Monocular Depth Estimation Using Neural Regression Forest

Anirban Roy, Sinisa Todorovic

Oregon State University

IEEE 2016 Conference on Computer Vision and Pattern Recognition

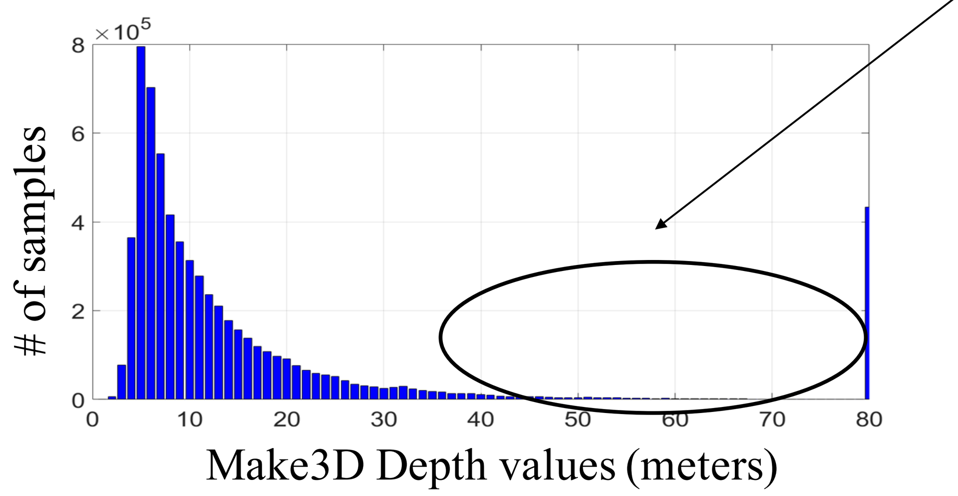
CVPR 2016

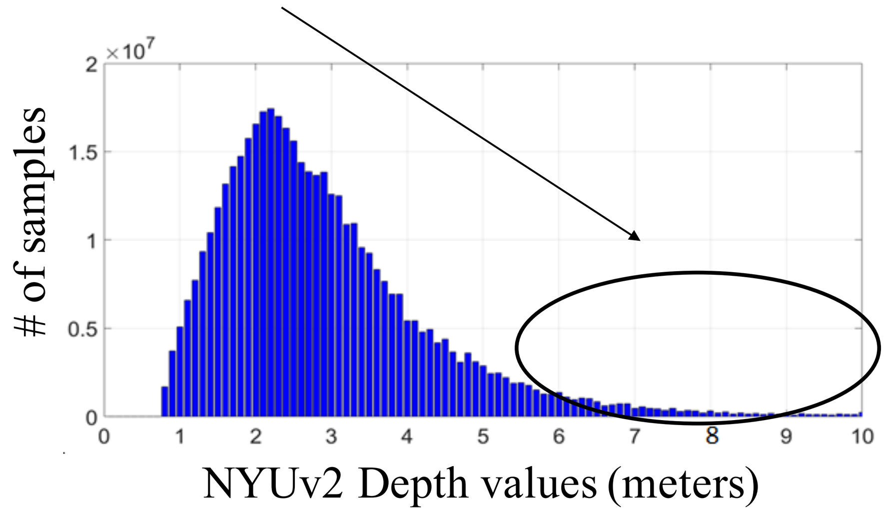
Problem statement

Input Image

Our depth prediction

Major challenge: Limited training data for some depths





Formulation

$$p_{\mathcal{F}}(d|\boldsymbol{x}) = \frac{1}{|\mathcal{F}|}\sum_{\mathcal{T}\in\mathcal{F}} p_{\mathcal{T}}(d|\boldsymbol{x})$$

Depth prediction of $\boldsymbol{x}$ by the forest

Depth prediction of $\boldsymbol{x}$ by a tree

$$p_{\mathcal{T}}(d|\boldsymbol{x}) = \sum_{l\in\mathcal{T}} p(d|l)P(l|\boldsymbol{x})$$

Depth prediction of $\boldsymbol{x}$ by leaf l

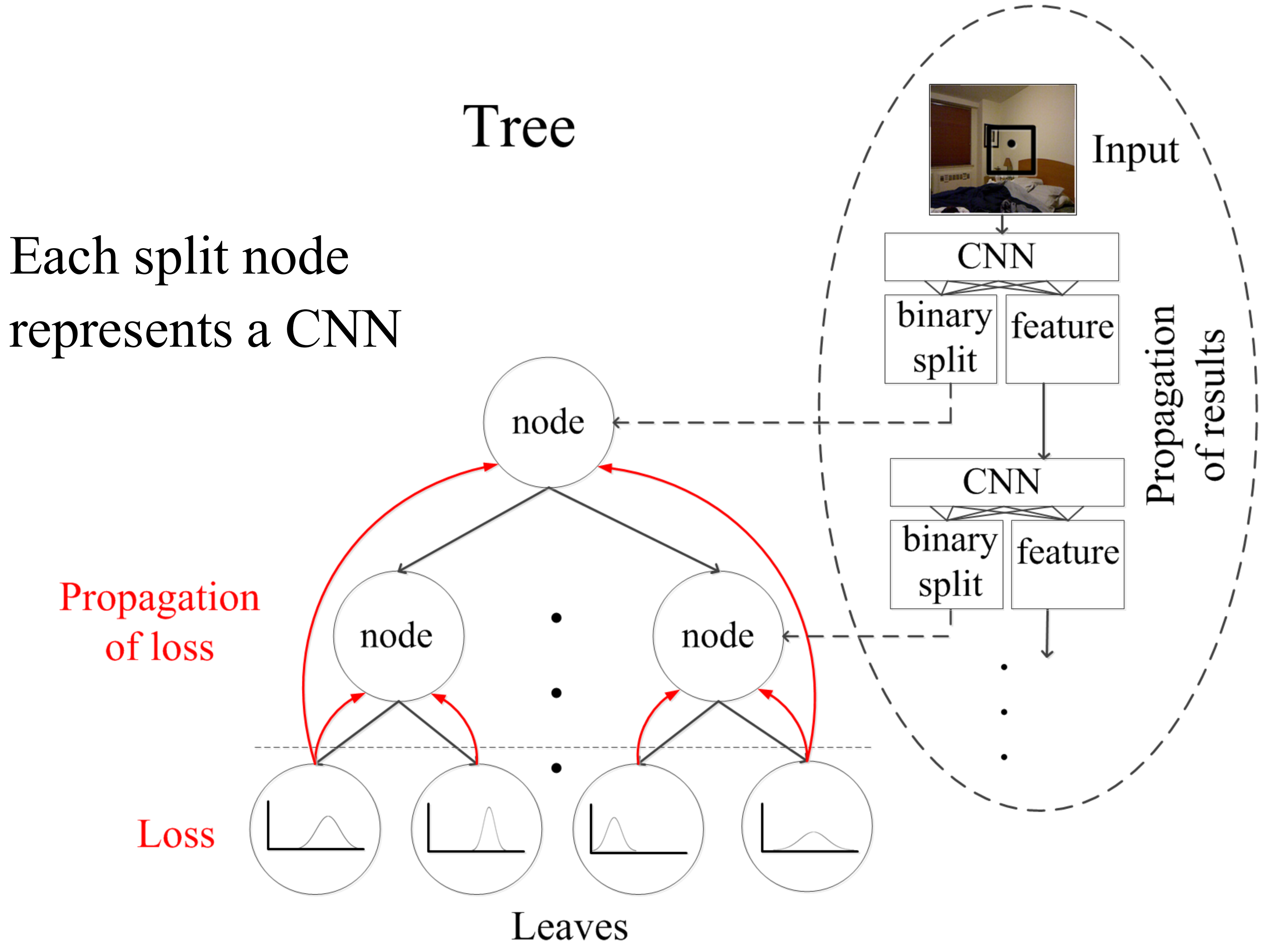
Probability of $\boldsymbol{x}$ reaching leaf l

$$P(l|\boldsymbol{x}) = \prod_{v\in\mathcal{V}} \underbrace{f_v(\boldsymbol{x})^{\mathbb{L}(l,v)}}\underbrace{(1-f_v(\boldsymbol{x}))^{\mathbb{R}(l,v)}}$$

Probability of $\boldsymbol{x}$ reaching leaf l

Split functions for left and right children of node v

Our Approach: Neural Regression Forest



Advantages:

- Binary decisions
- Shallow CNNs
- Parallel loss updates

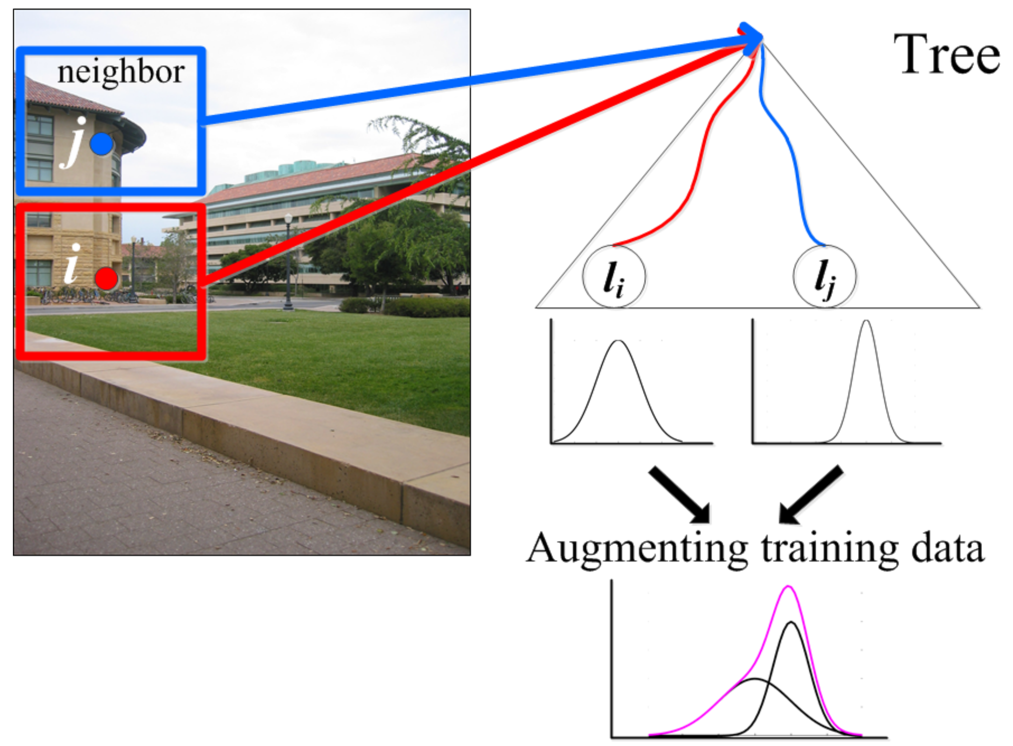
Augmenting Training Data

- Similar neighboring pixels

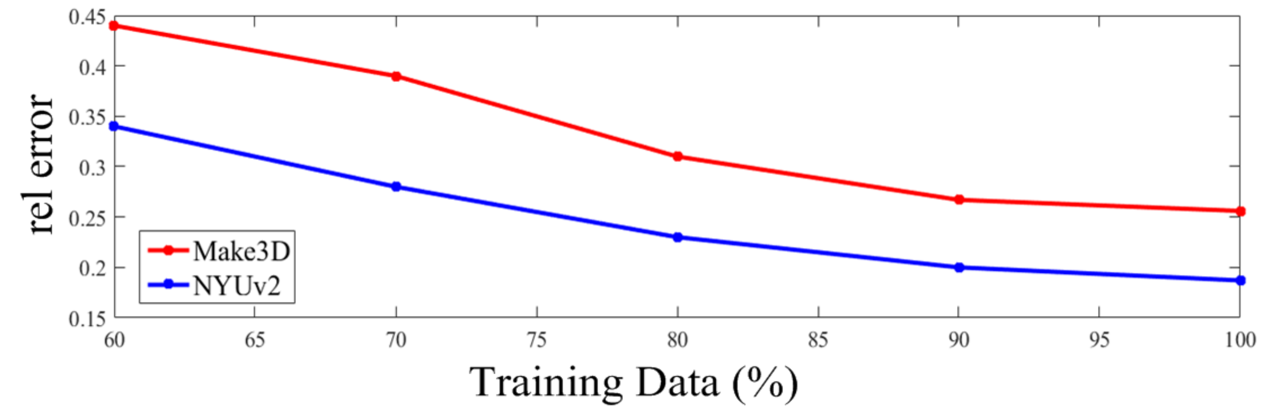
  ⬇

  Similar depths

- Similarity estimated by bilateral filtering



Evaluation of Robustness



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

Graceful increase in error when the training set size decreases

Results on the Benchmark Datasets

	Make3D		NYU v2	
	rel	rms	rel	rms
Liu et. al. 2014	0.338	12.60	0.335	1.06
Eigen et. al. 2014	-	-	0.215	0.907
Liu et. al. 2014	0.307	12.89	0.230	0.824
Lam et. al. 2015	0.364	-	-	-
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 error (rel):

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

Root mean square error (rms):

$\frac{1}{N}\sqrt{\sum(d^* - \hat{d})^2}$

Input Image | Ground Truth | Our Approach | Input Image | Ground Truth | Our Approach

Acknowledgment: NSF RI1302700. We appreciate Jose Roberto Alvarez and Bennett Wilburn for their feedback.