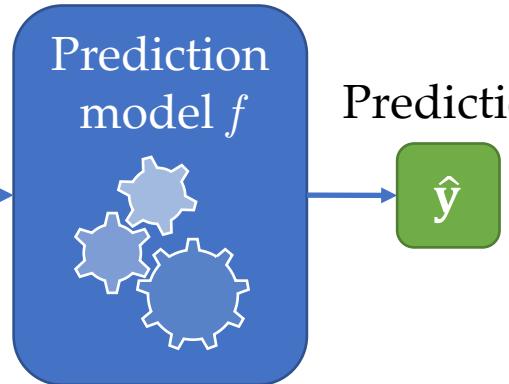


Input image to analyze

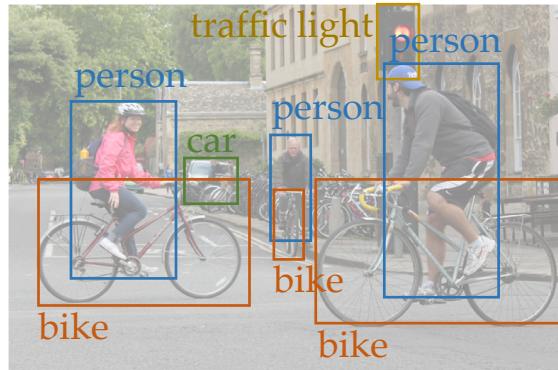


Examples of different tasks (i.e. types of prediction) in Computer Vision:

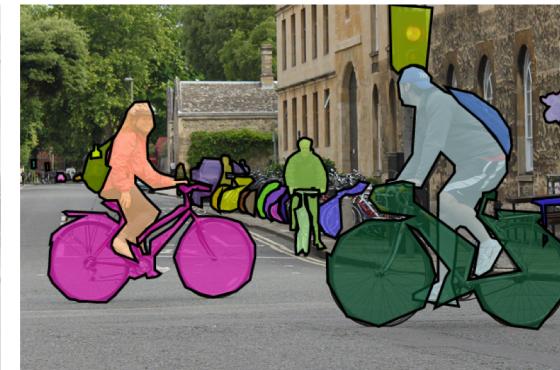
Classification

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

Object detection



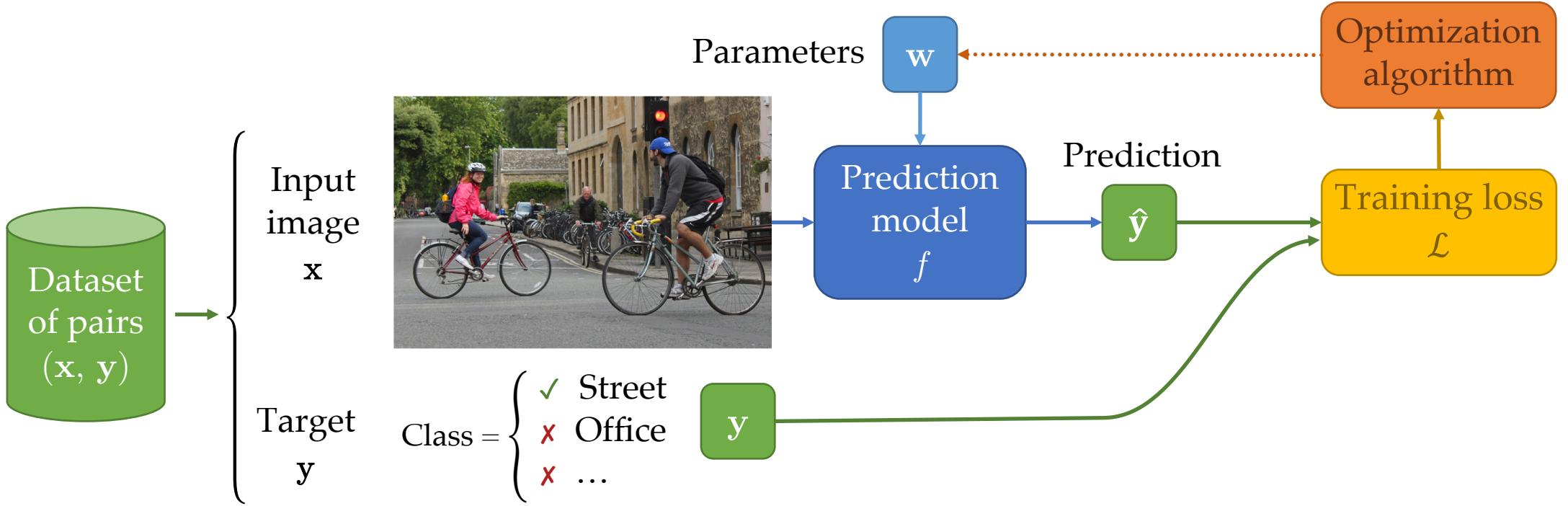
Segmentation



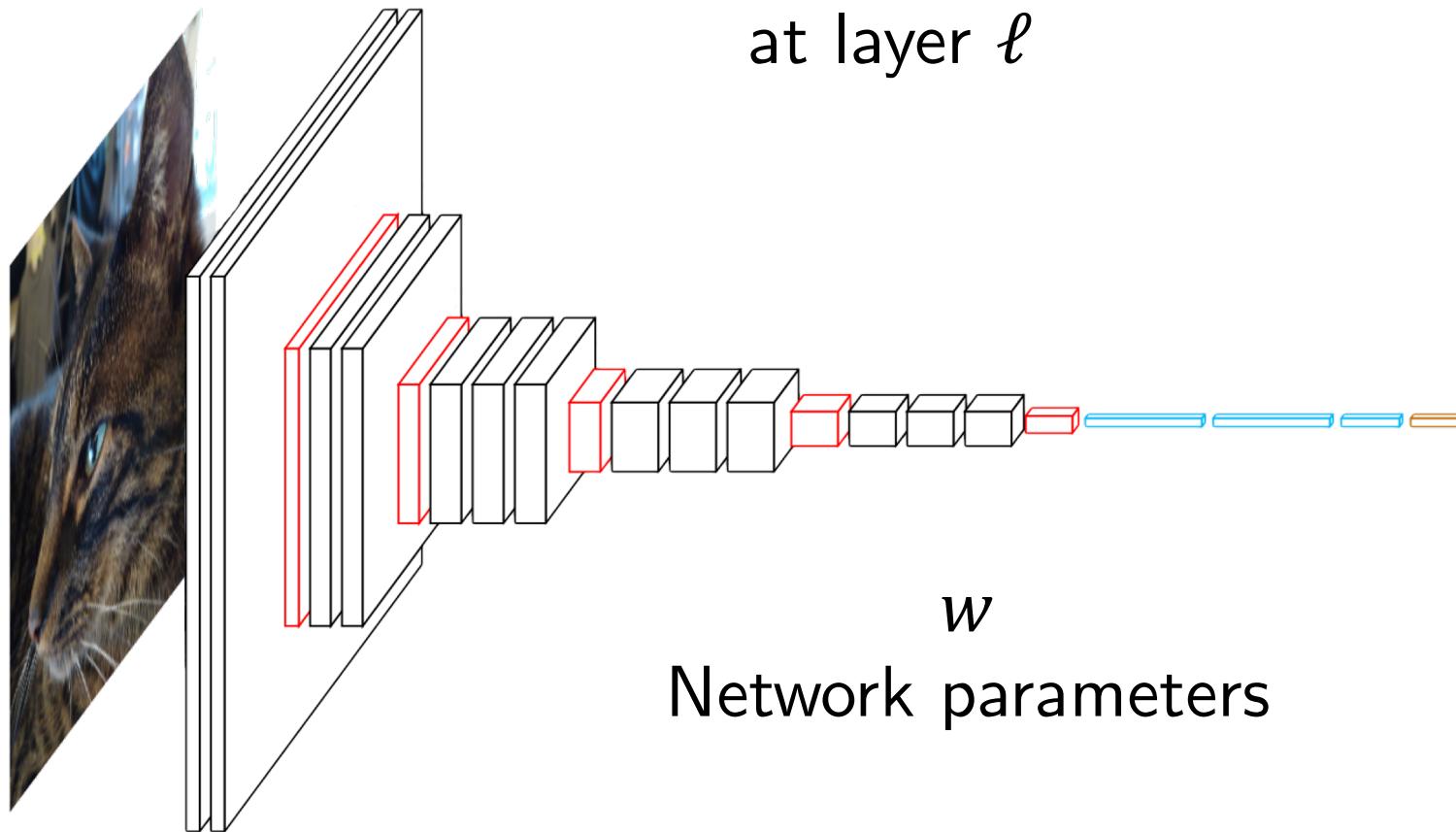
Captioning

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

...



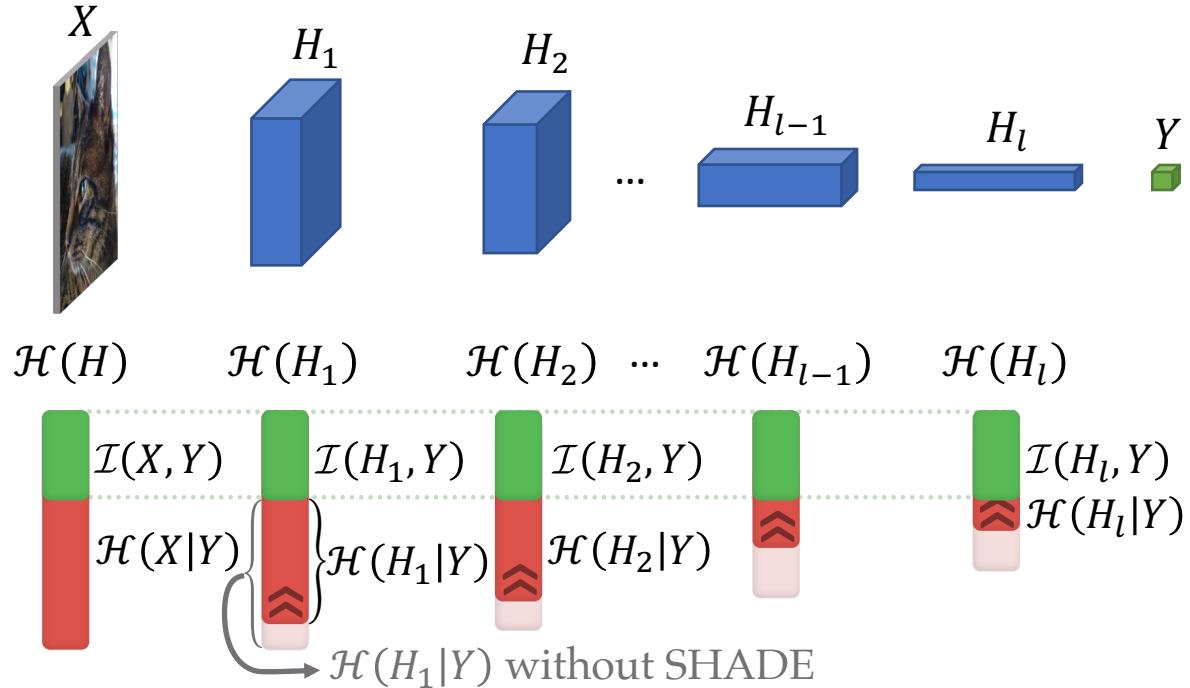
X
input image



$Y_\ell \quad \forall \ell$
Feature representation
at layer ℓ

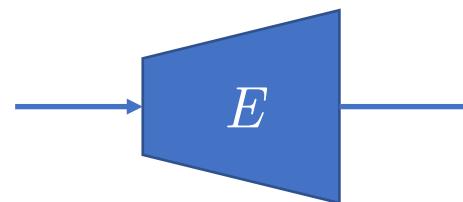
C
class label

w
Network parameters

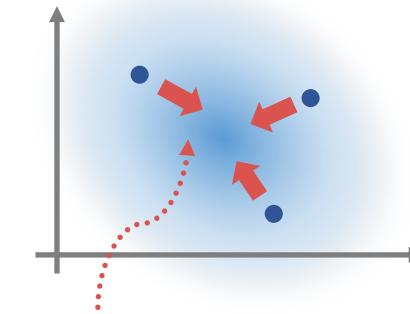


≈ Effect of SHADE: Reduction of conditional entropy

Input images

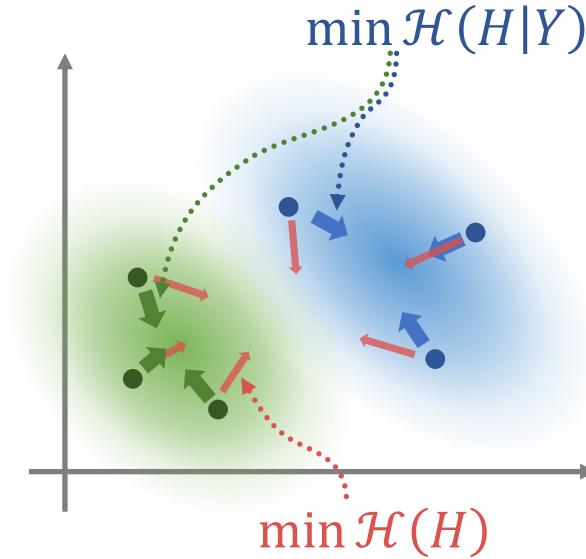
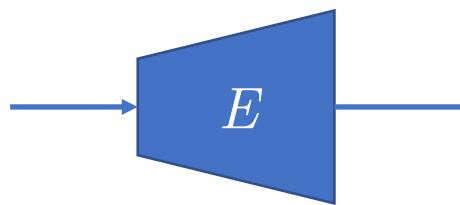


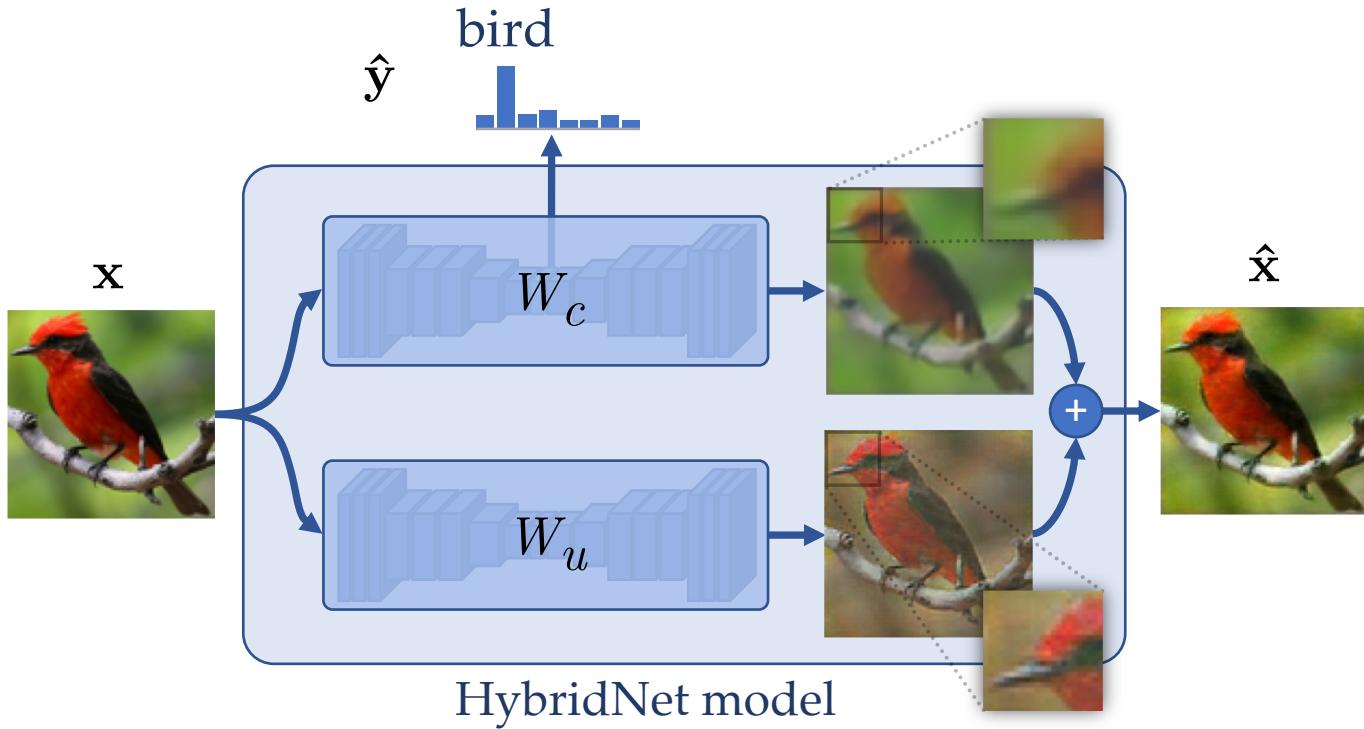
Feature representation

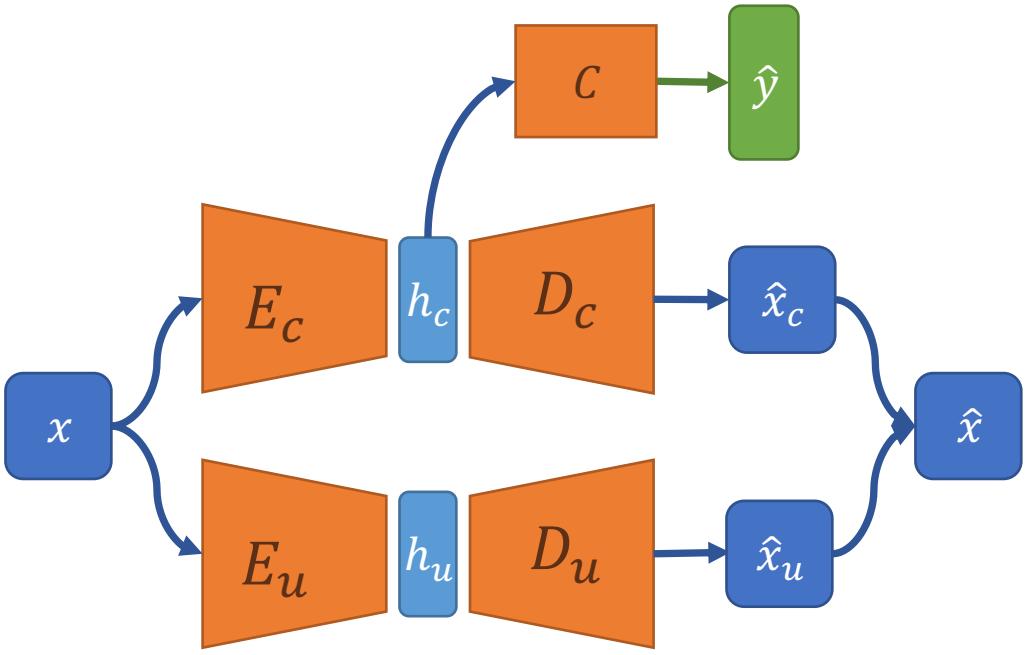


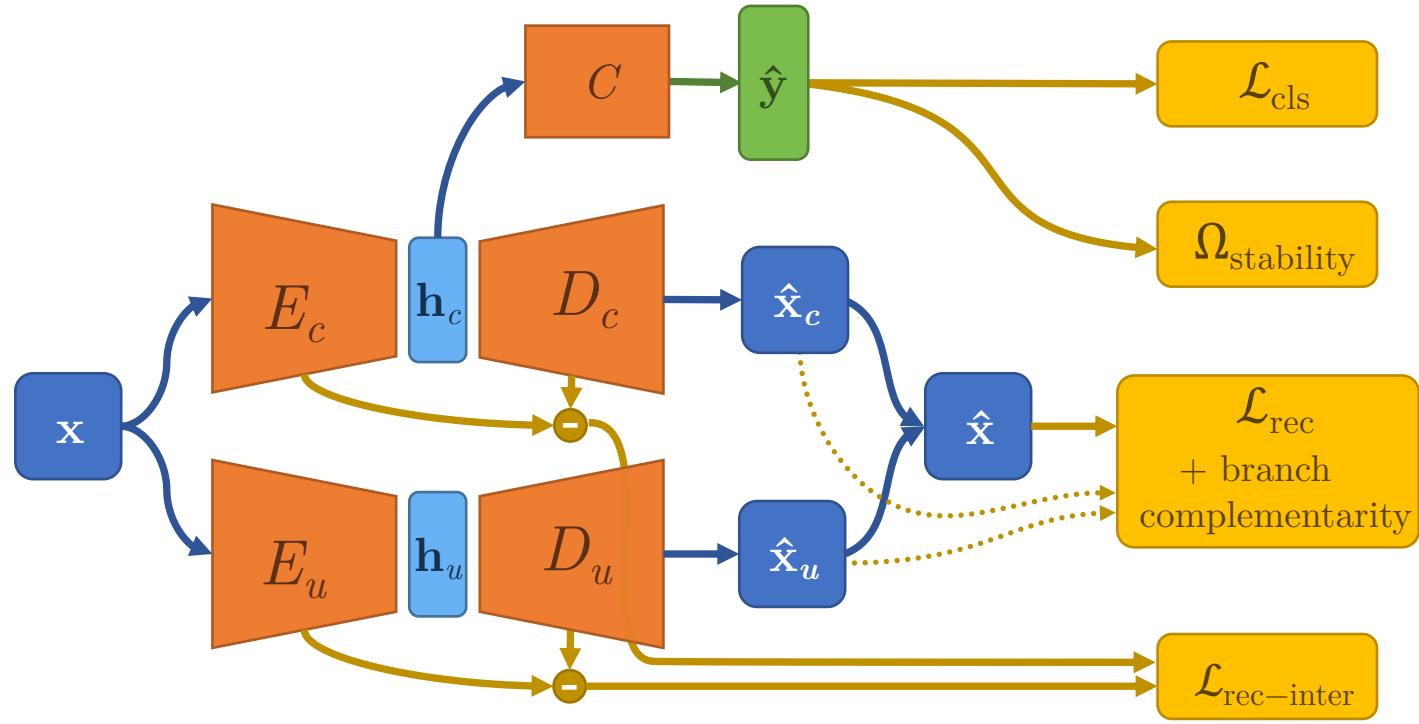
$$\min \mathcal{H}(H) \Rightarrow \max \mathcal{H}(X|H)$$

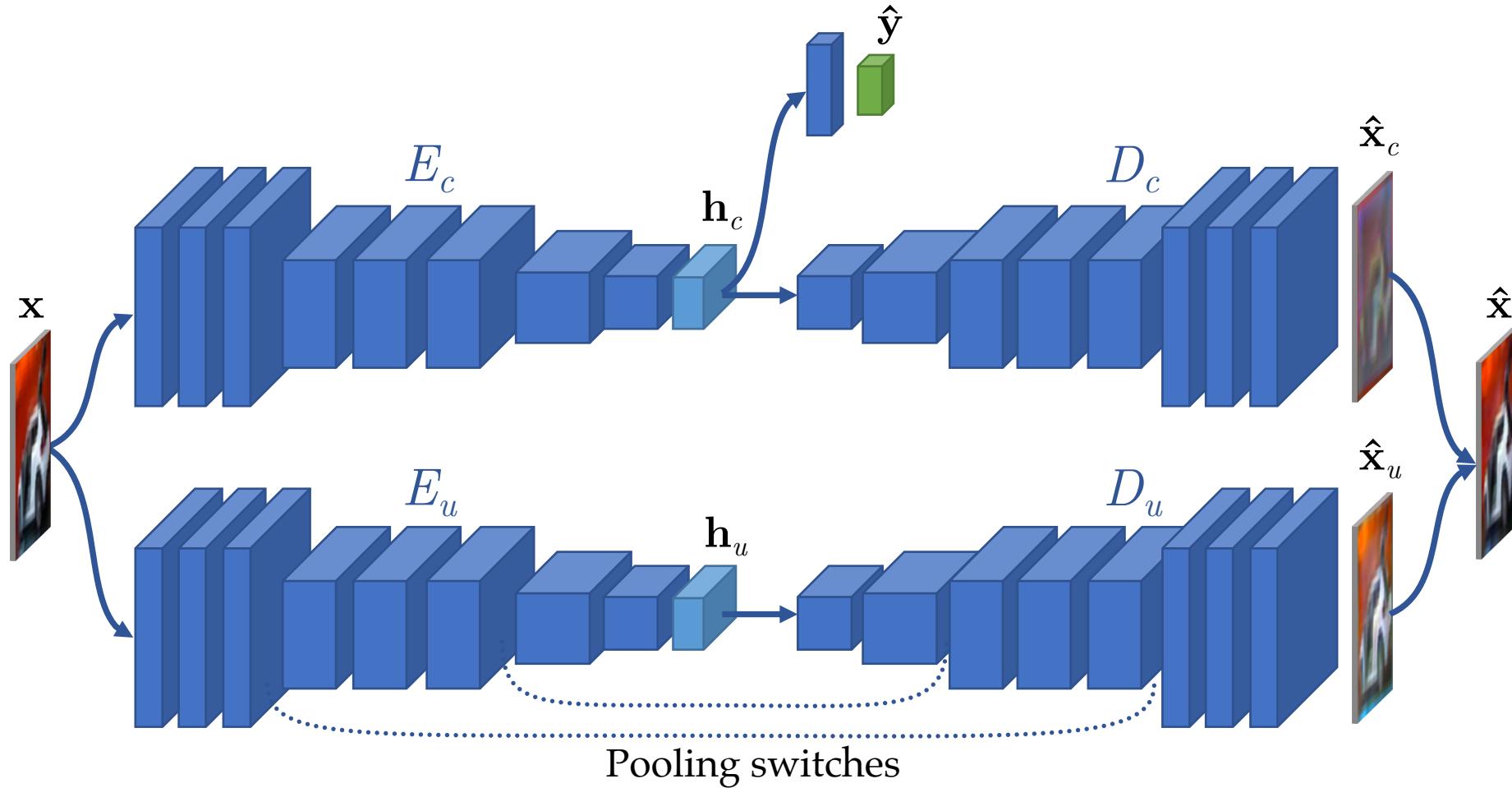
increases uncertainty in X given H

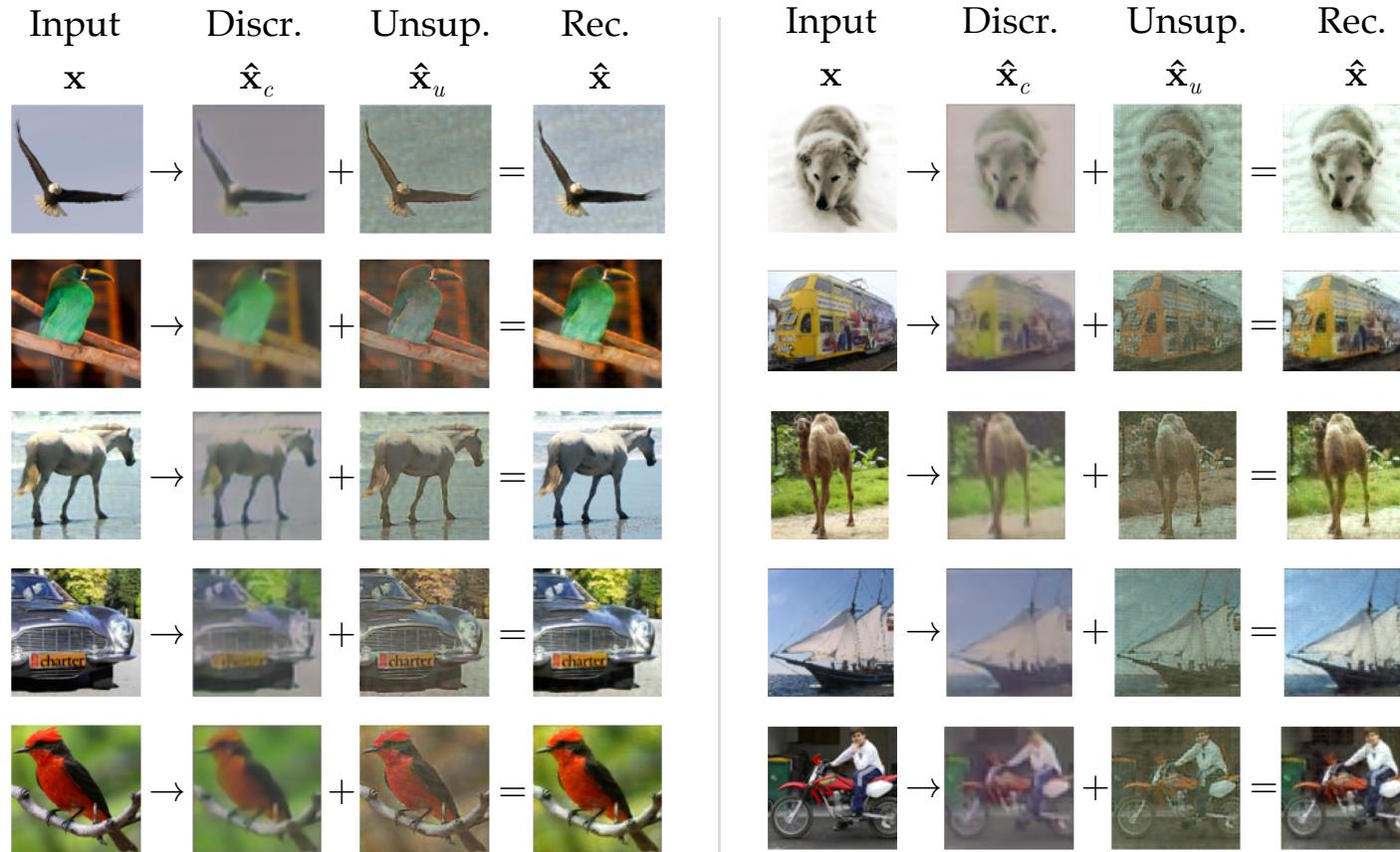




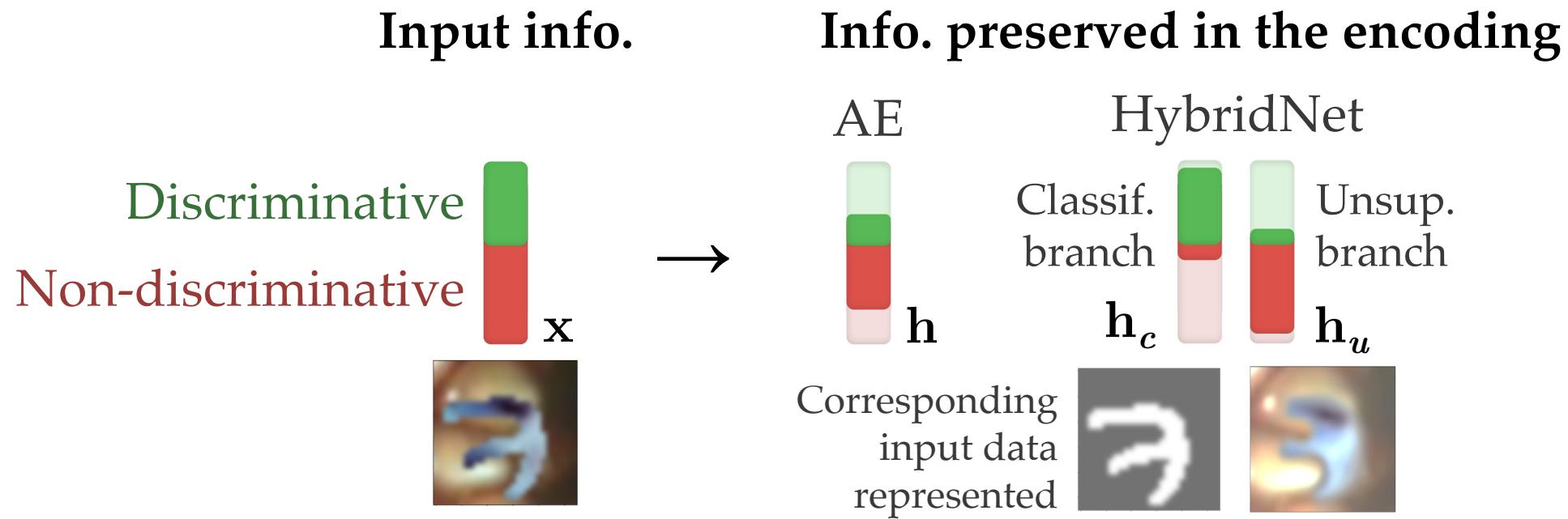


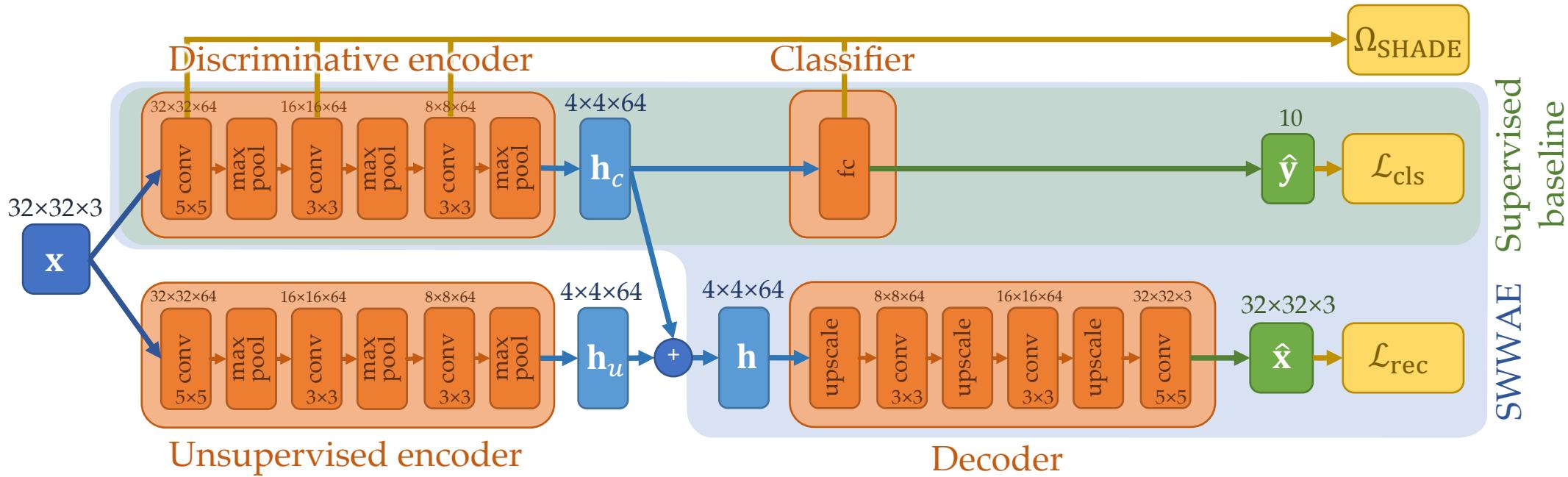




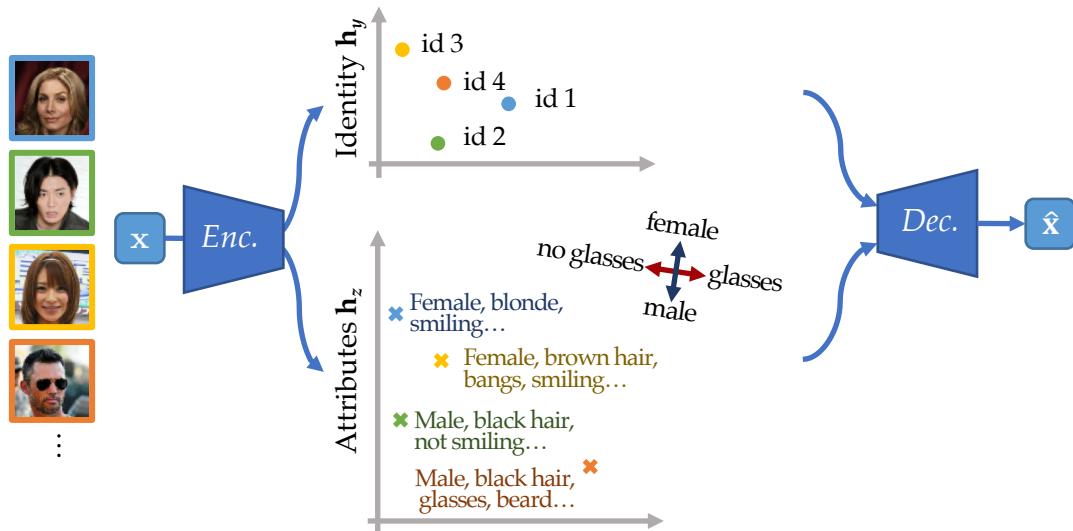


	Final rec. Intermed. rec. Complemen. Stability	Model accuracy	Visualisations											
			x	\hat{x}_c	\hat{x}_u	\hat{x}	x	\hat{x}_c	\hat{x}_u	\hat{x}	x	\hat{x}_c	\hat{x}_u	\hat{x}
(1)	✓	72.4												
(2)	✓ ✓	74.0												
(3)	✓ ✓ ✓	75.2												
(4)	✓ ✓ ✓ ✓	81.6												
(1)	✓	72.4												
(2)	✓ ✓	74.0												
(3)	✓ ✓ ✓	75.2												
(4)	✓ ✓ ✓ ✓	81.6												





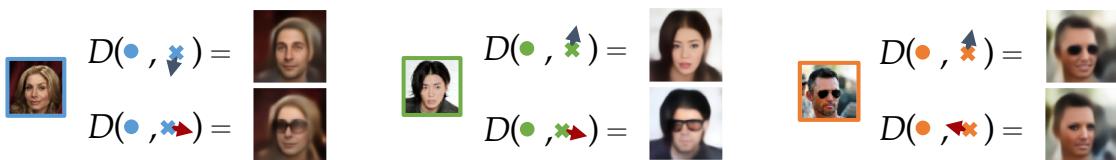
Latent representations of information domains

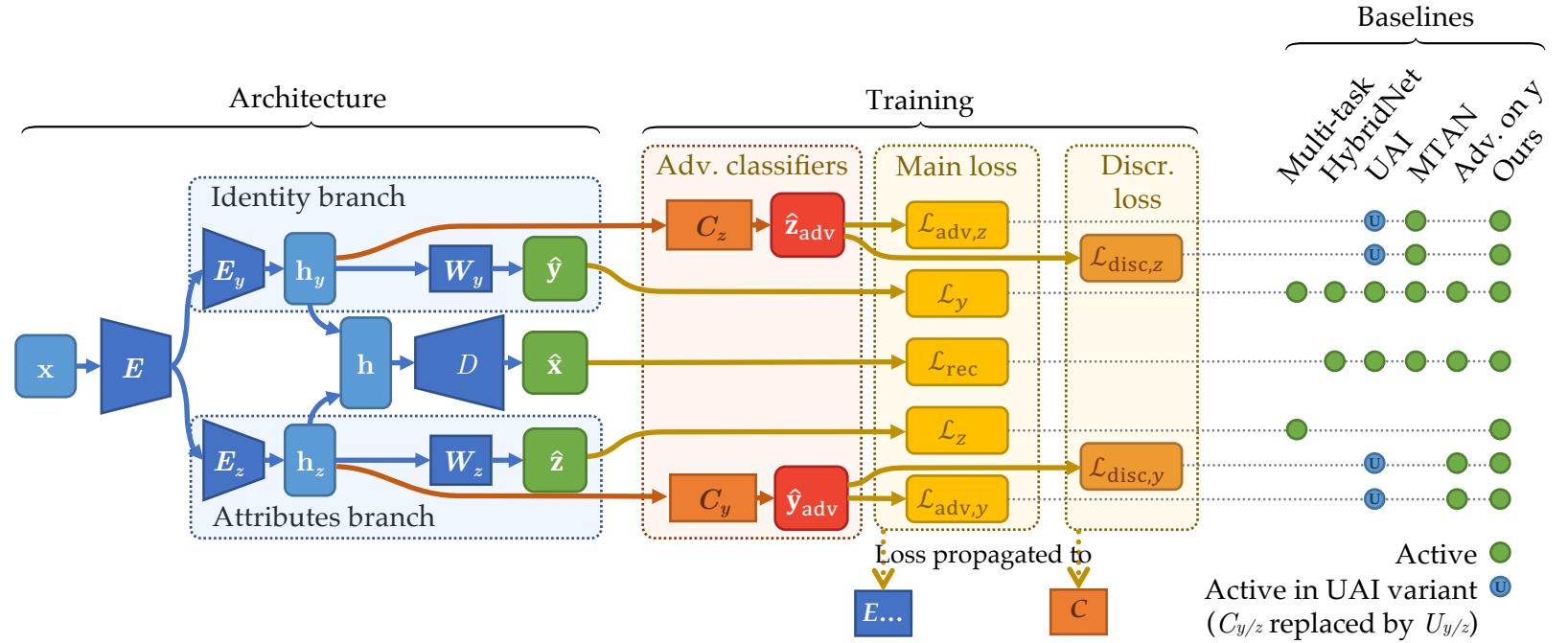


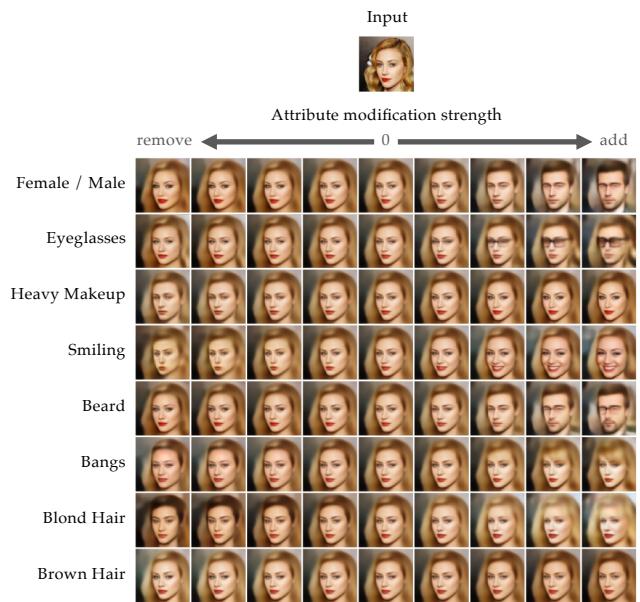
Disentangling

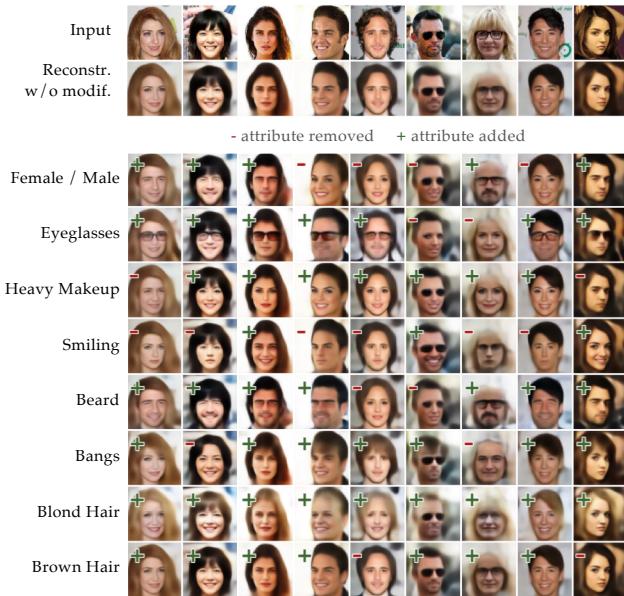


Image editing







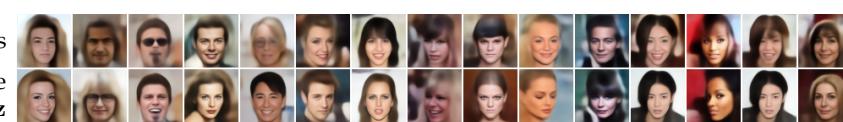


Initial images. $\mathbf{x}^{(\text{id})}$: Identity source / $\mathbf{x}^{(\text{attr})}$: Attribute source

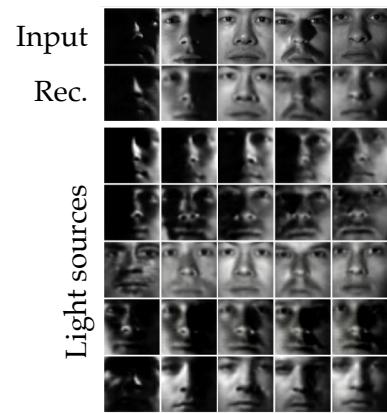


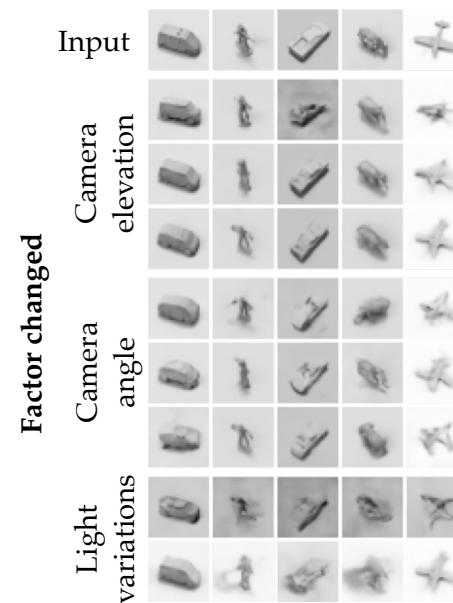
DualDis

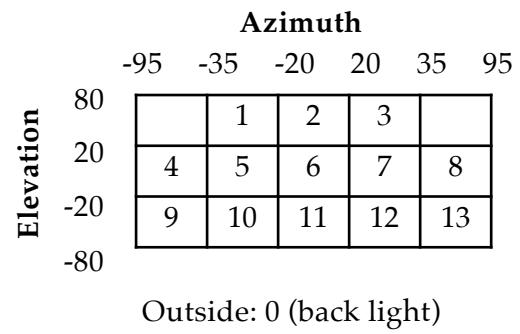
Baseline
without \mathbf{z}

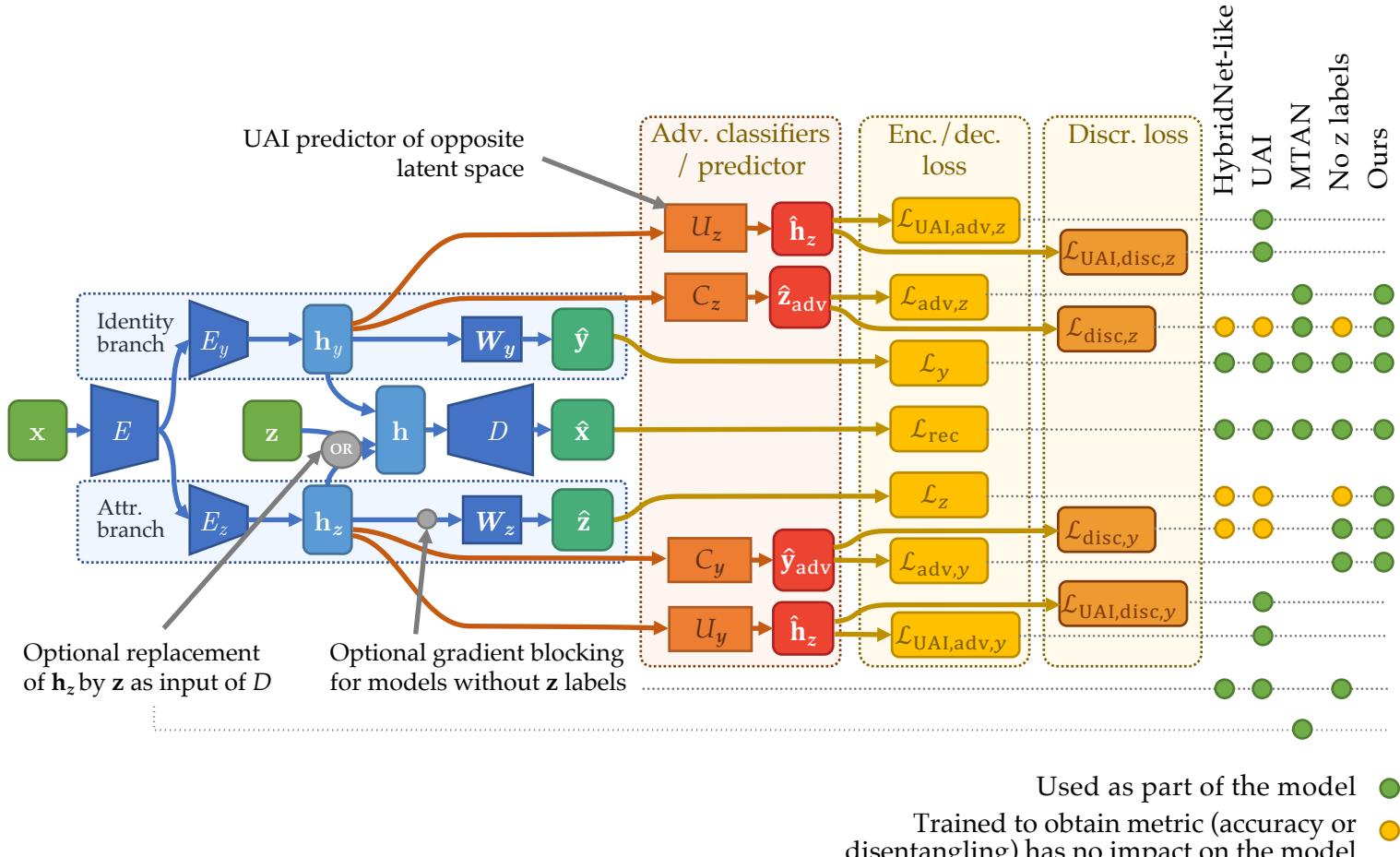


Generations from $(\mathbf{h}_y^{(\text{id})}, \mathbf{h}_z^{(\text{attr})})$ produced by DualDis and the baseline





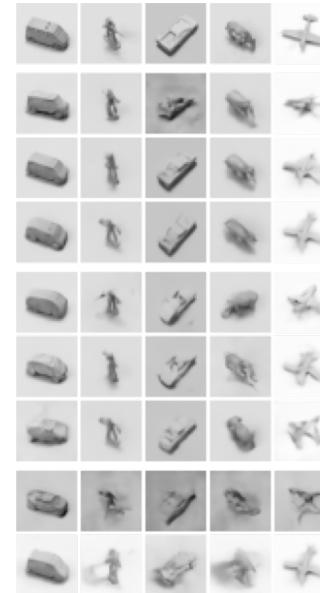


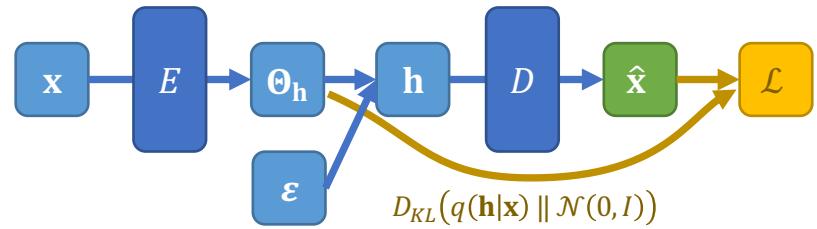


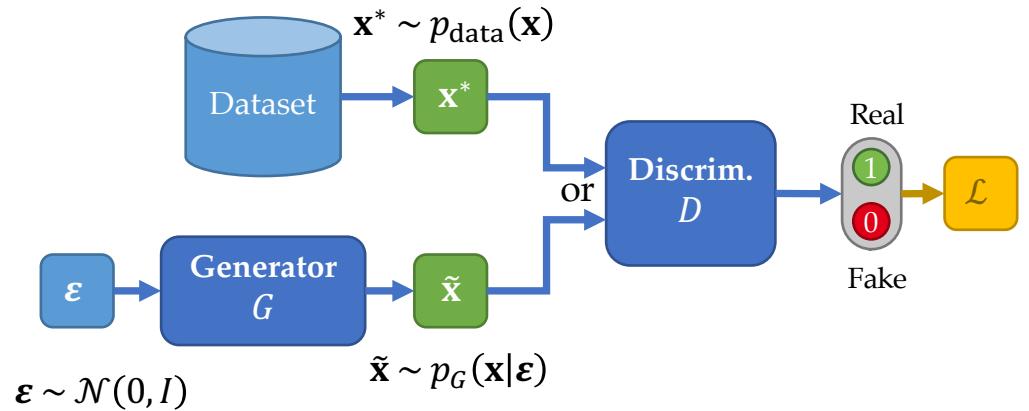
For "HybridNet-like + attr" and "UAI + attr", simply remove the gradient blocking before W_z

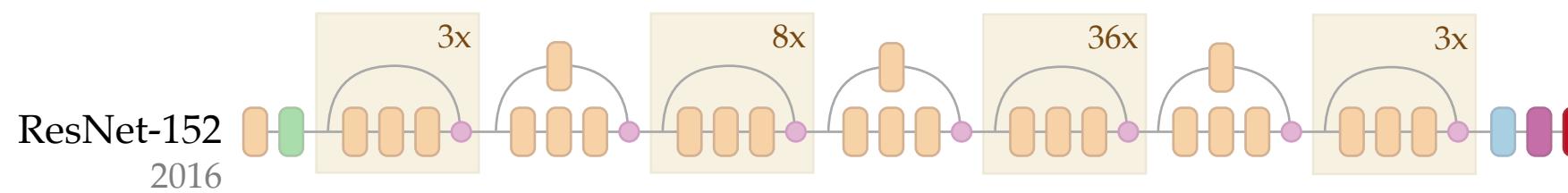
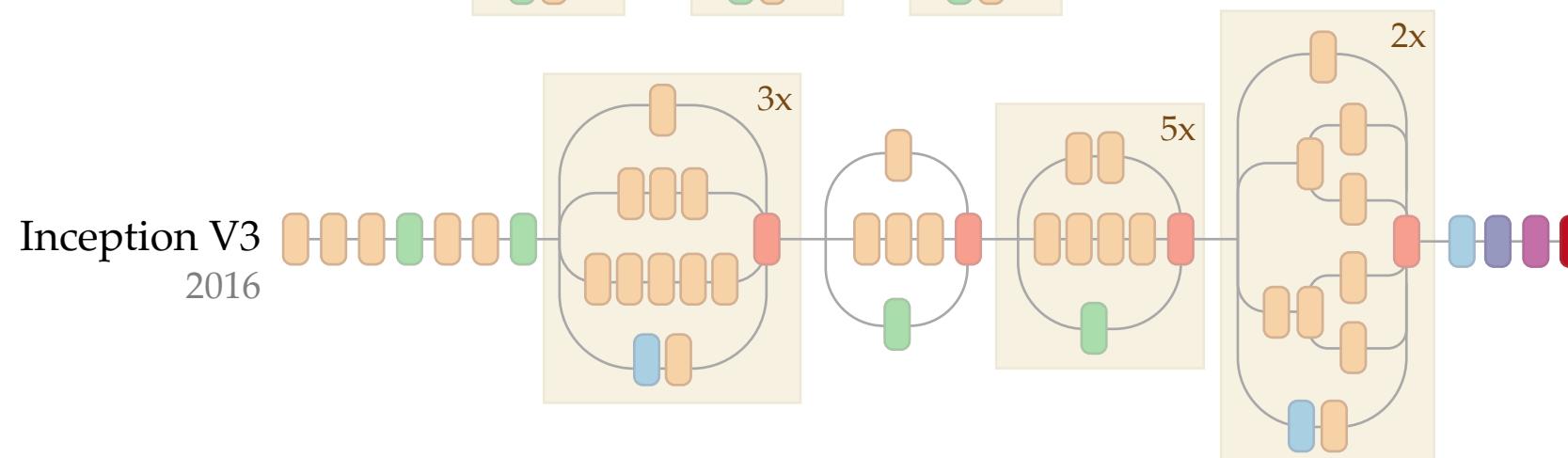
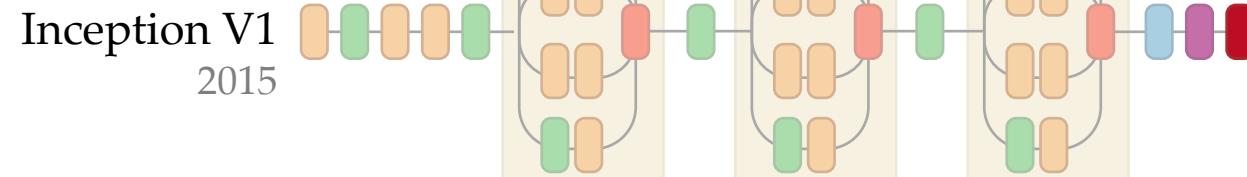
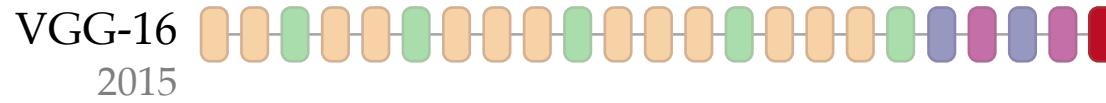
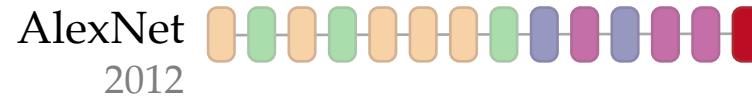
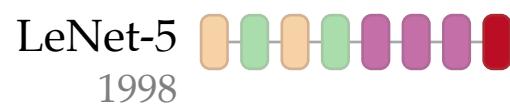


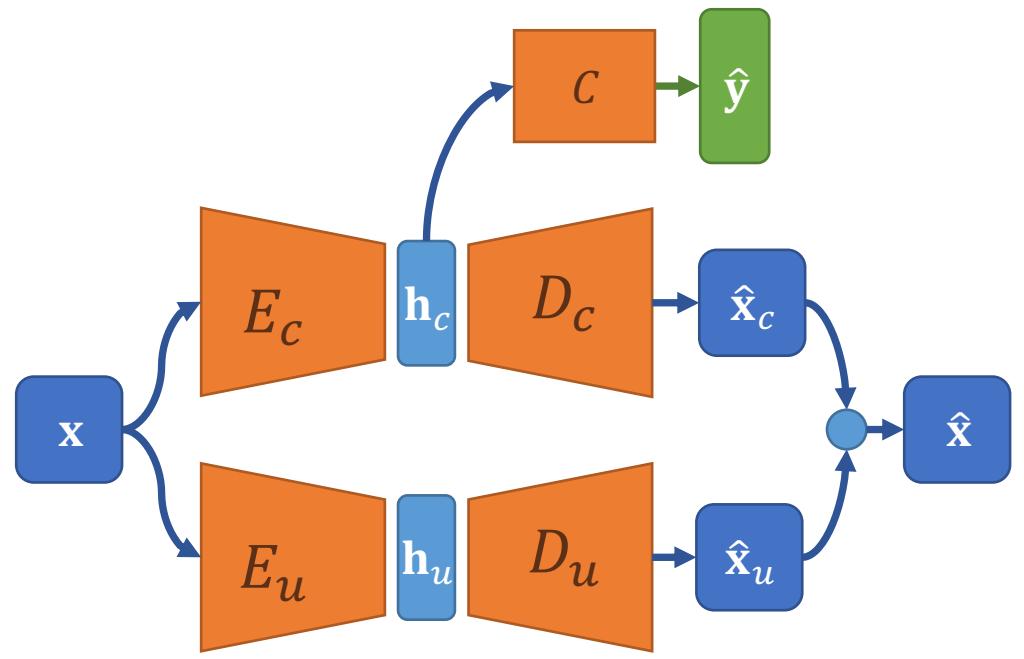
(a) Attribute (lightning source) editing on Yale

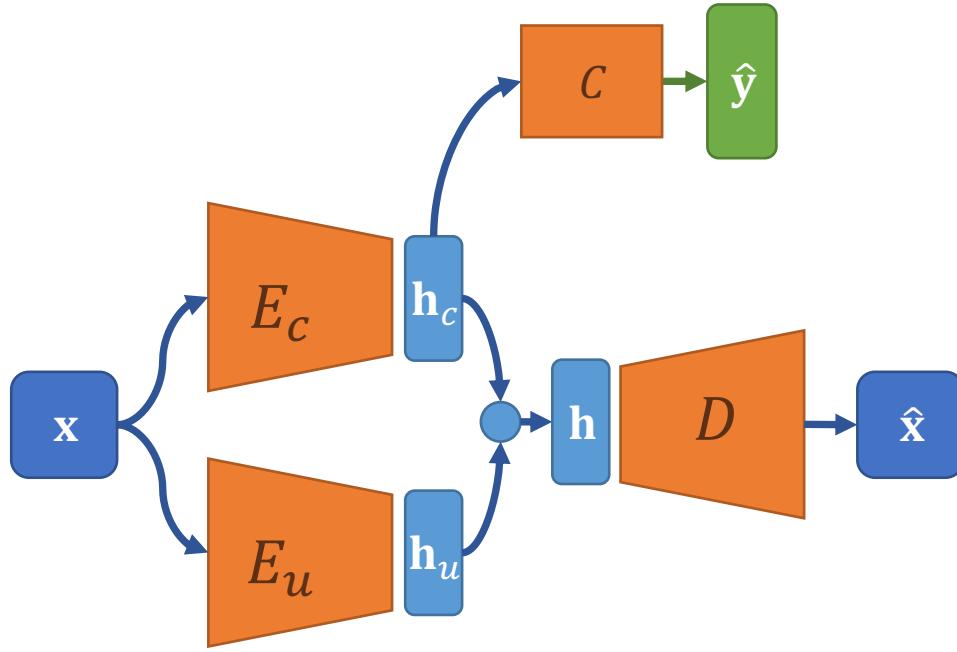


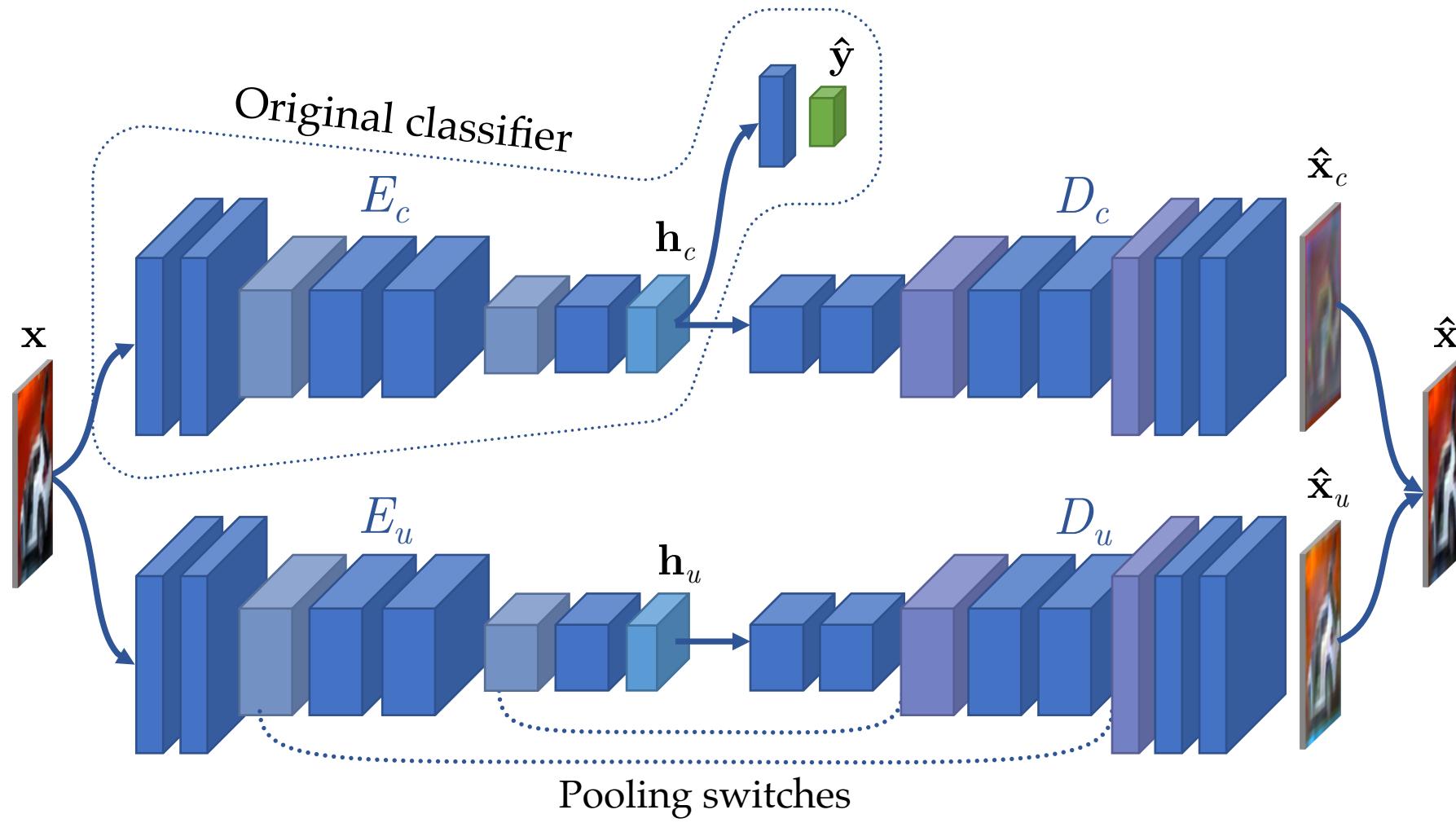








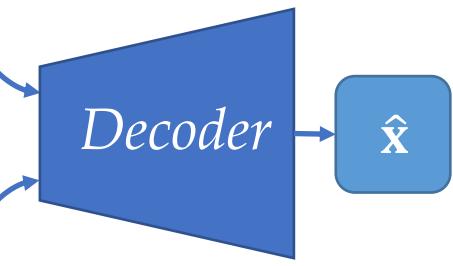
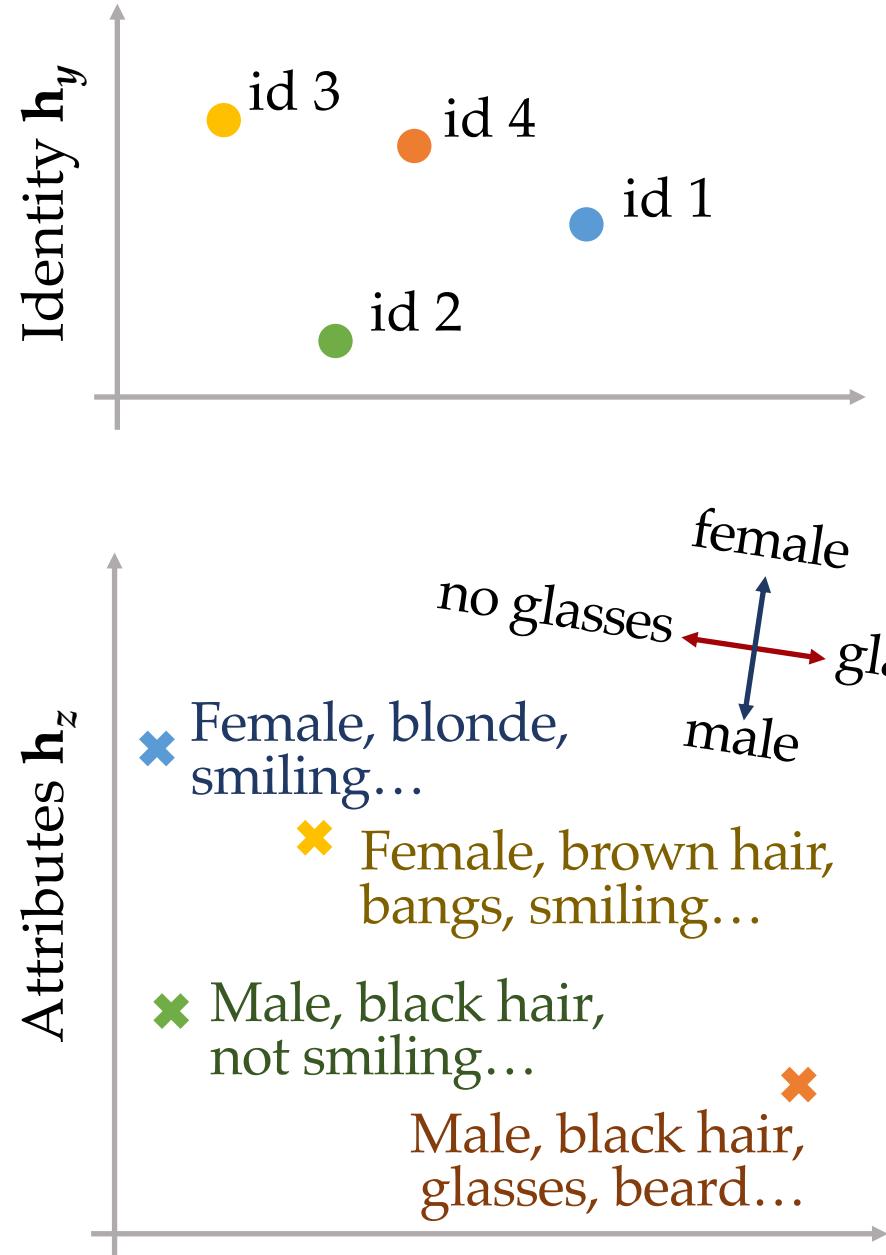
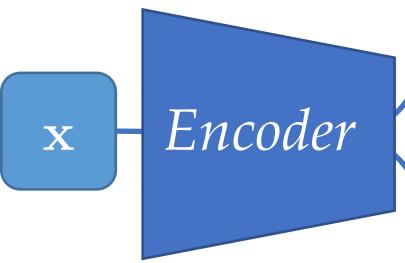




Convolution Max Pooling Upsampling / Unpooling

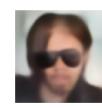


⋮





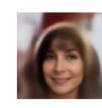
$$D(\bullet, \times) =$$

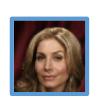


$$D(\bullet, \times) =$$



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





$$D(\bullet, \textcolor{blue}{\ddagger}) = \begin{matrix} \text{Portrait of a man with blonde hair, wearing a blue frame.} \end{matrix}$$

$$D(\bullet, \textcolor{red}{\star}) = \begin{matrix} \text{Portrait of a woman with blonde hair, wearing a black frame.} \end{matrix}$$



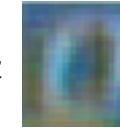
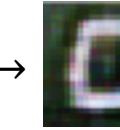
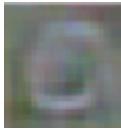
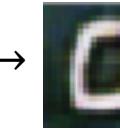
$$D(\bullet, \textcolor{green}{\ddagger}) = \begin{matrix} \text{Portrait of a woman with dark hair, wearing a green frame.} \end{matrix}$$

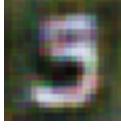
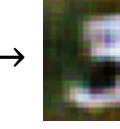
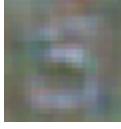
$$D(\bullet, \textcolor{red}{\star}) = \begin{matrix} \text{Portrait of a man with dark hair, wearing a black frame.} \end{matrix}$$

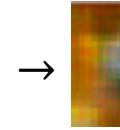
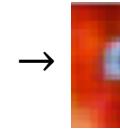


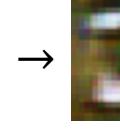
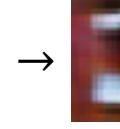
$$D(\bullet, \textcolor{orange}{\ddagger}) = \begin{matrix} \text{Portrait of a man with dark hair, wearing an orange frame.} \end{matrix}$$

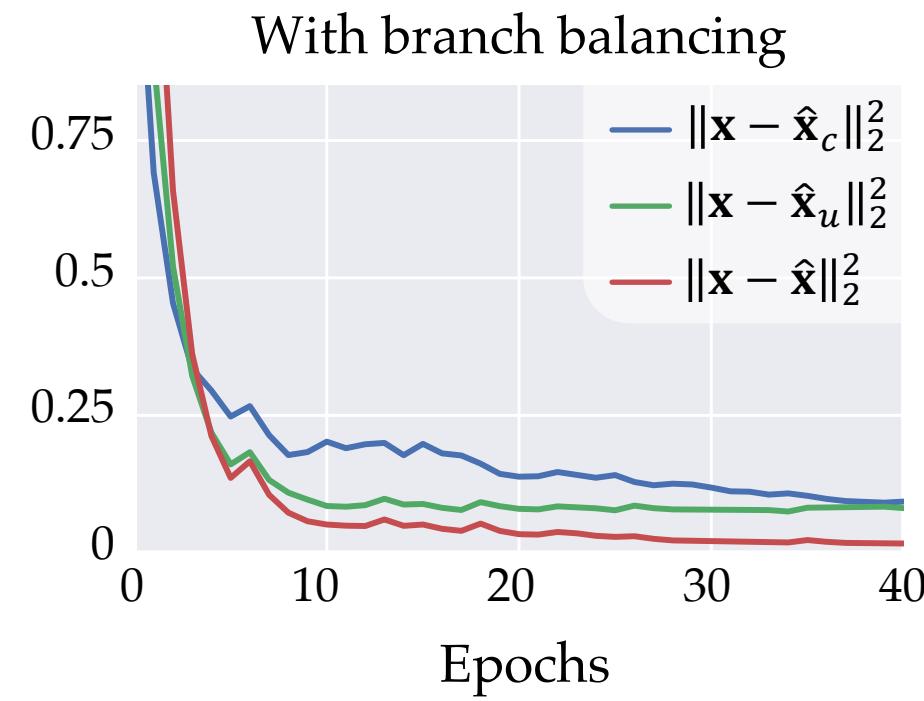
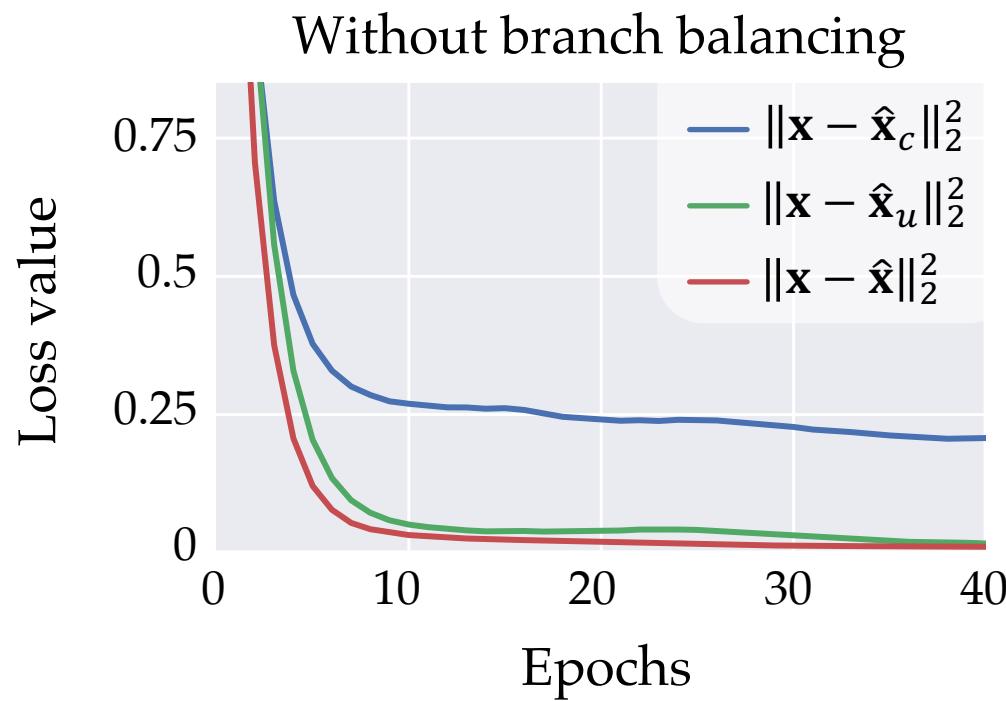
$$D(\bullet, \textcolor{red}{\star}) = \begin{matrix} \text{Portrait of a woman with dark hair, wearing an orange frame.} \end{matrix}$$

	Discrim. branch reconstr.	Unsup. branch reconstr.	Final reconstr.	Original input	
	$\hat{\mathbf{x}}_c$	$\hat{\mathbf{x}}_u$	$\hat{\mathbf{x}}$	\mathbf{x}	
No stability		&		\rightarrow	
SHADE stability		&		\rightarrow	

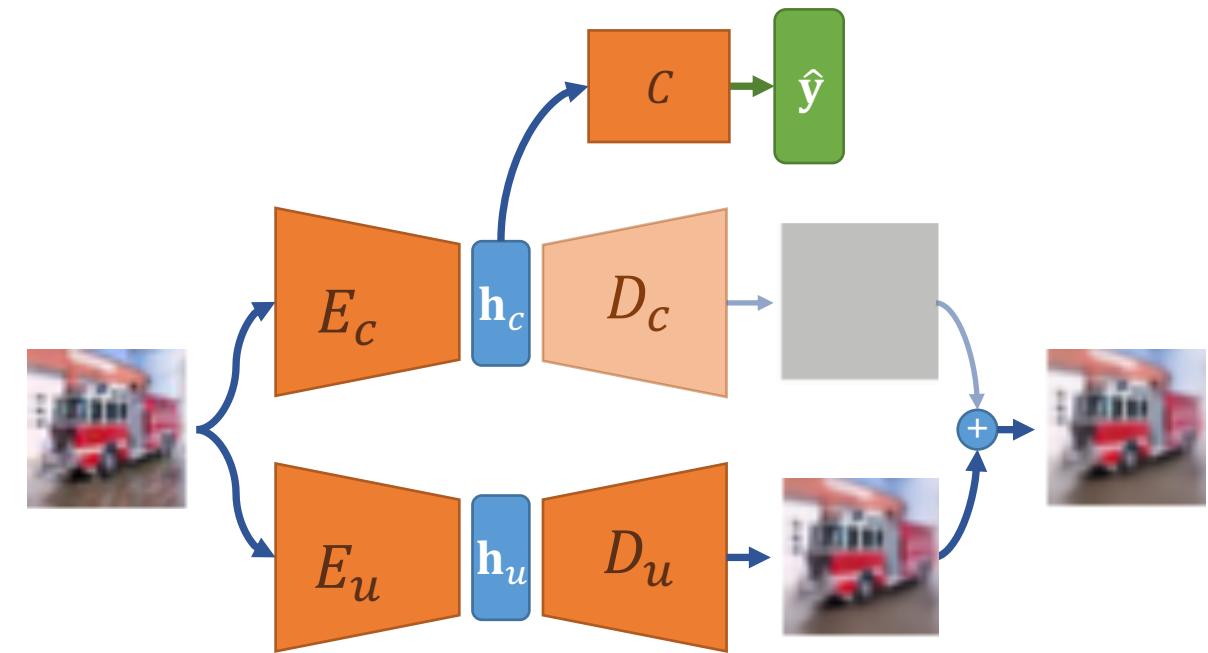
No stability		&		\rightarrow	
SHADE stability		&		\rightarrow	

	Discrim. branch reconstr.	Unsup. branch reconstr.	Final reconstr.	Original input	
	$\hat{\mathbf{x}}_c$	$\hat{\mathbf{x}}_u$	$\hat{\mathbf{x}}$	\mathbf{x}	
No stability		&		\rightarrow	
SHADE stability		&		\rightarrow	

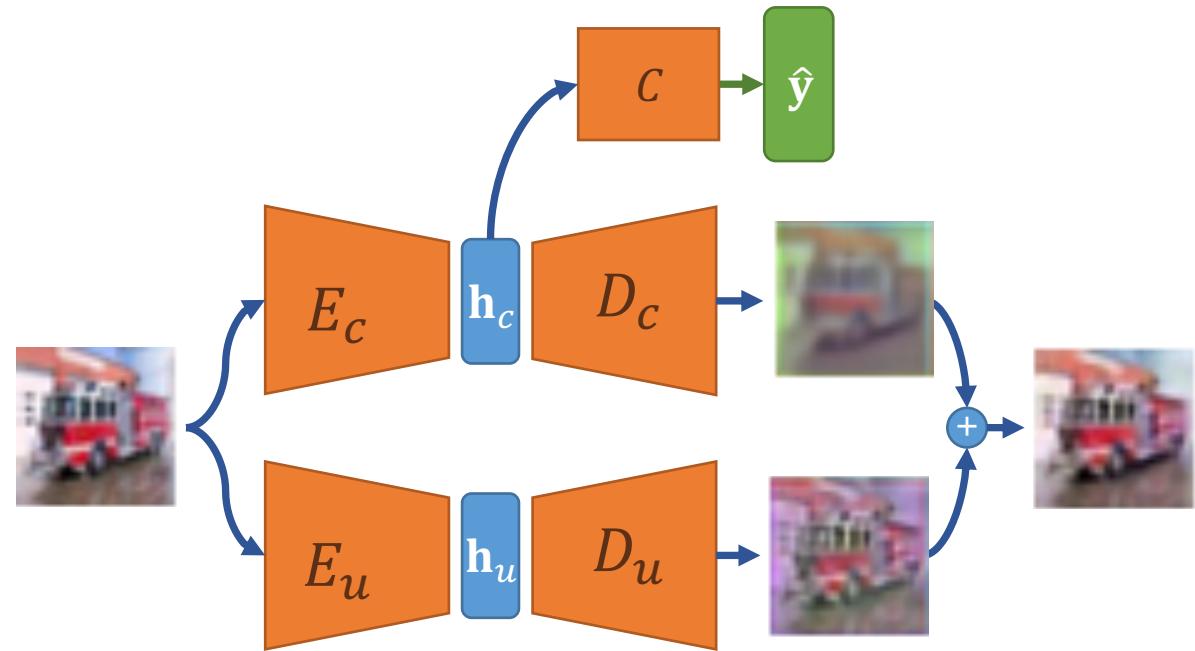
No stability		&		\rightarrow	
SHADE stability		&		\rightarrow	

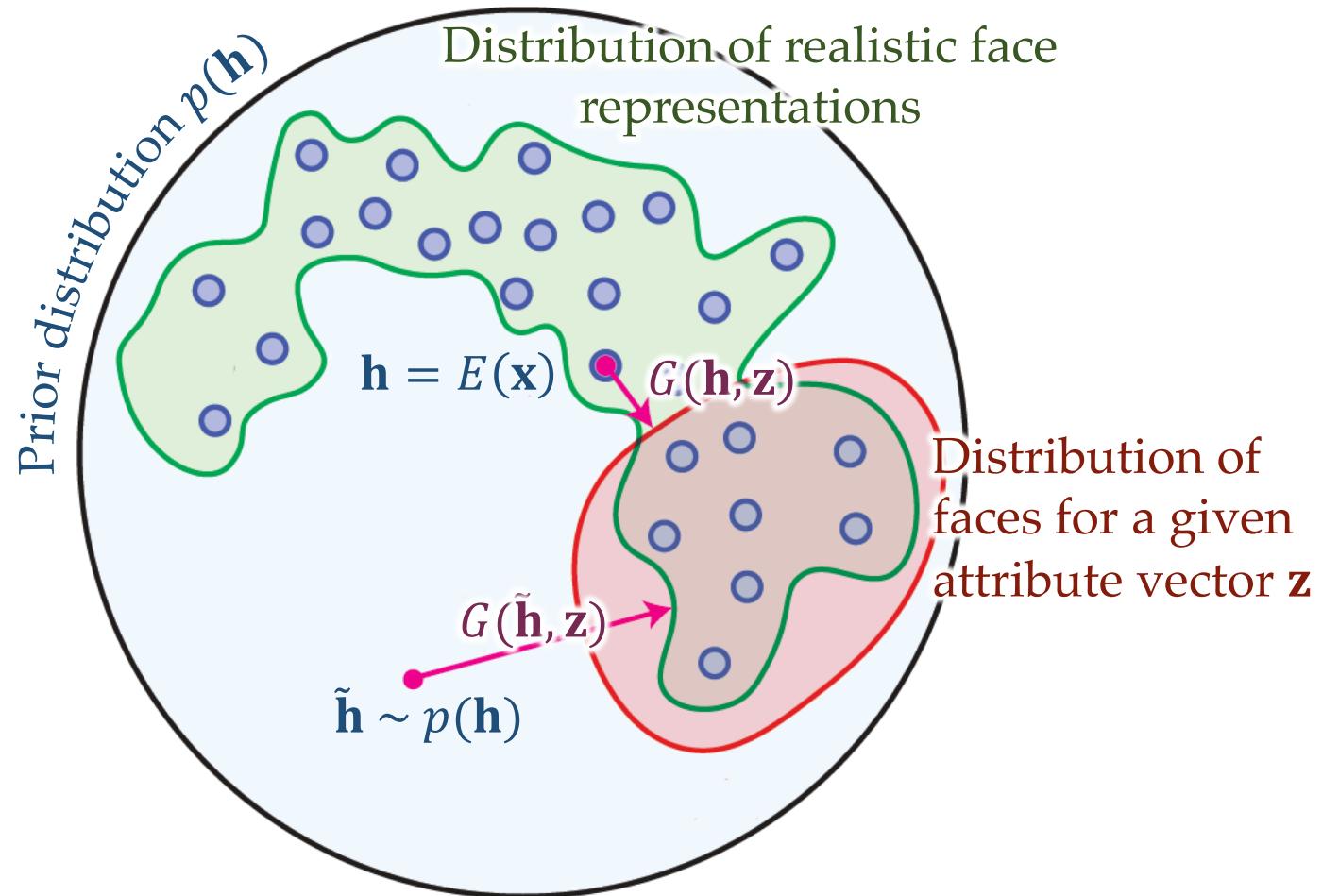


Without branch balancing



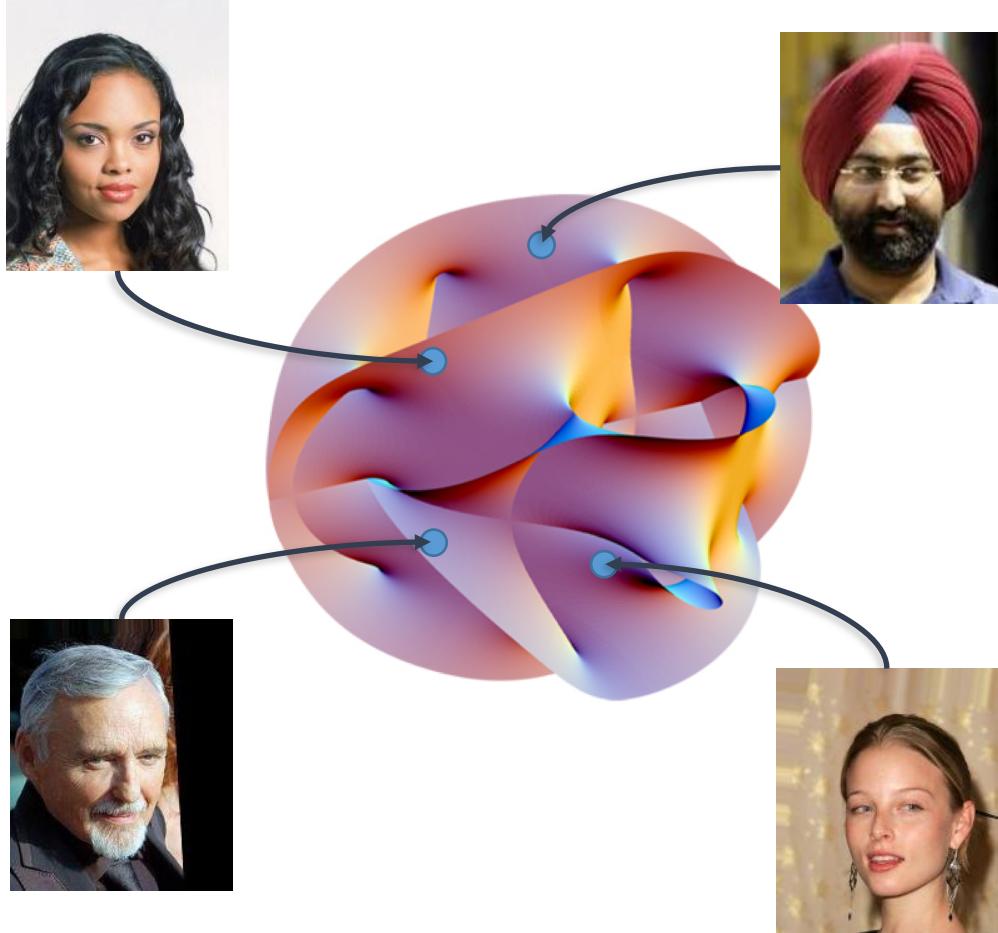
With branch balancing





Entangled manifold of input data

Illustrative view of the image manifold



Disentangled representation space

Schematic view of independent factors

