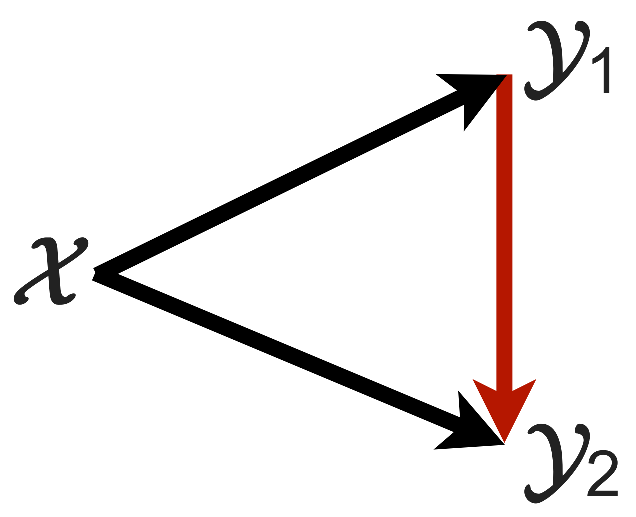
Robust Learning Through Cross-Task Consistency

A. Zamir*, A. Sax*, N. Cheerla, R. Suri, Z. Cao, J. Malik, L. Guibas

EPFL Swiss Federal Institute of Technology
Berkeley University of California
Stanford University

Baseline Learning

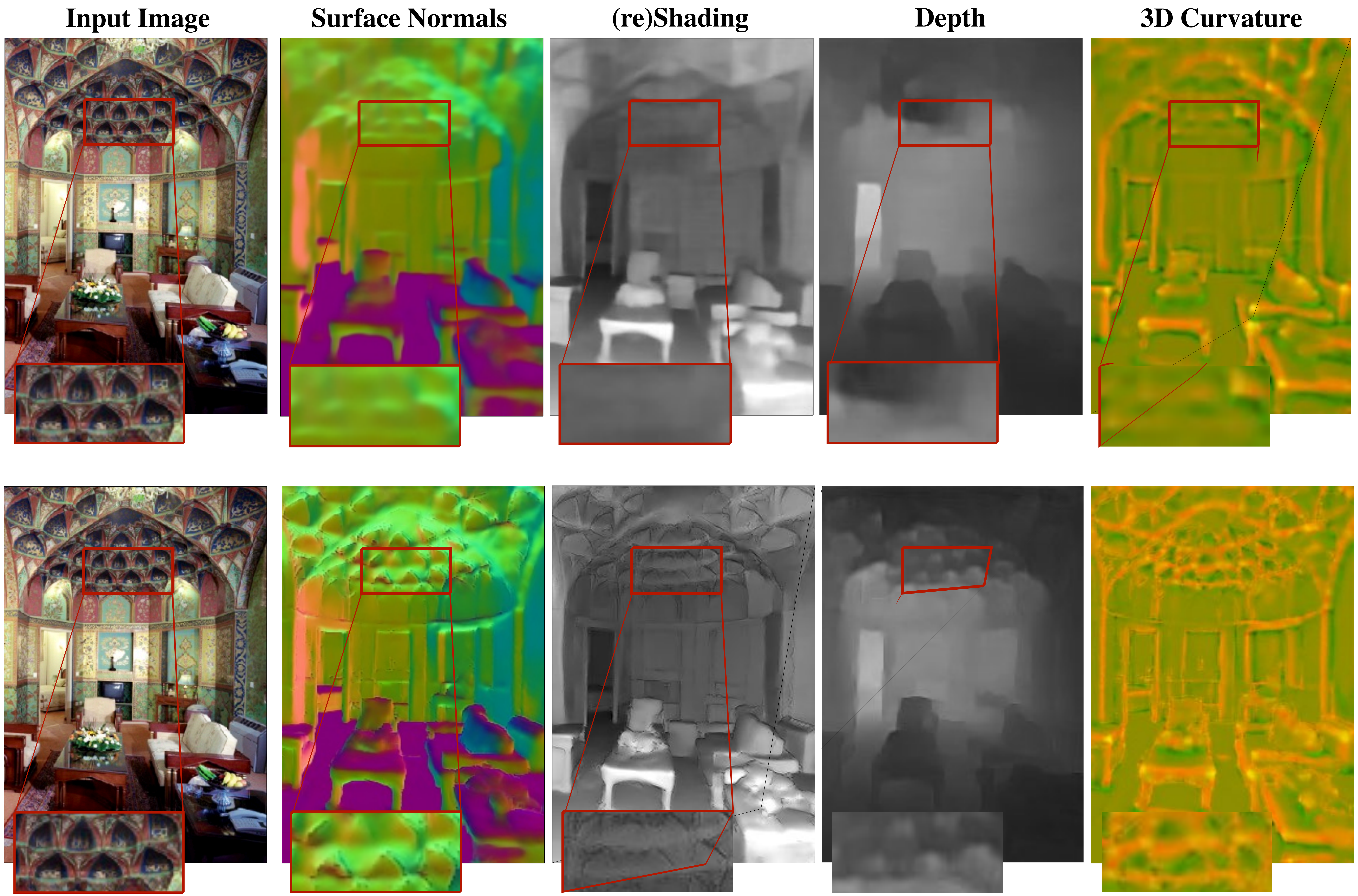


Cross-Task Consistent Learning

CVPR 2020

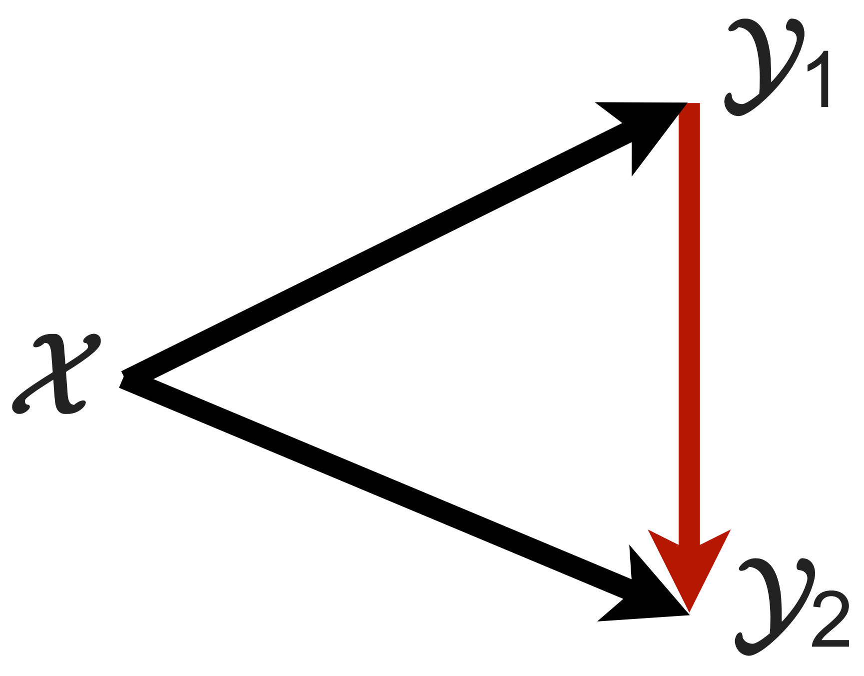
http://consistency.epfl.ch/

Cross-Task Consistency in Learning



Baseline
Learning

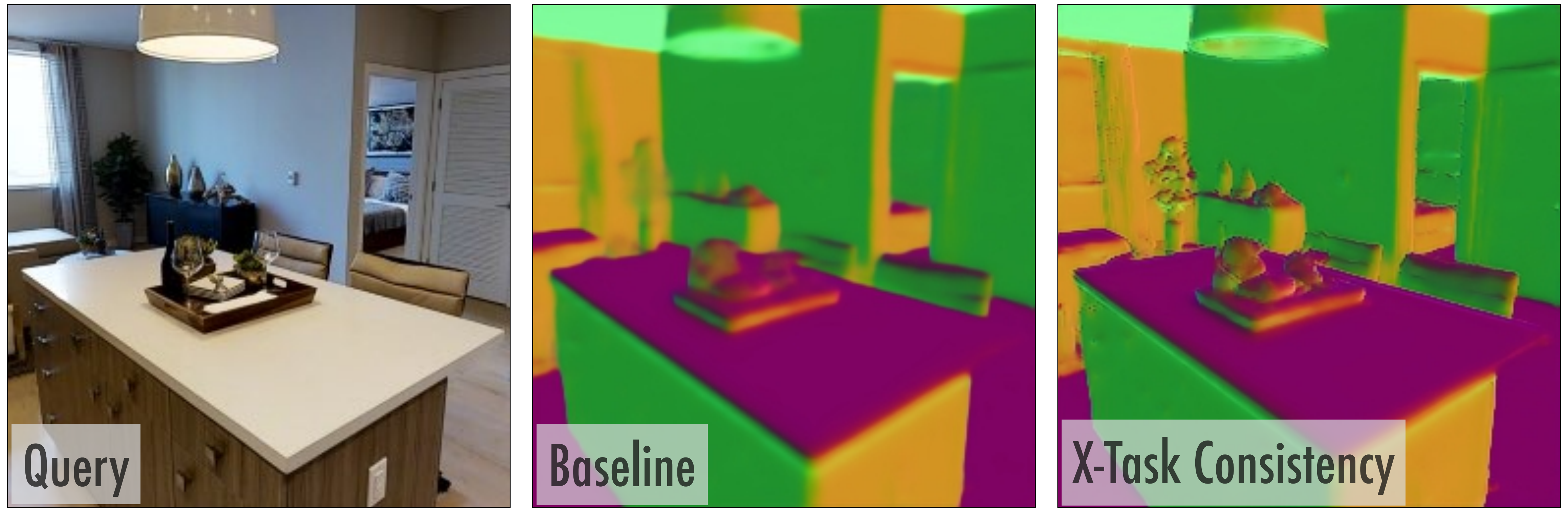
Cross-Task Consistent
Learning



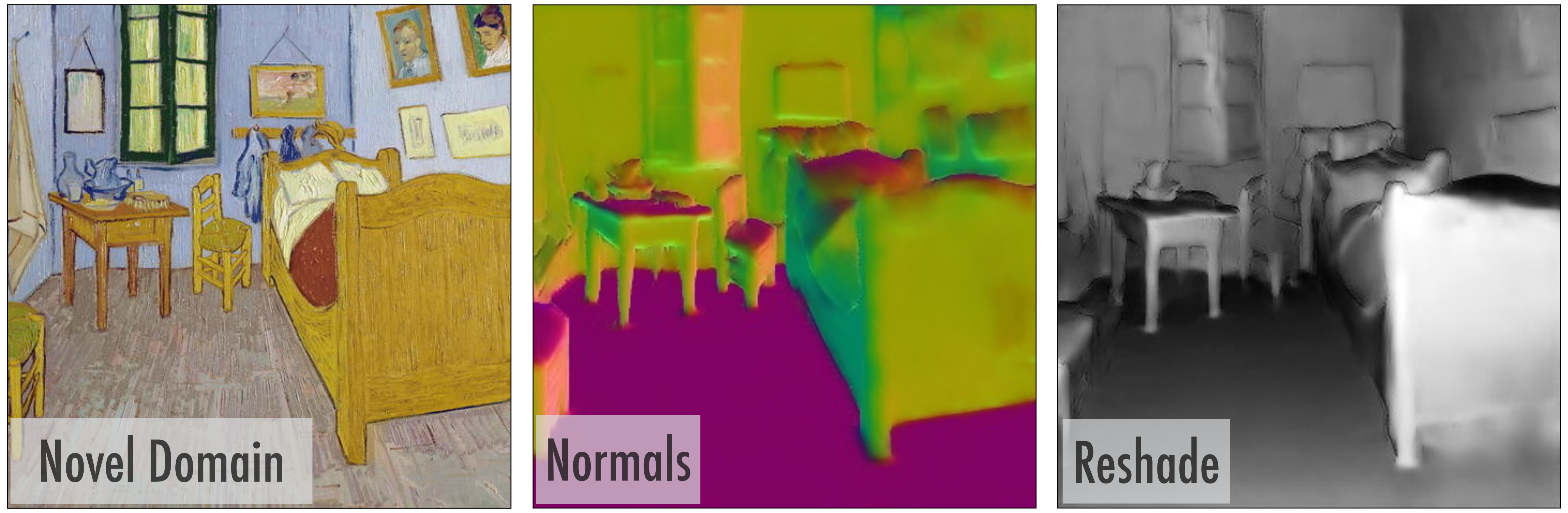
Cross-Task Consistency in Learning

Cross-Task Consistent Learning ($\mathcal{X} \to \mathcal{Y}_1,\ \mathcal{X} \to \mathcal{Y}_2,\ \mathcal{Y}_1 \to \mathcal{Y}_2$) vs Standard Learning ($\mathcal{X} \longrightarrow \mathcal{Y}$)

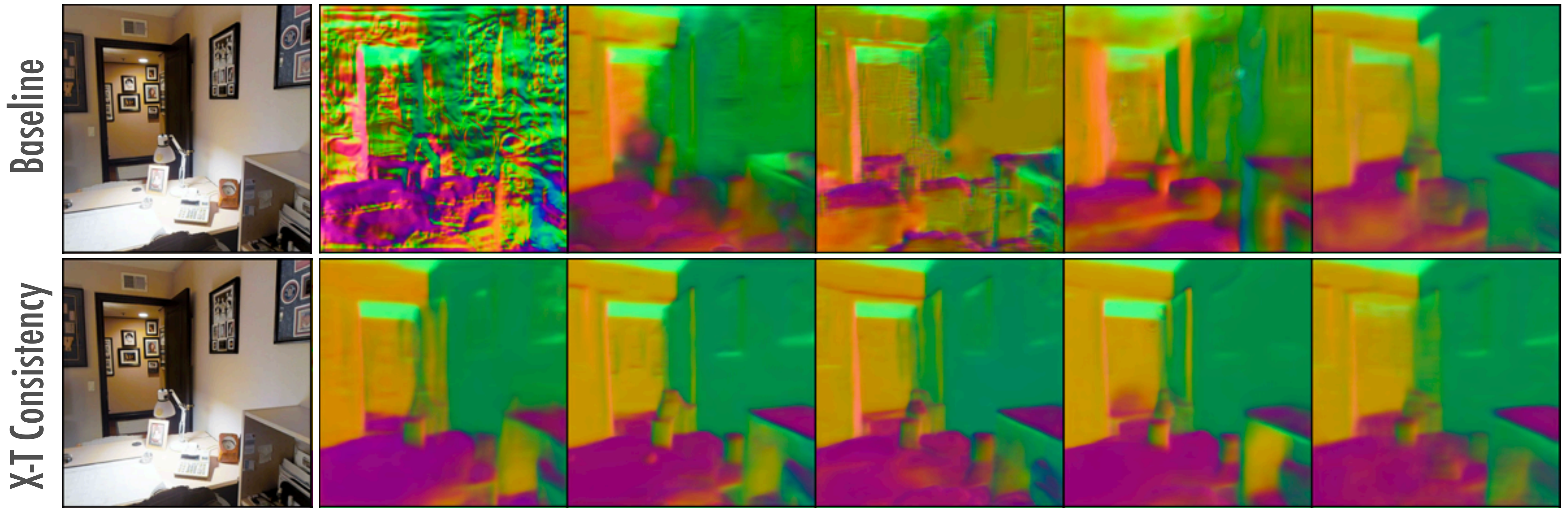
- More Accurate



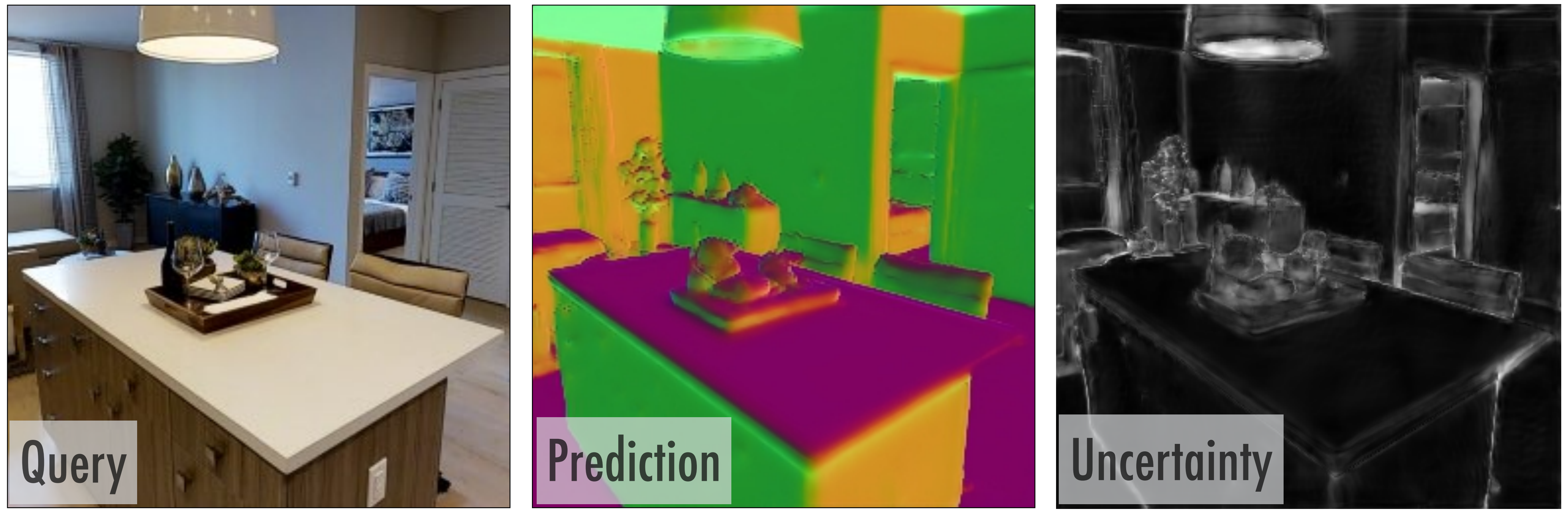
- Enhanced Generalization



- More Consistent



- Consistency Energy (Uncertainty)



Method

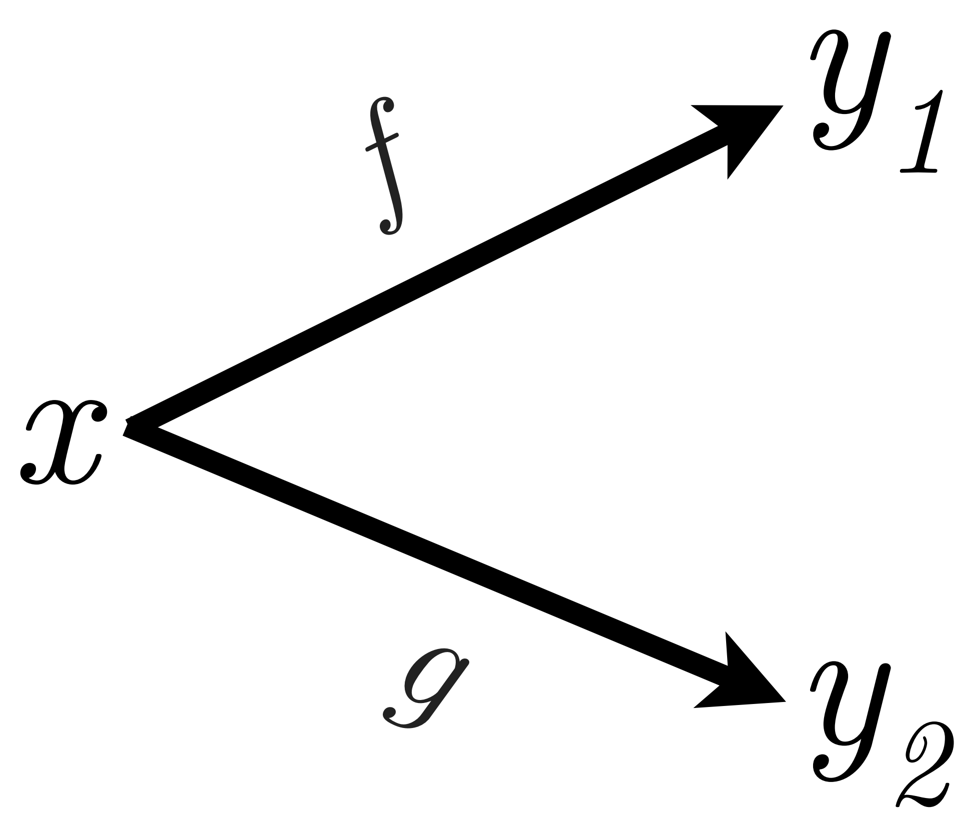
- Augments standard supervised learning with Cross-Task Consistency constraints
- Broadly applicable. Fully computational.
- Constraints (task relationships) are learned from the data
    - No need to differentiable or apriori given analytical task relationships.
- Based on `*Inference-Path Invariance*'.

Augmenting Learning with Cross-Task Consistency Constraints

$$\{x, y_1, y_2\}$$

(input) (output 1) (output 2)

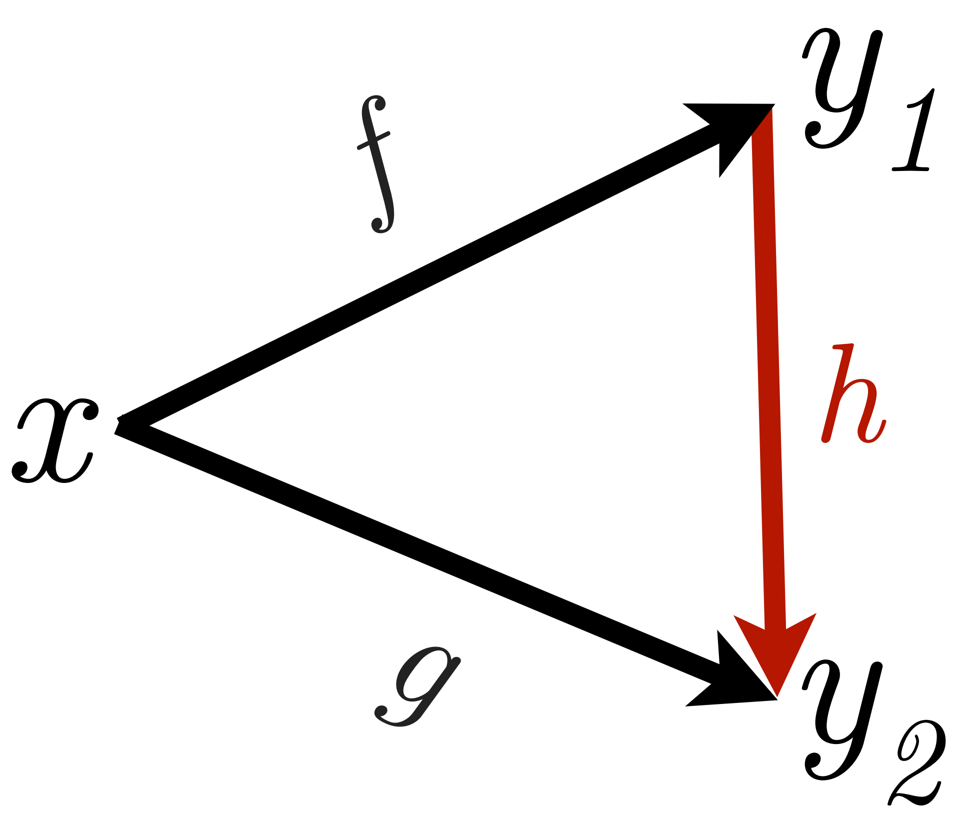
Standard Learning



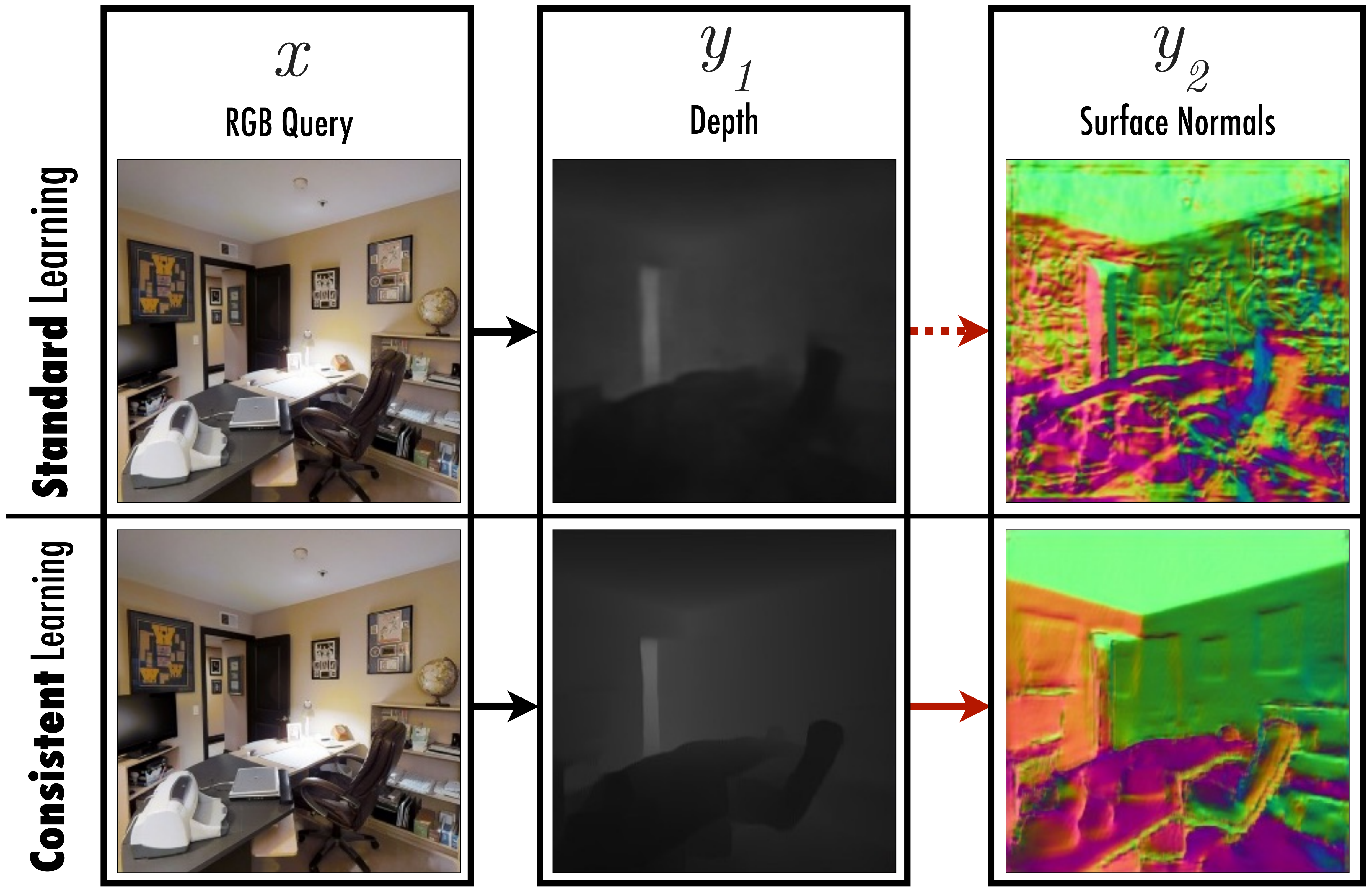
$$\mathcal{L}^f = |f(x) - y_1|$$

$$\mathcal{L}^g = |g(x) - y_2|$$

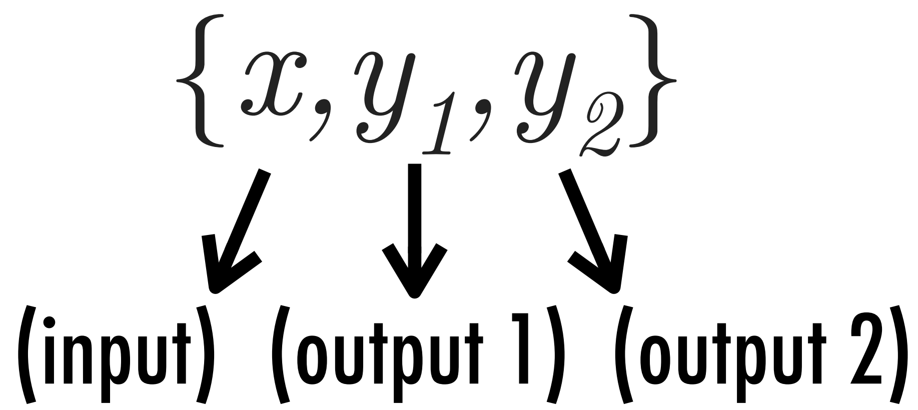
Triangle (3 domain) Consistent Learning



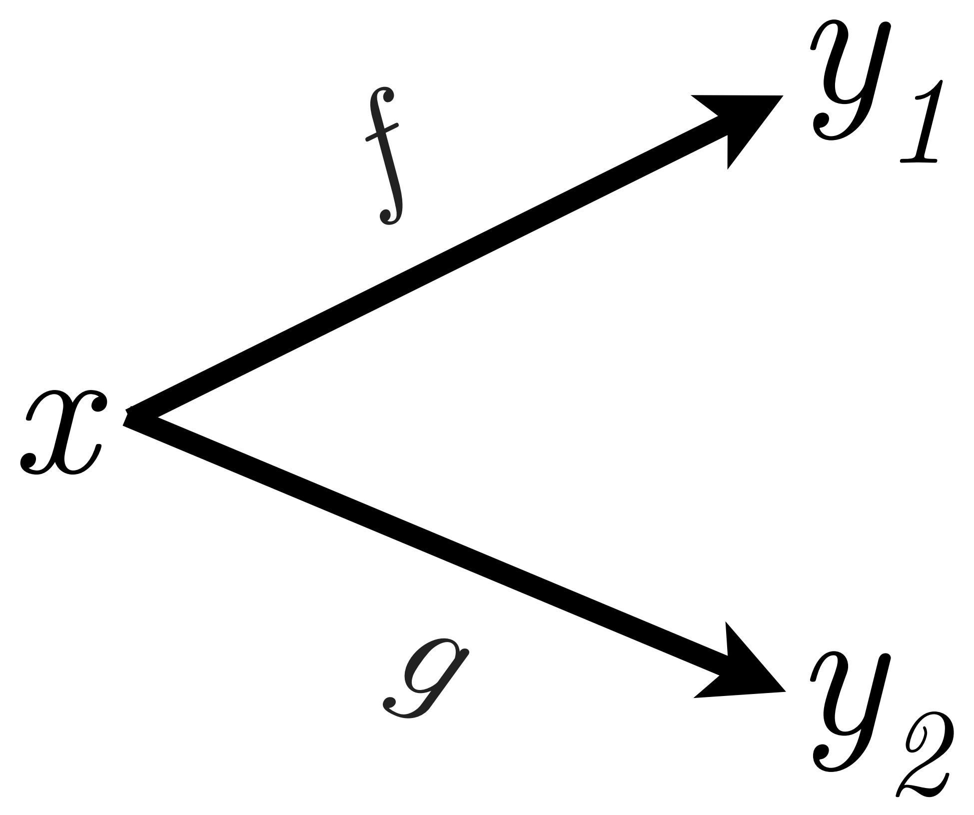
$$\mathcal{L}^{triangle} = \mathcal{L}_f + \mathcal{L}_g + \underbrace{|h \circ f(x) - g(x)|}_{\text{Cross-Task Consistency Term}}$$



Augmenting Learning with Cross-Task Consistency Constraints



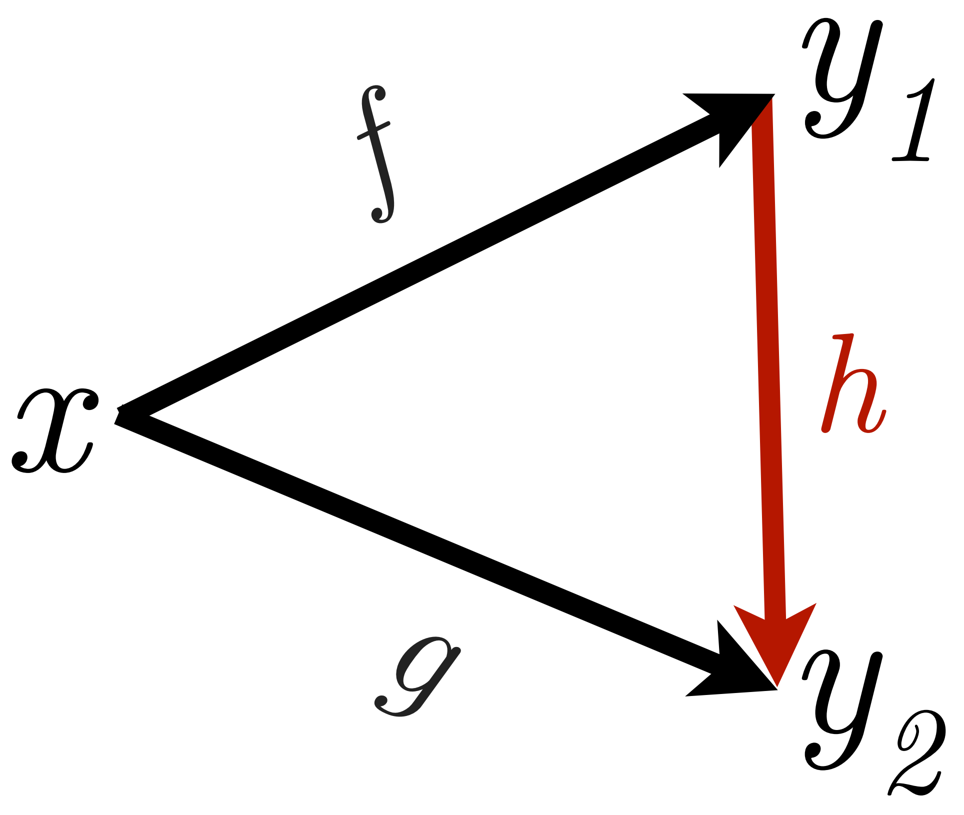
Standard Learning



$$\mathcal{L}^f = |f(x) - y_1|$$

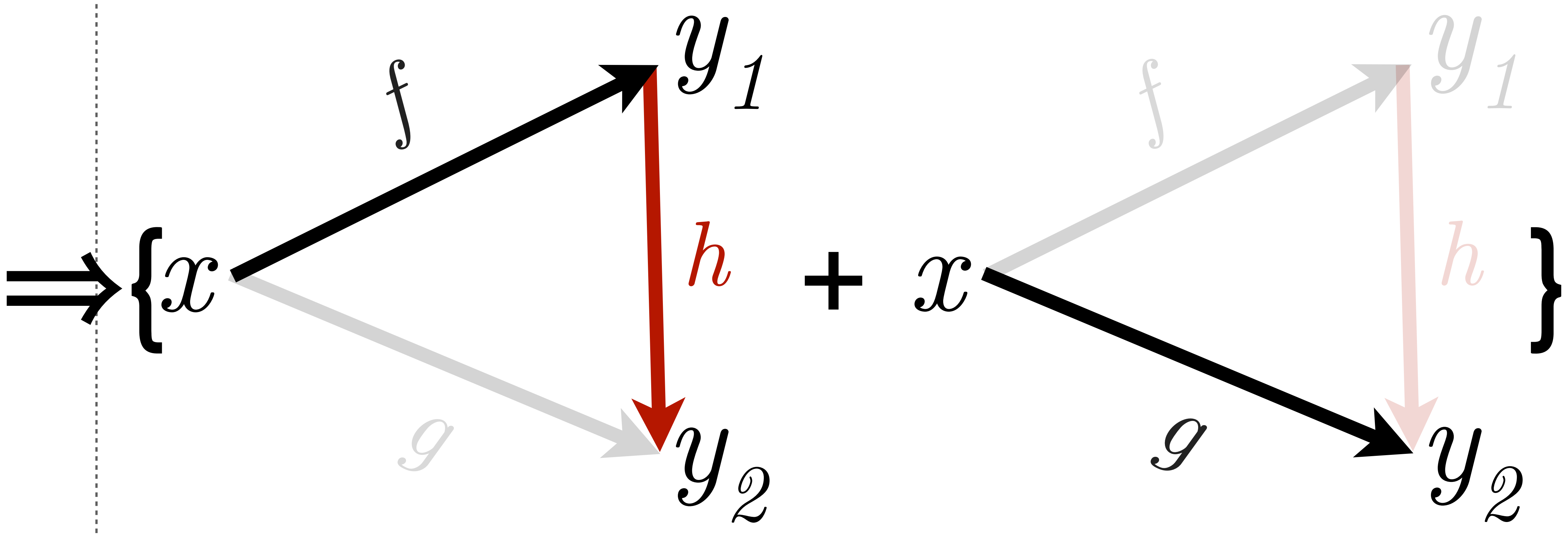
$$\mathcal{L}^g = |g(x) - y_2|$$

Triangle (3 domain) Consistent Learning



$$\mathcal{L}^{triangle} = \mathcal{L}_f + \mathcal{L}_g + |h \circ f(x) - g(x)|$$

Reconfigure into a "**Separable**" Loss



Augmenting Learning with Cross-Task Consistency Constraints

$\{x, y_1, y_2\}$

(input) (output 1) (output 2)

Standard Learning

$$\begin{tikzcd} & y_1 \\ x \arrow[ur, "f"] \arrow[dr, "g"'] & \\ & y_2 \end{tikzcd}$$

$$\mathcal{L}^f = |f(x) - y_1|$$

$$\mathcal{L}^g = |g(x) - y_2|$$

Triangle (3 domain) Consistent Learning

$$\begin{tikzcd} & y_1 \arrow[dd, "h"] \\ x \arrow[ur, "f"] \arrow[dr, "g"'] & \\ & y_2 \end{tikzcd}$$

$$\mathcal{L}^{triangle} = \mathcal{L}_f + \mathcal{L}_g + |h \circ f(x) - g(x)|$$

Reconfigure into a **"Separable" Loss**

$$x \xrightarrow{f} y_1 \xrightarrow{h} y_2$$

$$\mathcal{L}^{separate} = \mathcal{L}_f + |h \circ f(x) - y_2|$$

Reconfigure into a **"Perceptual" Loss**

$$x \xrightarrow{f} y_1 \xrightarrow{h} h(y_1)$$

$$\mathcal{L}^{perceptual} = \mathcal{L}_f + |h \circ f(x) - h(y_1)|$$

- Handles Ill-posed task relationships and imperfect cross-task networks
- Enables using pair training datasets (x, y_1) rather than triplets (x, y_1, y_2).

Augmenting Learning with Cross-Task Consistency Constraints

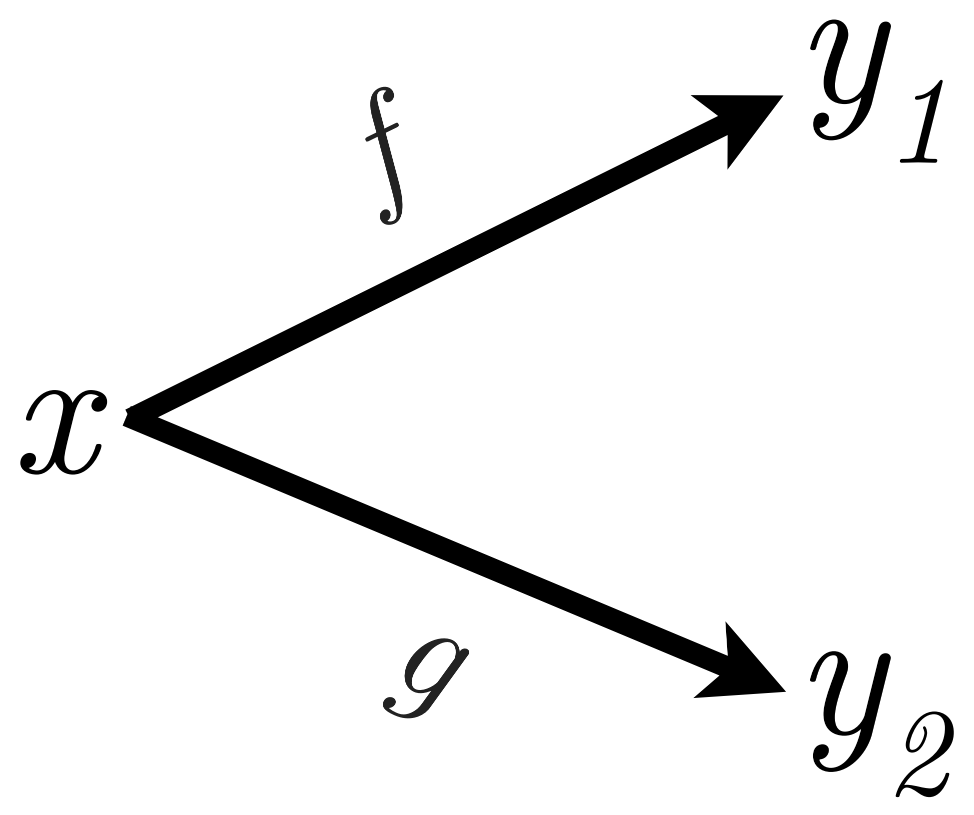
$\{x, y_1, y_2\}$

(input) (output 1) (output 2)

$\{x, y_1, y_2, \ldots, y_n\}$

(input) (output 1) (output n)

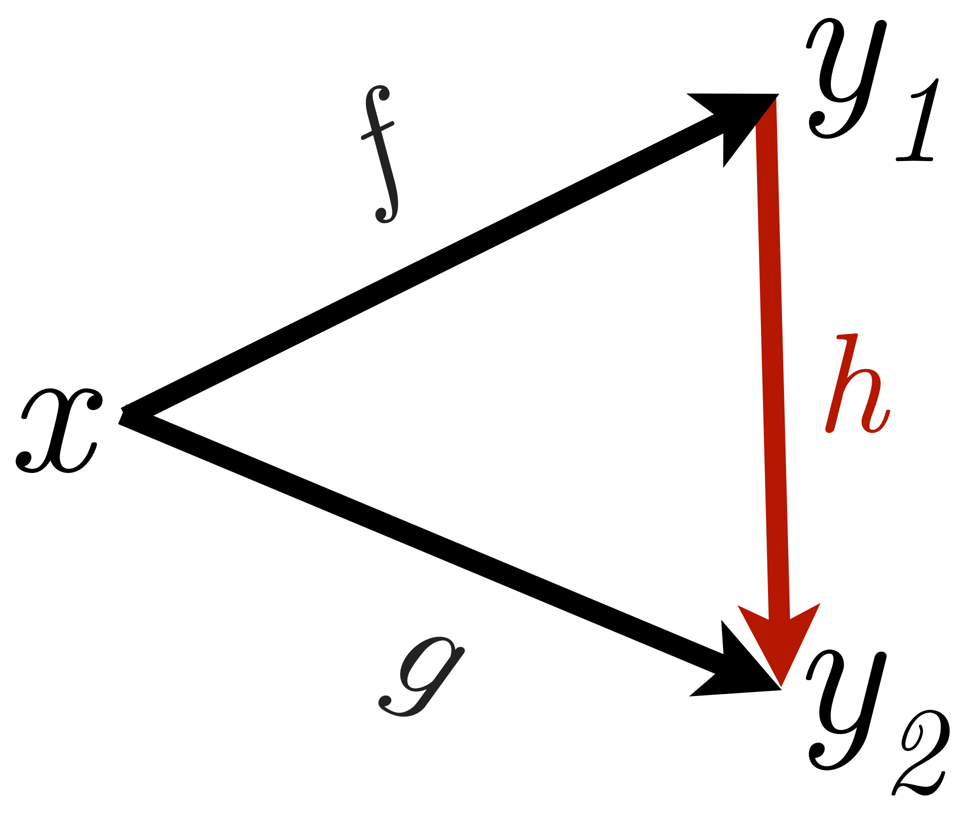
Standard Learning



$$\mathcal{L}^f = |f(x) - y_1|$$

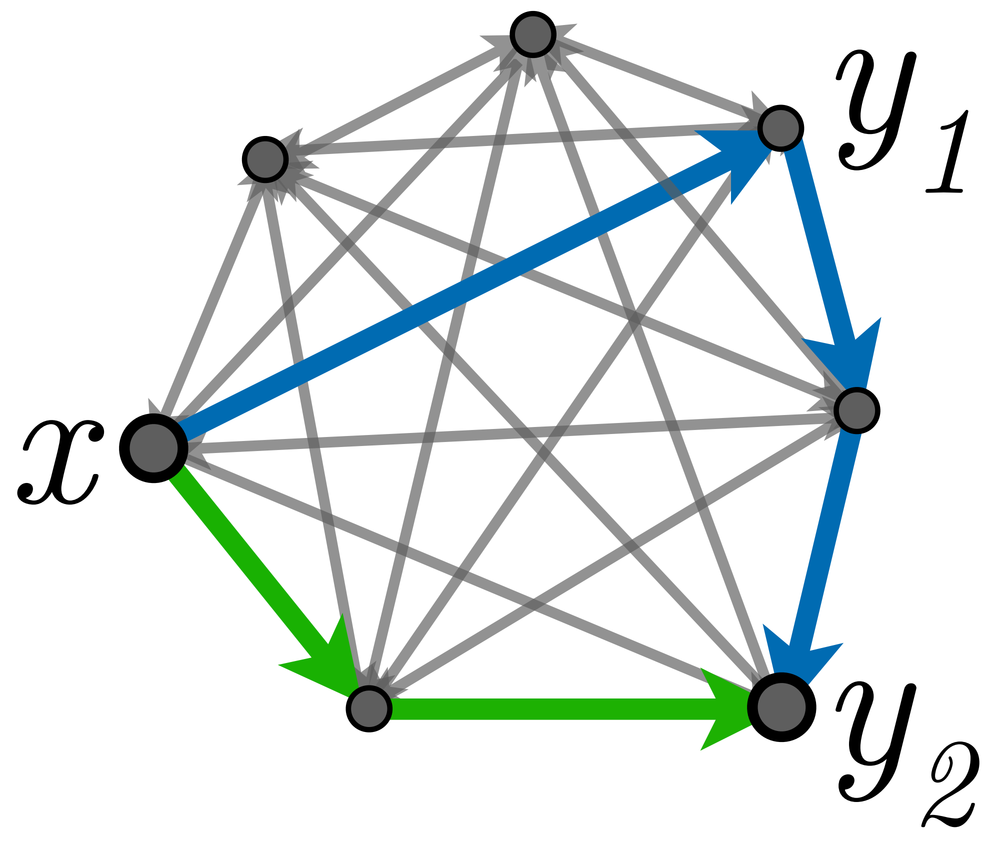
$$\mathcal{L}^g = |g(x) - y_2|$$

Triangle (3 domain) Consistent Learning



$$\mathcal{L}^{triangle} = \mathcal{L}_f + \mathcal{L}_g + |h \circ f(x) - g(x)|$$

Globally Consistent Learning



Global Consistency ≜ satisfying consistency constraints for <u>all feasible paths</u> in the graph.

Augmenting Learning with Cross-Task Consistency Constraints

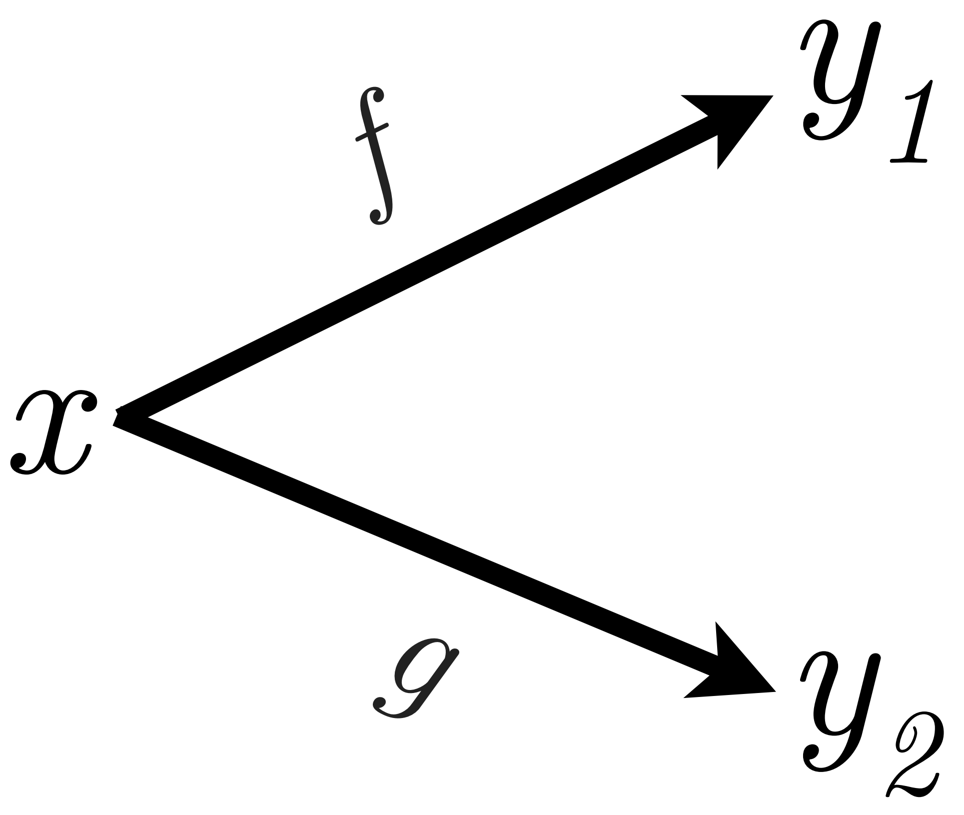
$\{x, y_1, y_2\}$

(input) (output 1) (output 2)

$\{x, y_1, y_2, ..., y_n\}$

(input) (output 1) (output n)

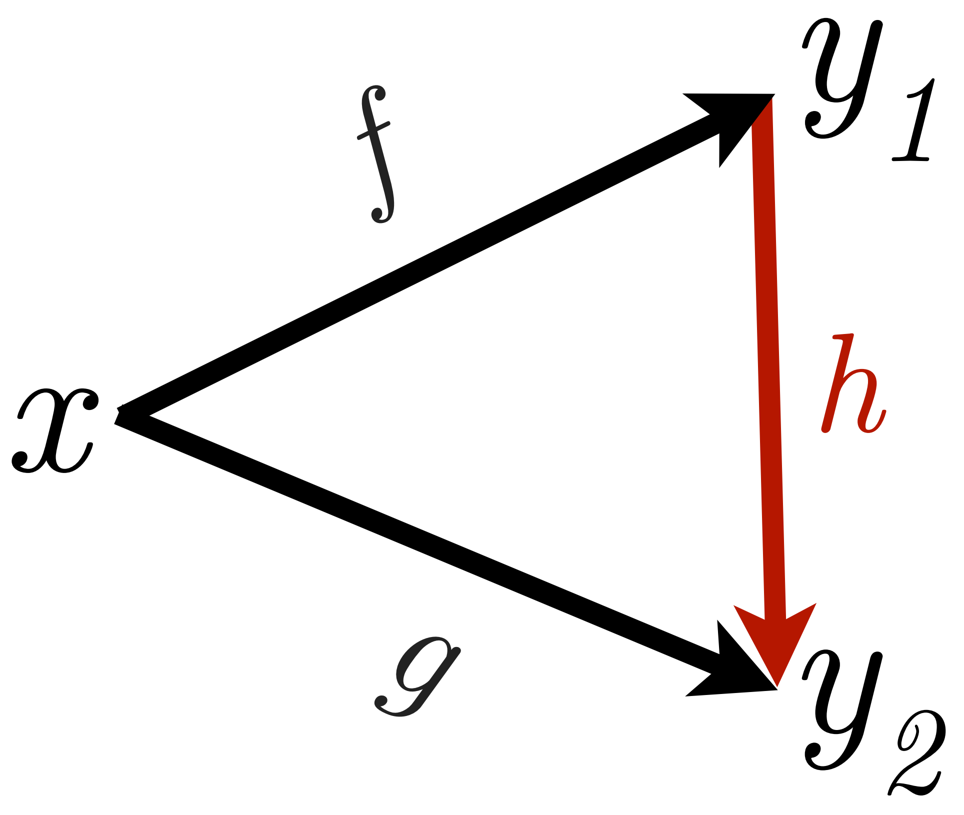
Standard Learning



$$\mathcal{L}^f = |f(x) - y_1|$$

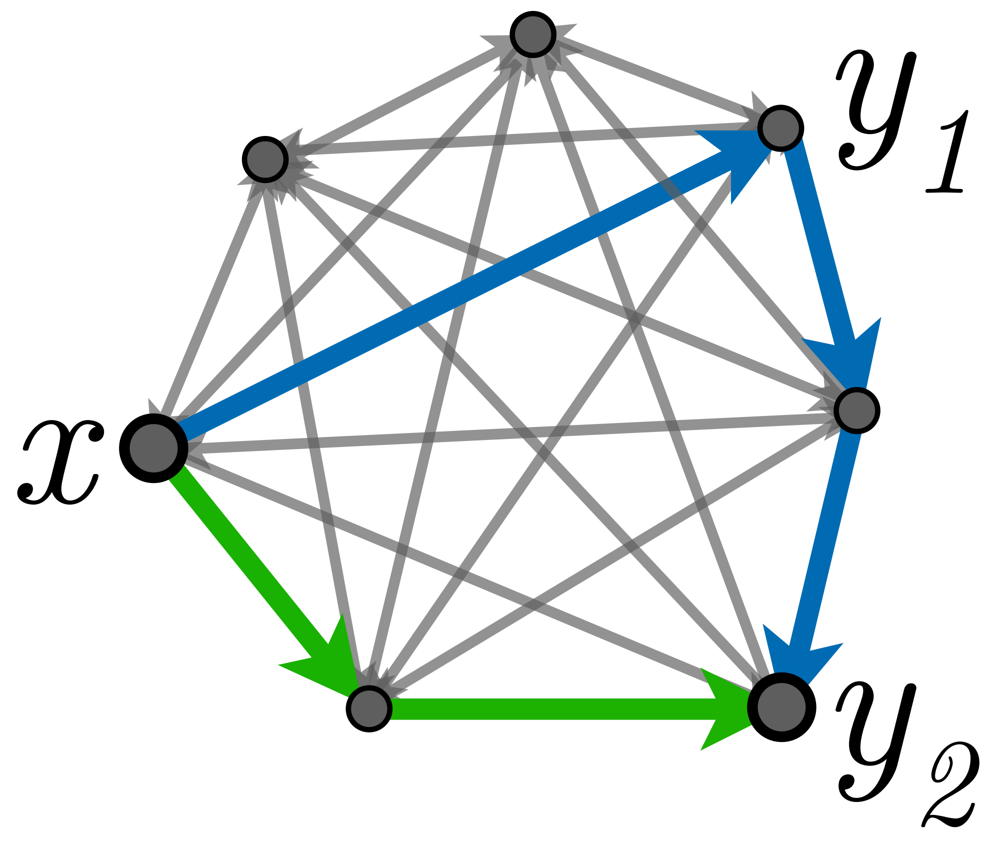
$$\mathcal{L}^g = |g(x) - y_2|$$

Triangle (3 domain) Consistent Learning



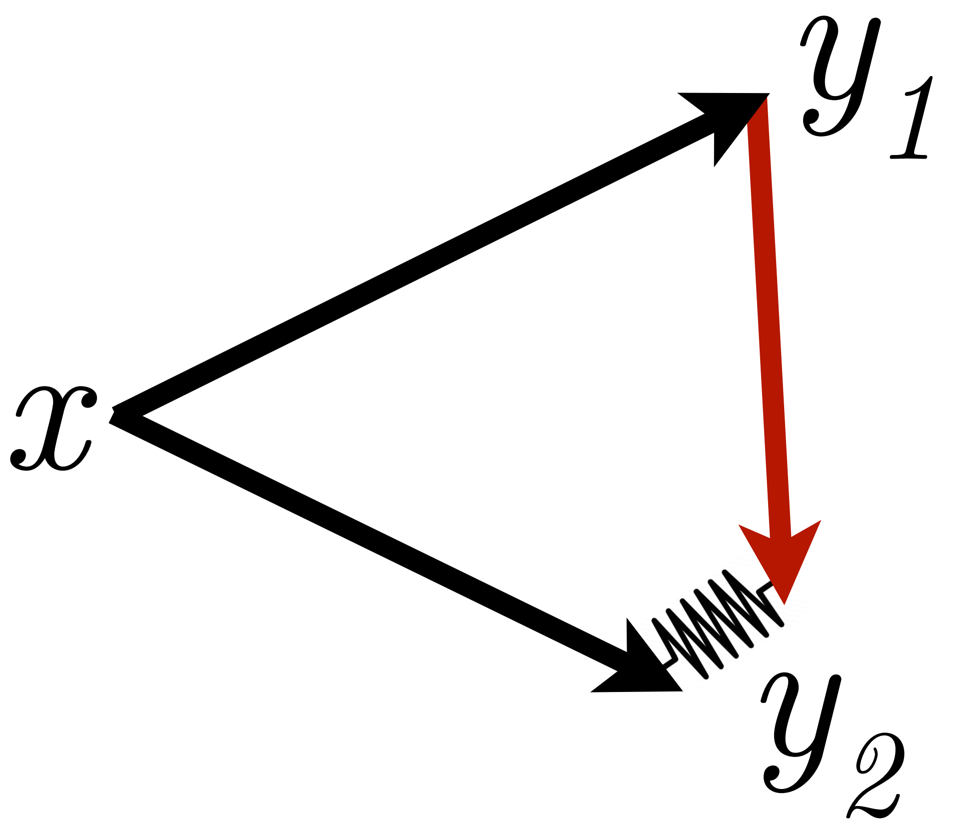
$$\mathcal{L}^{triangle} = \mathcal{L}_f + \mathcal{L}_g + |h \circ f(x) - g(x)|$$

Globally Consistent Learning



Global Consistency ≜ satisfying consistency constraints for <u>all feasible paths</u> in the graph.

Consistency Energy



- **Inconsistency** ≜ disagreement among paths.
- **Energy** ≜ total inconsistency/disagreement in the system.
- A informative test-time unsupervised quantity.

Frame-by-frame results on the test video of [47]. (Full video in extended slides on project webpage.)

Frame-by-frame results on the test video of [47]. (Full video in extended slides on project webpage.)

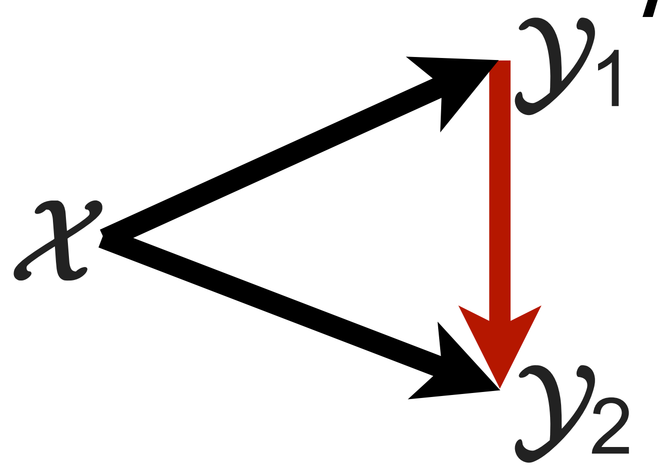
Frame-by-frame results on the test video of [47]. (Full video in extended slides on project webpage.)

Frame-by-frame results on the test video of [47]. (Full video in extended slides on project webpage.)

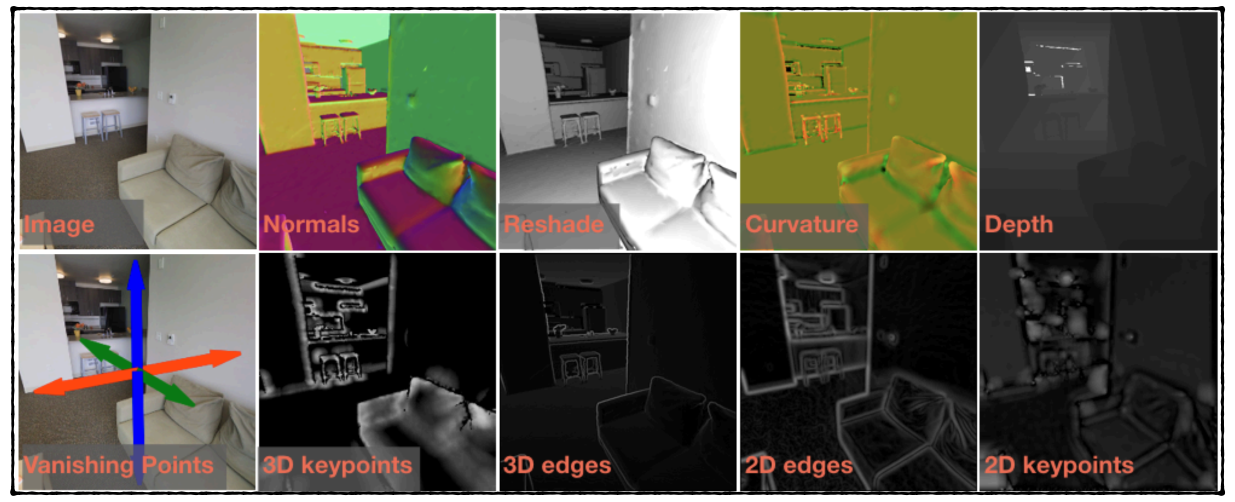
Experiments

Experiments

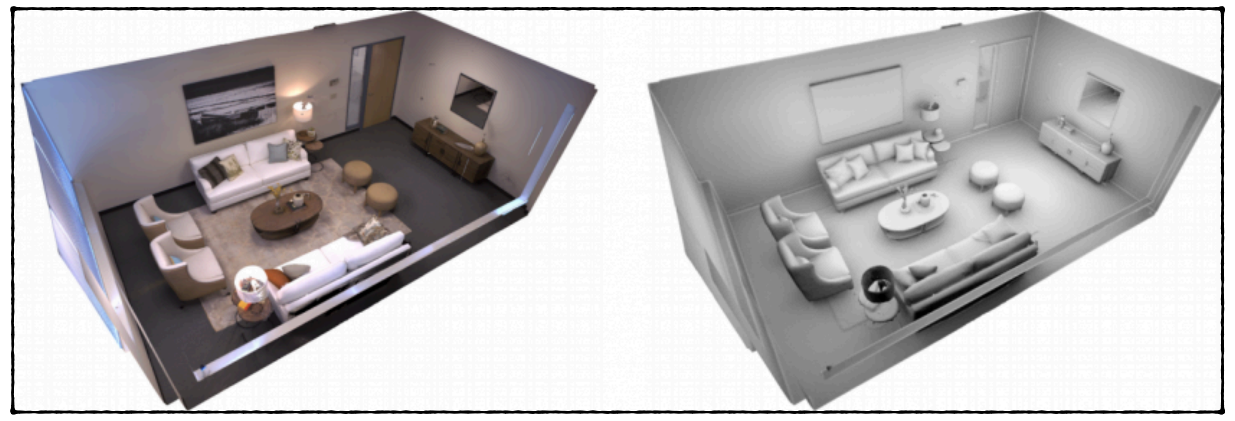
Cross-Task Consistency



- Datasets:
  - Taskonomy (Zamir et al. 2018)



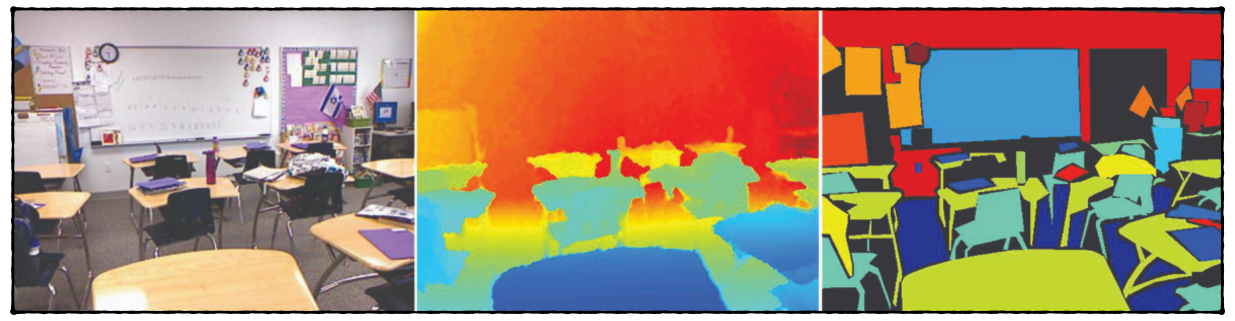
  - Replica (Straub et al. 2019)



  - CocoDoom (Mahendran et al. 2016)

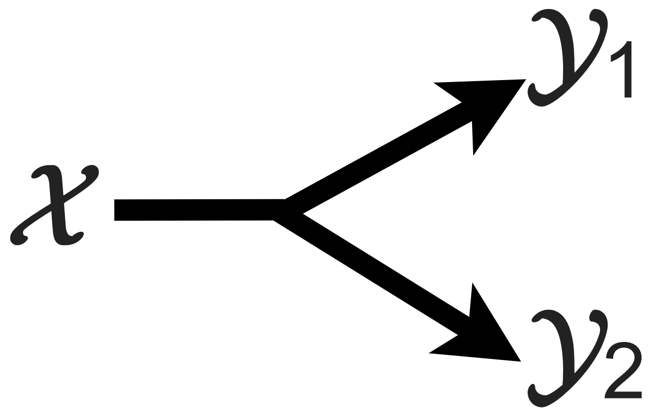
  - ApolloScape (Huang et al. 2018)

  - NYUv2 (Silberman et al. 2018)

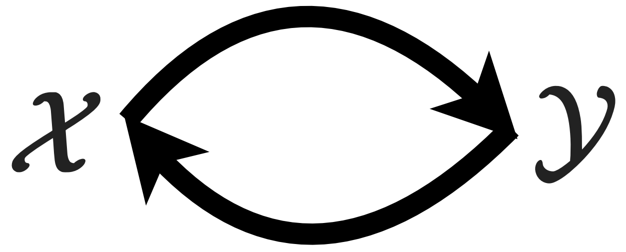


- Main Baselines:
  - Standard independent learning (L1 UNet)

  - Multi-task learning



  - Cycle-based consistency



  - Average estimator (blind guess)

  - Conditional GAN image translation (pix2pix, Isola et al. 2016)
  - GeoNet: analytical/curated consistency (Qi et al. 2018)

Experiments

- Evaluations:
  - **Accuracy** of Predictions
  - **Consistency** of Predictions
  - **Generalization** to out-of-training-domain data
  - **Energy** Utilities

Accuracy of Predictions

Accuracy of Predictions

Frame-by-frame results on the test video of [47]. (Full video in extended slides on project webpage.)

Accuracy of Predictions (out-of-training-domain data)

- 1: Bedroom in Arles, Van Gogh (1888)
- 2: Cotton Mill Girl, Lewis Hine (1908)
- 3: Chernobyl Pripyat Abandoned School (c. 2009)

Live Demo: http://consistency.epfl.ch/demo/

Accuracy of Predictions (all baseline comparisons)

Live Demo: http://consistency.epfl.ch/demo/

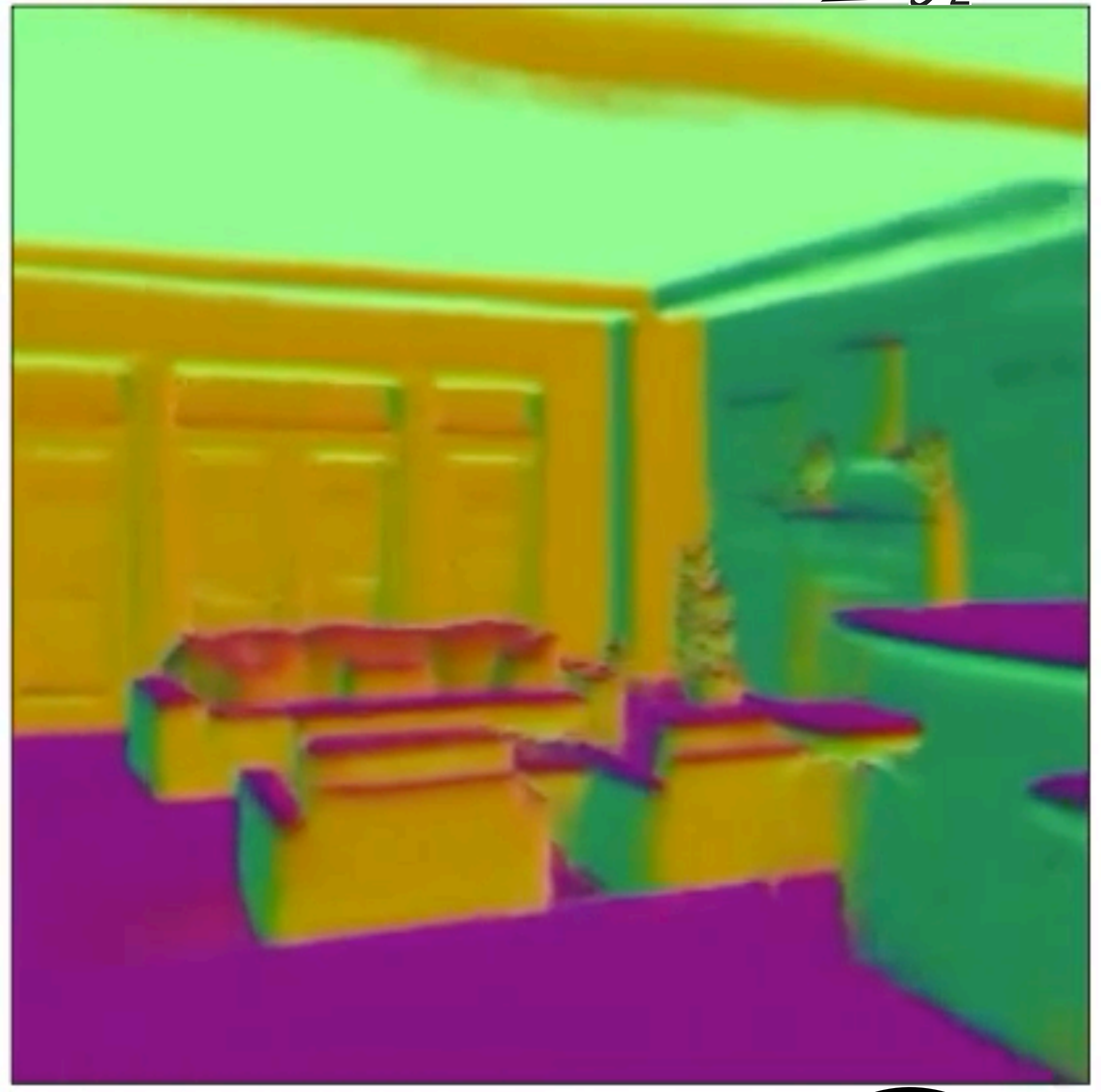
Accuracy of Predictions (all baseline comparisons)

Live Demo: http://consistency.epfl.ch/demo/

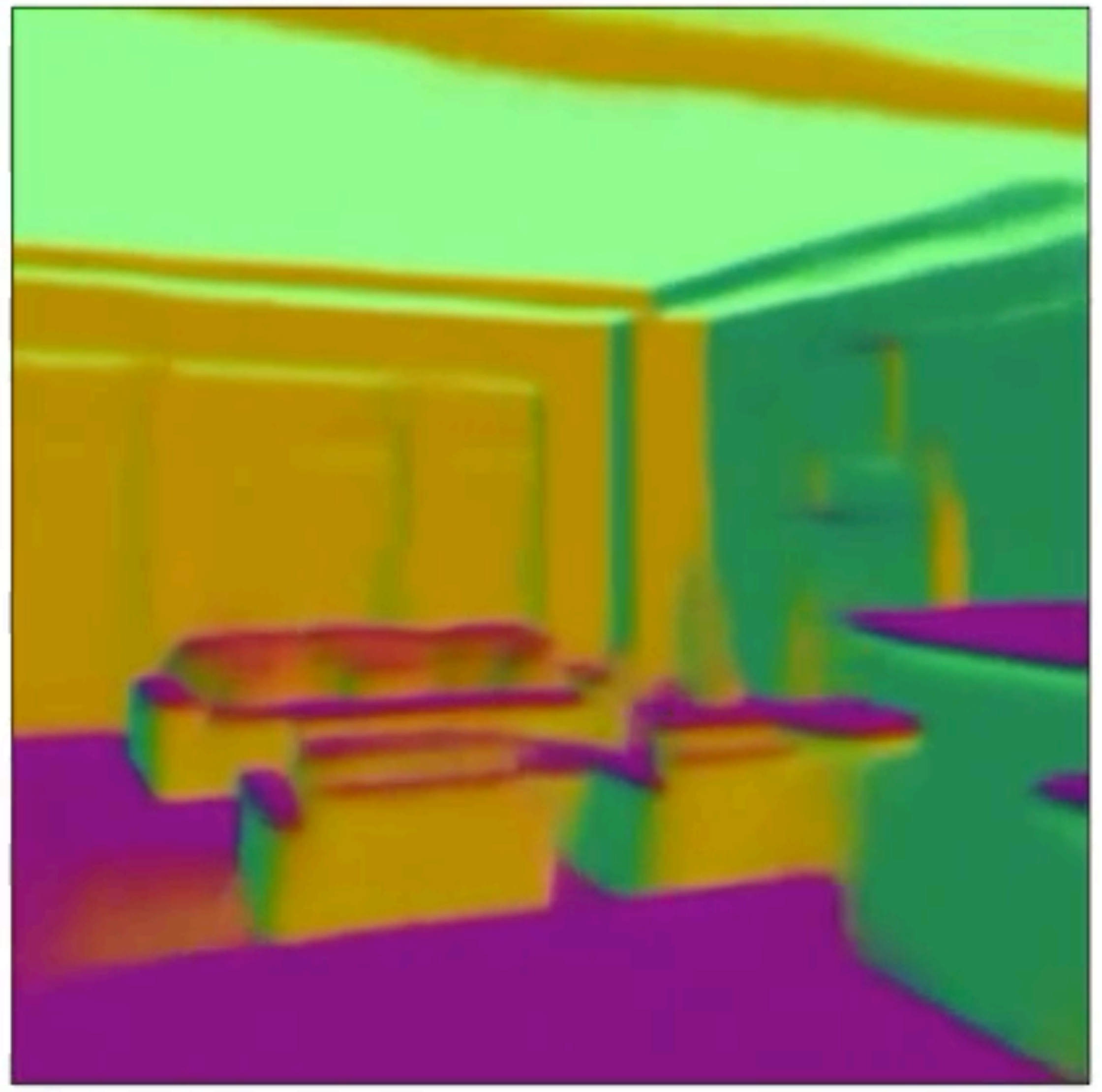
Accuracy of Predictions (all baseline comparisons)

Query

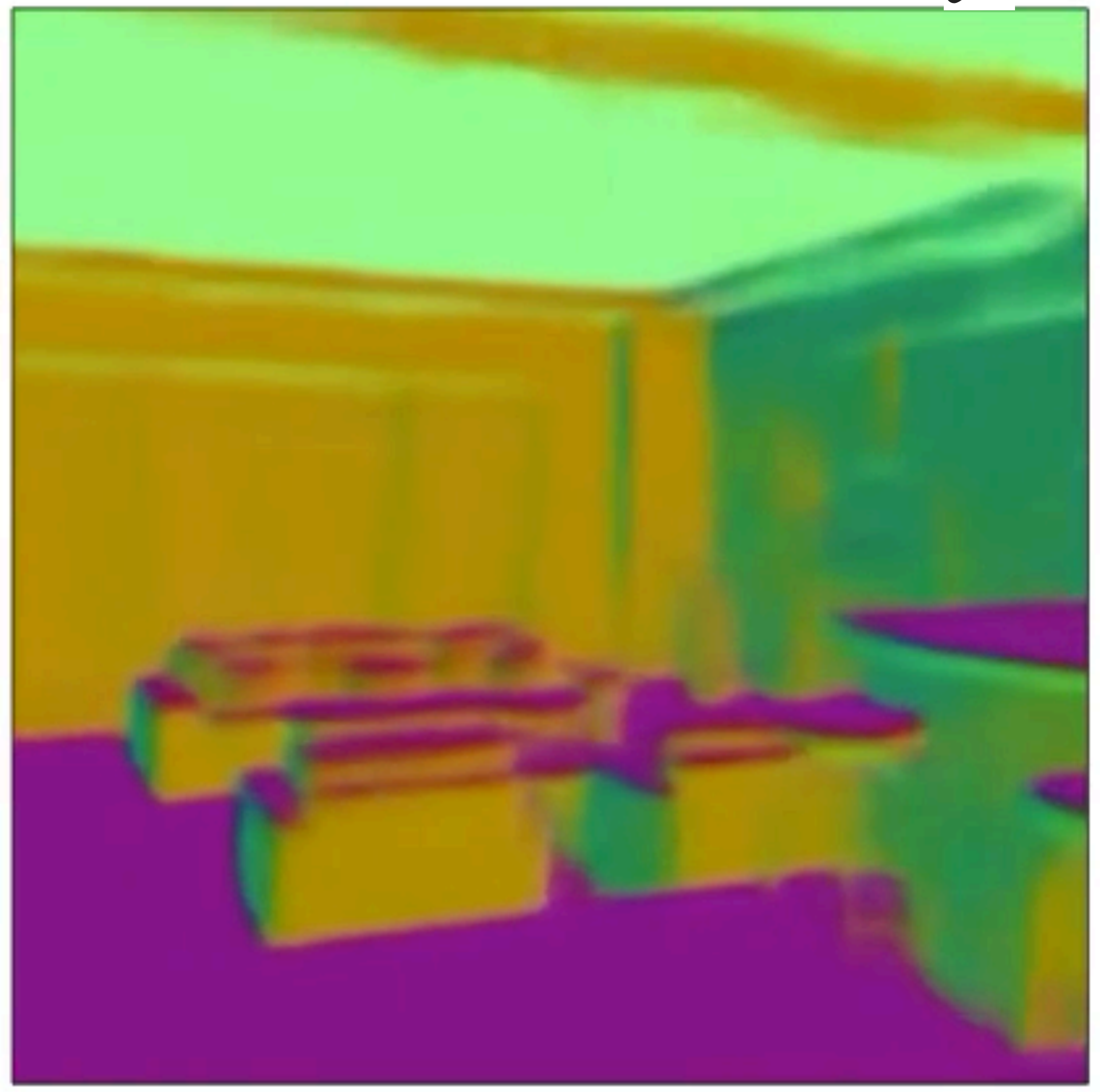
X-Task Consistency $x \to y_1, x \to y_2, y_1 \to y_2$



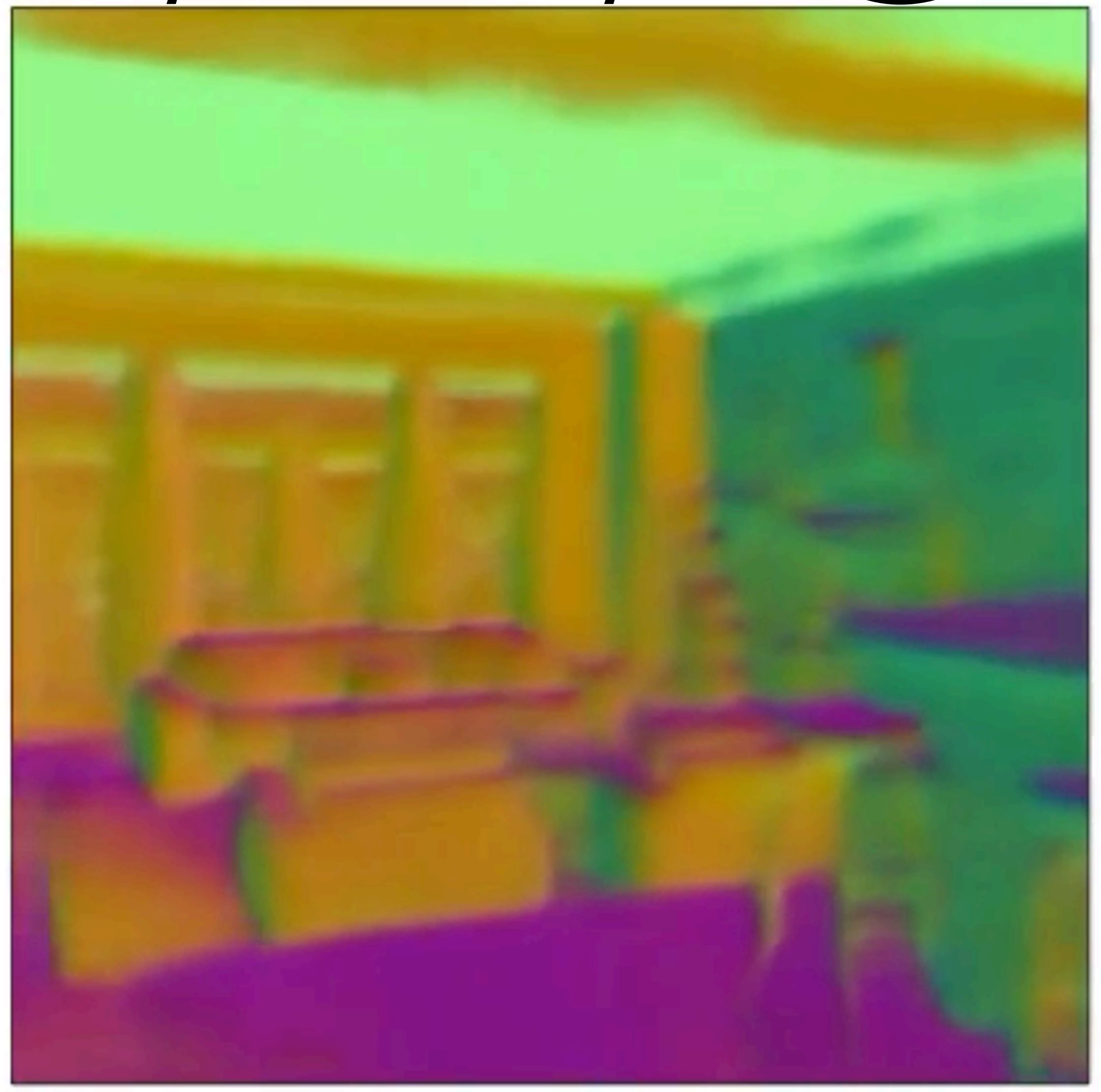
Baseline (L1 UNet) $x \longrightarrow y$



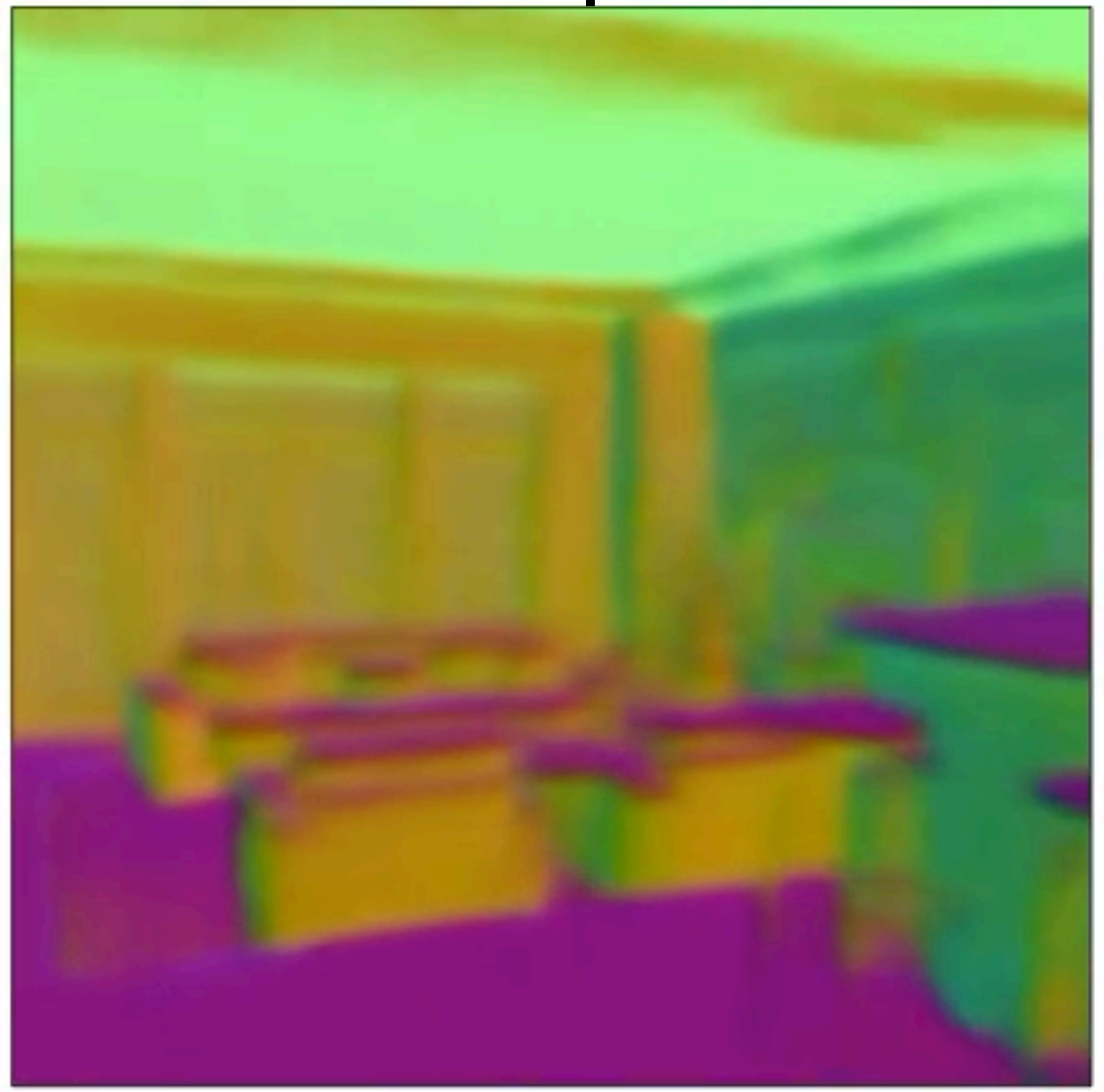
Multi-Task Network $x \to y_1, y_2$



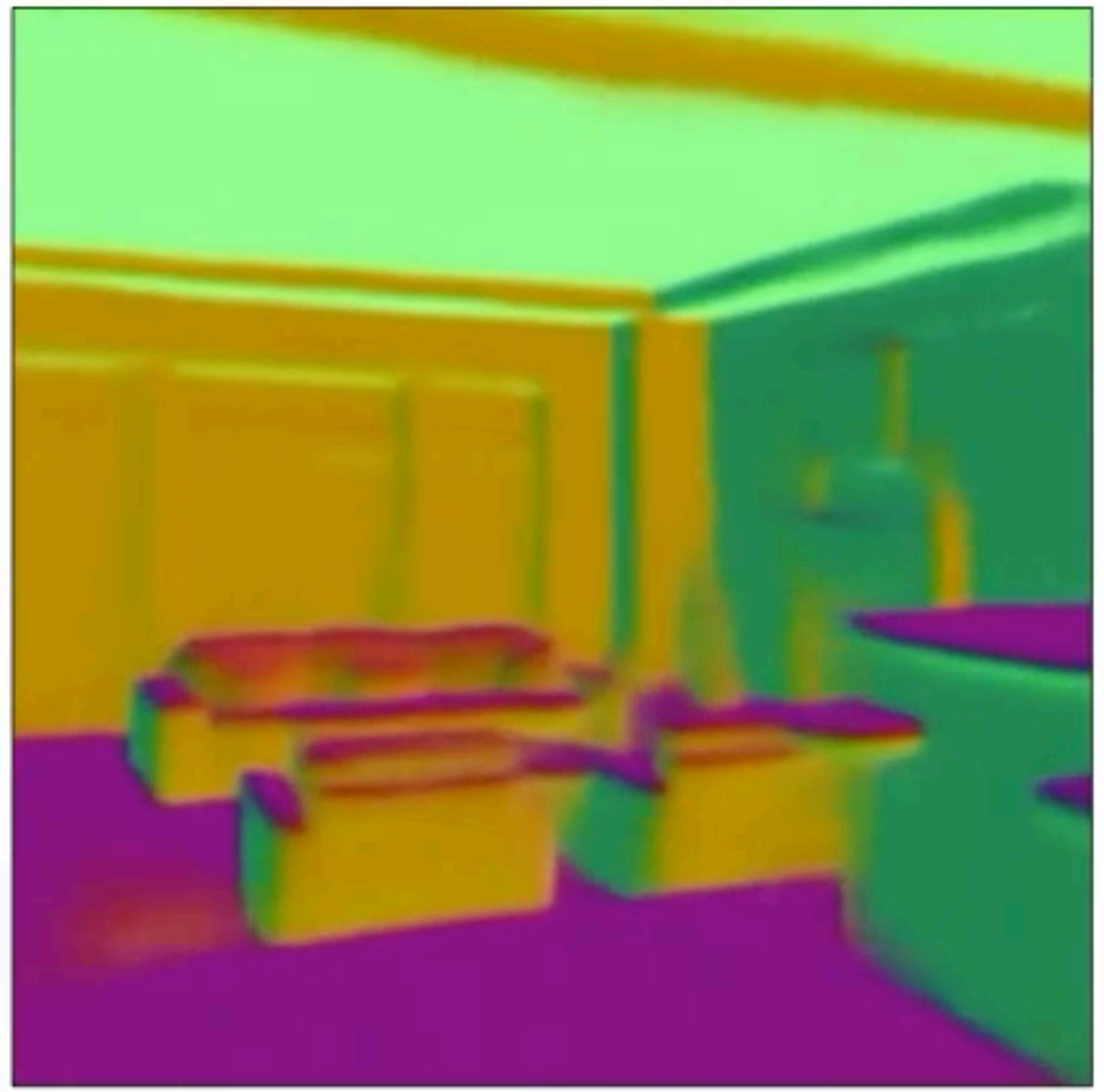
Cycle-Consistency $x \leftrightarrows y$



Pix2pix



GeoNet



Frame-by-frame results on the test video of [47].

(Full video in extended slides on project webpage.)

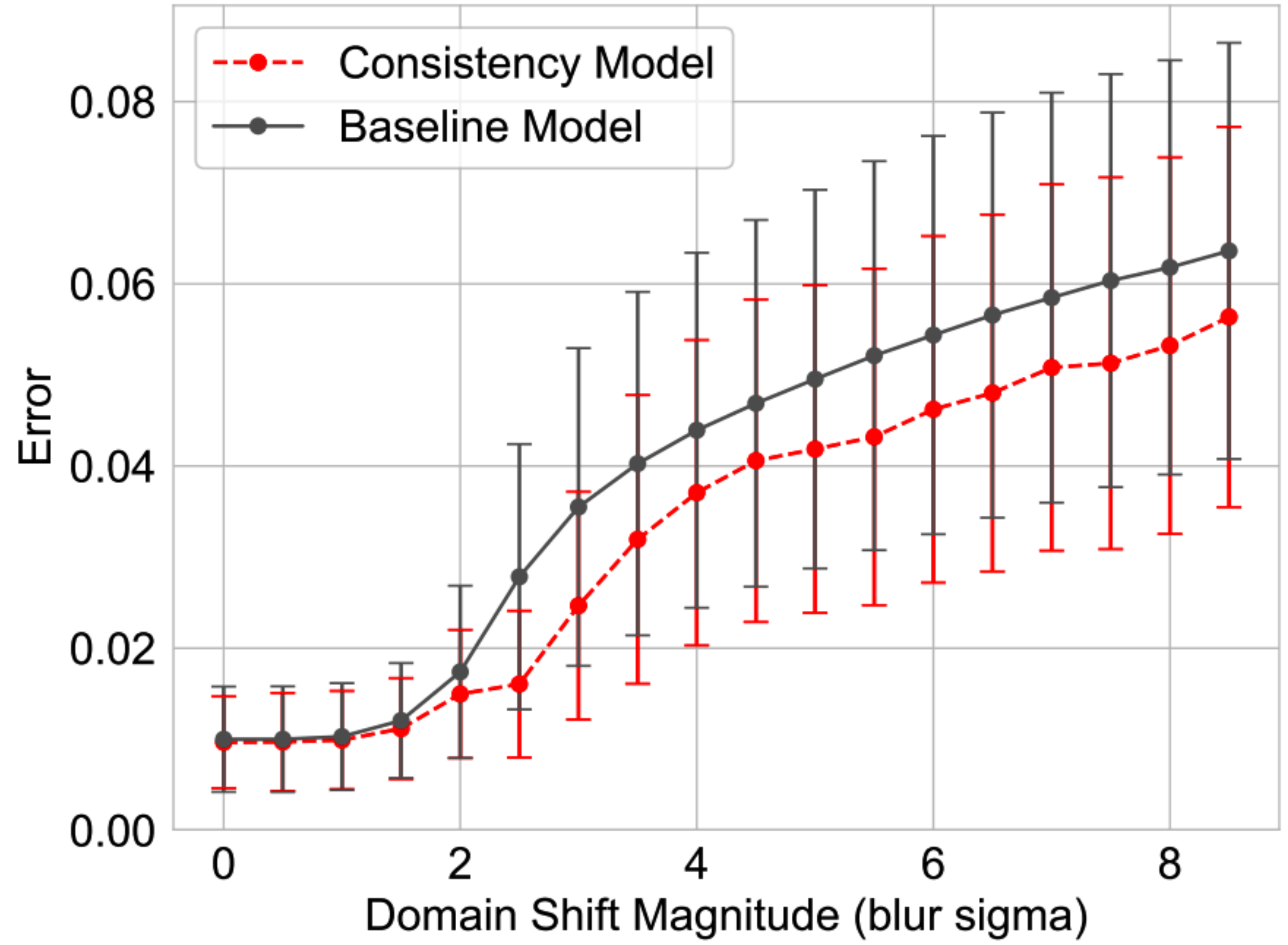
Accuracy of Predictions (quantitative)

Replica Dataset (high resolution 3D GT)

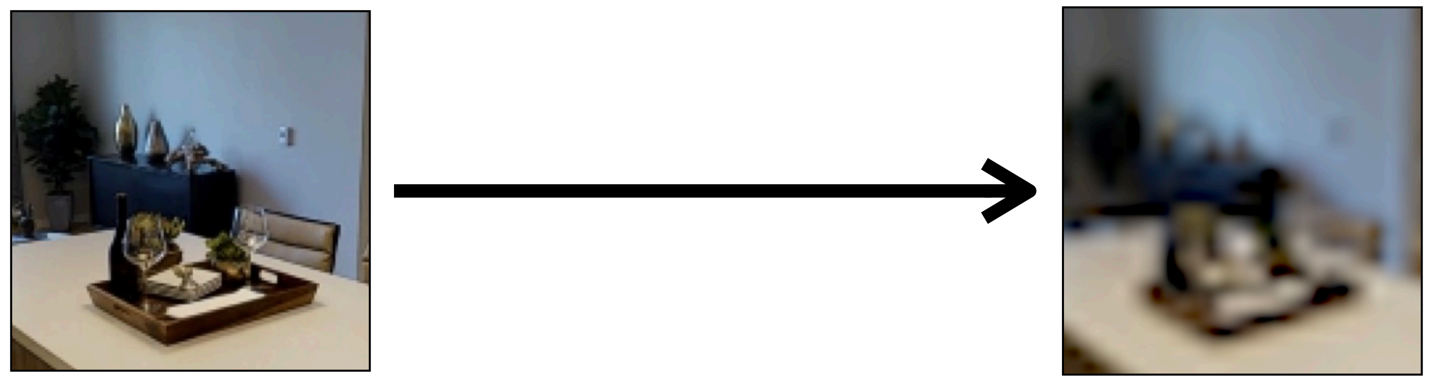
Taskonomy Dataset

Setup / Method	Replica Dataset									Taskonomy Dataset															
	Normals			**Depth**			**reShading**			**Normals**					**Depth**					**reShading**					**Semantic Segm.**
	Perceptual Err.		Direct	Perceptual Err.		Direct	Perceptual Err.		Direct	Perceptual Err.				Direct	Perceptual Err.				Direct	Perceptual Err.				Direct	Direct
	Depth	reShade	ℓ_1 Err.	Norm.	reShade	ℓ_1 Err.	Norm.	Depth	ℓ_1 Err.	Depth	reShade	Curv.	Edge(2D)	ℓ_1 Err.	Norm.	reShade	Curv.	Edge(2D)	ℓ_1 Err.	Norm.	Depth	Curv.	Edge(2D)	ℓ_1 Err.	X-Entropy (↓)
Blind Guess	4.75	33.31	16.02	22.23	19.94	4.81	15.74	5.14	16.45	7.39	38.11	3.91	12.05	17.77	22.37	27.27	7.96	12.77	7.07	19.96	7.14	3.53	12.62	24.85	
Taskonomy Networks	3.73	11.07	6.55	18.06	15.39	3.72	8.70	3.85	11.43	7.19	22.68	3.68	10.70	7.54	18.82	20.83	6.65	14.10	4.55	11.72	4.69	3.54	11.19	16.58	
Multi-Task	5.58	22.11	6.03	15.30	16.14	2.44	7.24	3.36	10.32	8.78	27.32	3.65	10.16	7.07	17.18	19.55	7.54	13.67	2.81	9.19	3.54	3.56	10.75	11.61	
GeoNet (original)	6.23	19.34	7.48	13.88	14.03	4.01	×	×	×	7.71	27.35	**3.32**	**9.09**	9.58	15.44	18.73	4.03	10.78	4.07	×	×	×	×	×	×
Cycle Consistency	5.65	22.39	7.13							8.81	30.33	3.84	10.26	8.68											
Baseline Perceptual Loss	4.88	15.34	4.99							8.59	23.98	3.41	10.01	6.17											
Pix2Pix	4.52	19.03	7.70							8.12	26.23	3.83	10.33	9.40											
Baseline UNet (ℓ_1)	4.69	13.15	4.96	10.47	12.99	1.99	6.90	2.74	**9.55**	8.17	20.94	3.41	9.98	**5.95**	13.62	15.68	7.31	12.61	**2.27**	9.58	3.38	3.78	10.85	**10.45**	0.246
GeoNet (updated)	4.62	12.79	**4.70**	10.47	12.75	1.83	×	×	×	8.18	20.84	3.40	9.99	**5.91**	13.77	15.76	7.52	12.67	**2.26**	×	×	×	×	×	×
X-Task Consistency	**2.07**	**9.99**	**4.80**	**7.01**	**11.21**	**1.63**	**5.50**	**1.96**	**9.22**	**4.32**	**12.15**	**3.29**	9.50	6.08	**9.46**	**12.66**	**3.61**	**9.82**	**2.29**	**7.13**	**2.51**	**3.28**	**9.38**	**10.52**	**0.237**

Generalization & Adaptation to out-of-training-domain data (quantitative)

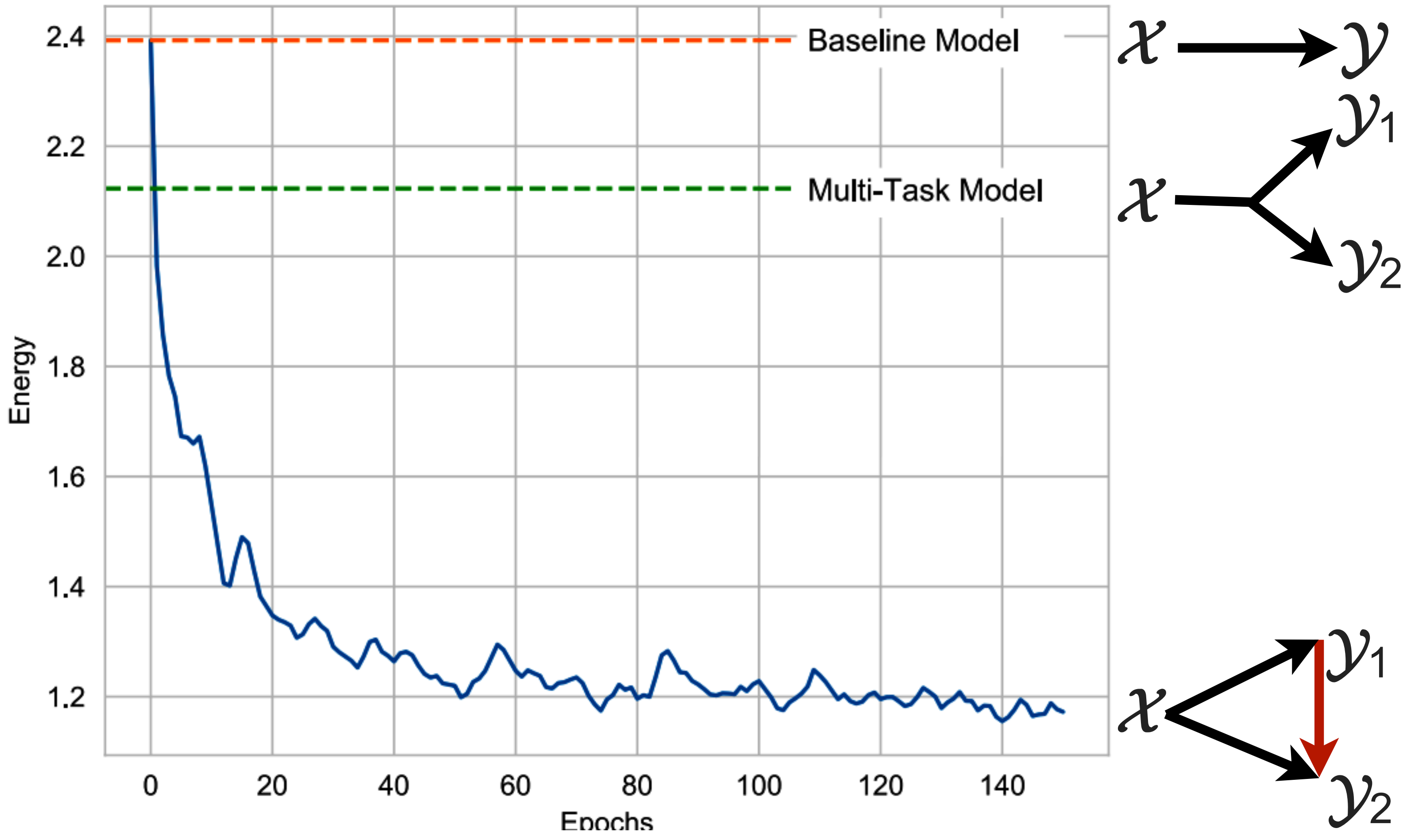


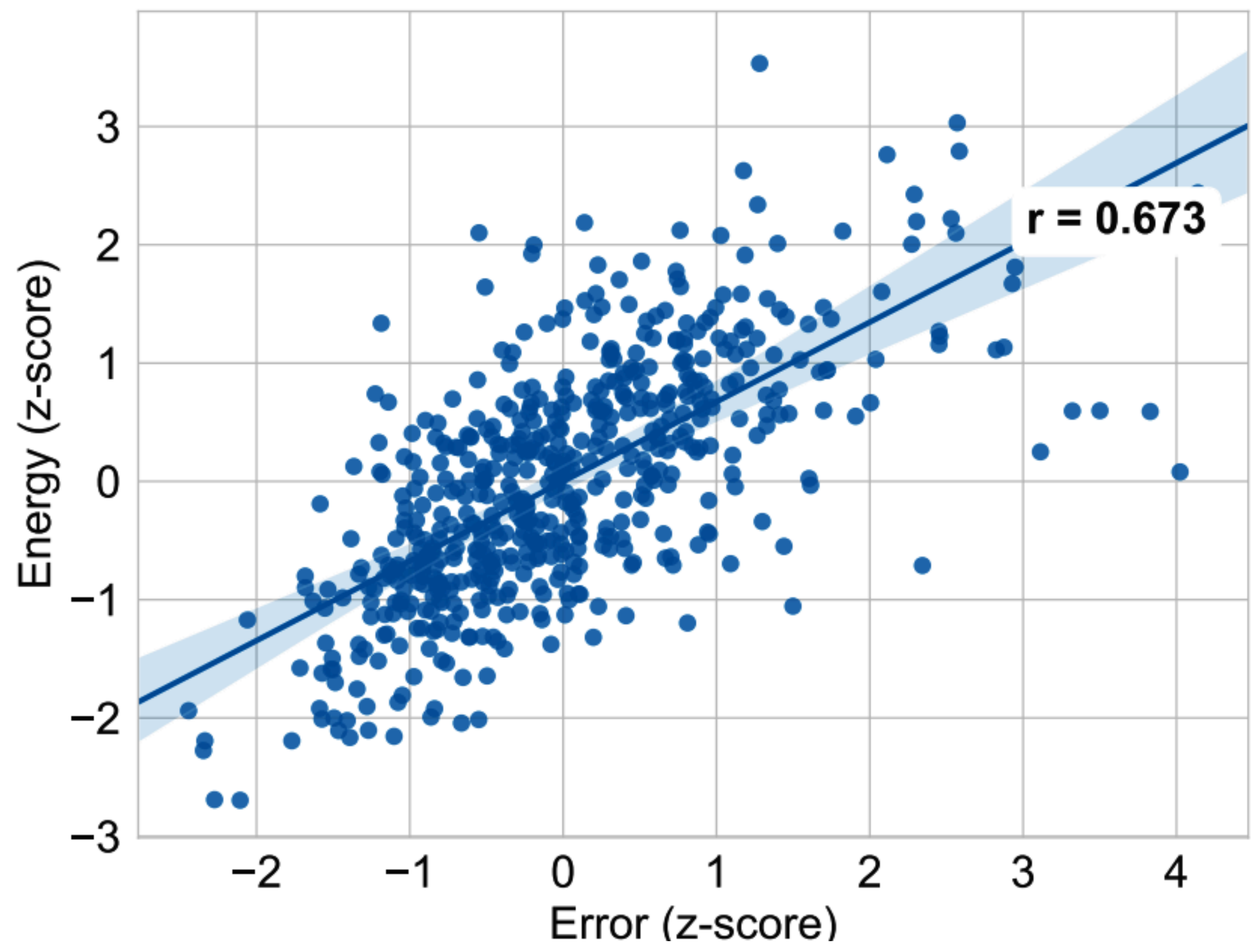
Novel Domain	Error (Post-Adaption)			Error (Pre-Adaptation)	
	# images	Consistency	Baseline	Consistency	Baseline
Gaussian blur (Taskonomy)	128	17.4 (+14.7%)	20.4	46.2 (+12.8%)	53.0
	16	22.3 (+8.6%)	24.4		
CocoDoom	128	18.5 (+19.2%)	22.9	54.3 (+15.8%)	64.5
	16	27.1 (+24.5%)	35.9		
ApolloScape	8	40.5 (+11.9%)	46.0	55.8 (+5.5%)	59.1



Consistency Energy Utilities

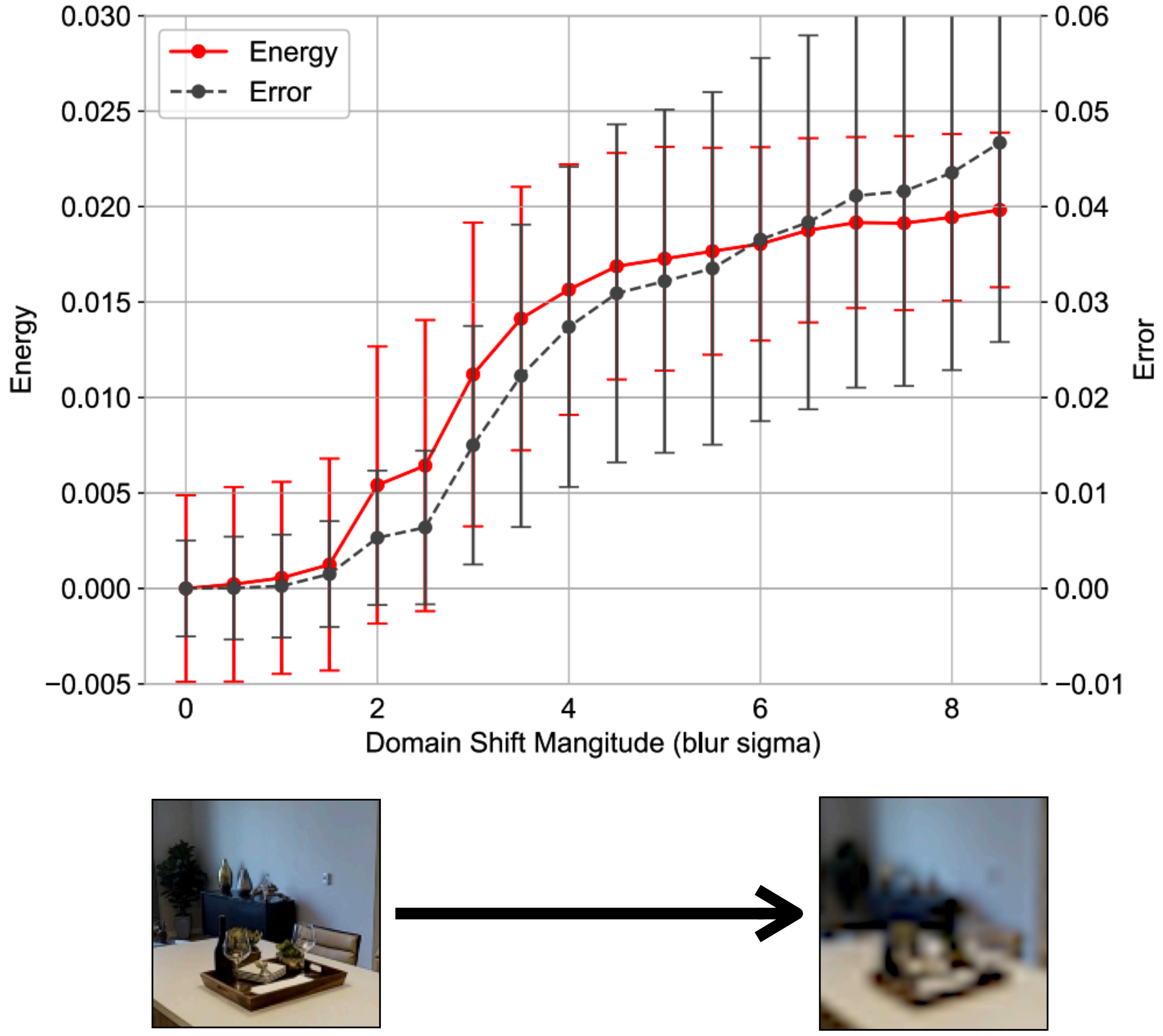
Energy (Inconsistency) During Training

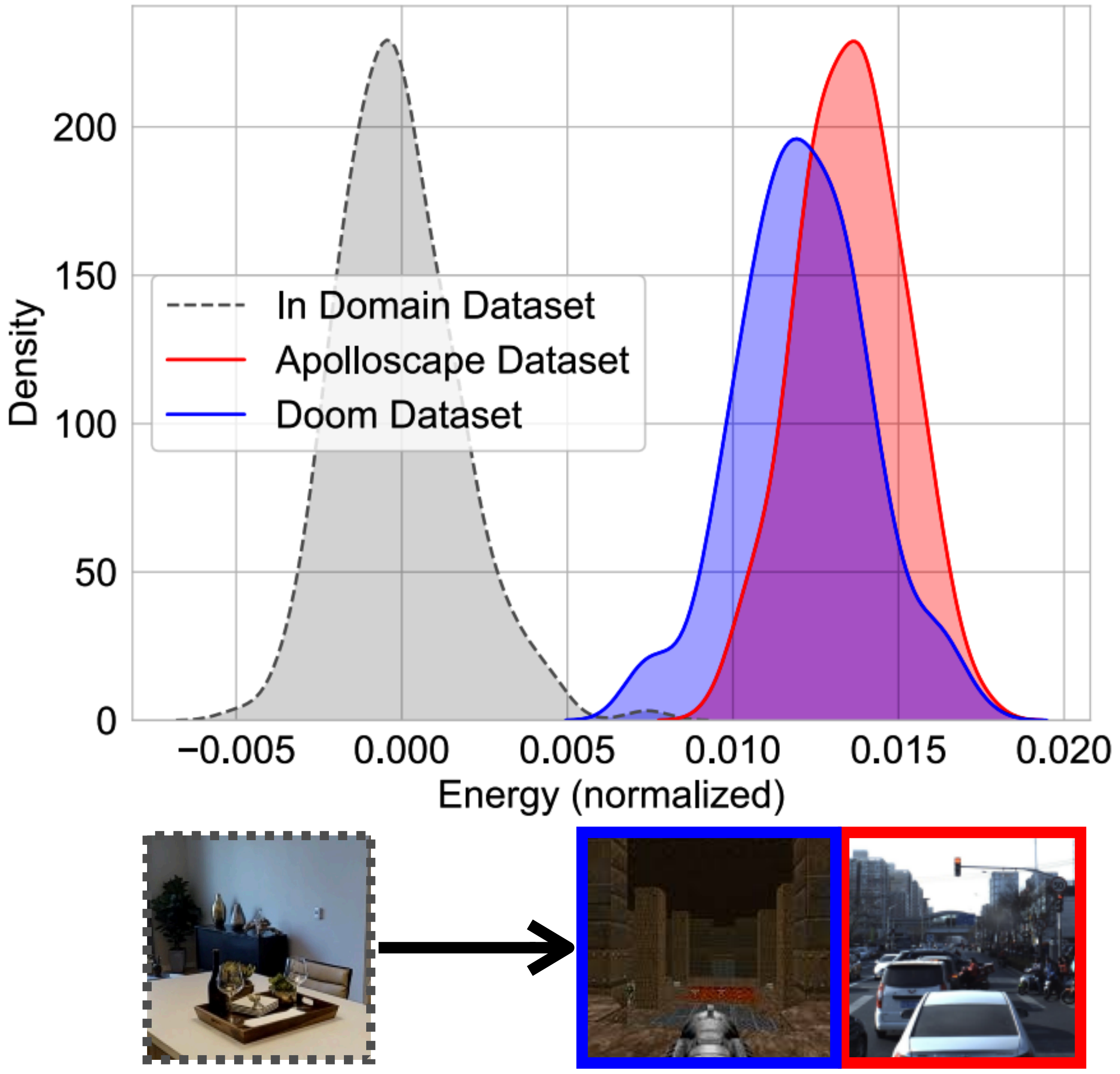




Energy vs Supervised Error (confidence estimator)

Energy vs Continuous Domain Shift

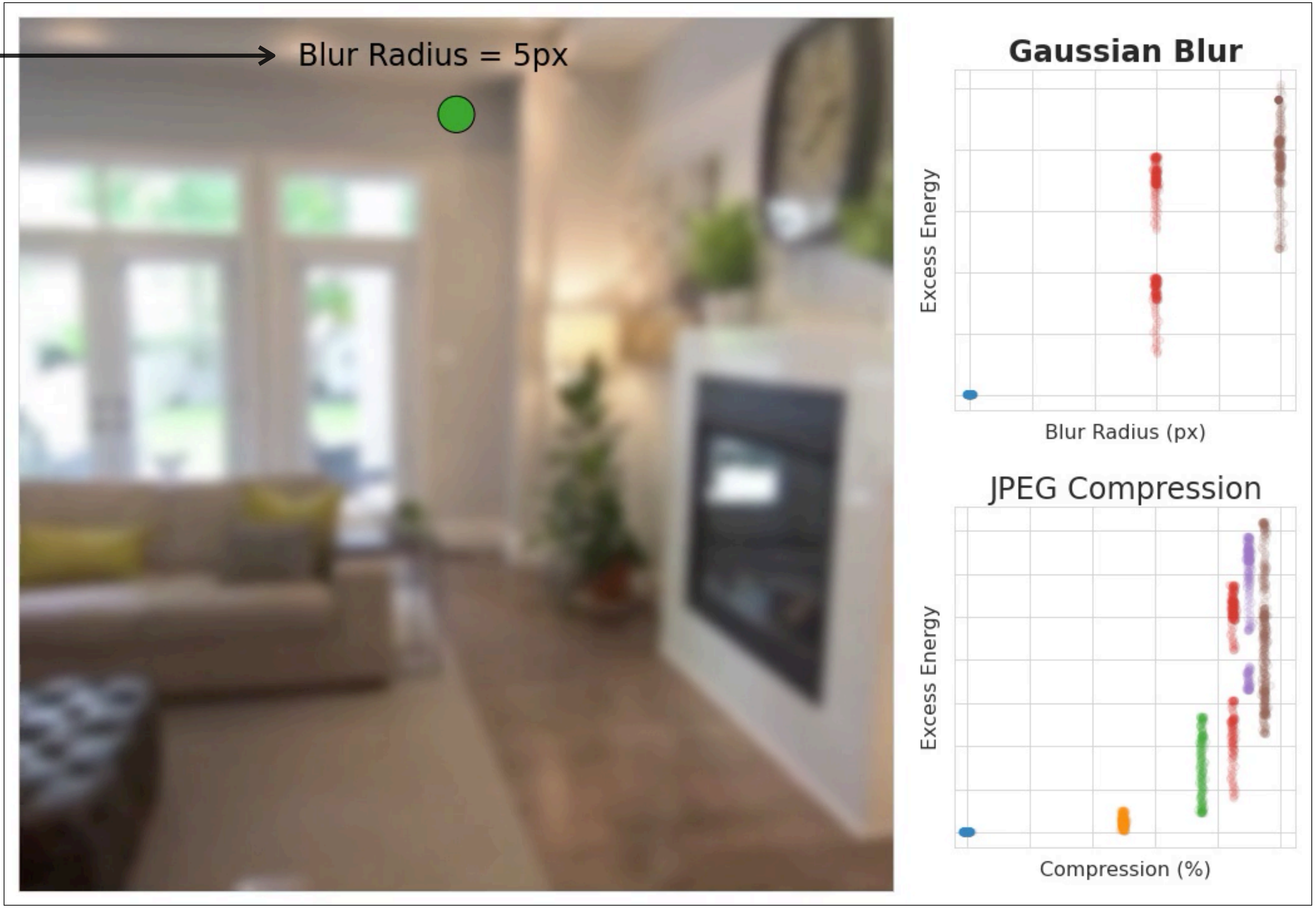




Energy vs Discrete Domain Shift

Consistency Energy as Domain Change Indicator

Energy vs Domain Shift



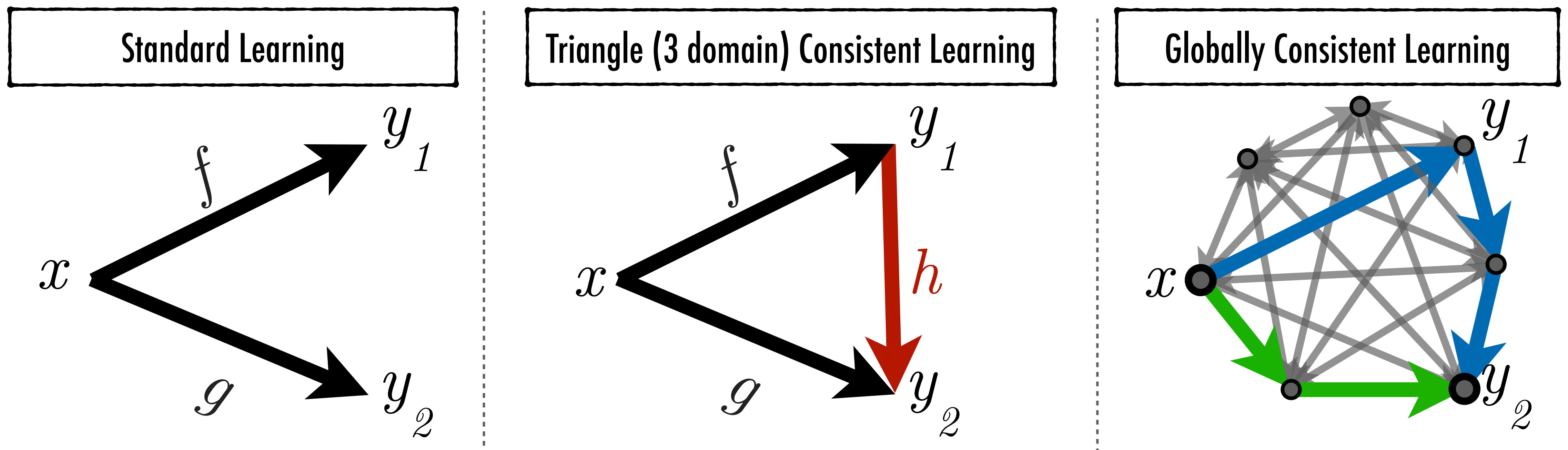
Frame-by-frame results on the test video of [47].

(Full video in extended slides on project webpage.)

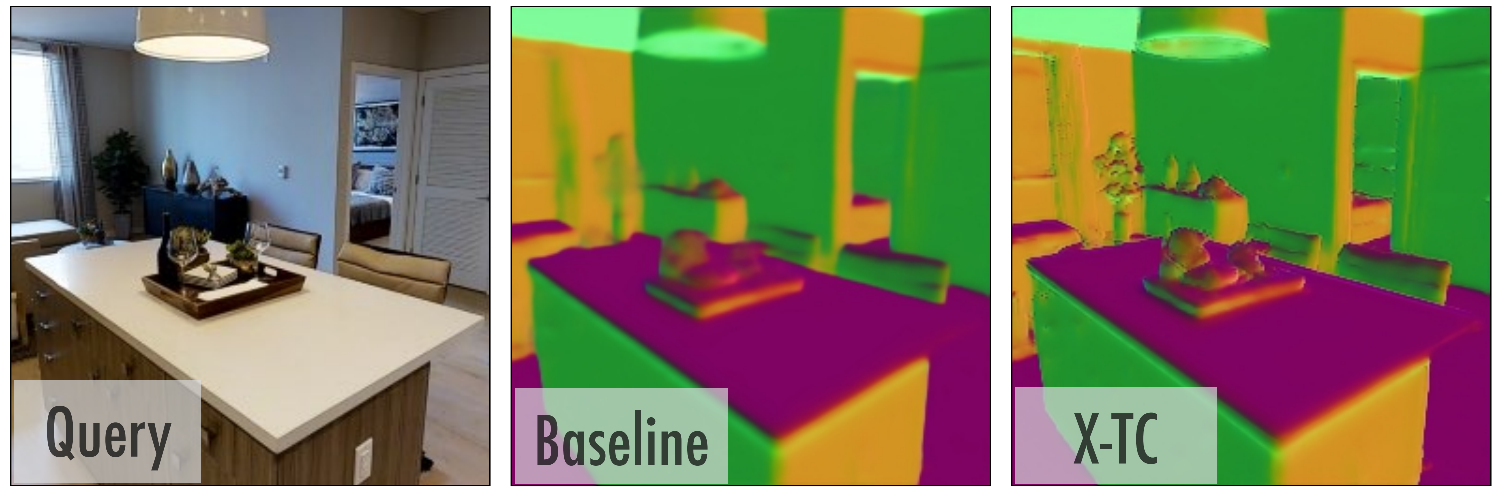
Summary

Summary

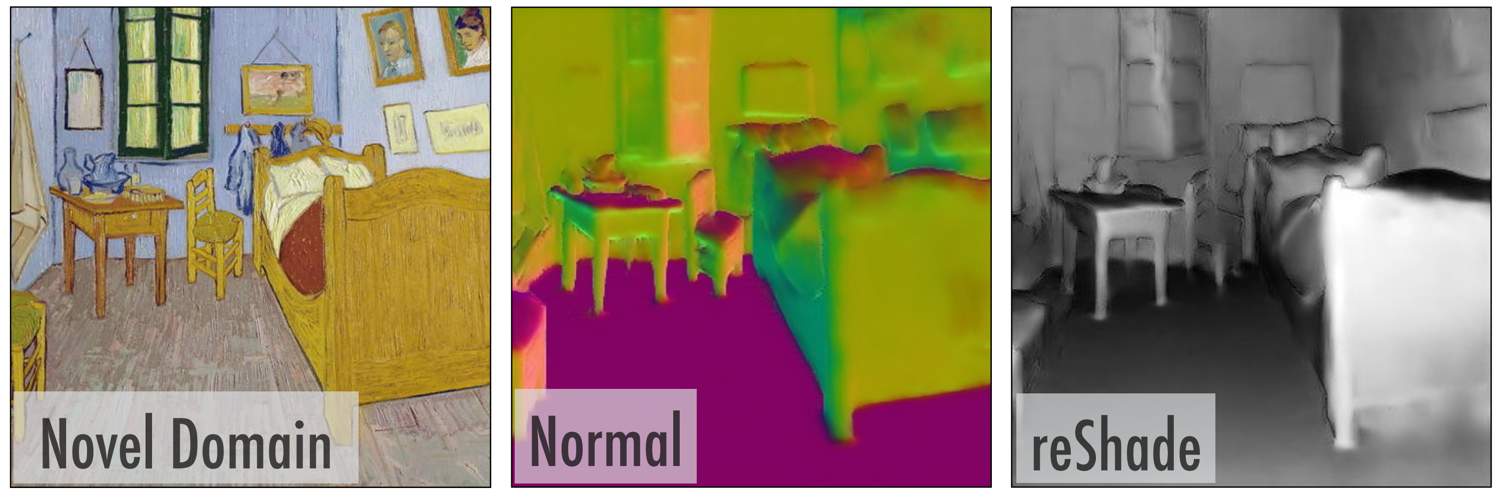
- A general method for augmenting standard supervised learning with Cross Task Consistency constraints.
- Based on ` Inference-Path Invariance’ concept.
- Constraints learned from data. No need to differentiable or apriori given analytical task relationships.
- Fully computational. Broadly applicable.



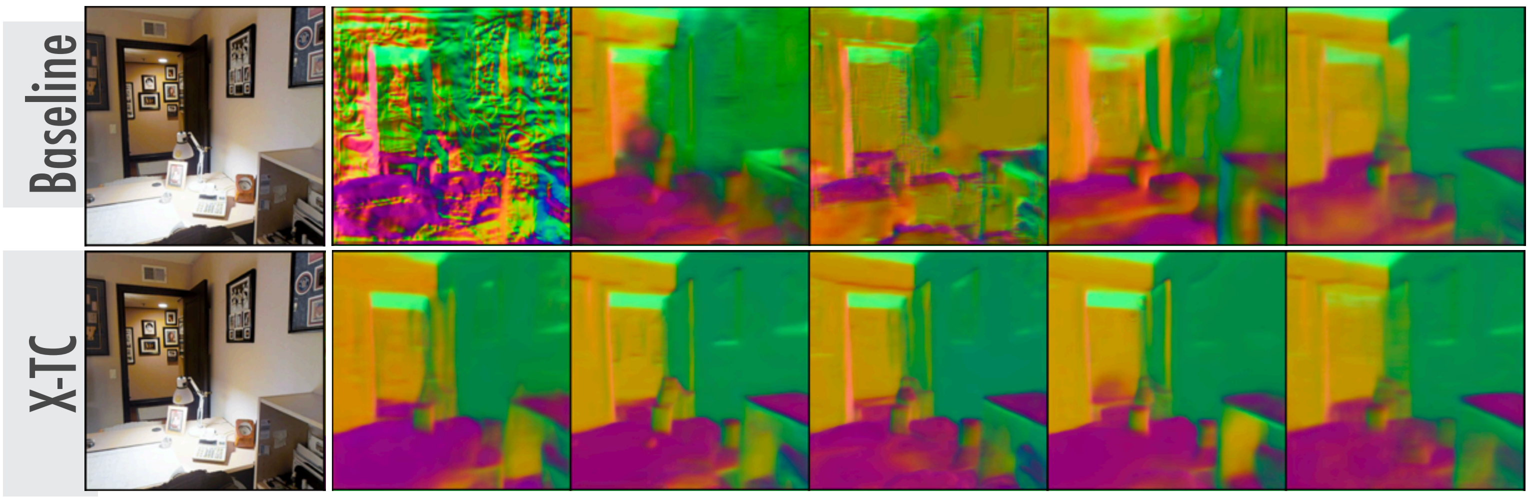
- Results in:
  - More accurate prediction



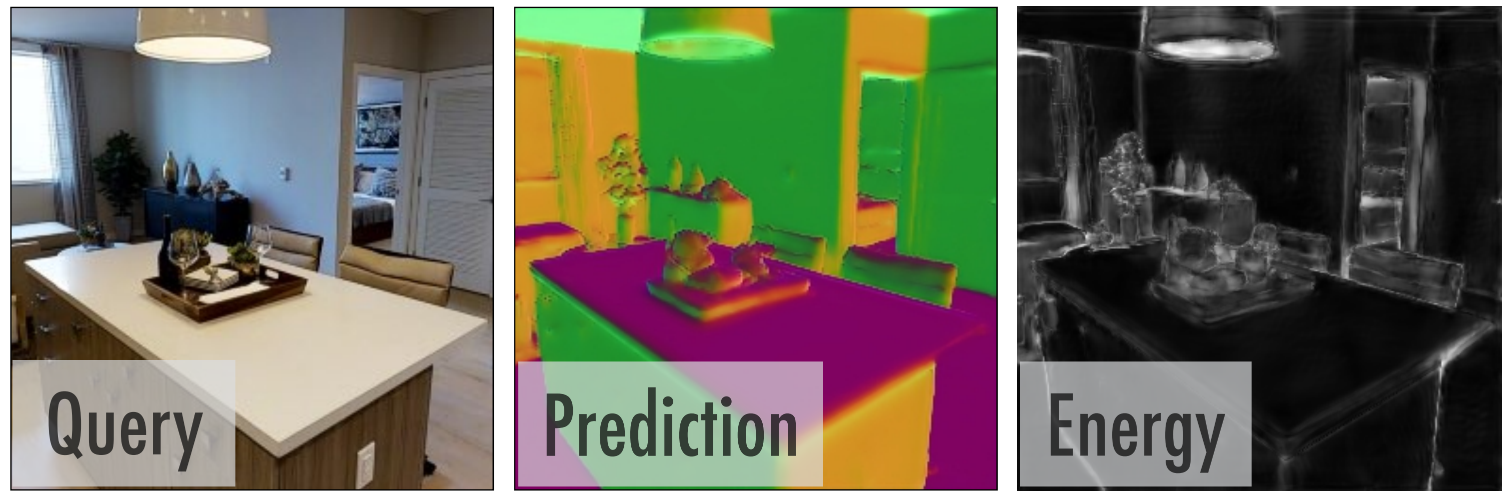
  - Enhanced generalization and adaptation to new data



  - More consistent predictions



- Consistency Energy: An informative test-time unsupervised quantity.



Robust Learning Through Cross-Task Consistency

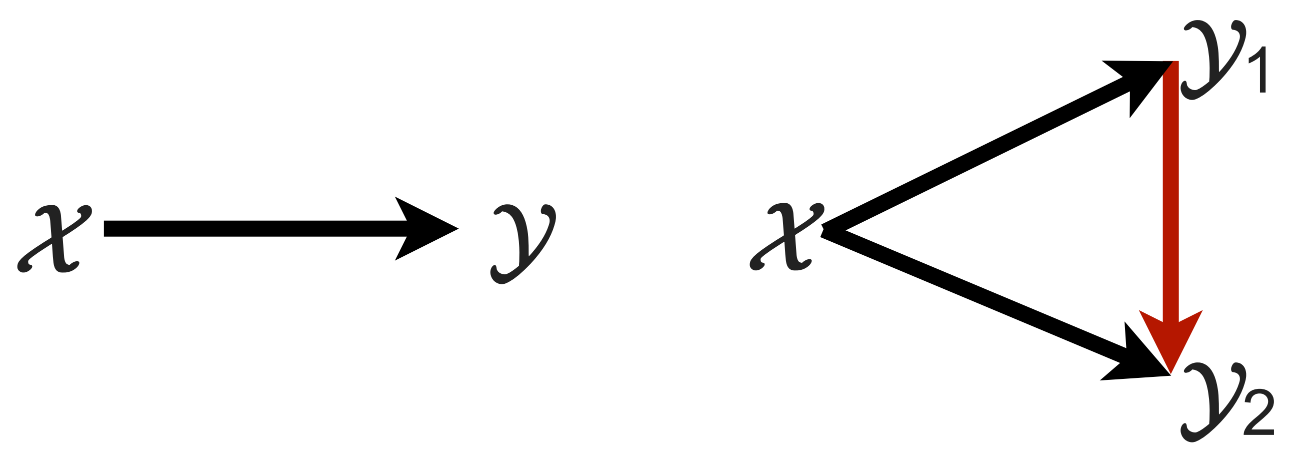
A. Zamir*, A. Sax*, N. Cheerla, R. Suri, Z. Cao, J. Malik, L. Guibas

EPFL Swiss Federal Institute of Technology

Berkeley UNIVERSITY OF CALIFORNIA

STANFORD UNIVERSITY

http://consistency.epfl.ch/



Baseline Learning

Cross-Task Consistent Learning

CVPR 2020