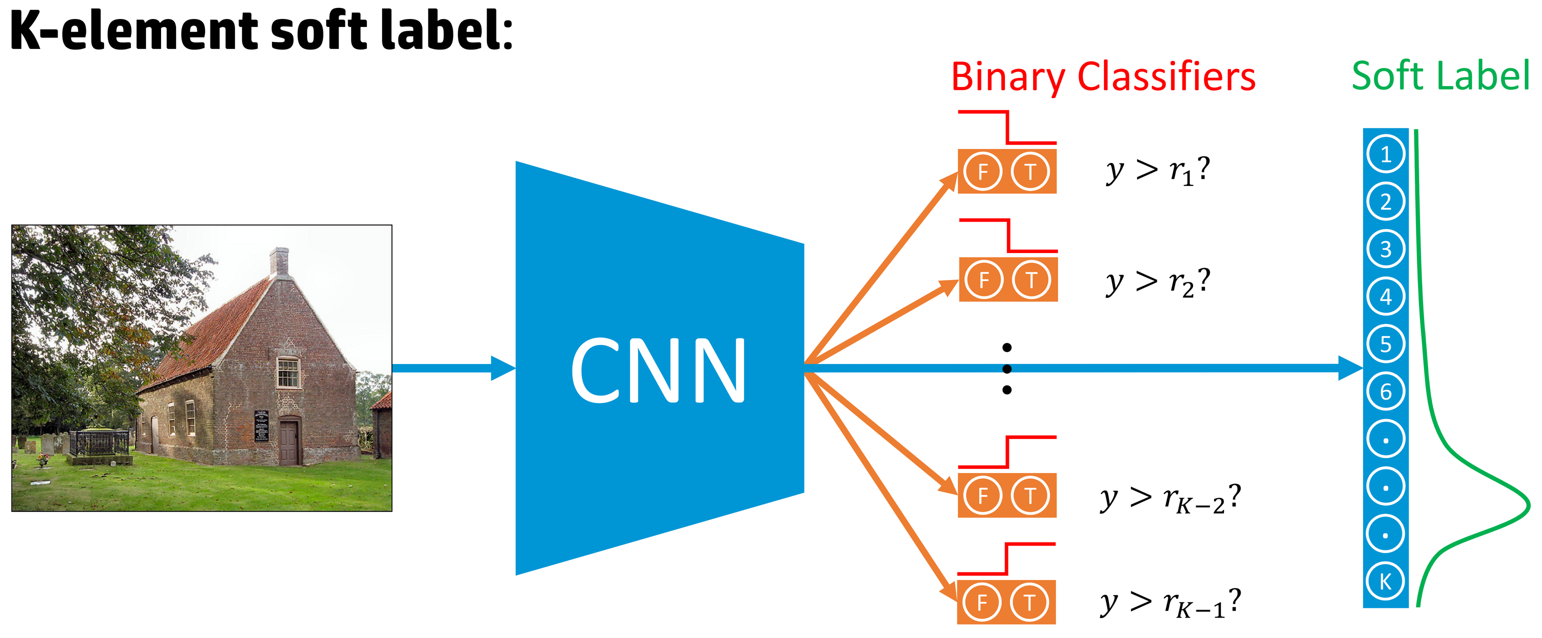
# Soft Labels for Ordinal Regression

Raúl Díaz and Amit Marathe
{raul.diaz.garcia, amit.marathe}@hp.com

## 1. Objective

Our goal is to substitute the **1-hot label** representations in ordinal regression problems. Instead of building an ensemble of **K-1 binary classifiers** that distinguish rank thresholds, we use the same default, fully-connected output layer of an off-the-shelf classification CNN with a **K-element soft label**:




## 2. Method

Our **S**oft **Ord**inal labels (**SORD**) encode metric penalties *à la Softmax*:

$$y_i = \frac{e^{-\phi(r_t, r_i)}}{\sum_{k=1}^{K} e^{-\phi(r_t, r_k)}} \quad \forall r_i \in \mathcal{Y}$$

- $\mathcal{Y} = \{r_1, r_2, \ldots, r_K\}$ is the list of ordinal ranks in the problem.
- $\phi$ penalizes how far a given rank is from the true value $r_t$.

When using Softmax and a cross-entropy loss, the computation of the gradient with respect to the output probits $p_i$ is simply:

$$\frac{\partial L}{\partial p_i} = -\frac{e^{-\phi(r_t, r_i)}}{e^{o_i + logC}}$$

**SORD** trains the network to learn features so that the outputs $o_i$ for each rank match the interclass distance $\phi$, reaching the minima when:

$$o_i + logC = -\phi(r_t, r_i) \quad \forall r_i \in \mathcal{Y}$$

**SORD Features:**

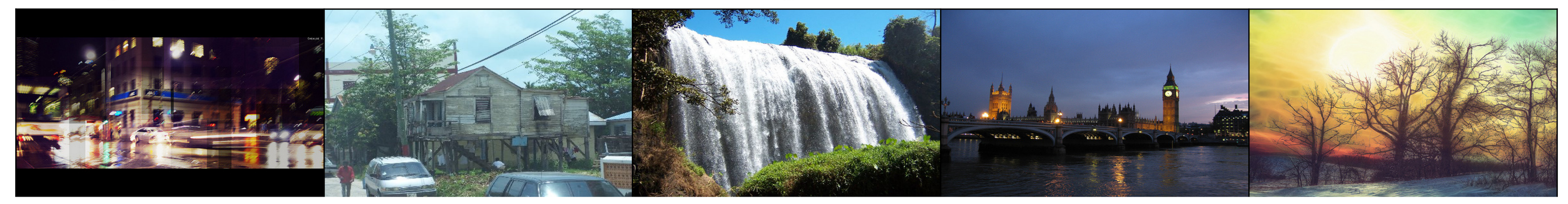
- General-purpose: a wild variety of tasks can use SORD.
- Simple to implement: just two lines of code.
- No new CNN architectures: an off-the-shelf network is sufficient.
- Inference can be done with either argmax or expected value.
- Continuous domains: $\phi$ is valid when $r_t \notin \mathcal{Y}$. This smoothly balances the SORD label without needing discretization or quantization.

## 3. Image Ranking: inputs with equally spaced labels

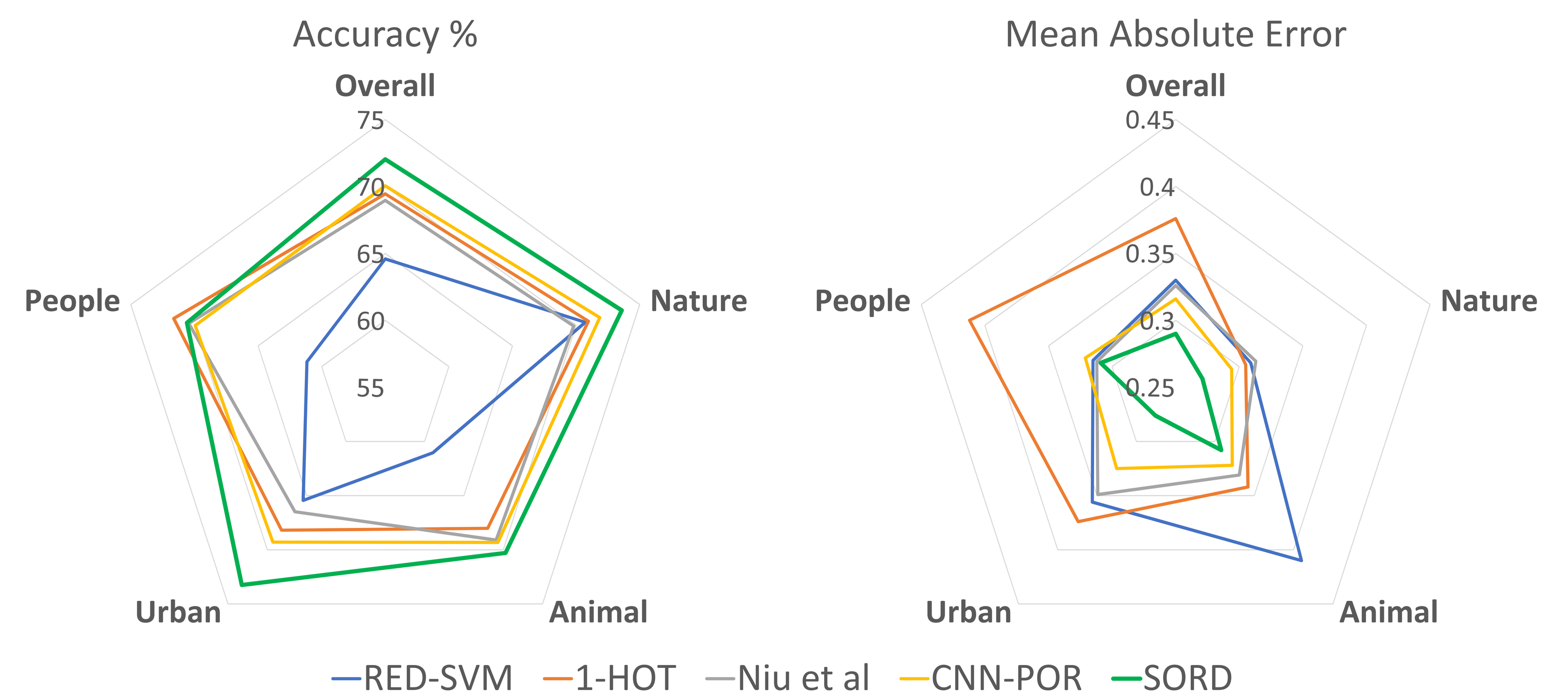
We use an off-the-shelf **VGG-16** and the metric:

$$\phi(r_t, r_i) = |r_t - r_i|$$

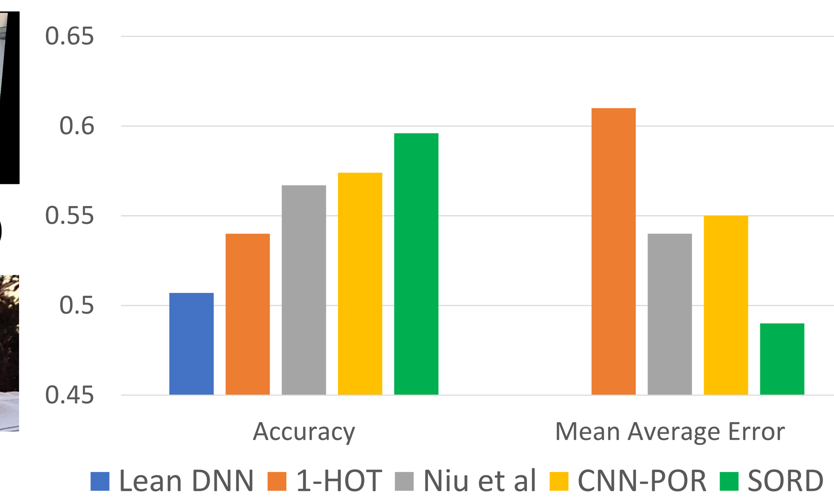
**Image Aesthetics dataset:** 15K images labeled in 5 ordinal categories



1 - Unacceptable | 2 - Flawed | 3 - Ordinary | 4 - Professional | 5 - Exceptional




**Adience dataset:** 26K images labeled in 8 age groups

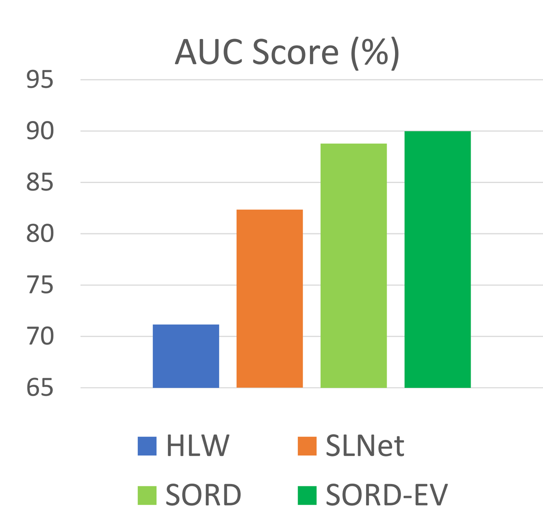



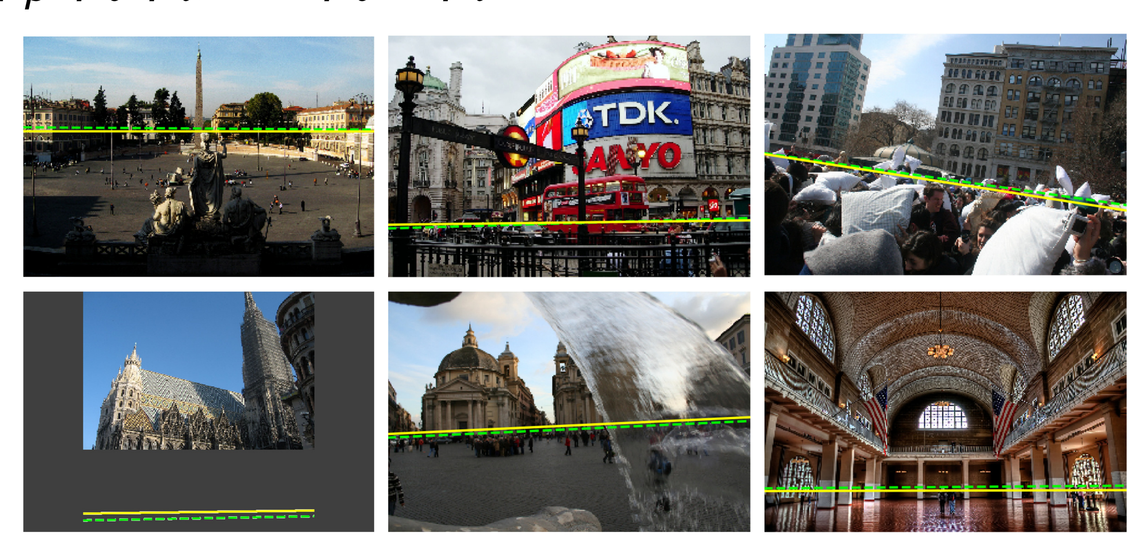

## 4. Horizon Lines in the Wild: 100K images from SfM models

We learn the two continuous domain line parameters by interpolating 100 bins from the training data, using **Resnet50** and the metrics:

$$\phi_\theta(\theta_t, \theta_i) = \min(\|\theta_t - \theta_i\|^2, \|(\theta_t - \theta_i - \pi) \ mod \ 2\pi\|^2)$$

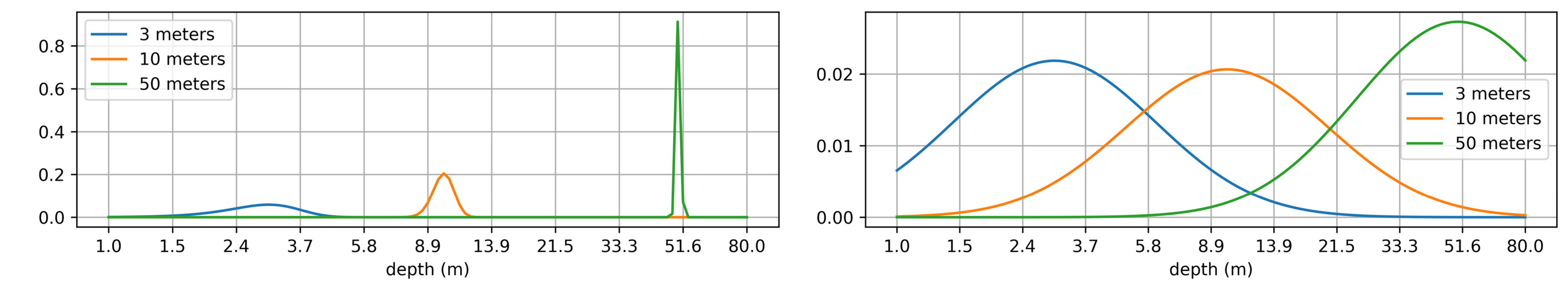
$$\phi_\rho(\rho_t, \rho_i) = \|\rho_t - \rho_i\|^2$$



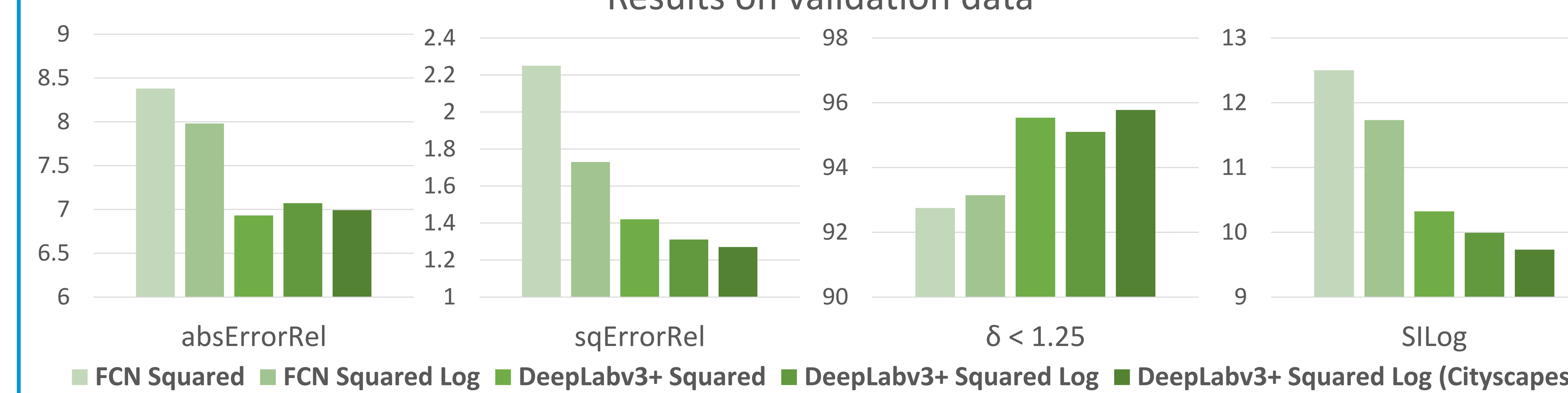
## 5. Monocular Depth Estimation on KITTI

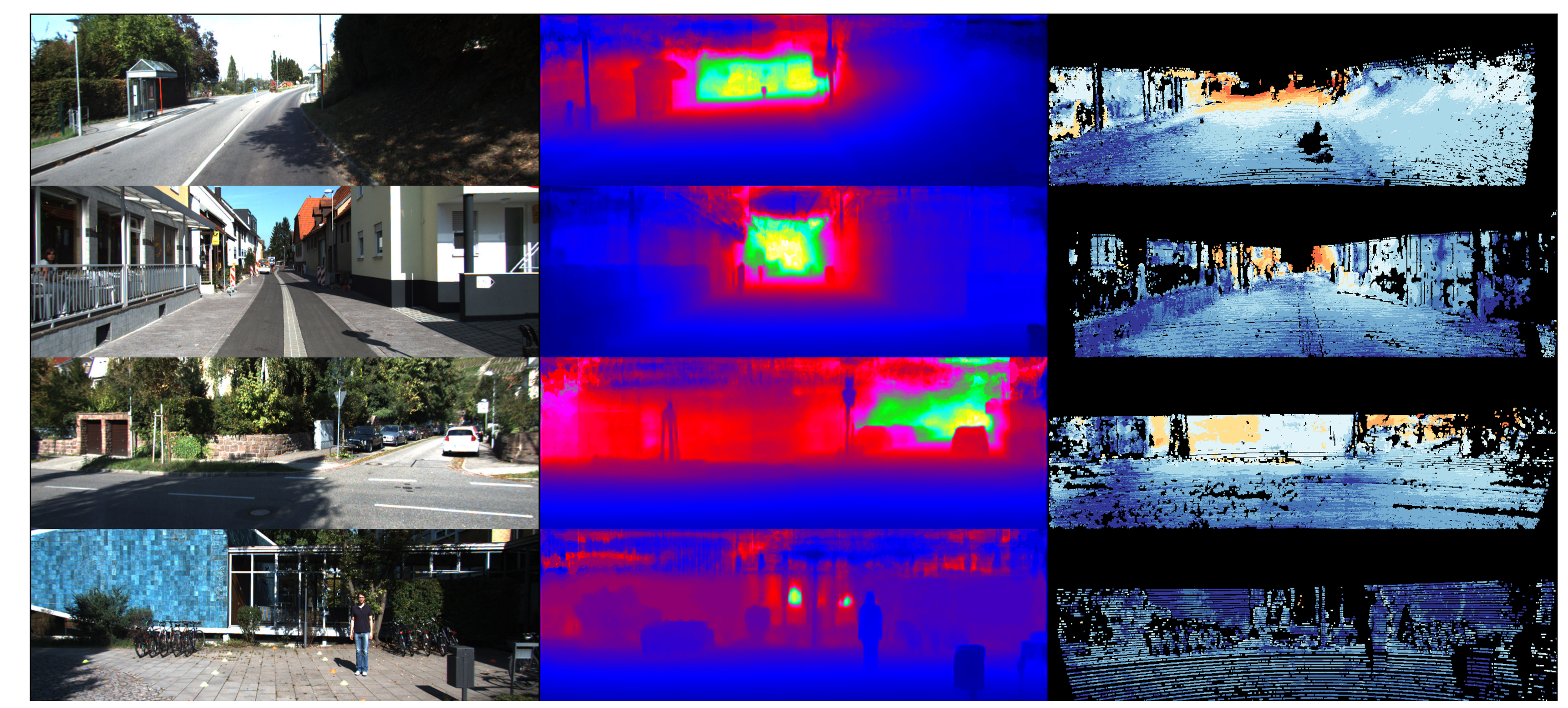
SORD can estimate depth using segmentation networks, adapting a scale-increasing discretization method (SID) with K=120 intervals:




$$\phi(r_t, r_i) = \|r_t - r_i\|^2 \qquad \phi(r_t, r_i) = \|\log r_t - \log r_i\|^2$$

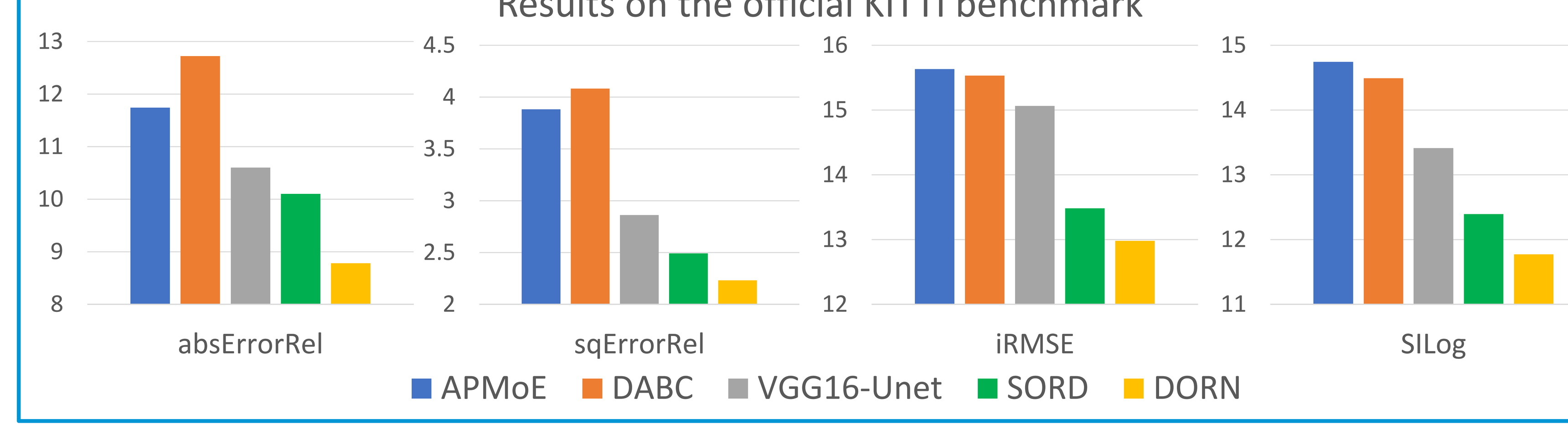
Results on validation data



Input Image | Depth Prediction | Error Map

Results on the official KITTI benchmark




## Selected References

[1] H. Fu, M. Gong, C. Wang, K. Batmanghelich, and D. Tao. Deep ordinal regression network for monocular depth estimation. In CVPR, 2018.

[2] J.-T. Lee, H.-U. Kim, C. Lee, and C.-S. Kim. Semantic line detection and its applications. In ICCV, 2017.

[3] Y. Liu, A. Wai Kin Kong, and C. Keong Goh. A constrained deep neural network for ordinal regression. In CVPR, 2018.

[4] S. Workman, M. Zhai, and N. Jacobs. Horizon lines in the wild. In BMVC, 2016.