Monocular Depth Estimation Using Neural Regression Forest

Anirban Roy, Sinisa Todorovic

Oregon State University

Monocular Depth Estimation

Input Image

Our prediction

Monocular Depth Estimation

Input Image

Our prediction

- Prior Work:

  - Graphical models [Saxena et. al., IJCV 07; Saxena et. al., PAMI 2009; Liu et. al., ICCV 10; Liu et. al., CVPR 2014; Batra et. al., CVPR 12; Lam et. al., CVPR 15]

  - Deep learning framework [Eigen et. al., NIPS 14; Liu et. al., CVPR 14; Liu et. al., CVPR 15]

Monocular Depth Estimation

Input Image

Our prediction

- Major challenge:

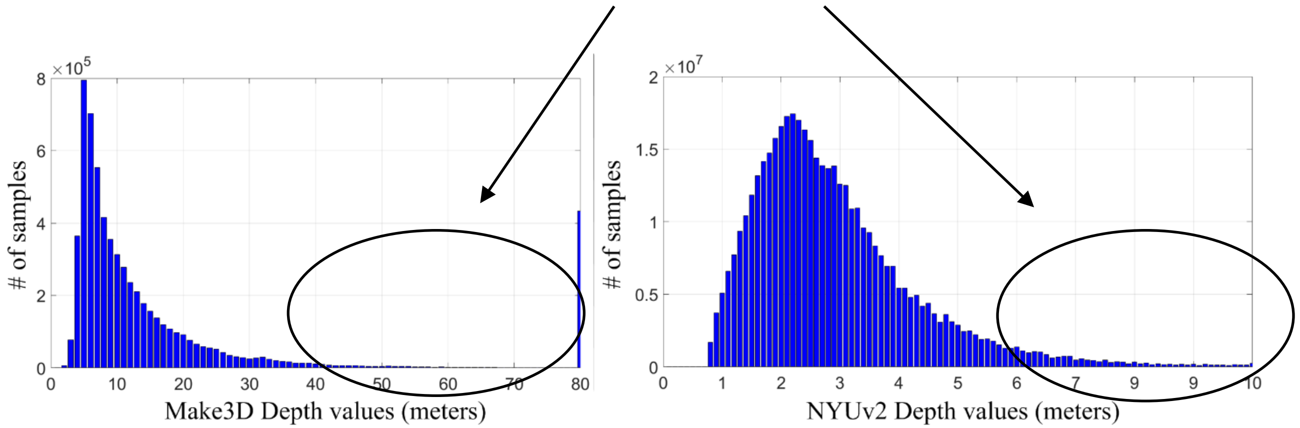
  Limited data for some depths

- Depth range:

  Make 3D [0 - 80] meters

  NYU v2   [0 -10] meters

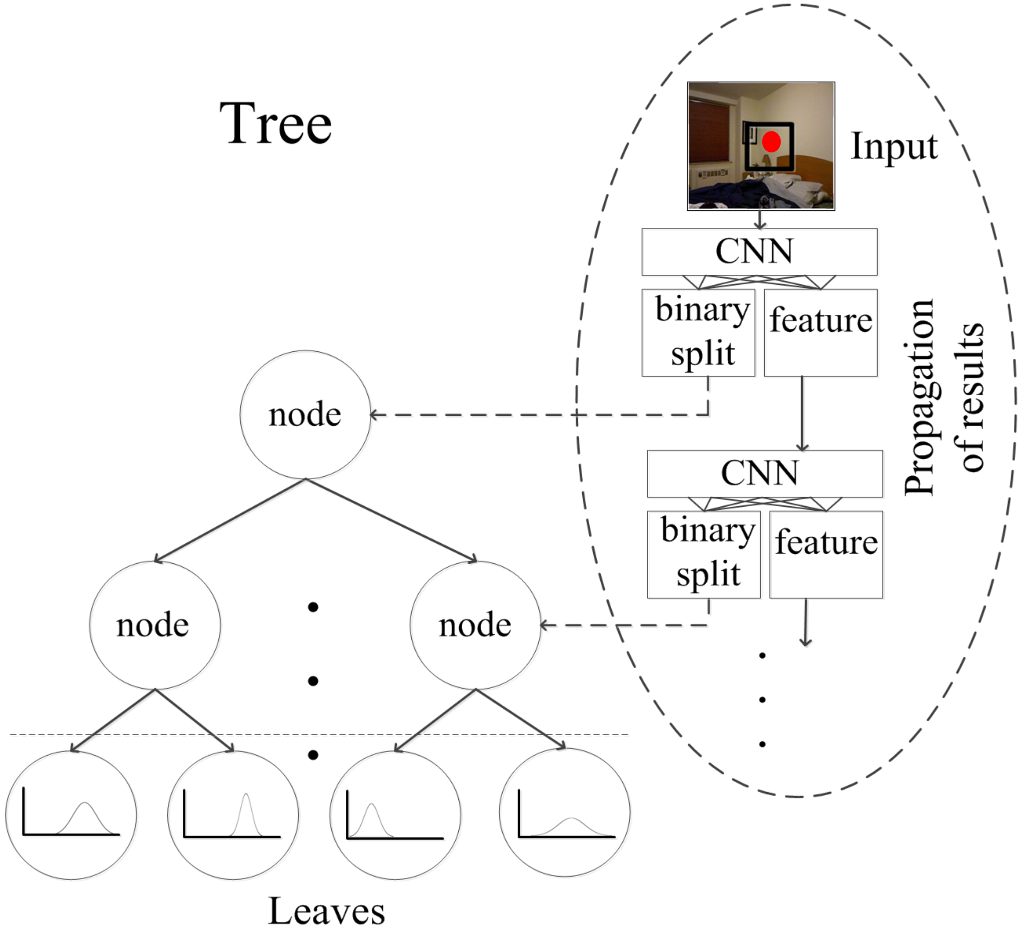
Limited training data



Depth histogram for Make3D and NYU v2

Our Approach: Neural Regression Forest

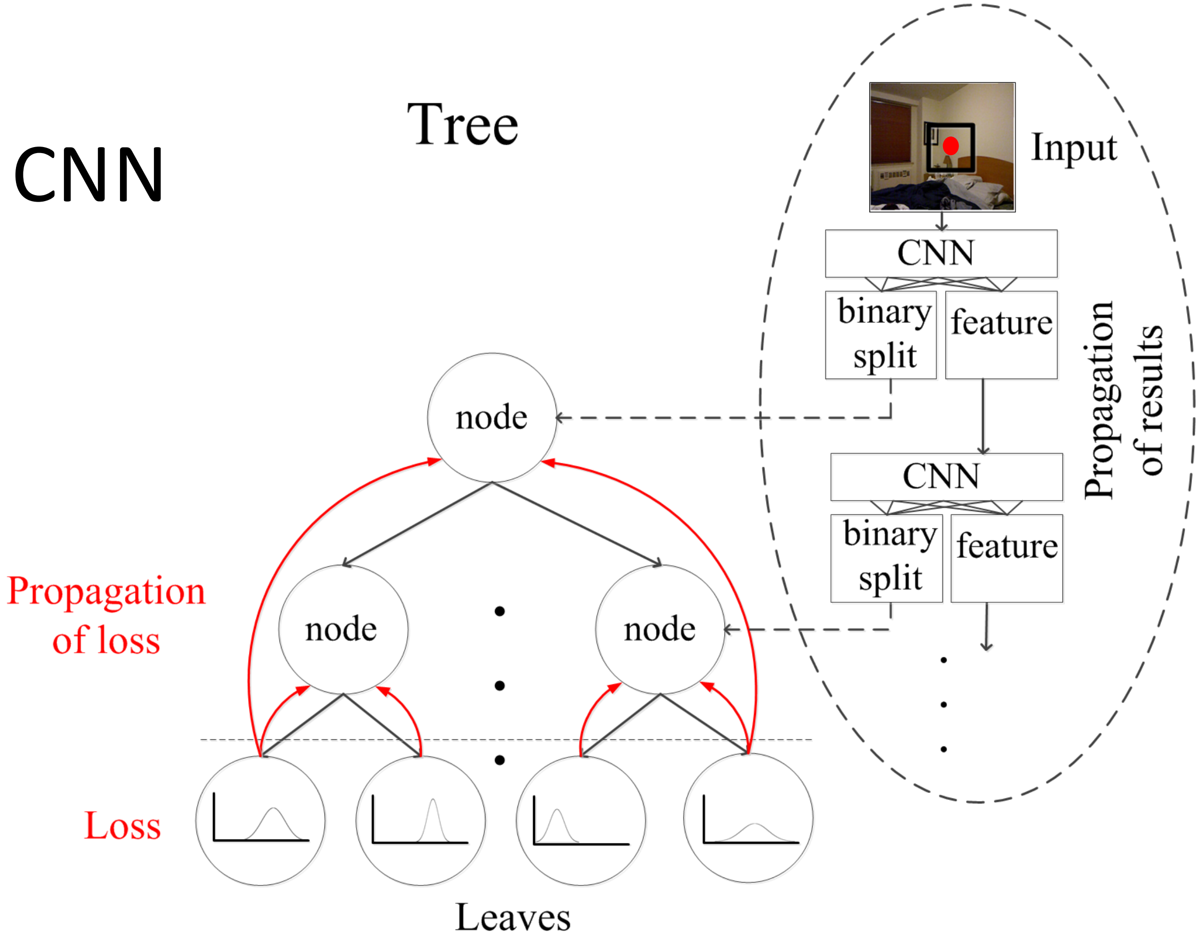
- Each split node represents a CNN



Neural Regression Forest

- Each split node represents a CNN

- Advantages:
  - Simpler binary decisions
  - Shallow CNNs
  - Parallel loss updates



Augmenting Training Data

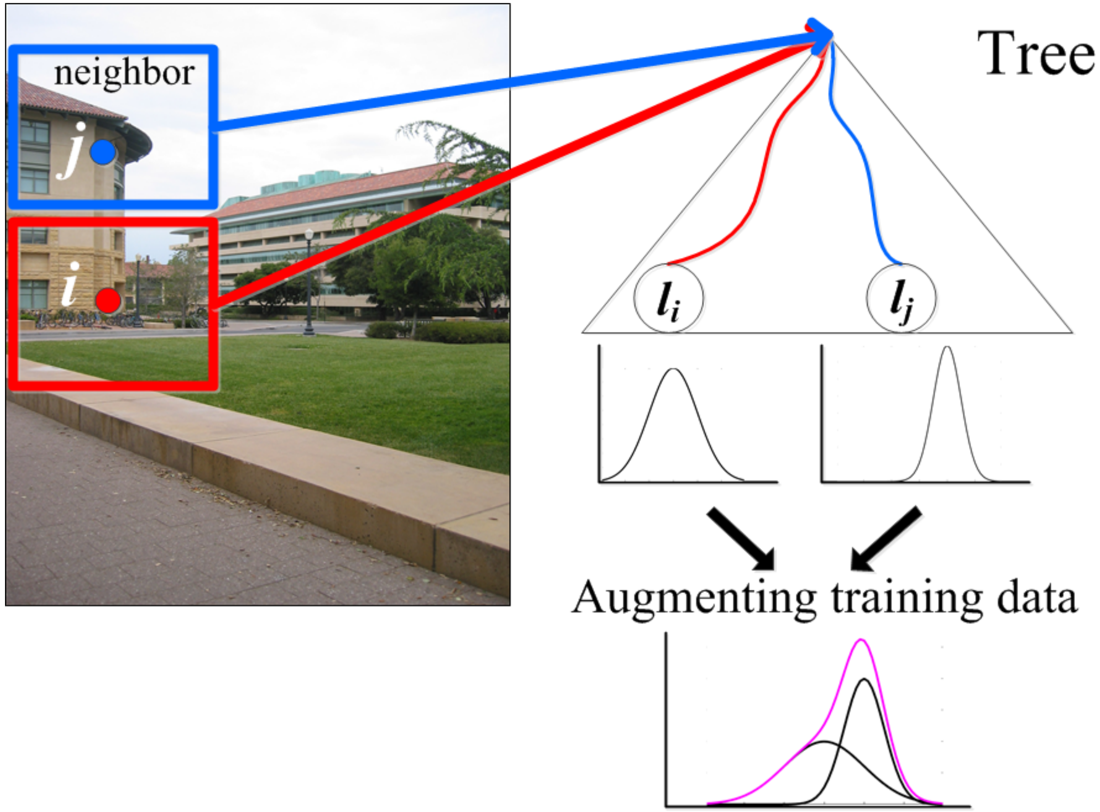
- Similar neighboring pixels

  ⬇

  Similar depths

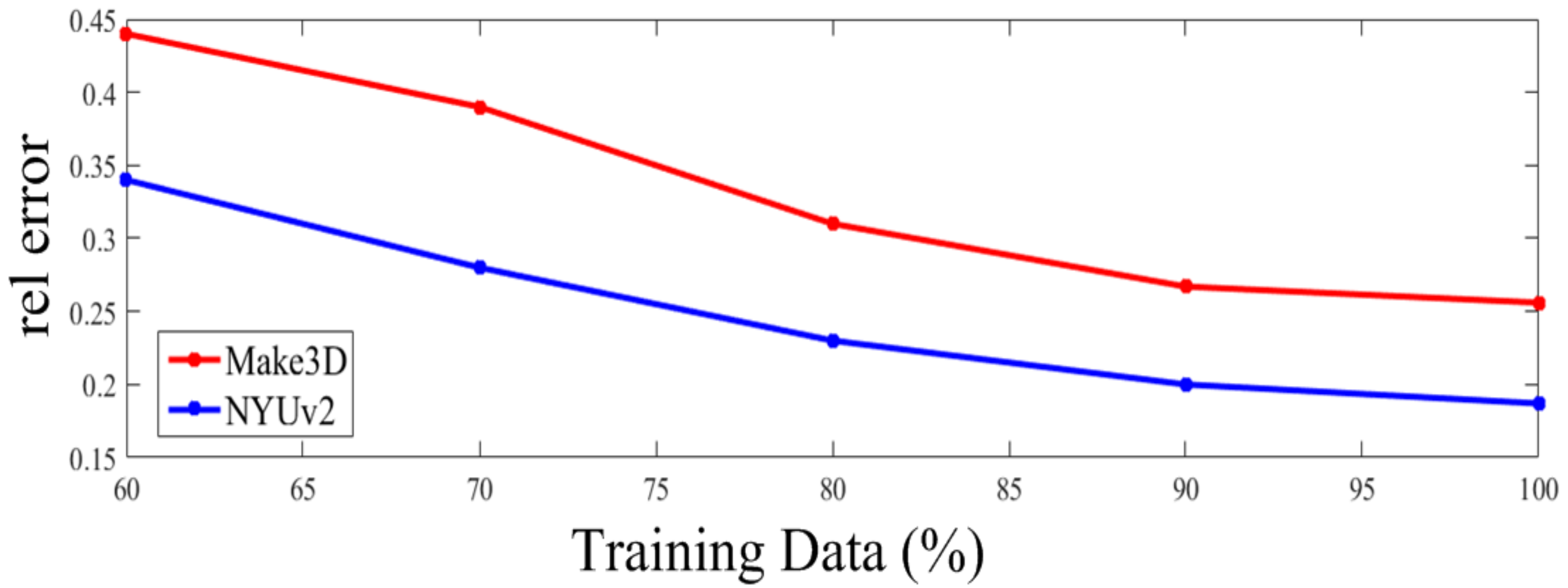
- Similarity in the image estimated efficiently by bi-lateral filtering

Prediction for i th pixel



More robust and smoother depth prediction

Evaluation of Robustness



Relative (rel): $\frac{1}{N}\sum(|d^* - \hat{d}|/d^*)$

Graceful increase in error for a reduction in the amount of training data

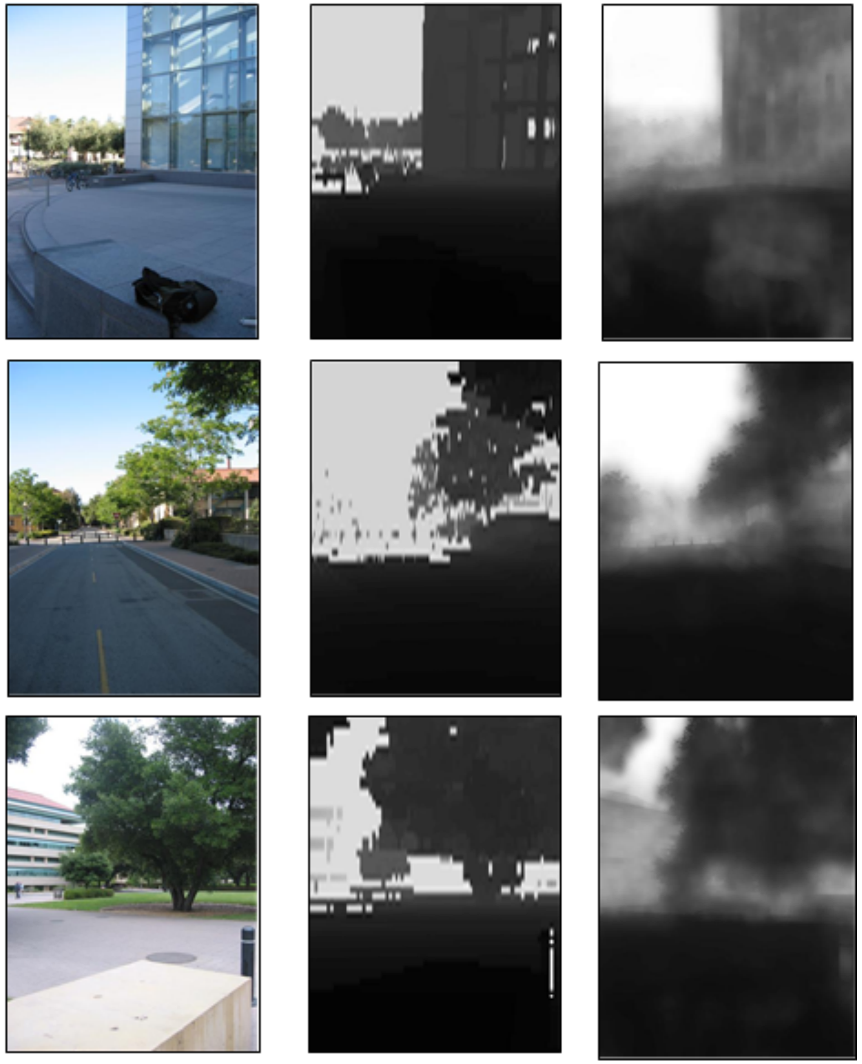
Results on the Benchmark Datasets

	Make3D		NYU v2	
	rel	rms	rel	rms
Saxena et. al. 2009	0.370	-	0.349	1.214
Batra et. al. 2012	0.362	15.8	-	-
Liu et. al. 2014	0.338	12.60	0.335	1.06
Karsch et. al. 2014	0.361	15.10	0.35	1.2
Lam et. al. 2015	0.364	-	-	-
Liu et. al. 2010	0.379	-	-	-
Eigen et. al. 2014	-	-	0.215	0.907
Liu et. al. 2014	0.307	12.89	0.230	0.824
Zhuo et. al. 2015	-	-	0.305	1.04
Ours	**0.26**	**12.40**	**0.187**	**0.744**
Error reduction	**- 0.04**	**-0.2**	**- 0.2**	**- 0.08**

Relative (rel): $\frac{1}{N}\sum(|d^* - \hat{d}|/d^*)$

Root mean square (rms): $\frac{1}{N}\sqrt{\sum(d^* - \hat{d})^2}$

Results on the Make3D outdoor Images



Input Image | Ground Truth | Our Approach

Results on the NYUv2 Indoor Images

Input Image | Ground Truth | Our Approach