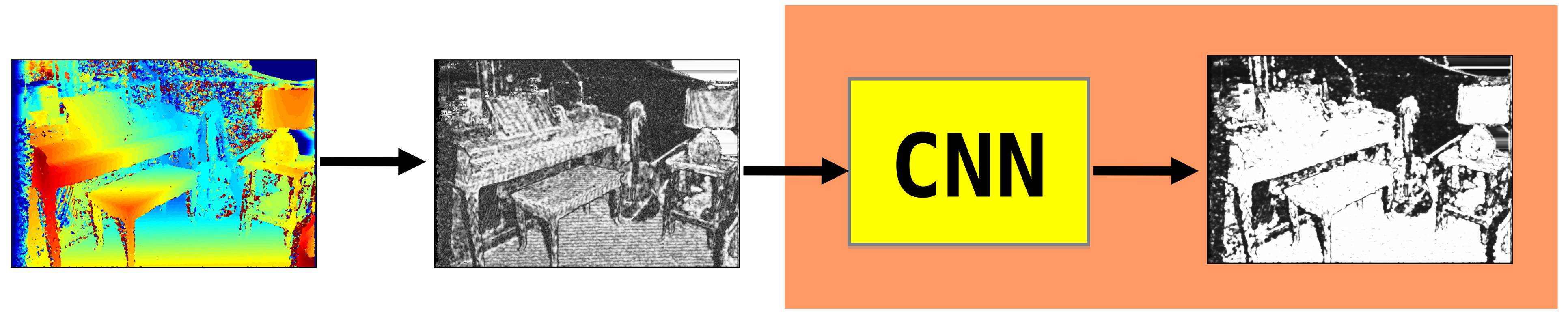
Learning to predict stereo reliability enforcing local consistency of confidence maps

Matteo Poggi, Stefano Mattoccia
University of Bologna

Confidence measures and local consistency

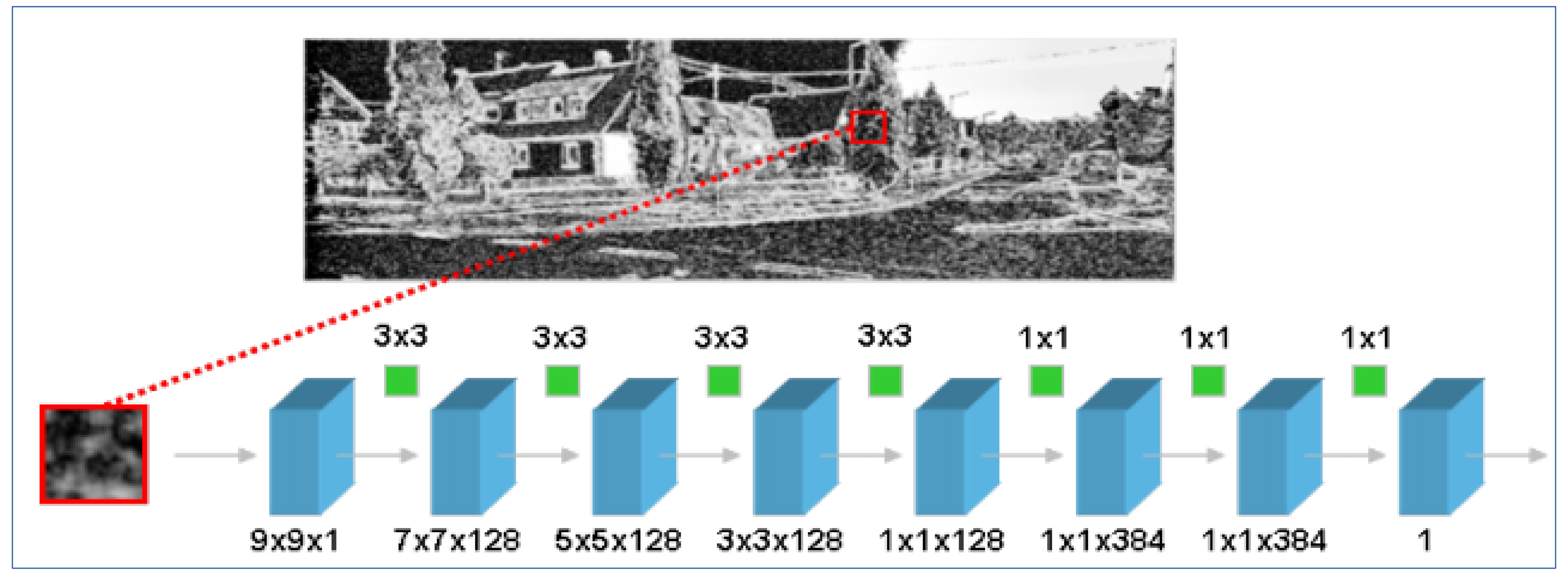
In stereo matching, confidence measures provide a per-pixel estimation of the correctness of the assigned disparity

- Our proposal: **using confidence prediction from the neighboring pixel to improve the effectiveness of the measure**
- Processing confidence maps (e.g., depicted by PKR measure), we obtain a new map (e.g., we call, in this case, PKR+)



Enforcing local consistency with deep learning

- Patch-based CNN trained to enforce local consistency for a single confidence measure



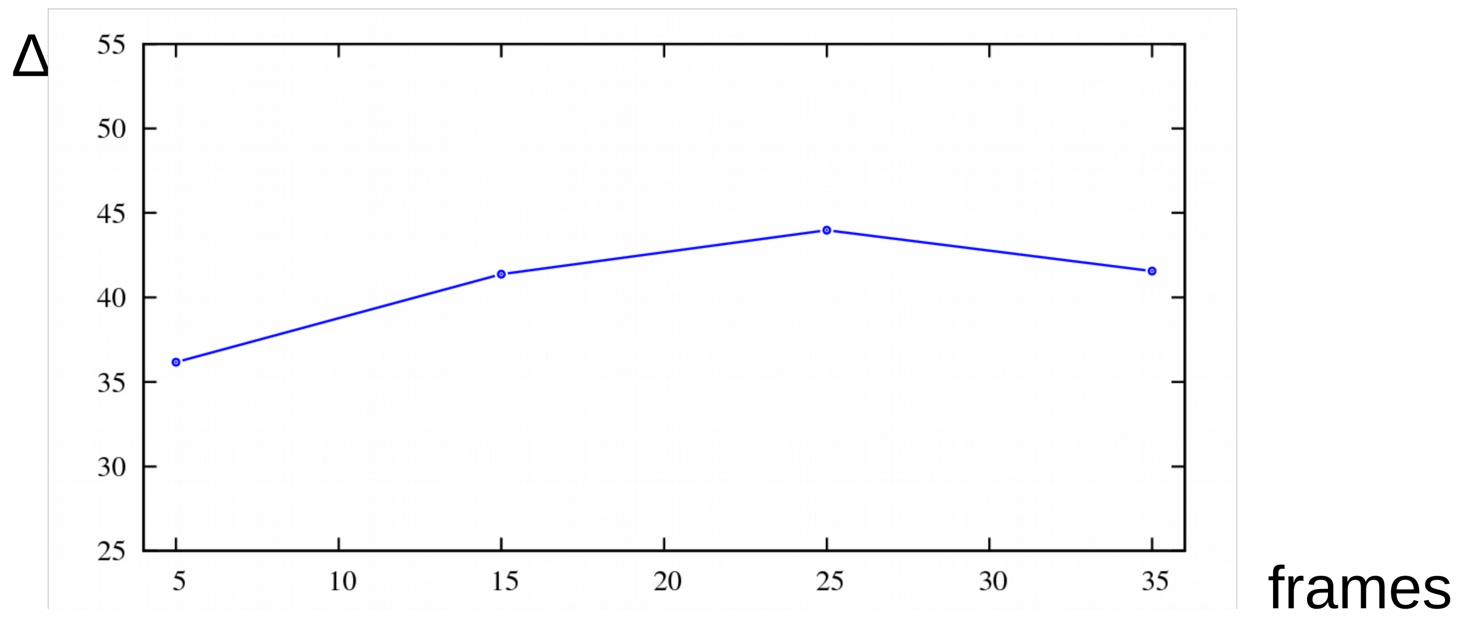
- Perceptive field: 9x9
- Four convolutional layer (3x3), each extracting 128 feature maps
- Two convolutional layers (1x1), each extracting 384 feature maps
- Final regression neuron
- Trained by Stocastic Gradient Descent (SGD), minimizing binary cross entropy as loss function (BCE)

Building the training set

- Compute disparity maps and confidence maps, according to given stereo algorithm and confidence measure, on stereo images with available ground-truth (e.g., from KITTI or Middlebury datasets)
- For each pixel, compare the computed disparity with the available ground-truth
- Fixing a threshold T, label as correct (1) pixels having a disparity error lower than T, or wrong (0) if greater.
- Train the network on the obtained binary labels

Impact of the amount of training data

- We train on a subset of images from KITTI 2012 (20 images, T=3)
- Each pixel with available ground-truth represent a training sample
- Training with 5, 15, 25 and 35 images (about 0.7, 1.5, 2.0, 2.7 and 3.5 million samples)
- Evaluation of the networks on Middlebury v3
- Δ= (AUC – AUC+)/(AUC - AUCopt)
- (Train on Middlebury v3. → only 1.2% better)

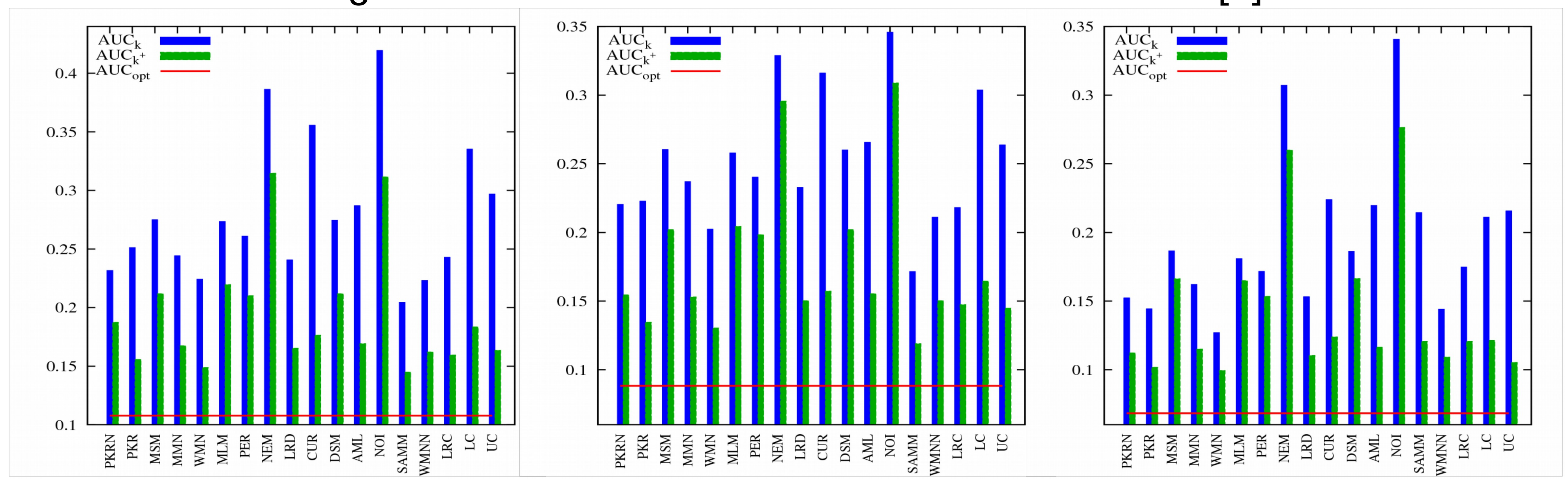


Source code and trained networks
http://vision.disi.unibo.it/~mpoggi/code.html

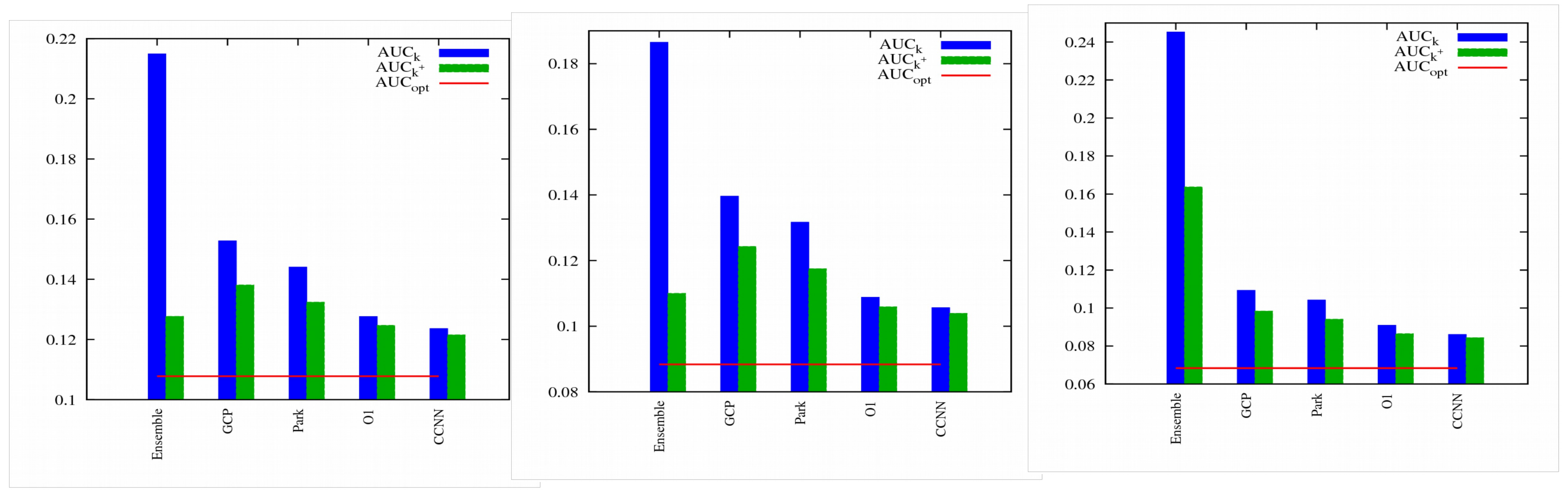
Acknowledgment: the Titan X Pascal used for this research was donated by the NVIDIA Corporation

Experimental results

- **Training on KITTI 2012** (20 images, T=3), evaluation on the **KITTI 2012**, **KITTI 2015** and **Middlebury v3** datasets. Stereo algorithm **AD-CENSUS** (and **MC-CNN**, in the paper)
- Testing on **18 confidence measures from literature** [1]



- Testing on **5 state-of-the-art machine-learning techniques**: Enseble [2], GCP [3], Park [4], O1 [5], CCNN [6]



- Exploiting local consistency with our network **always improves the input confidence measure in term of AUC** from Δ=75% to 9%

References

[1] Hu and Mordohai, PAMI 2012
[2] Haeusler et al., CVPR 2013
[3] Spyropoulos et al., CVPR 2014
[4] Park and Yoon, CVPR 2015
[5] Poggi and Mattoccia, 3DV 2016
[6] Poggi and Mattoccia, BMVC 2016