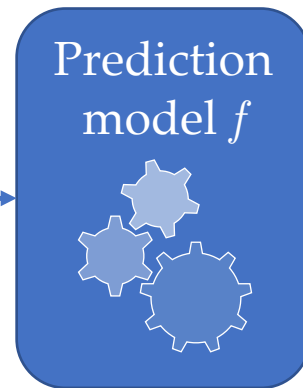


Input image to analyze



Prediction



Examples of different tasks (i.e. types of prediction) in Computer Vision:

Classification

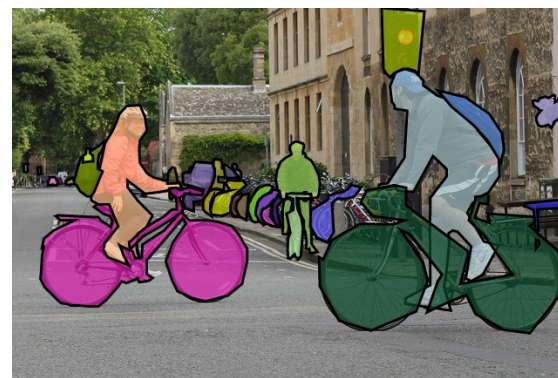
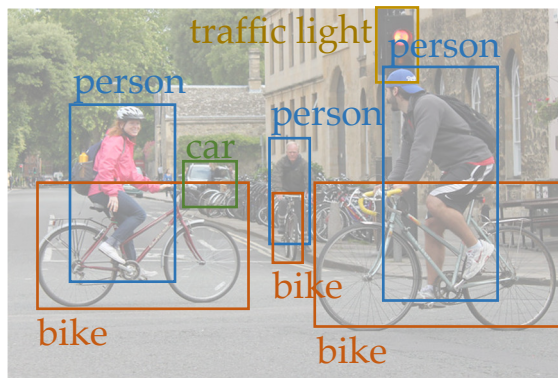
Object detection

Segmentation

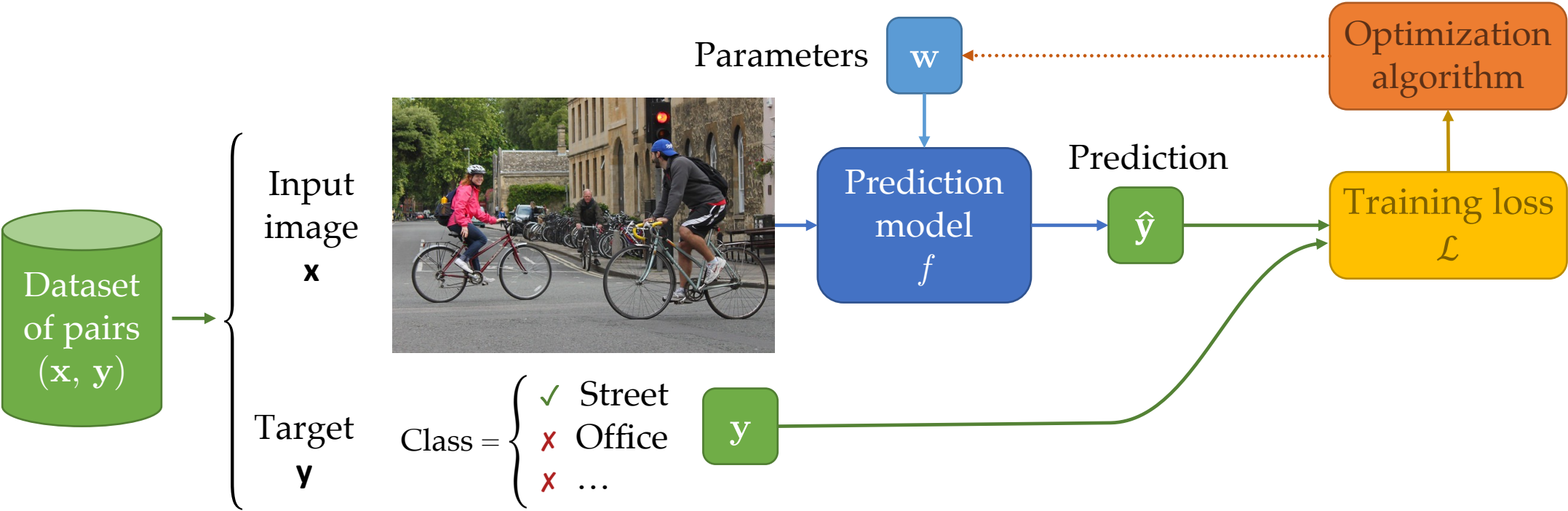
Captioning

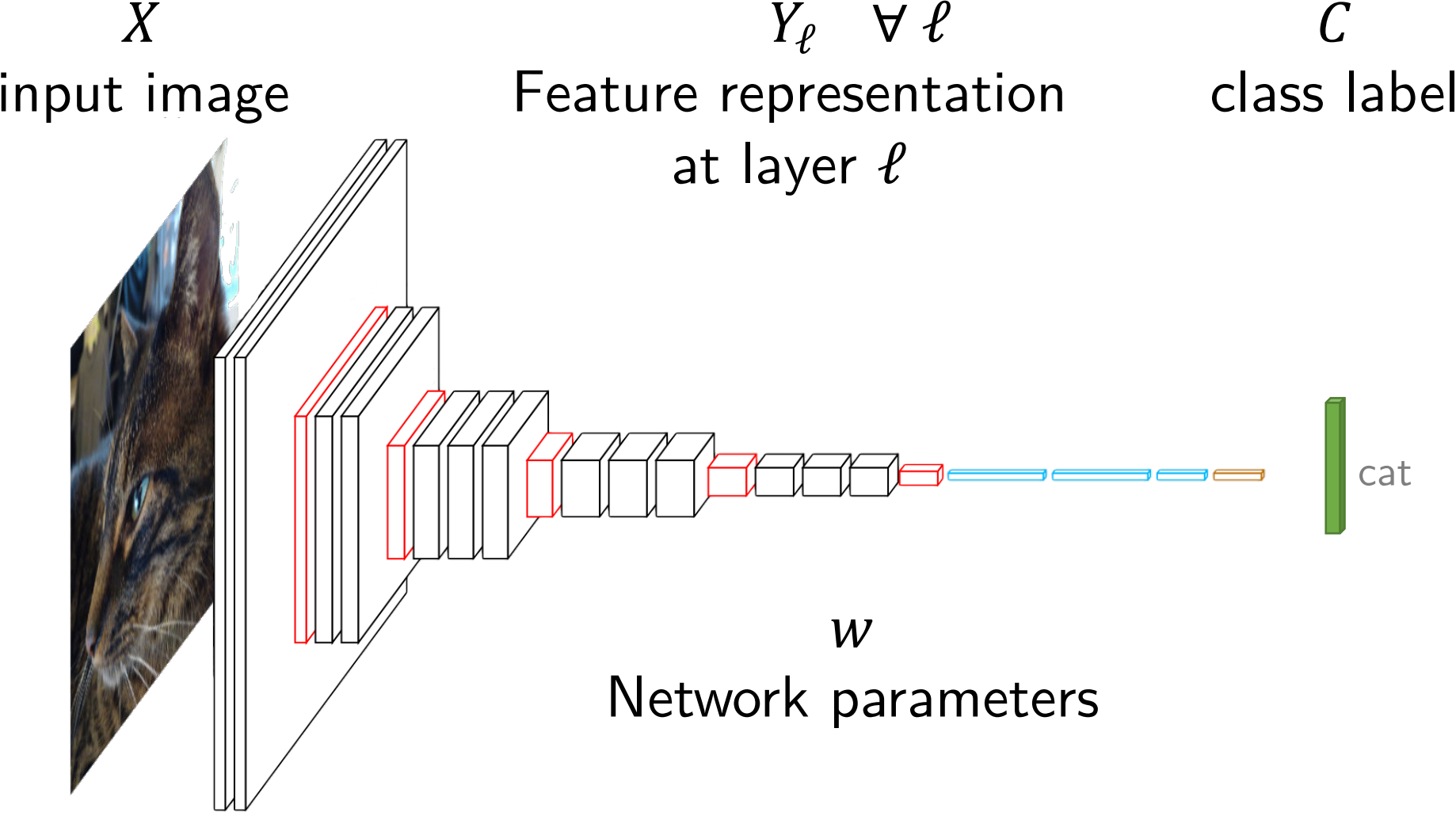
...

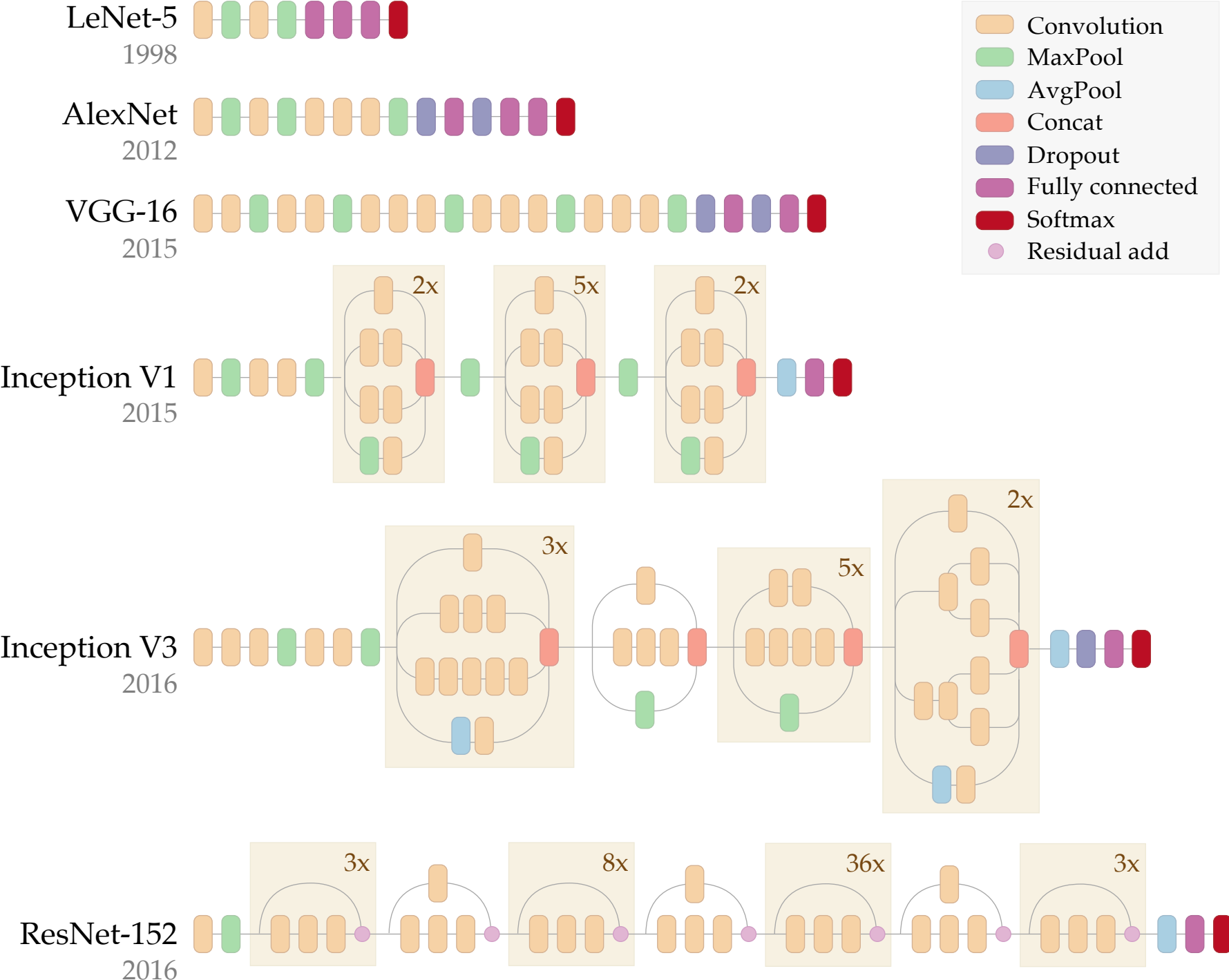
- ✓ Street
- ✗ Office
- ✗ Bedroom
- ✗ ...

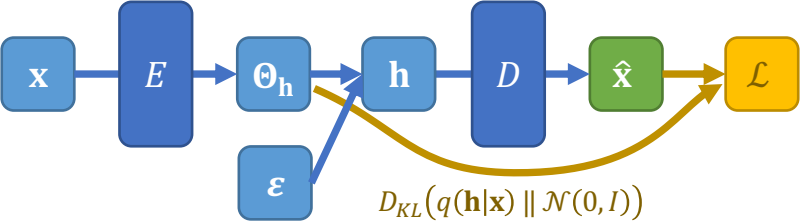


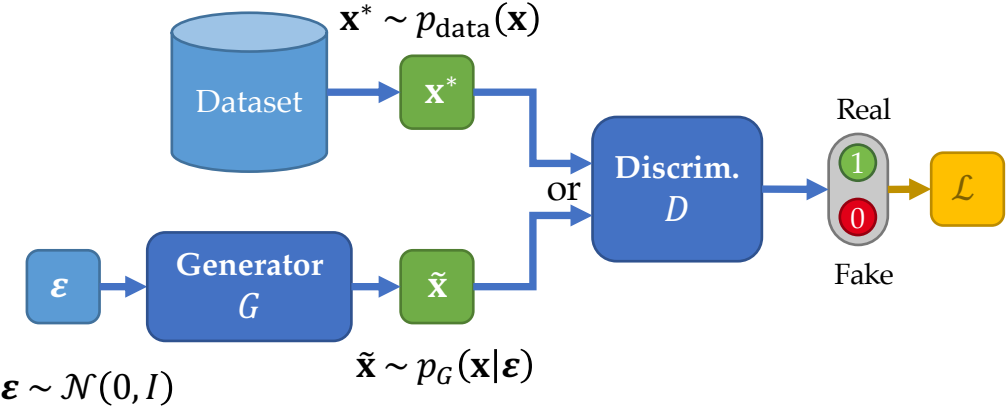
*a girl in a pink jacket
on a bicycle passes a
man in a blue cap on
a bicycle.*

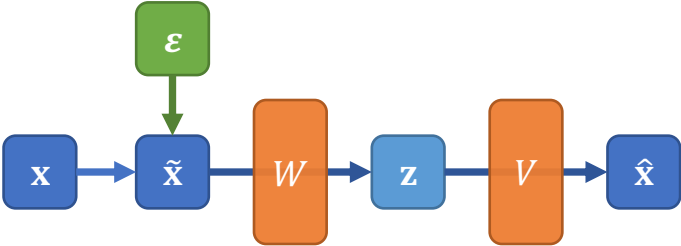












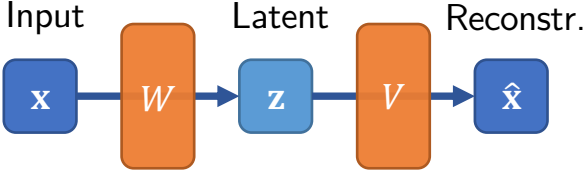
Input

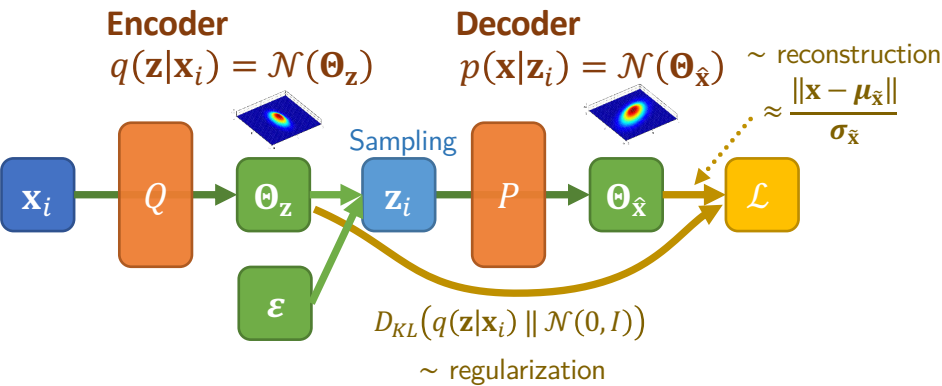
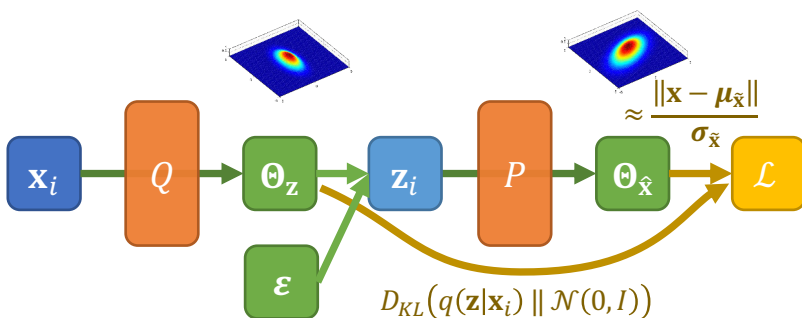
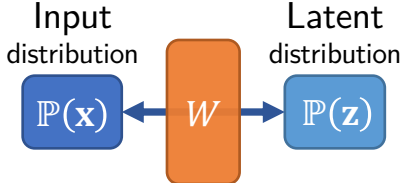


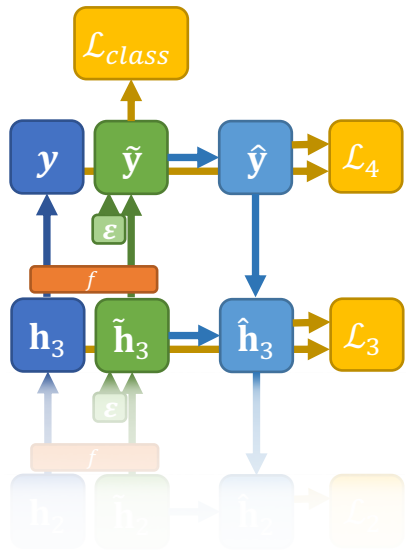
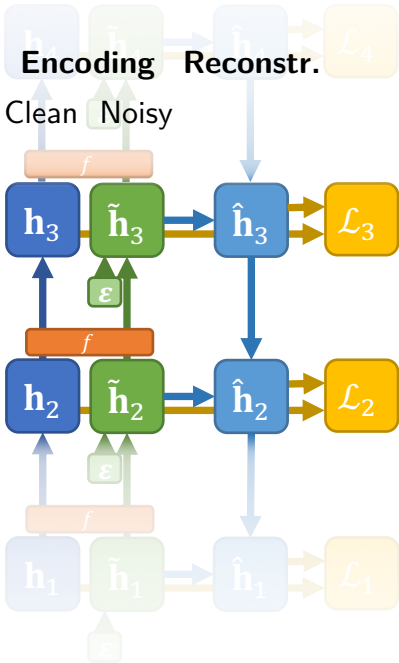
Latent

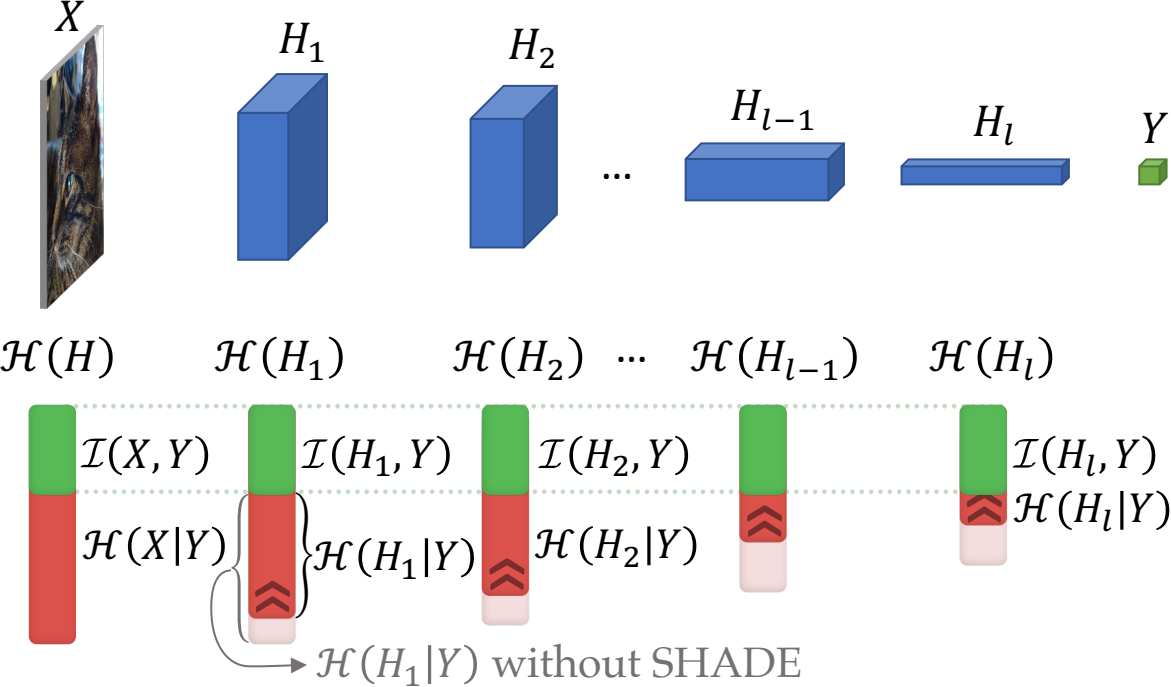


Reconstr.



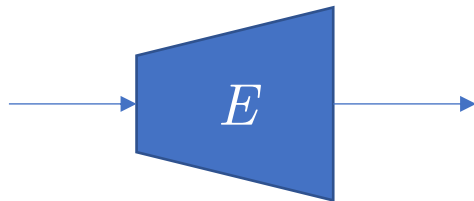




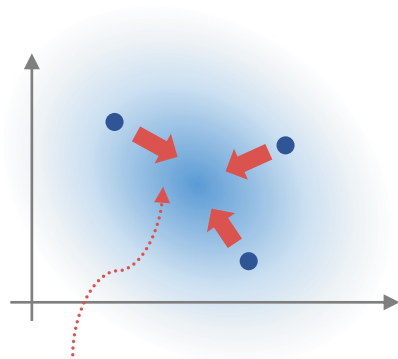


\rightrightarrows Effect of SHADE: Reduction of conditional entropy

Input images



Feature representation

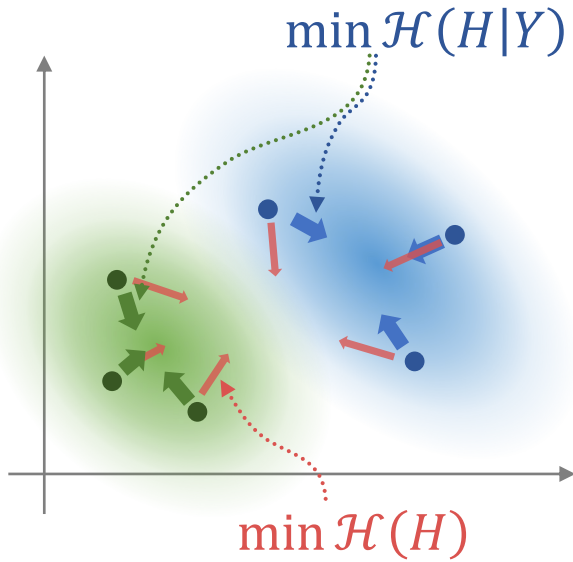
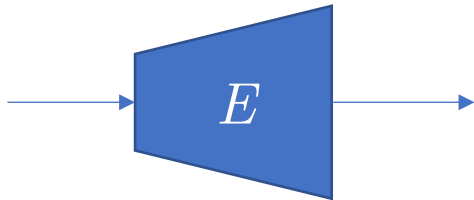


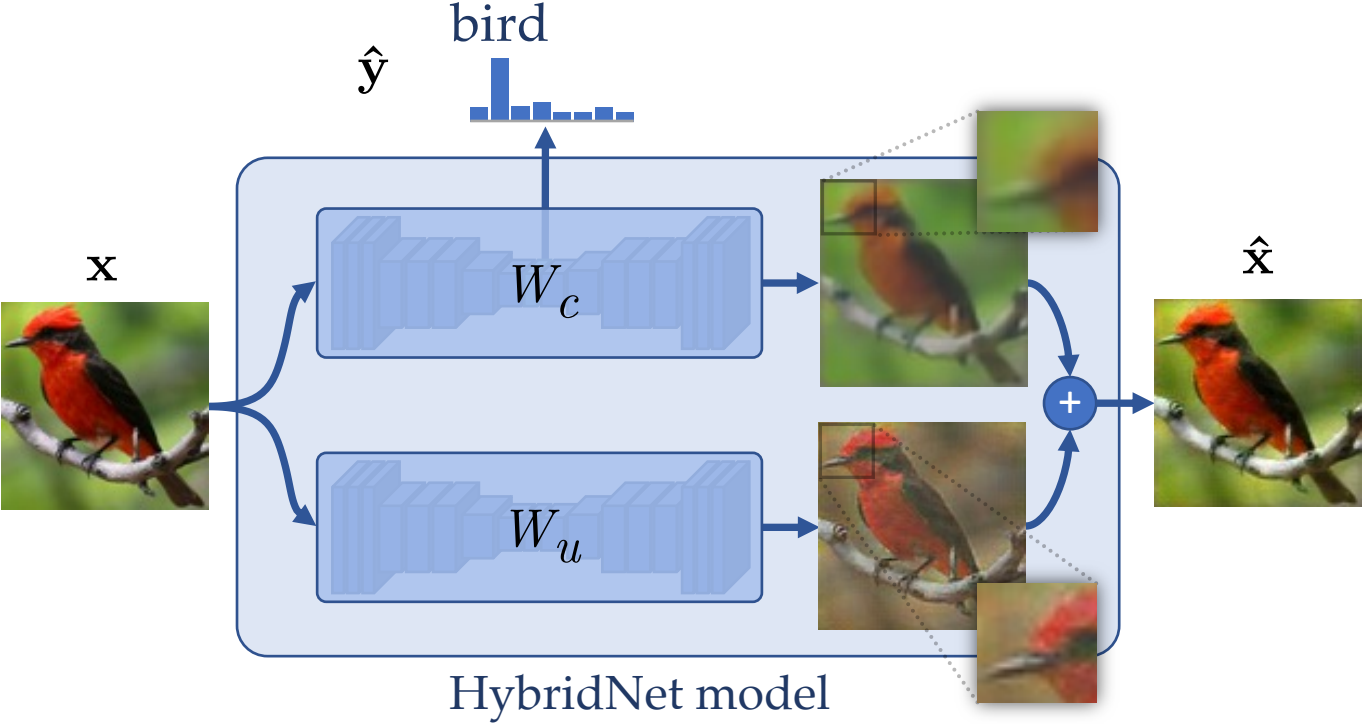
$\min \mathcal{H}(H) \Rightarrow \max \mathcal{H}(X|H)$
increases uncertainty in X given H

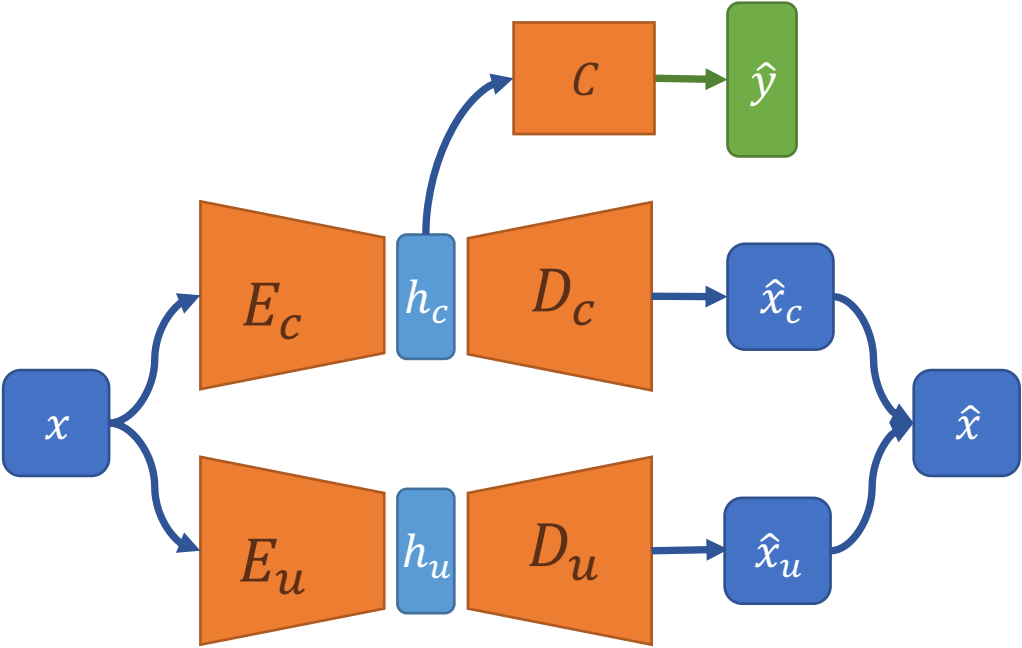
cats

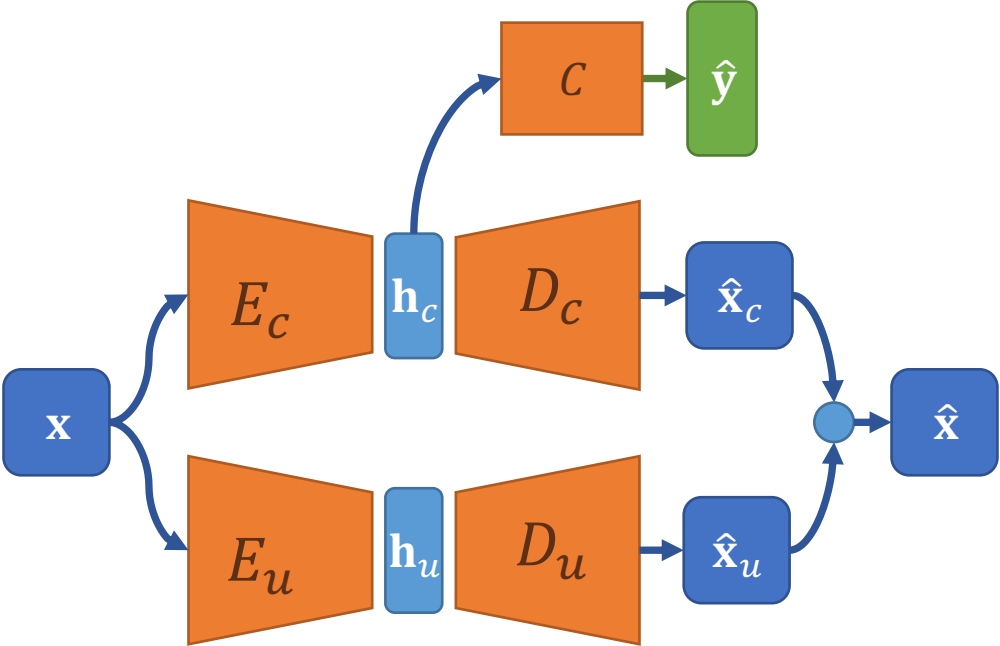


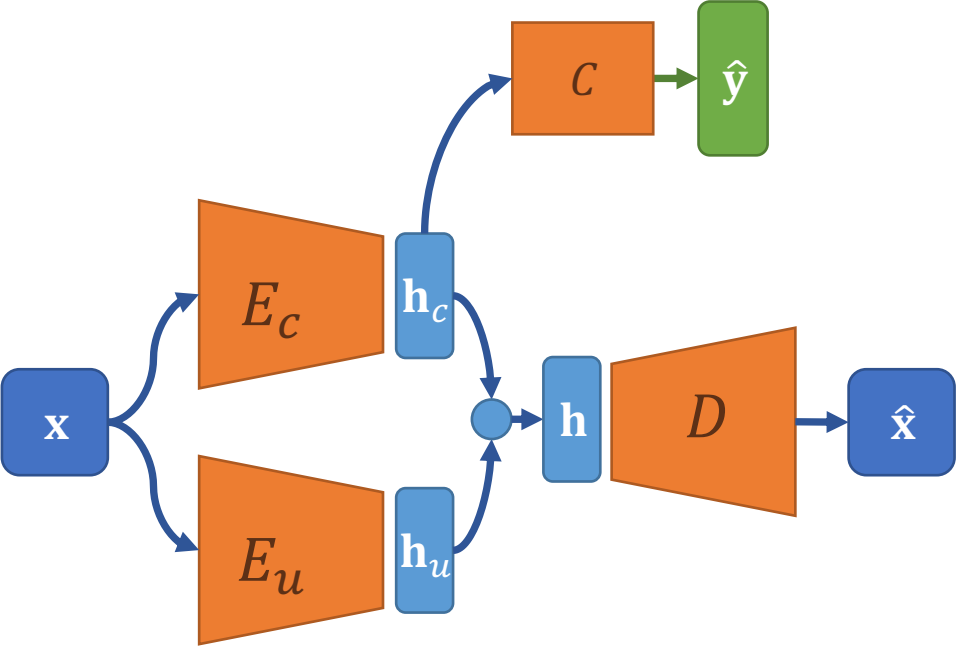
cars

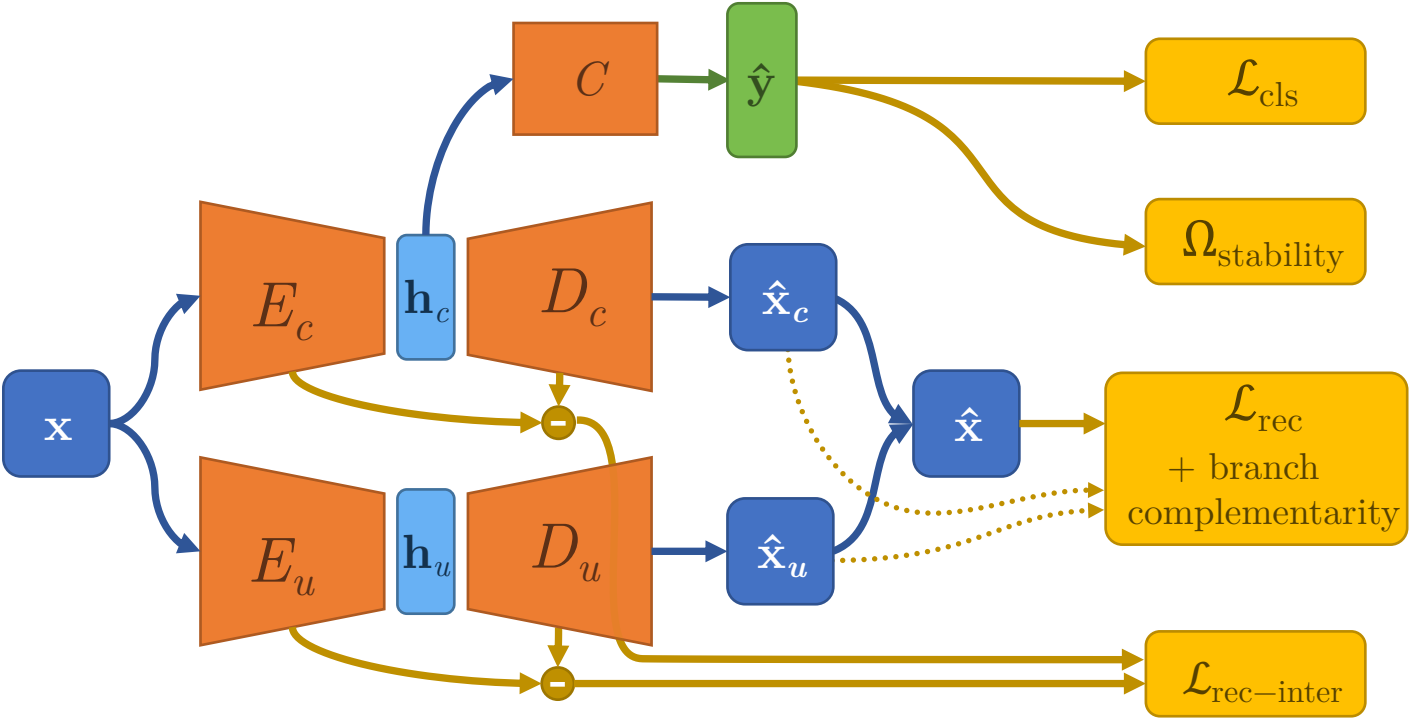


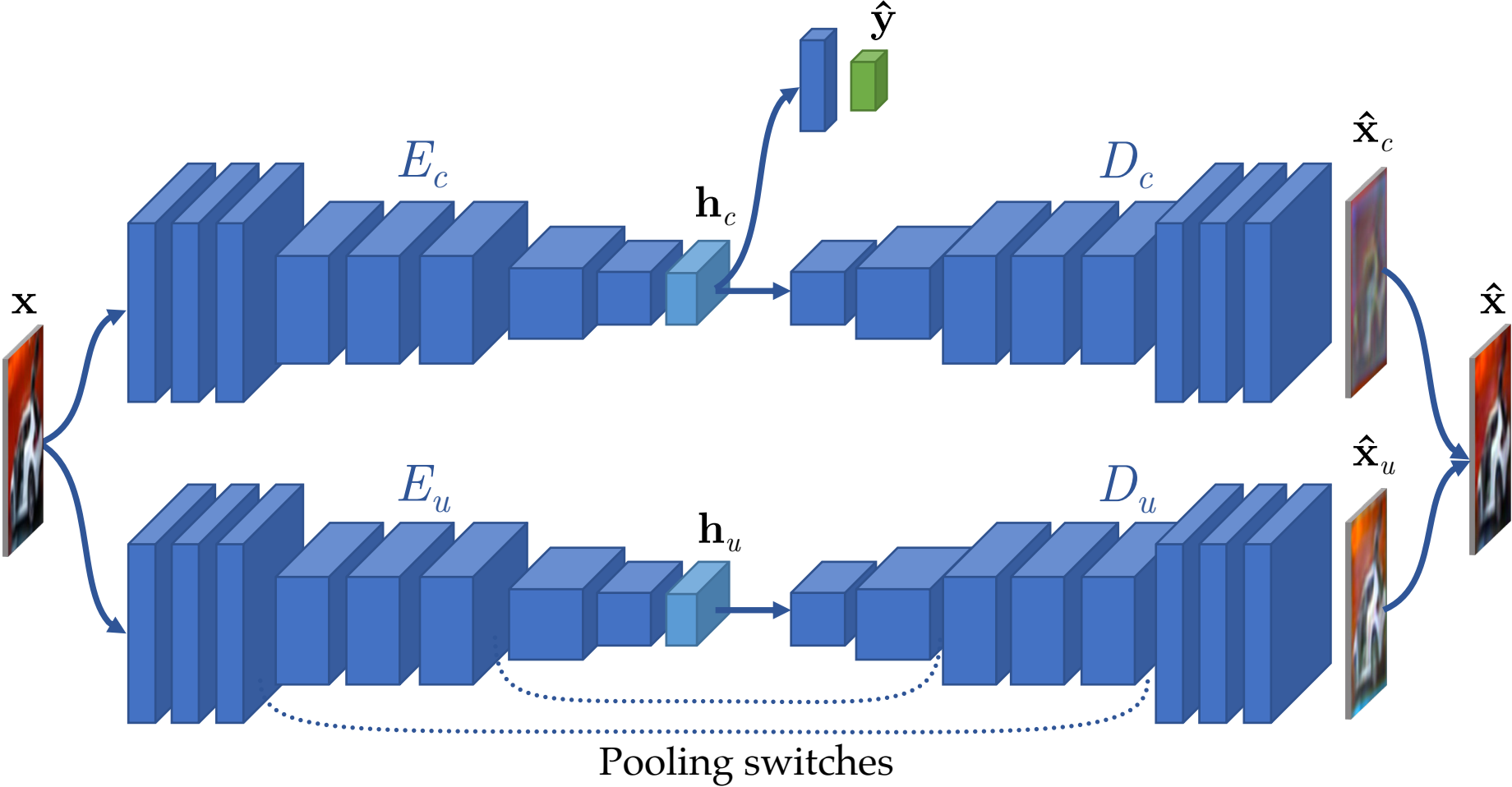












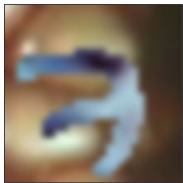
Input info.

Info. preserved in the encoding

Discriminative
Non-discriminative



x



AE



h

Corresponding
input data
represented

HybridNet

Classif.
branch



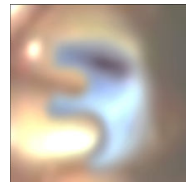
h_c

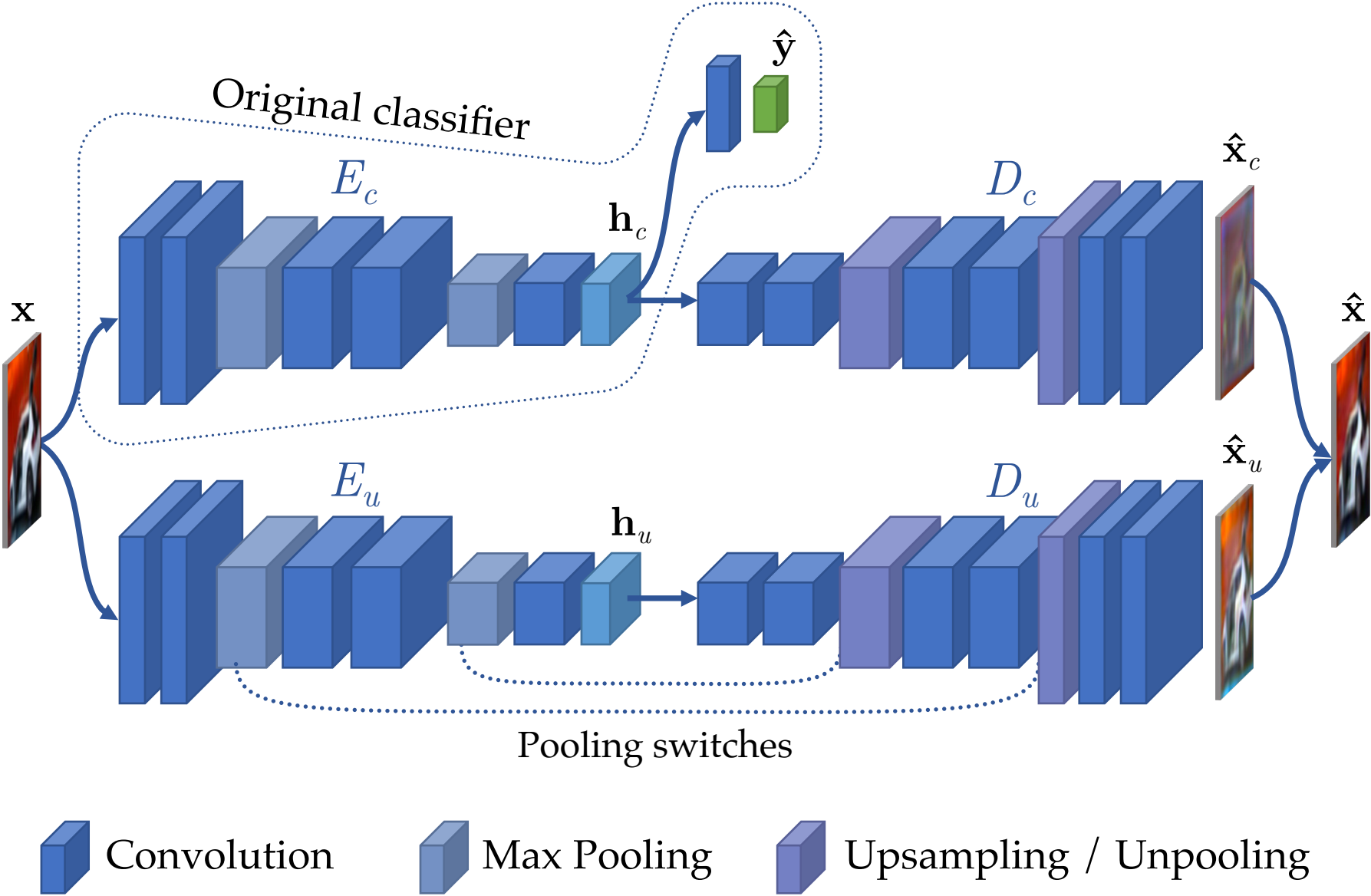


Unsup.
branch

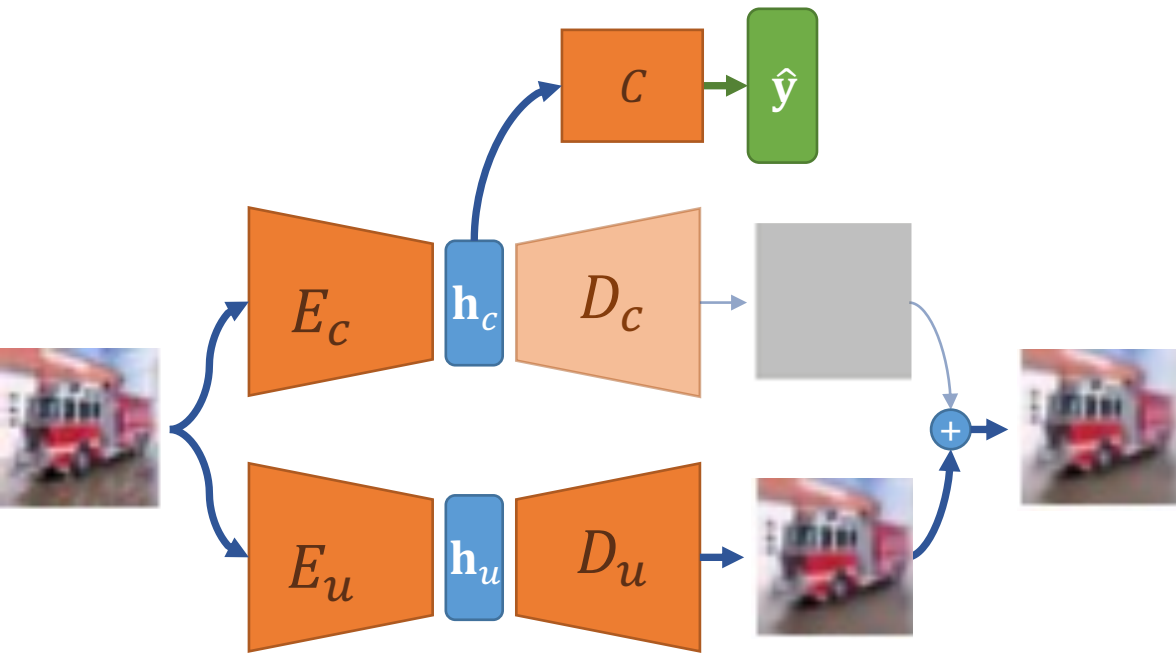


h_u

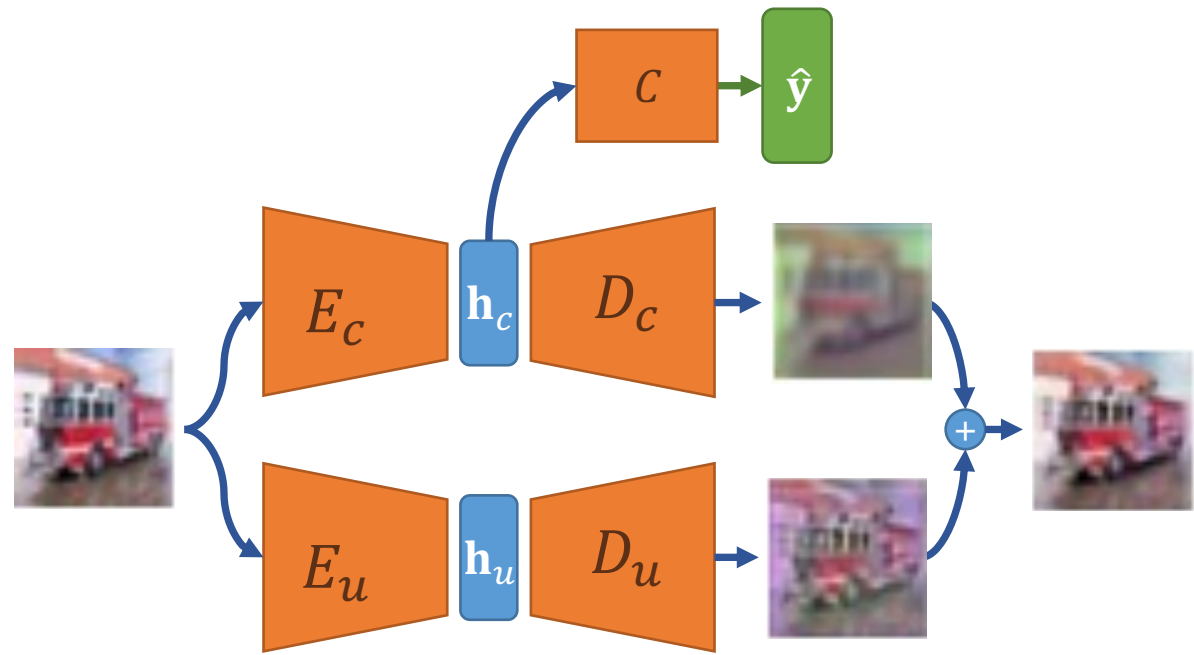


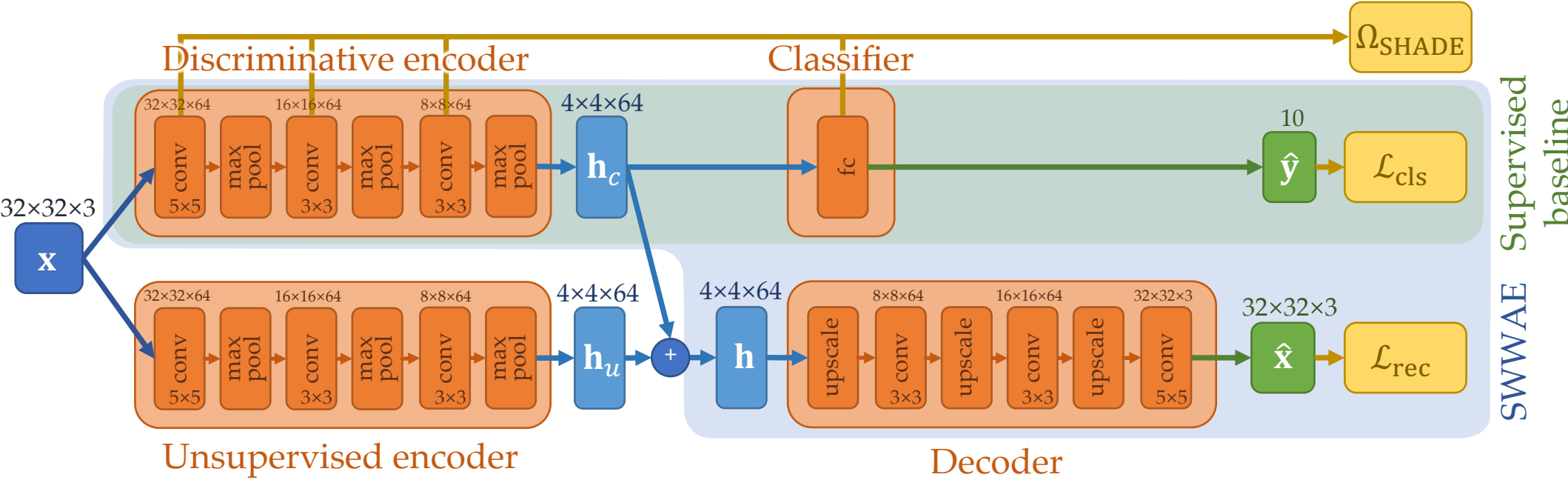


Without branch balancing



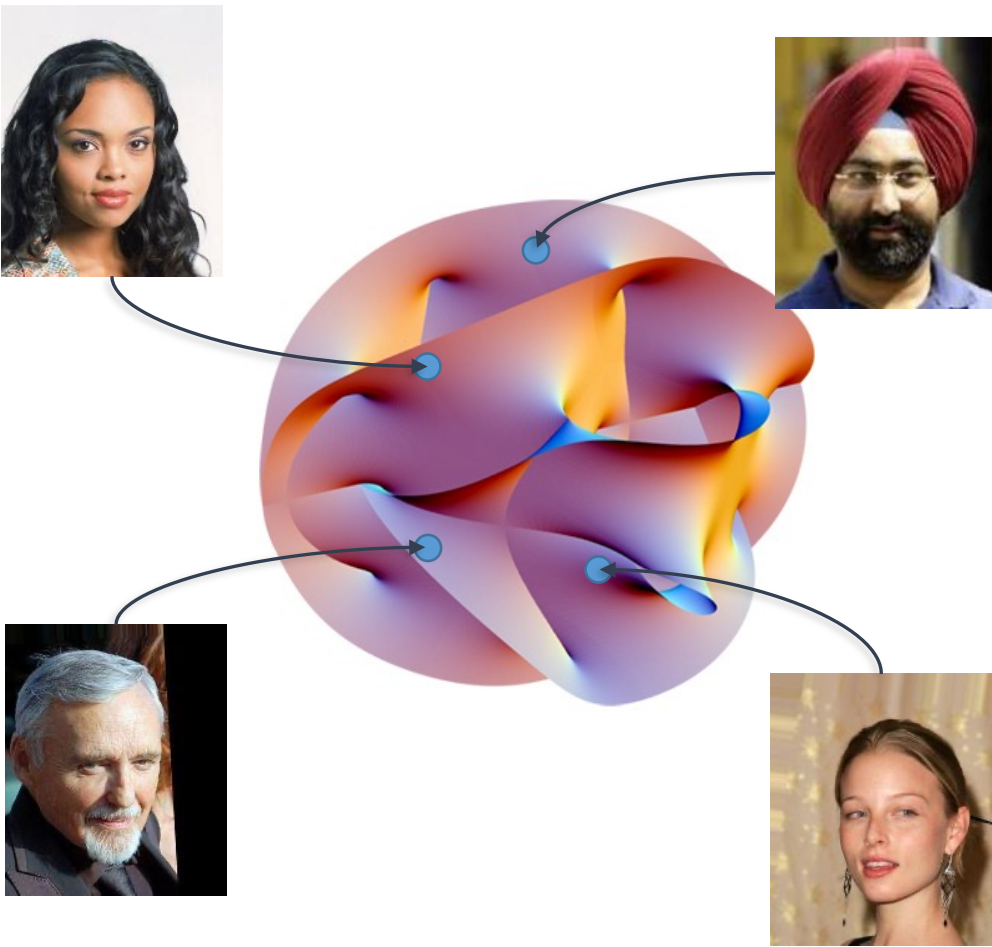
With branch balancing





Entangled manifold of input data

Illustrative view of the image manifold



Disentangling
model

Disentangled representation space

*Schematic view of
independent factors*



Latent representations of information domains



Disentangling

$$D(\bullet, \times) =$$

$$D(\bullet, \times) =$$

$$D(\bullet, \times) =$$

Image editing

$$D(\bullet, \times) =$$

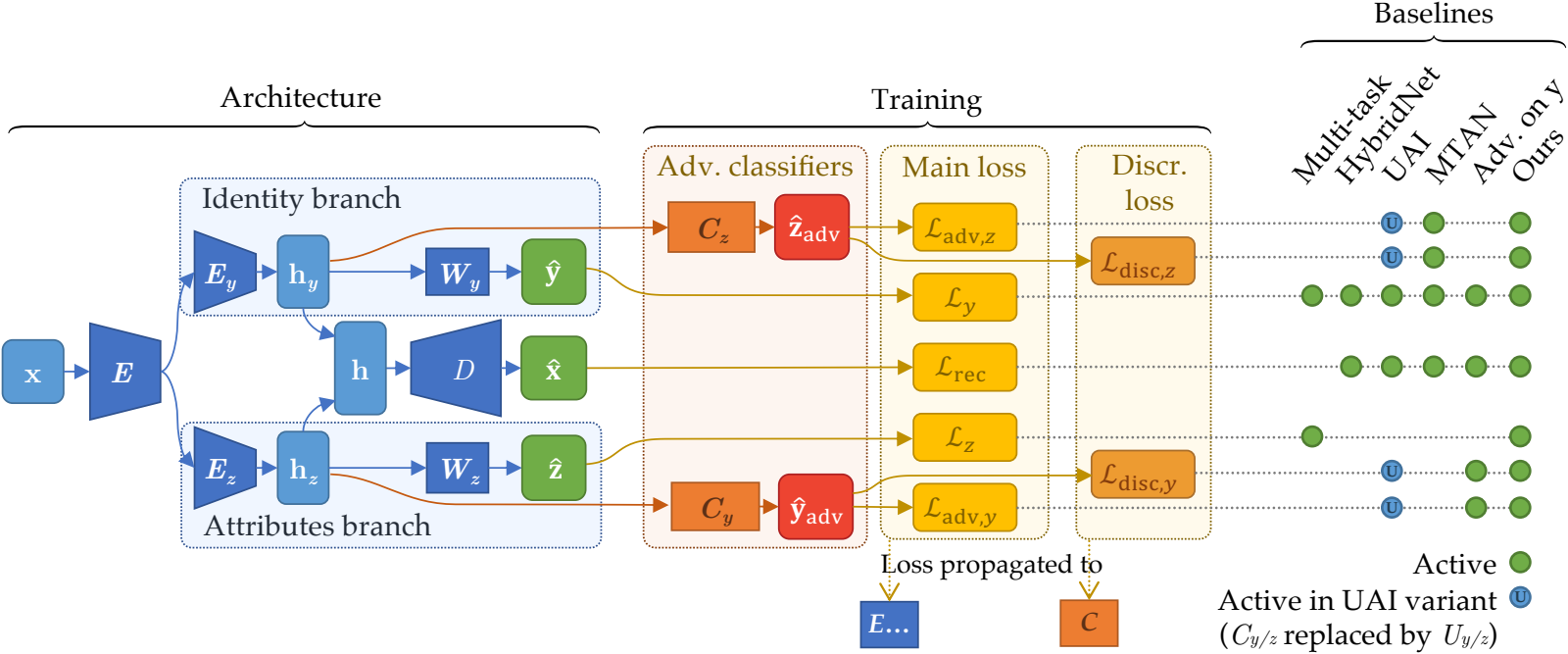
$$D(\bullet, \times) =$$

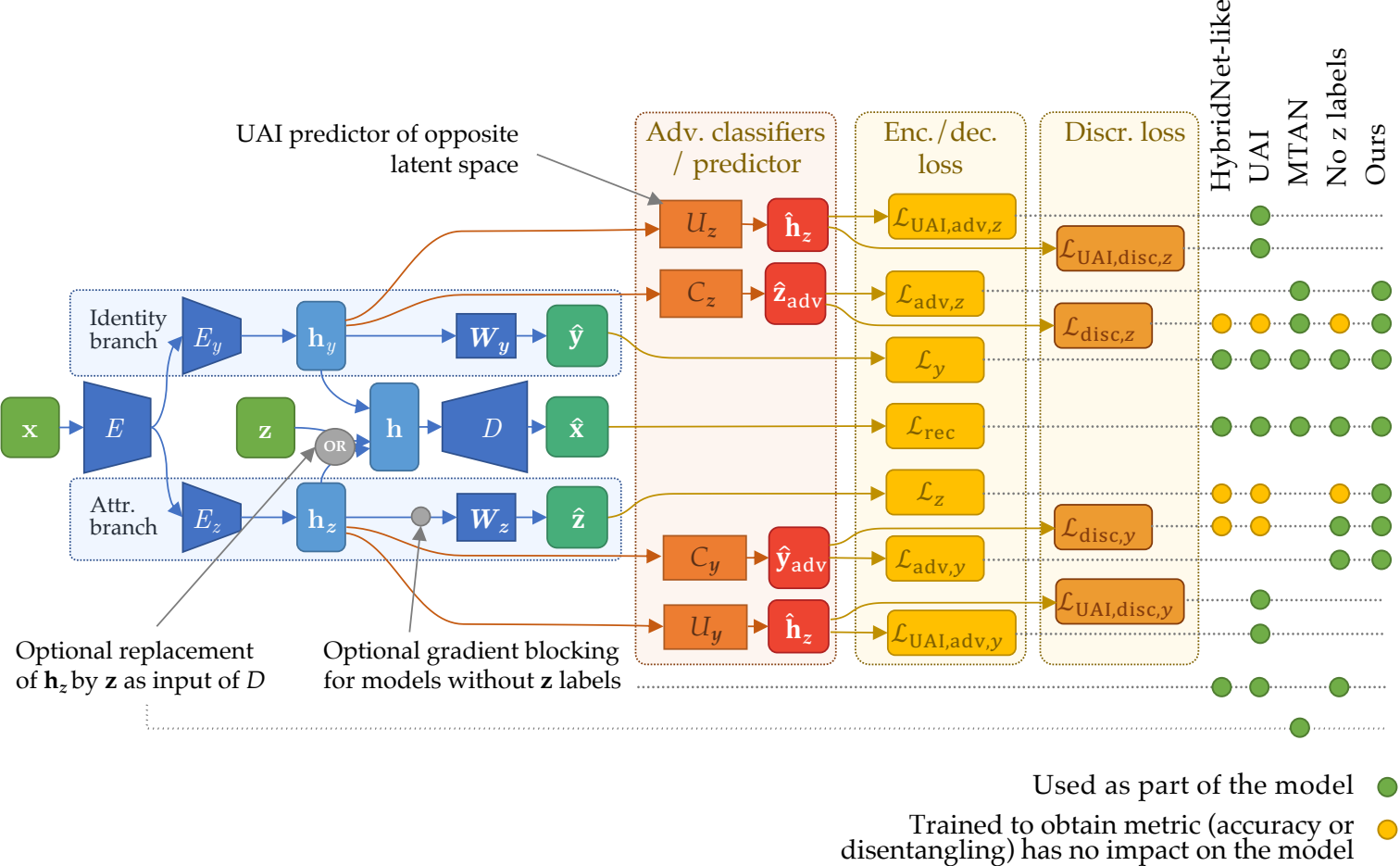
$$D(\bullet, \times) =$$

$$D(\bullet, \times) =$$

$$D(\bullet, \times) =$$

$$D(\bullet, \times) =$$





For “HybridNet-like + attr” and “UAI + attr”, simply remove the gradient blocking before W_z

