Discriminator Network with corresponding kernel size (k), number of feature maps (n) and stride (s) indicated for each convolutional layer.

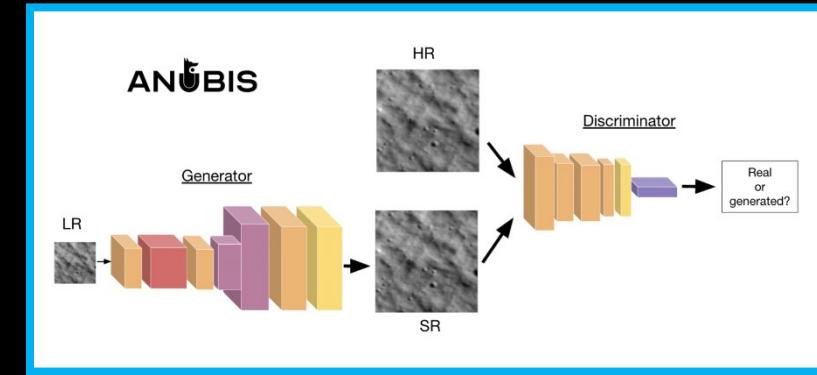
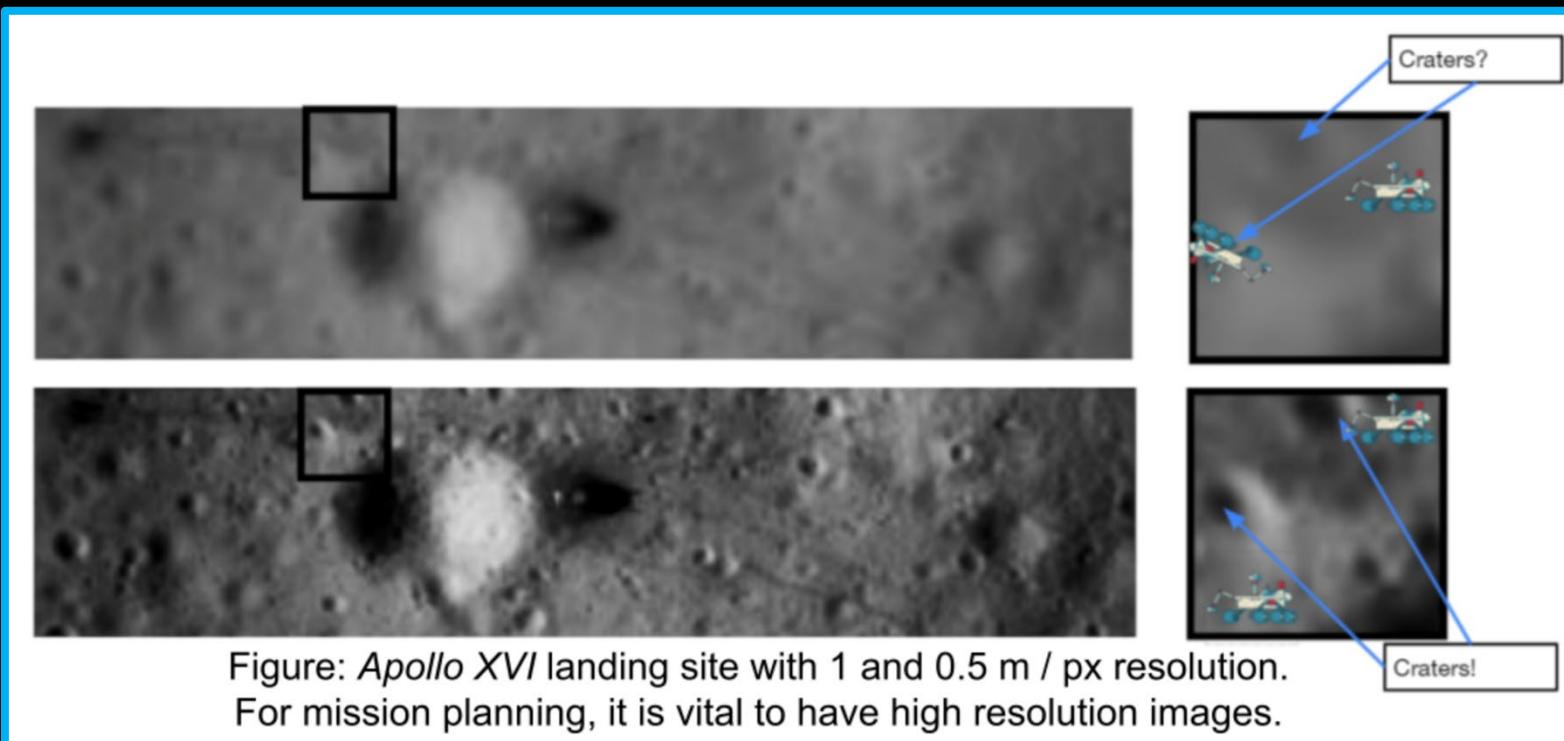
SuperResolution



SRGAN from 2017

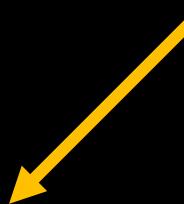
Topic II: SuperResolution Examples

Real-world example: Super resolution of Lunar maps



SuperResolution: Art examples

Models such as the Progressive Growing GANs jump between the resolutions of the generated data:

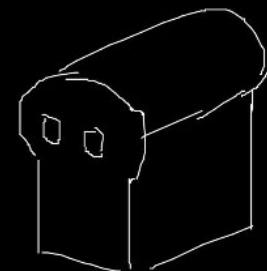


Instead we can use Super Resolution

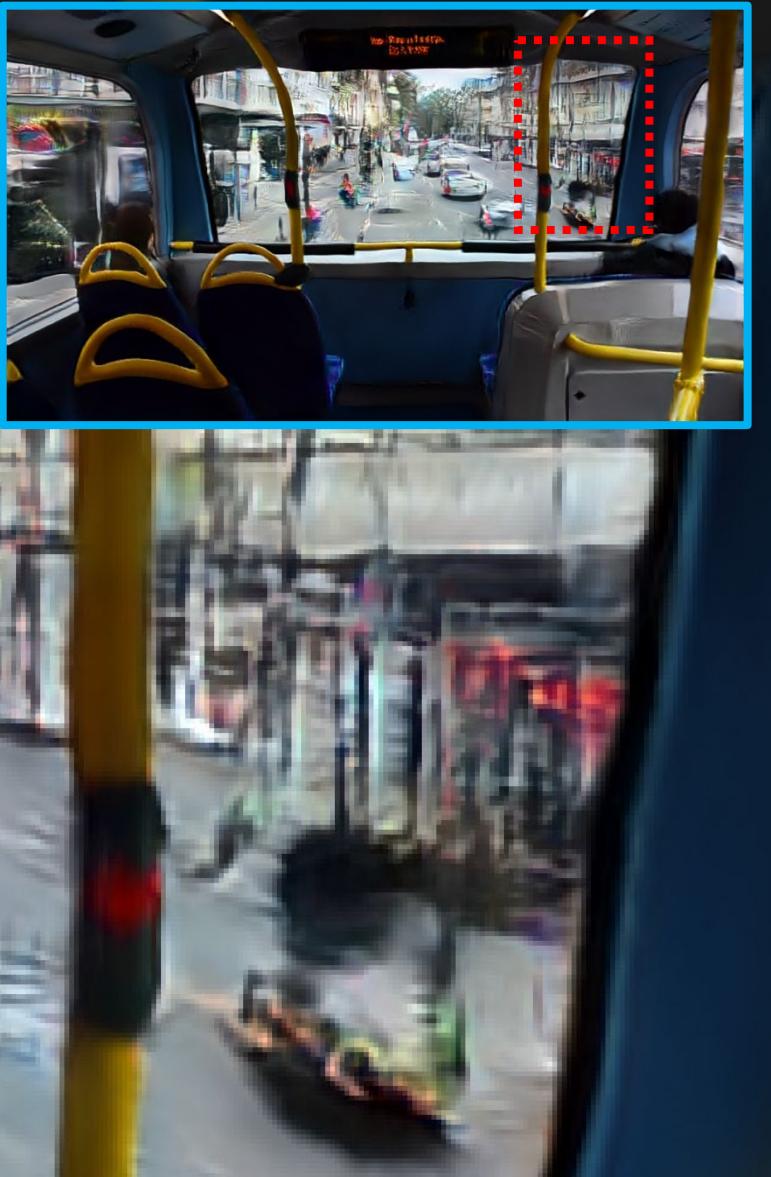
Earlier epochs (while lower res) may look better



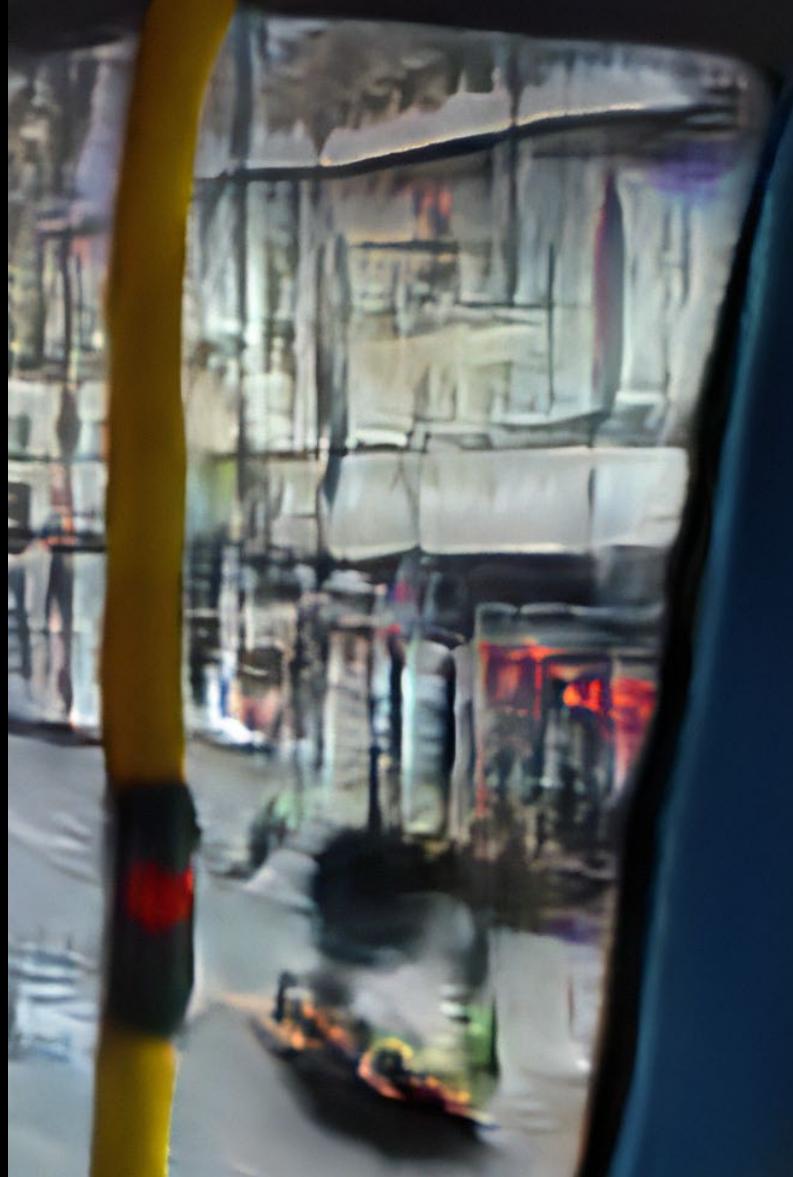
Than the model just few iterations after



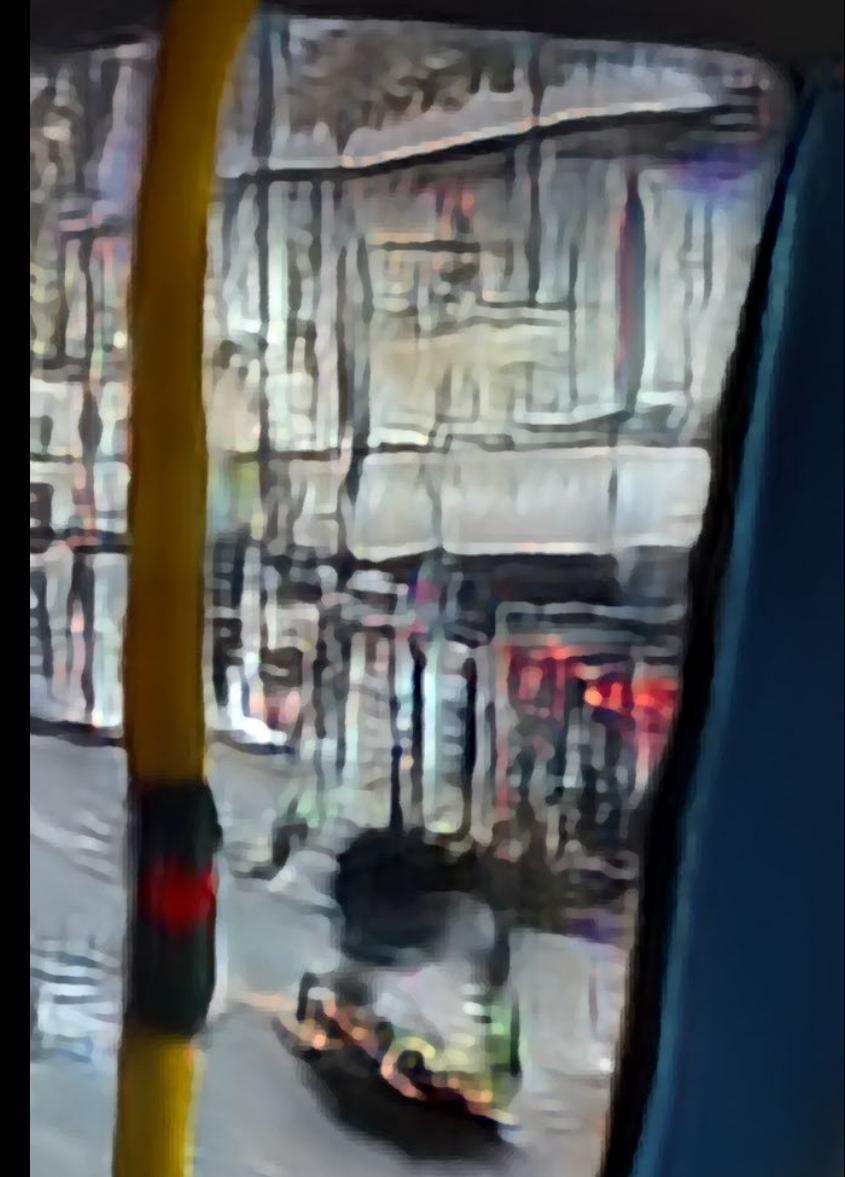
Or some models may produce limited resolution outputs



Original



RealSR (4x)



Anime4k* (4x)

*) Real-time: <https://github.com/bloc97/Anime4K> / The promise of real-time would mean running SuperRes while playing some low-res video. *Compute > Connection*.

SuperResolution: Art examples



Restoration process:

- Higher **resolution**
- Higher **fps** (interpolate between frames)
- Add **colour** ?

Example DeOldify video:
youtu.be/EqbOhqXHL7E



SuperResolution: Art examples

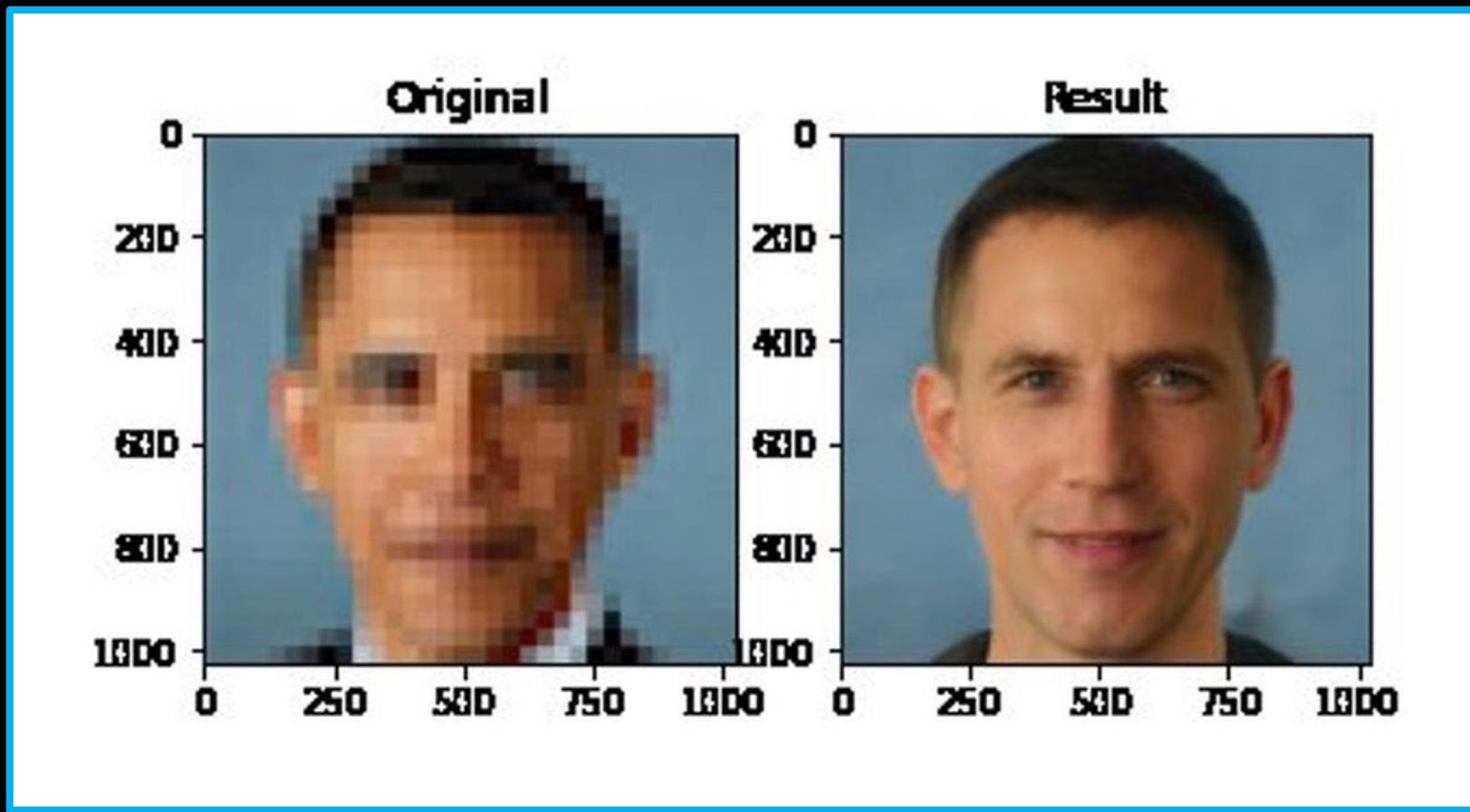
Upscaling game resources



ESRGAN, [X. Wang](#), 2018

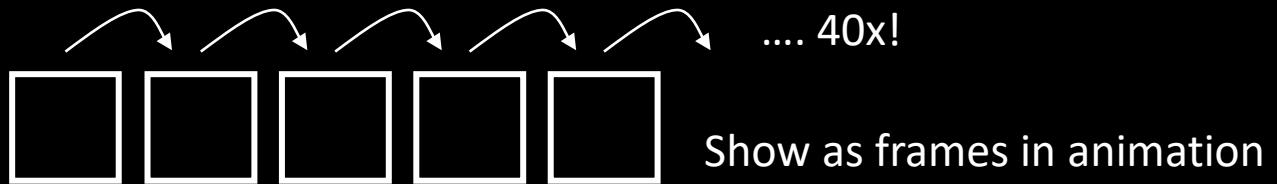
Mass Effect Legendary Edition: “We ran all the original, uncompressed source art through an AI up-res program along with some other custom batch tools.” - [2021/02/02/bioware-used-ai-upscaling-to-remaster-mass-effects-original-textures/](https://www.bioware.com/news/bioware-used-ai-upscaling-to-remaster-mass-effects-original-textures/)

Topic II: SuperResolution Limitations



SuperResolution: Limitations

- **Complication:**
Conceptually we are mixing our data
with noise from someone else's
datasets

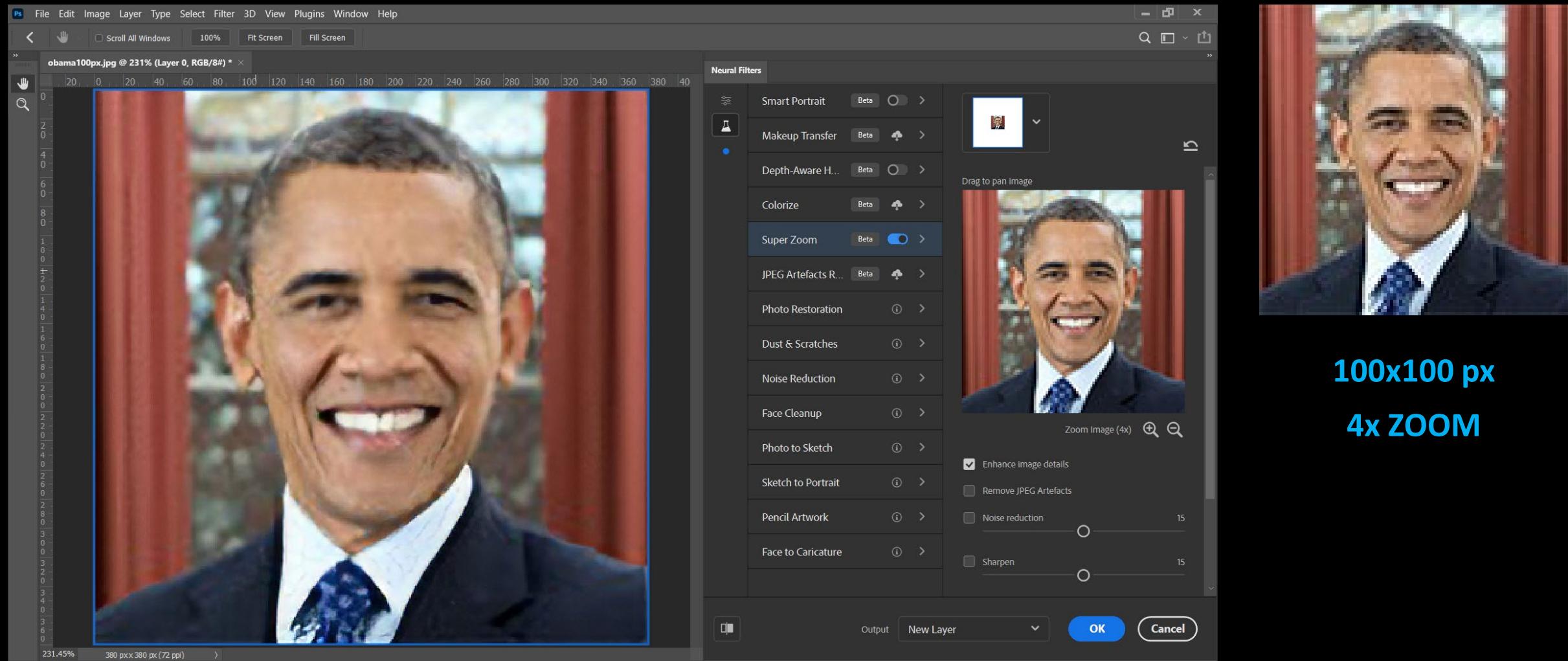


**Exploring different SuperResolution
models aesthetics: [demo example](#)**

SuperResolution: Easy tools

- **Applications:**
 - There are some free repositories with access to pre-trained models (video2x)
 - Or these are making it into professional tools (+- expansive usually)
 - Photoshop / Camera raw: [helpx.adobe.com/uk/camera\(raw\)/using/enhance.html](https://helpx.adobe.com/uk/camera(raw)/using/enhance.html)
- **Video2x** (free application, [repo](#))
 - Colab notebooks ([link](#))
 - Windows application ([zip](#))
 - **Models:** Waifu2X, SRMD, RealSR, Anime4KCPP
 - **Input:** image / [gif](#) / [video](#)





Photoshop Neural Filters

**100x100 px
4x ZOOM**

Summary from the lecture

- **CycleGAN:**
 - **Model architecture** with two generators (there and back) and one discriminator
 - **Example** learning between two domains (not paired)
- **Super Resolution:**
 - **Models** as GAN structure
 - **Examples** from research, art and industry ... easy to use **Applications**
 - **Limitations** tricky! (*What happens to the reality?*)

Links and additional readings:

- **Papers:**
 - **CycleGAN**: [Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks](#) by Jun-Yan Zhu*, Taesung Park*, Phillip Isola, Alexei A. Efros. In ICCV 2017.
 - **pix2pix**: [Image-to-Image Translation with Conditional Adversarial Networks](#) by Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. In CVPR 2017.
 - **SuperResolution** blog post: [blog.paperspace.com/image-super-resolution/](#)

End of the lecture

*) PS: follows material for the practical session ...

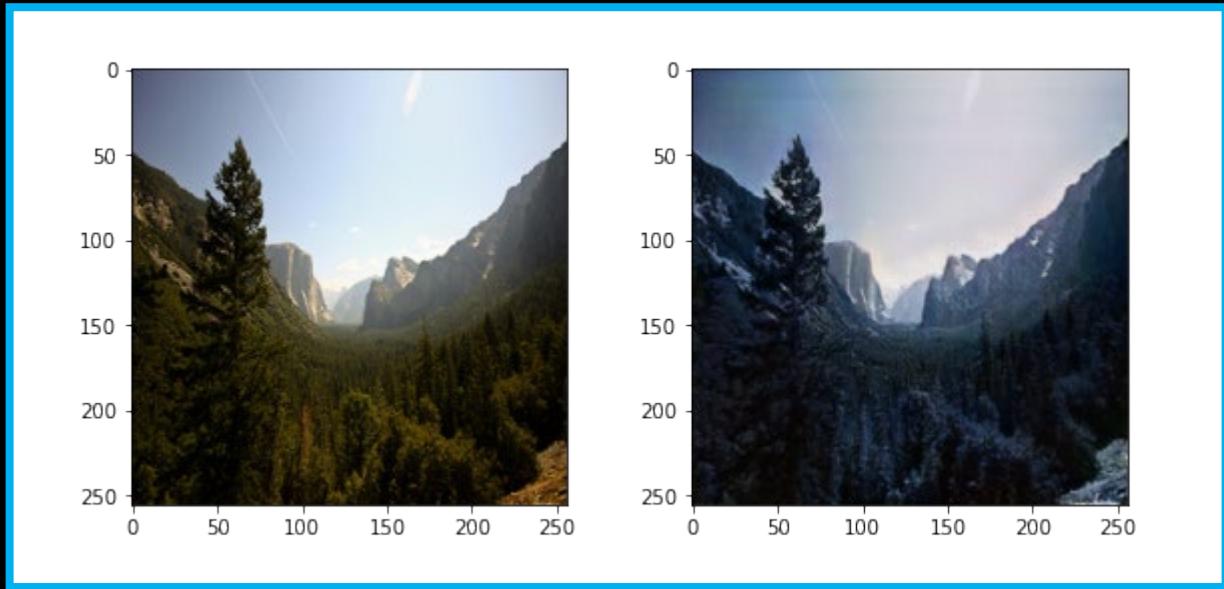
AI for the Media

Week 7, CycleGAN + SuperRes



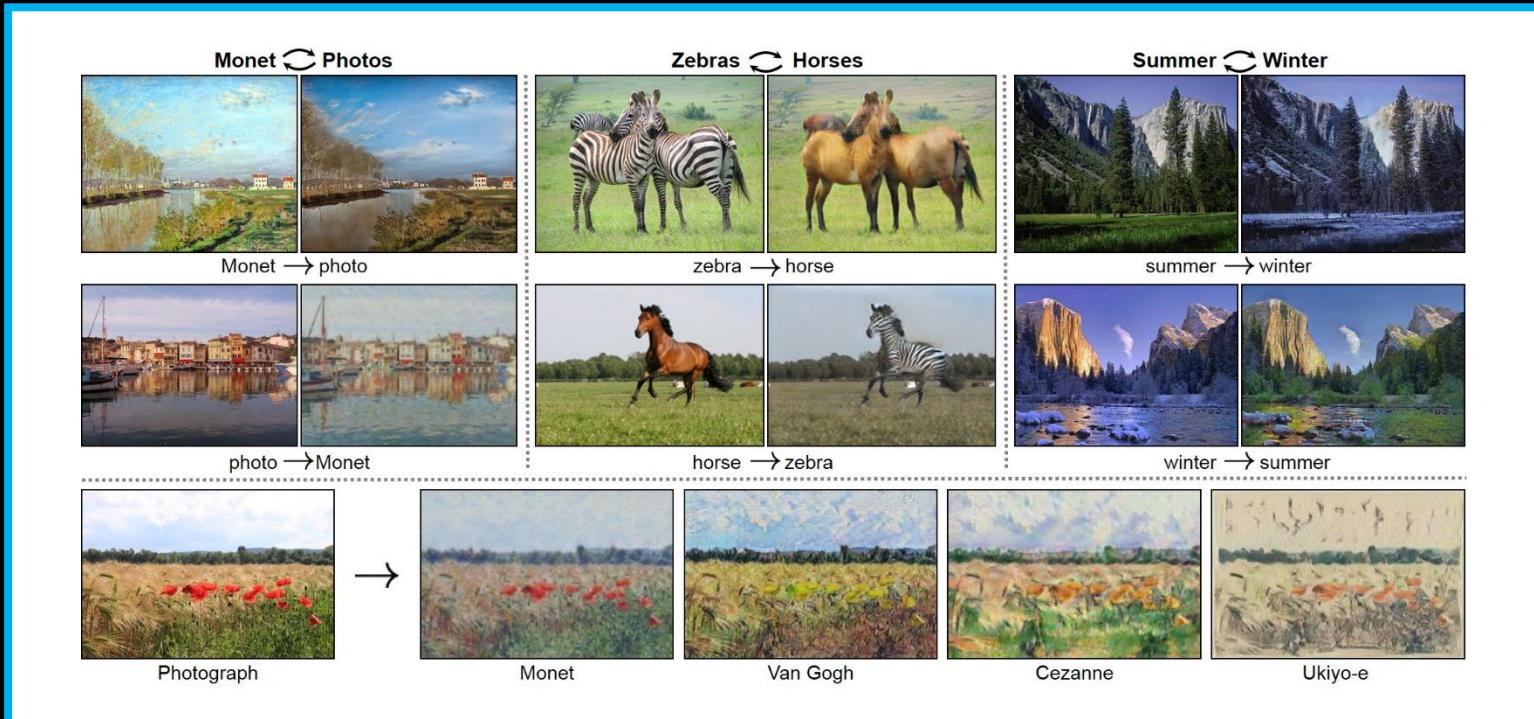
Practical: CycleGANs

Examples from the practical codebase



Summer2Winter

All models



Links

Starter code:

Opens directly in Colab

[w07_cycleGAN_inference.ipynb](#)

Interaction

Copy your team's results here:

<https://docs.google.com/document/d/19wwnPe9qWI7wkVYpcgTBRT6ydFKcauDJZn21atiVVDE/edit?usp=sharing>

The end