DeepRob

Lecture 22
Domain Adaption
University of Michigan | Department of Robotics

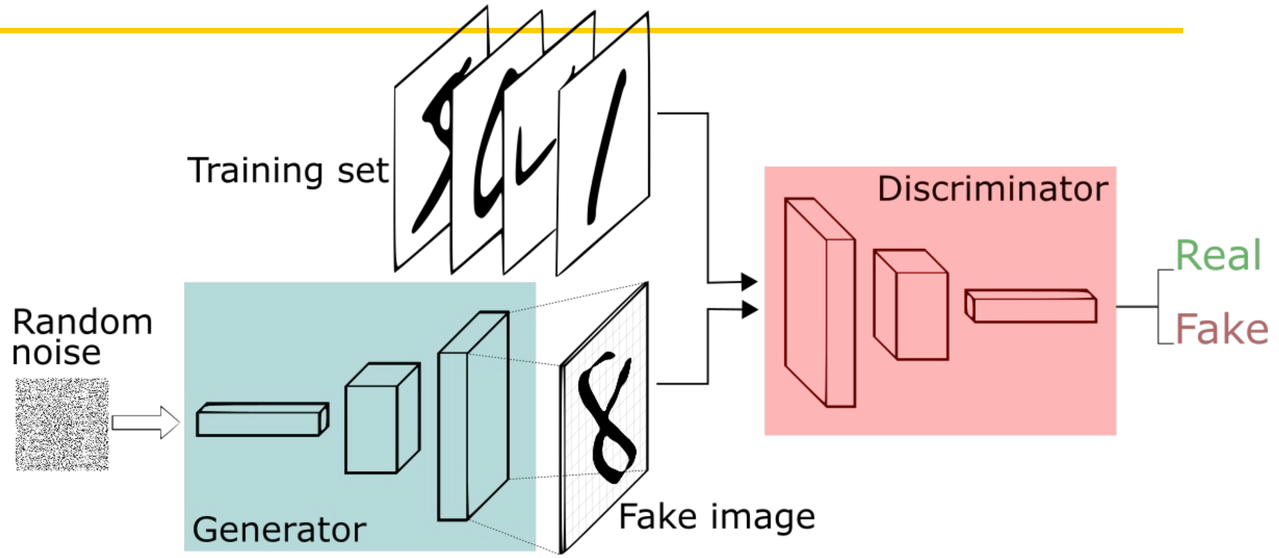
Recall: Unsupervised Learning

Data: x

Just data, no labels!

Goal: Learn some underlying hidden *structure* of the data

GANs



VAEs

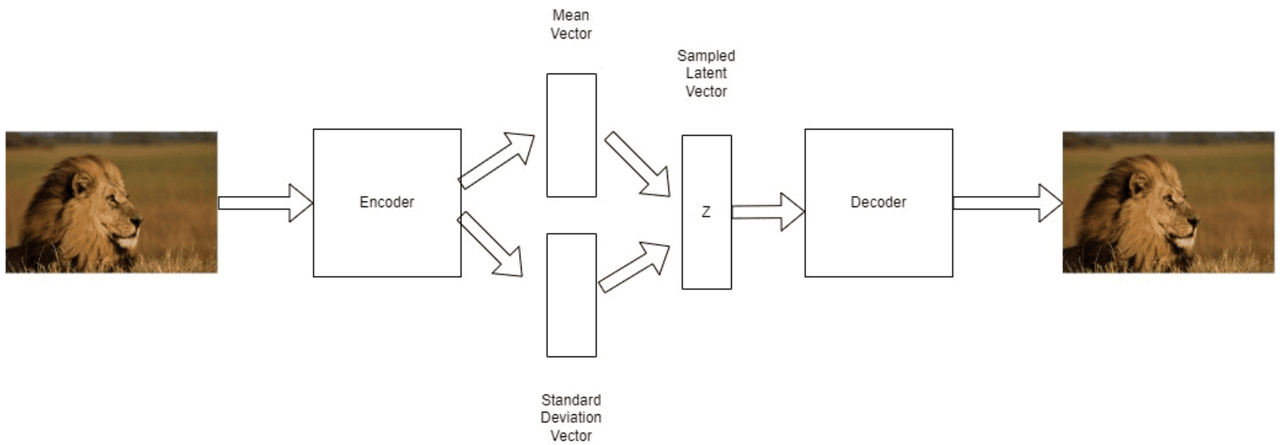


Image-to-Image translation

Unsupervised Image-to-Image Translation Networks, NVIDIA

Image-to-Image translation

CycleGAN

https://junyanz.github.io/CycleGAN/

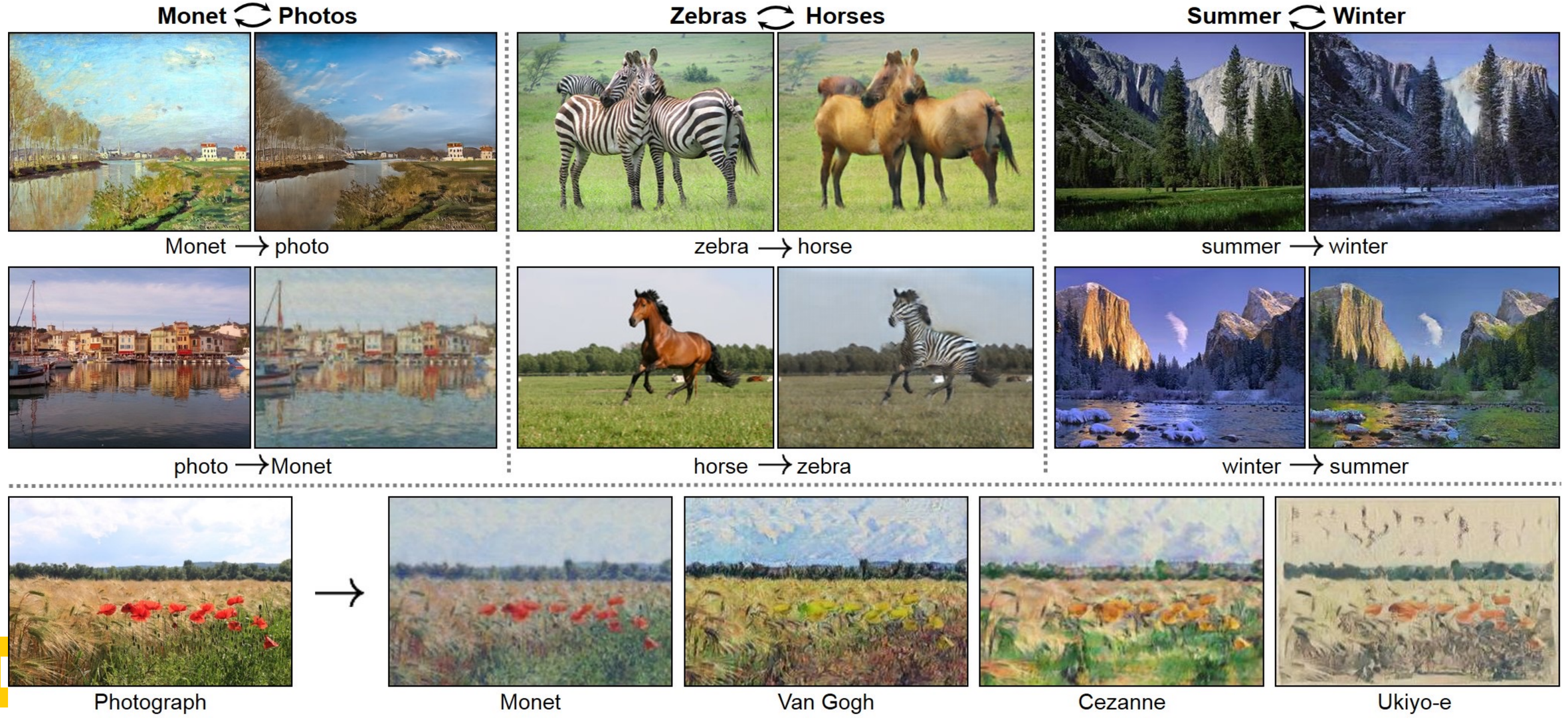
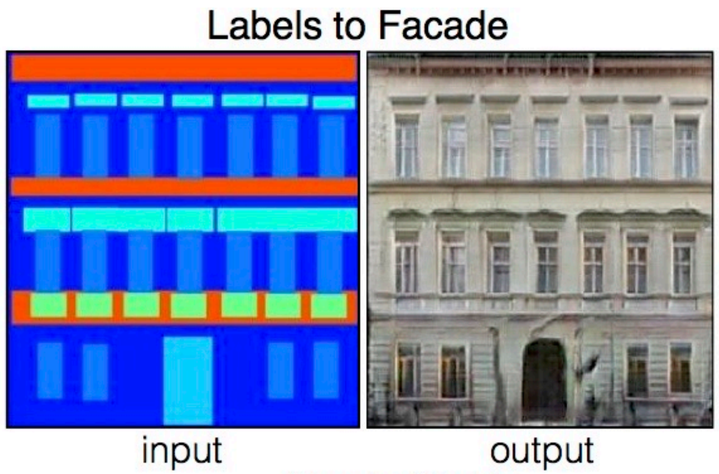
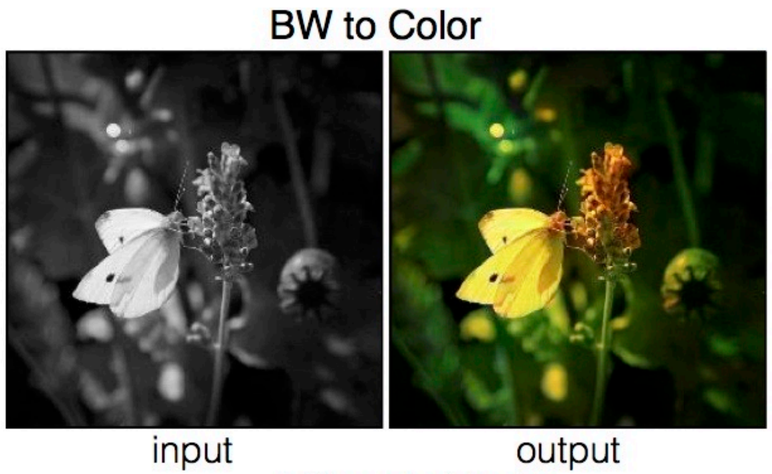


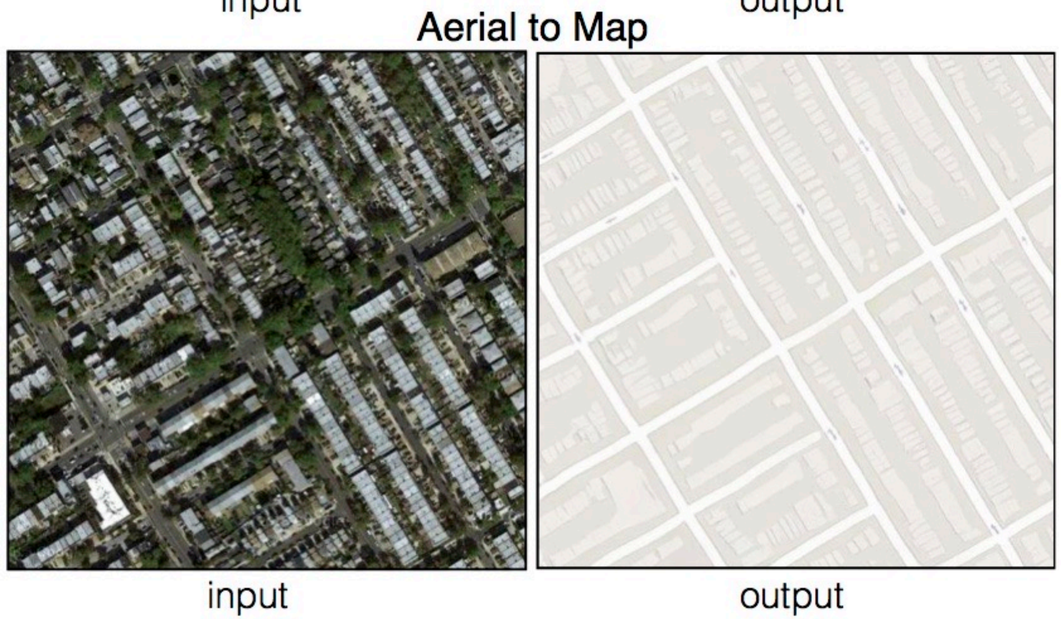
Image-to-Image translation

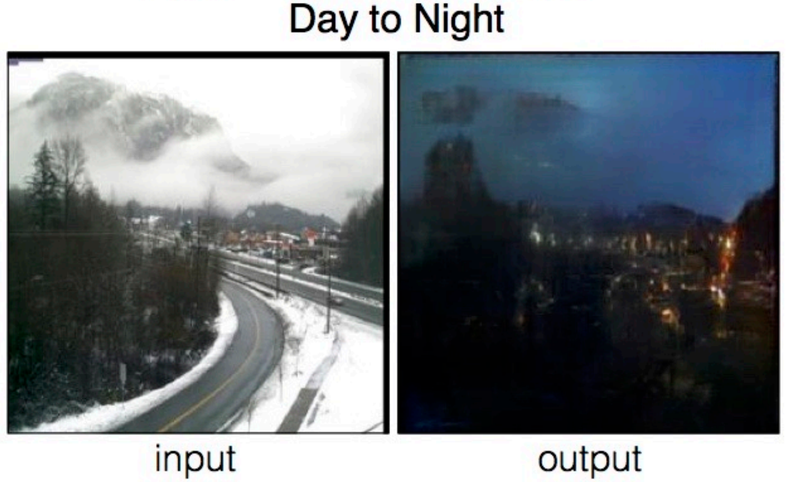
Pix2Pix

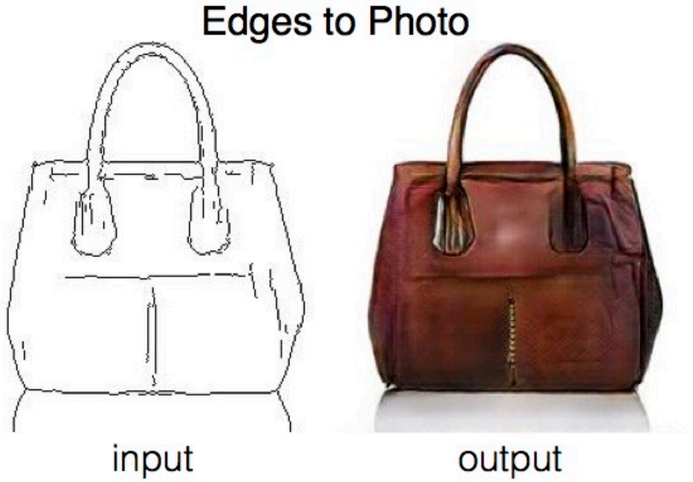
https://www.tensorflow.org/tutorials/generative/pix2pix











Extension: InstructPix2Pix w/ LLMs

https://www.timothybrooks.com/instruct-pix2pix

Domain Adaptation Task

e.g., C->FC

source dataset $S = \{x_s^i, y_s^i\}_{i=1}^{N_s}$

Source image

(labeled)

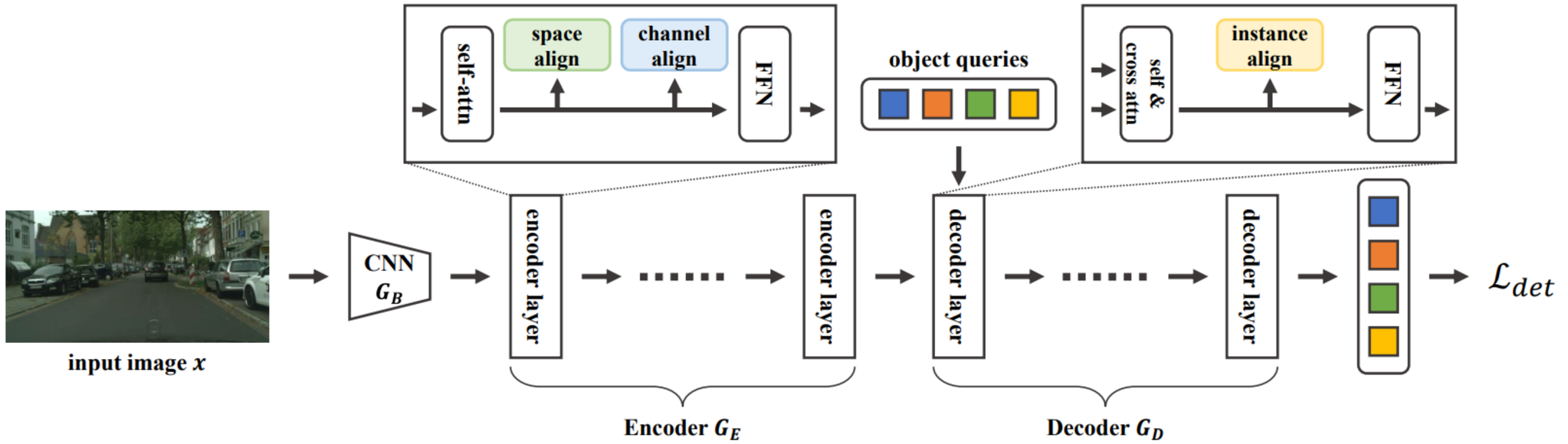
target dataset $T = \{x_t^i\}_{i=1}^{N_t}$

Target image

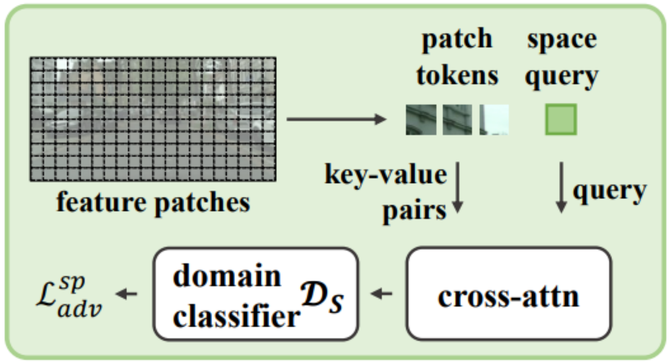
(unlabeled)

https://people.ee.ethz.ch/~csakarid/SFSU_synthetic/

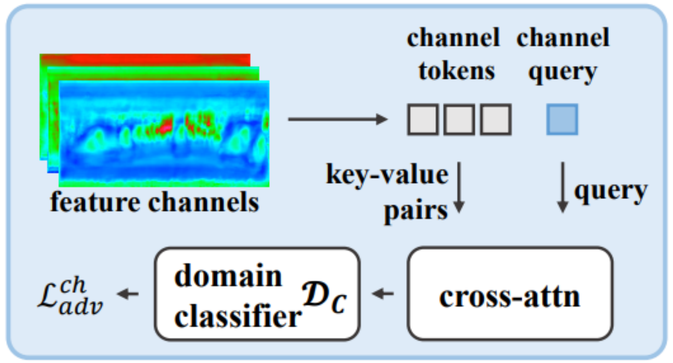
Adversarial Query Transformers (AQT)



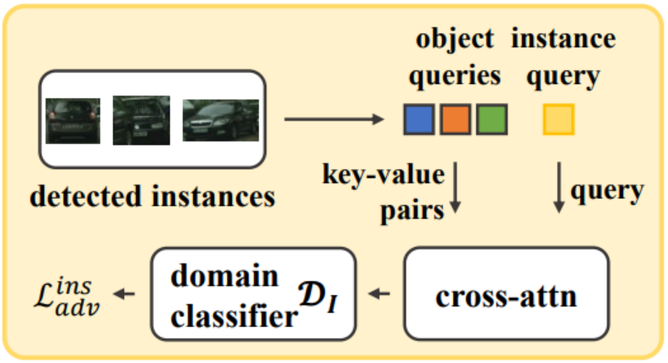
(a) Overview of our method AQT



(b) Space-level feature alignment module



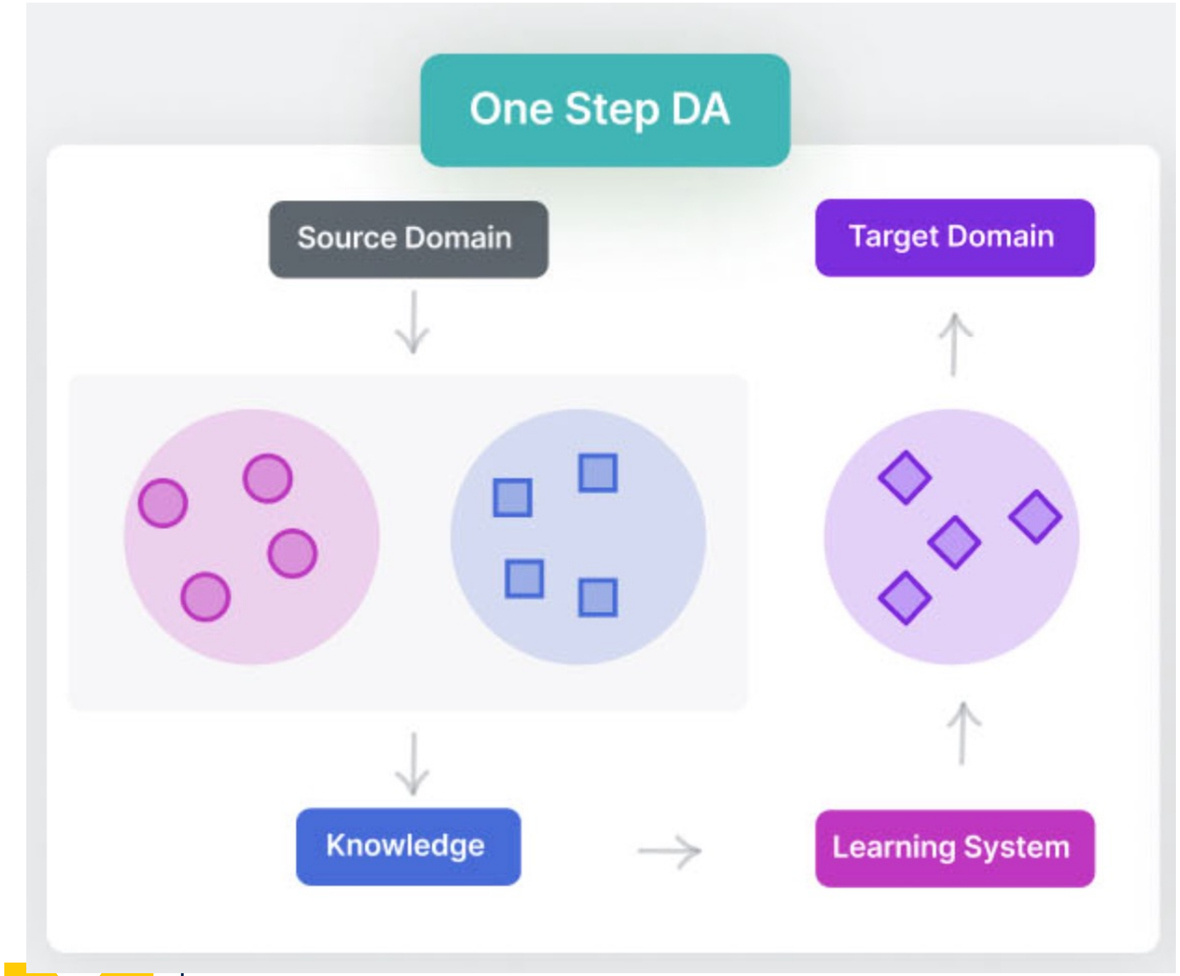
(c) Channel-level feature alignment module

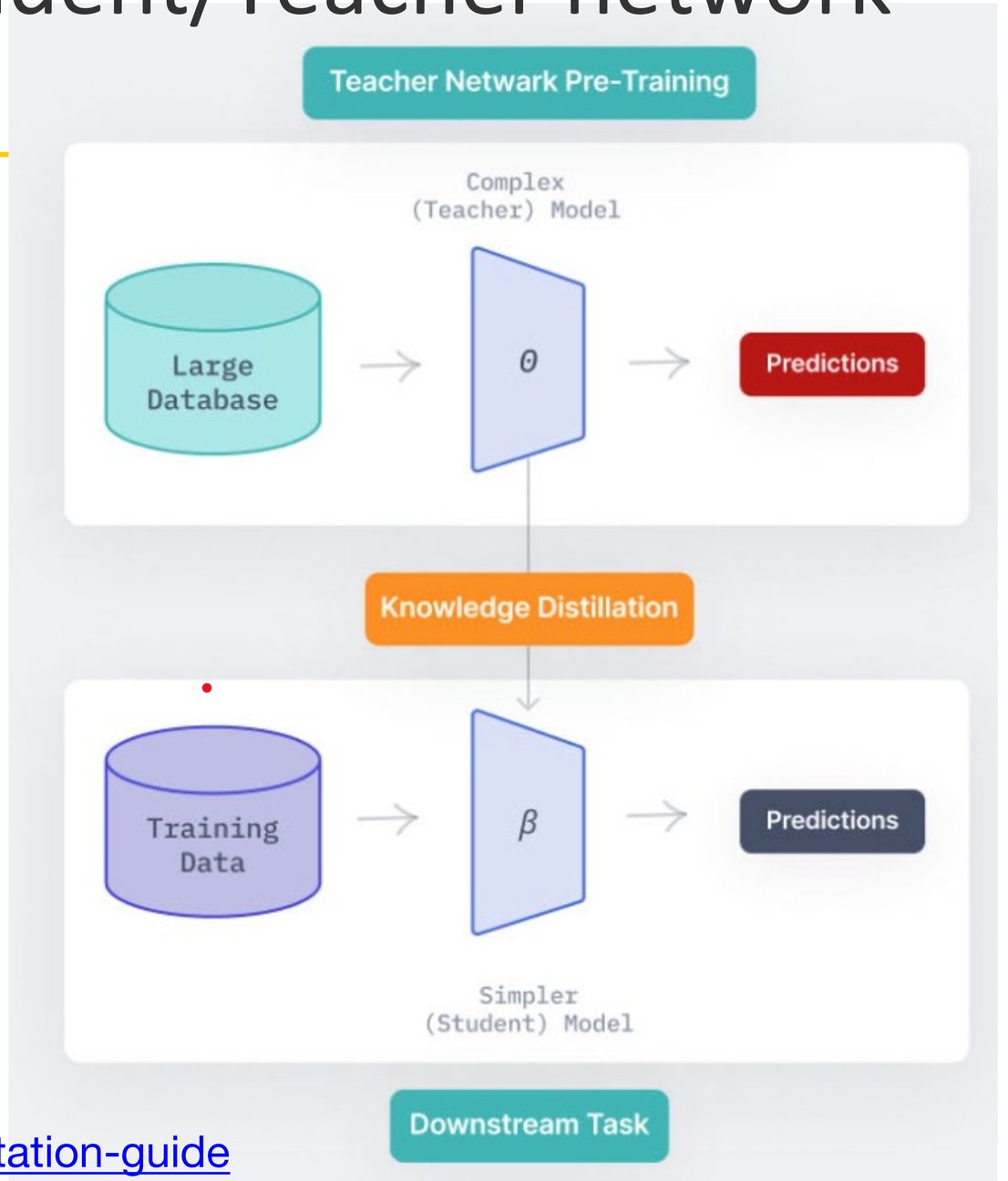


(d) Instance-level feature alignment module

https://www.ijcai.org/proceedings/2022/0136.pdf

Knowledge Distillation – Student/Teacher network





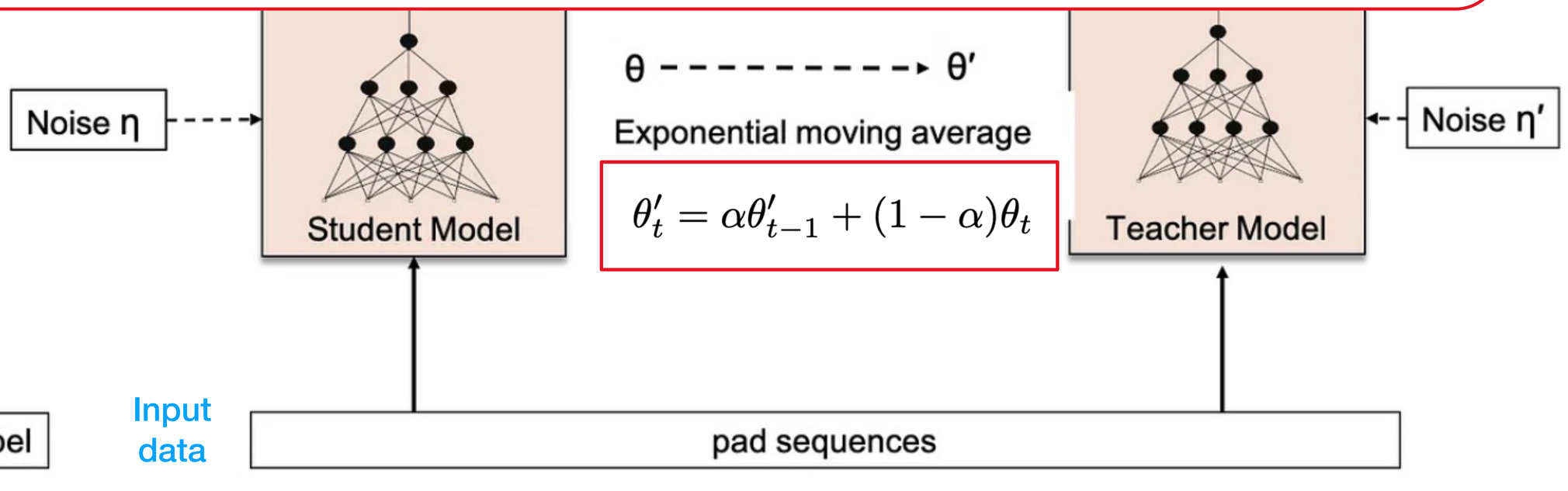
https://www.v7labs.com/blog/domain-adaptation-guide

EMA “Mean Teacher”

$$J(\theta) = \mathbb{E}_{x,\eta',\eta}\left[\|f(x,\theta',\eta') - f(x,\theta,\eta)\|^2\right]$$

Why?

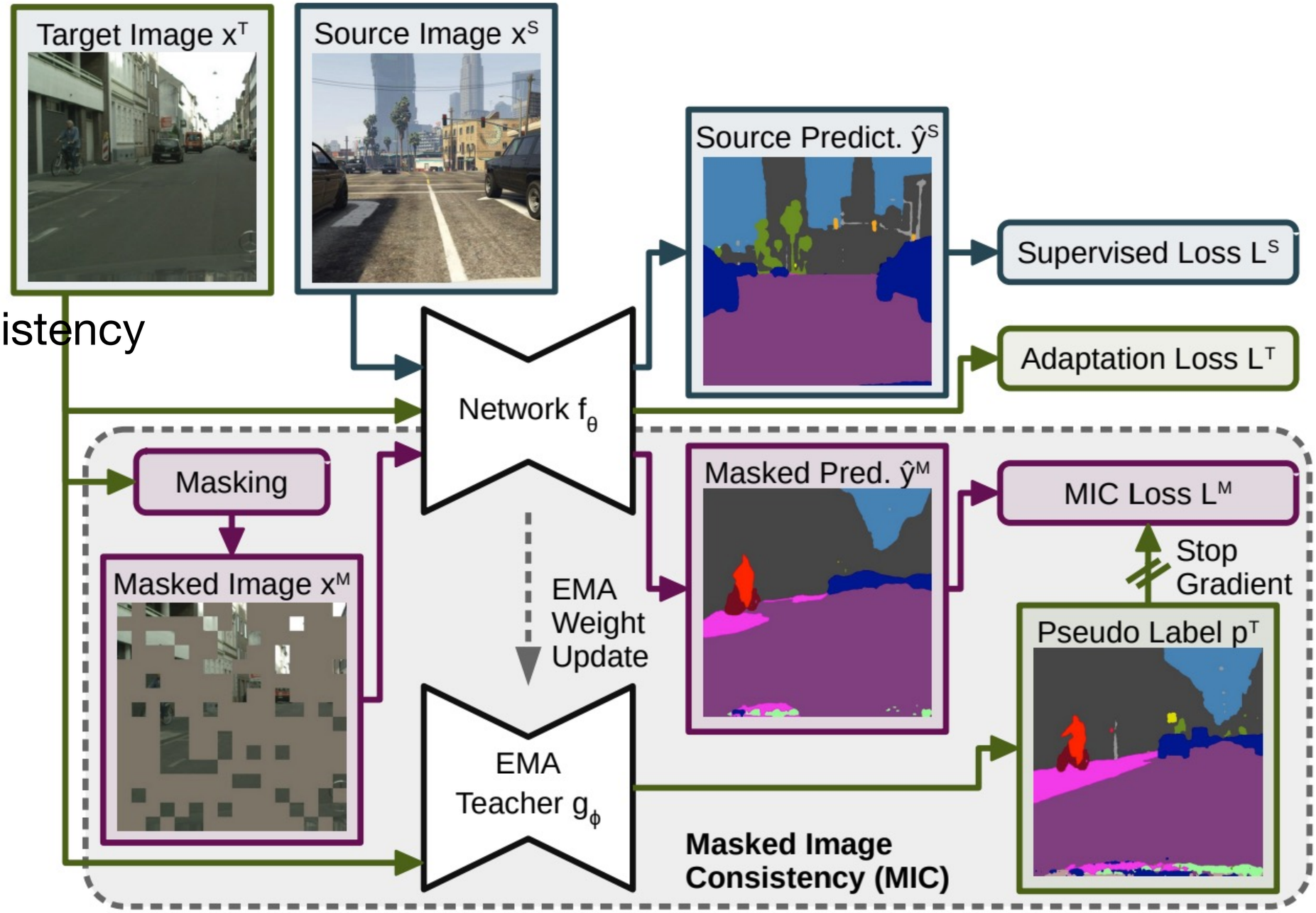
Usually, simpler student, larger teacher model



MIC

Masked Image Consistency

https://arxiv.org/pdf/2212.01322v2.pdf



DeepRob

Lecture 22
Domain Adaption
University of Michigan | Department of Robotics