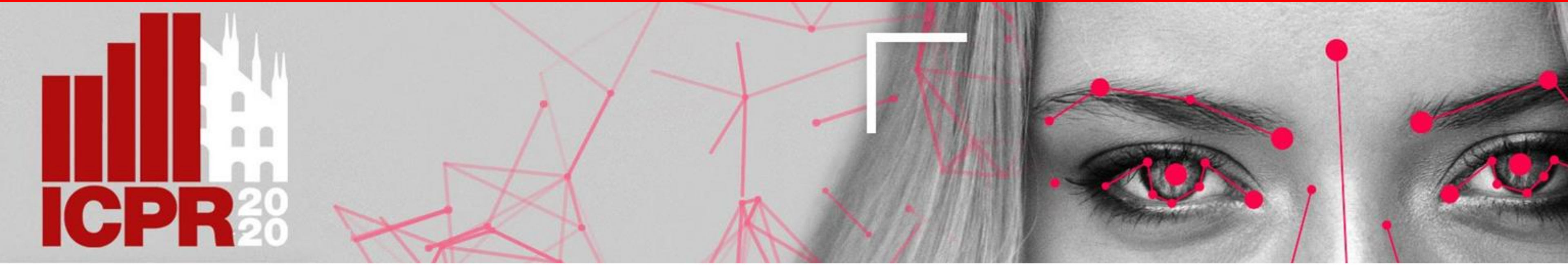


Leveraging a weakly adversarial paradigm for joint learning of disparity and confidence estimation

Matteo Poggi*, Fabio Tosi, Filippo Aleotti, Stefano Mattoccia

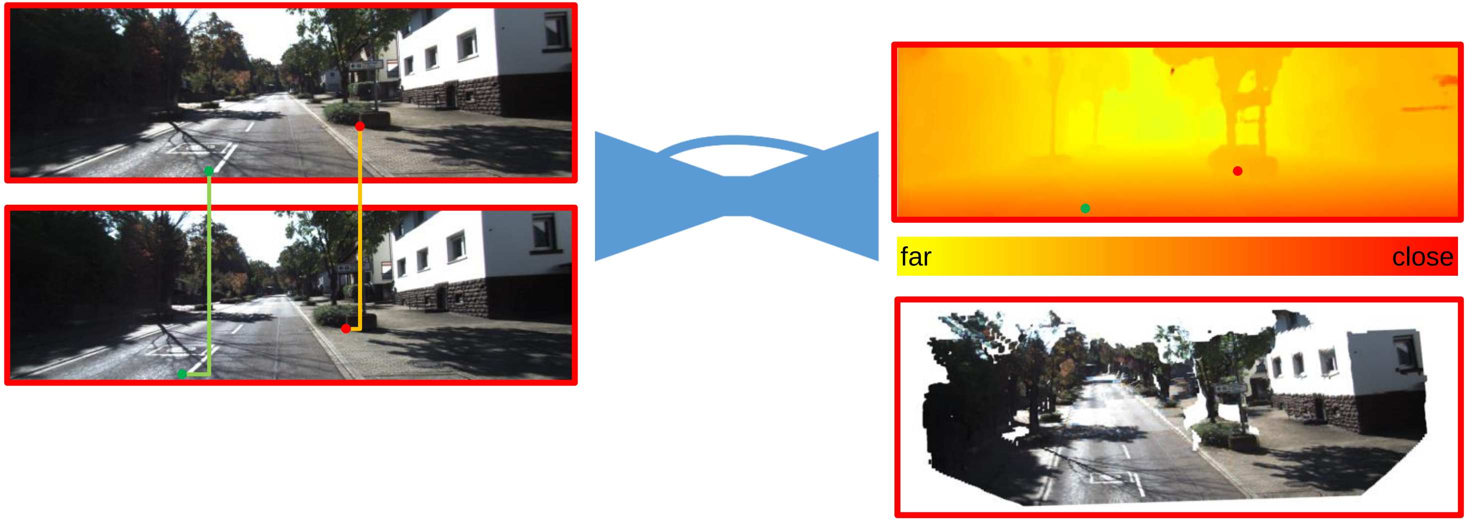
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Stereo matching

Stereo matching is one of the most popular image-based techniques to infer depth

Given two images, the horizontal displacement (**disparity**) is computed for each pixel
Neural networks **excel** at this

Depth is **triangulated** by knowing the camera parameters



Stereo matching and confidence

In parallel to the rapid evolution of stereo matching solutions, estimating the **confidence** [1] of such algorithms as gained raising popularity

This task consists into assigning to each pixel a **score** that ranks how reliable the pixel itself is in the entire disparity map

Stereo matching and confidence

Why is it important? For instance, to detect **wrong matches**

Detecting the wrong matches produced by a stereo algorithm is crucial for higher-level reasoning (e.g., **obstacle avoidance**)

This task can be learned by **deep networks** as well

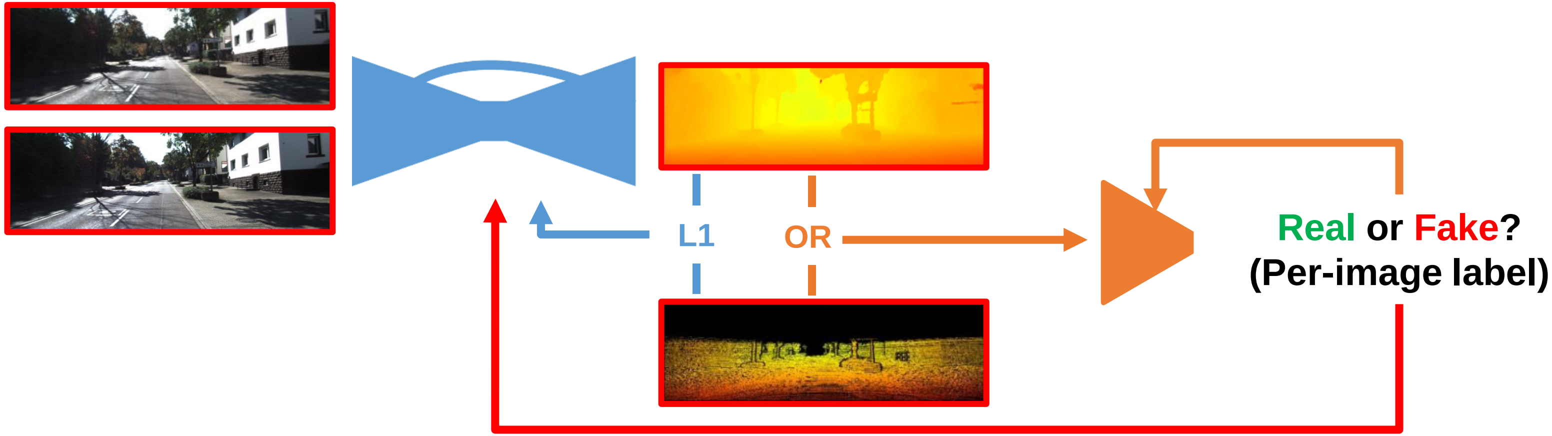
Joint disparity and confidence inference

We propose to learn in synergy the two tasks, as a **competition** between the two networks

A naïve strategy could consist into considering the disparity network as a **generator**

Then, a **discriminator** to distinguish between ground truth and estimated disparity maps

This would produce a single label (no per-pixel knowledge) and is not suited for sparse ground truth

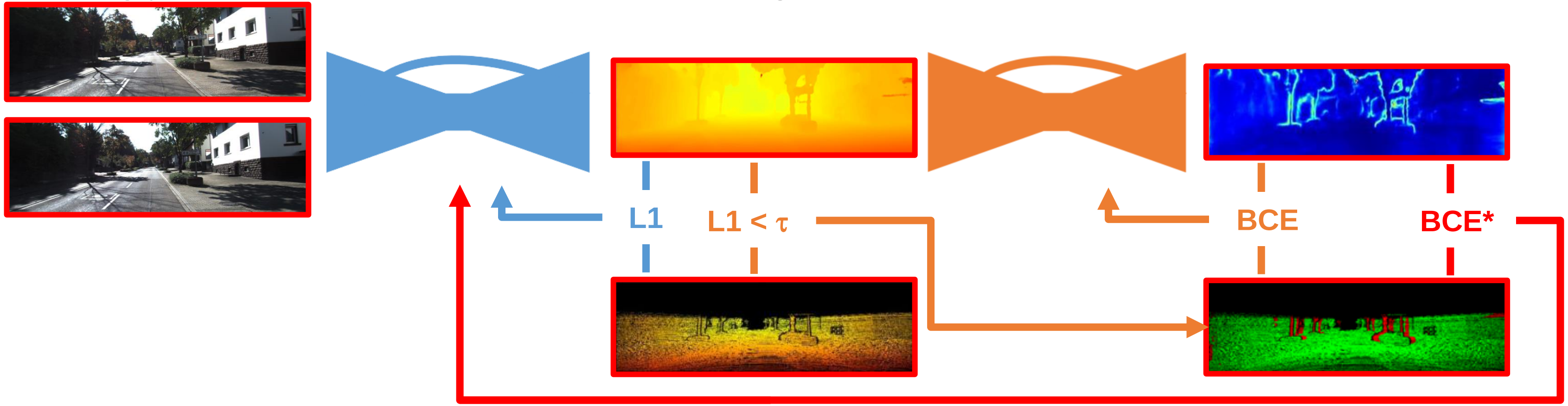


Weakly adversarial networks

We adopt a **weakly adversarial** strategy

The discriminator produces pixel-level labels (i.e., **confidences**) and is trained to distinguish correct disparities from outliers, on valid pixels only

Moreover, ad **adversarial term** over outliers forces the generator to produce better disparities, making wrong pixels **fewer and fewer**, thus weakening the adversarial term itself



Experimental results

Experimental setup

Baseline stereo network (generator): PSMNet [2]

Confidence network (discriminator): ConfNet [3]

Competitors:

- Join frameworks: Reflective confidence estimation [4], Heteroscedastic uncertainty modeling [5]
- Confidence estimators: CCNN [6], ConfNet [3], LGC-Net [3]

Experimental results

Training on KITTI 2012, testing on KITTI 2015

Disparity estimation

Model	>2(%) Noc	>2(%) All	>3(%) Noc	>3(%) All	>4(%) Noc	>4(%) All	>5(%) Noc	>5(%) All	MAE Noc	MAE All
PSMNet [5]	5.850	6.490	2.736	3.131	1.911	2.186	1.561	1.765	1.163	1.203
Heteroscedastic-PSMNet [41]	5.871	6.562	2.903	3.439	2.047	2.487	1.675	2.052	1.087	1.164
Reflective-PSMNet [33]	**5.670**	**6.209**	2.736	3.108	1.936	2.216	1.585	1.804	1.325	1.369
WAN-PSMNet (ours)	5.687	6.246	**2.681**	**3.062**	**1.885**	**2.176**	**1.528**	**1.762**	**0.972**	**1.025**

Confidence estimation

Estimator	AUC_{opt}	AUC	AUCM
CCNN	0.398	1.265	0.867
ConfNet	0.398	2.282	1.884
LGC-Net	0.398	1.059	0.661
Heteroscedastic	0.395	0.955	0.560
Reflective	0.450	1.250	0.800
WAN	0.358	0.908	**0.550**

Training on Middlebury trainingQ, testing additionalQ

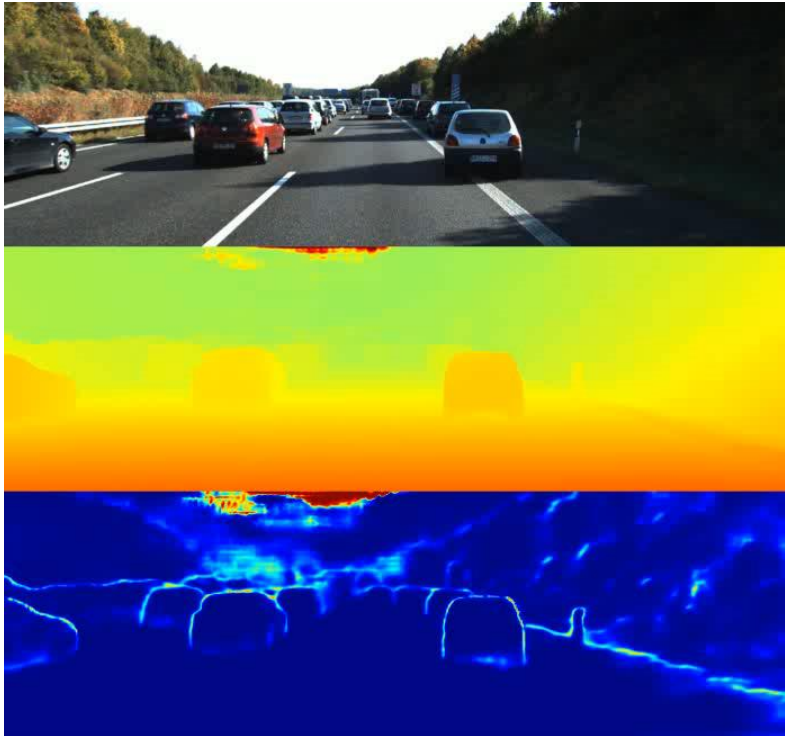
Disparity estimation

Model	>1(%)	>2(%)	>4(%)	MAE
PSMNet [5]	26.121	14.547	8.536	1.920
Heteroscedastic-PSMNet [41]	33.458	18.887	11.722	2.874
Reflective-PSMNet [33]	26.002	14.689	7.159	1.911
WAN-PSMNet (ours)	**25.496**	**14.476**	**7.132**	**1.906**

Confidence estimation

	AUC_{opt}	AUC	AUCM
CCNN	0.046	0.217	0.176
ConfNet	0.046	0.248	0.207
LGC-Net	0.046	0.194	**0.148**
Heteroscedastic	0.090	0.363	0.273
Reflective	0.045	0.166	0.191
WAN	0.041	0.194	0.153

Qualitative results



Video available at:
https://www.youtube.com/watch?v=Zk2IIlWKy78

References

[1]: M. Poggi, F. Tosi and S. Mattoccia, ICCV 2017 – «Quantitative evaluation of confidence measures in a machine learning world»

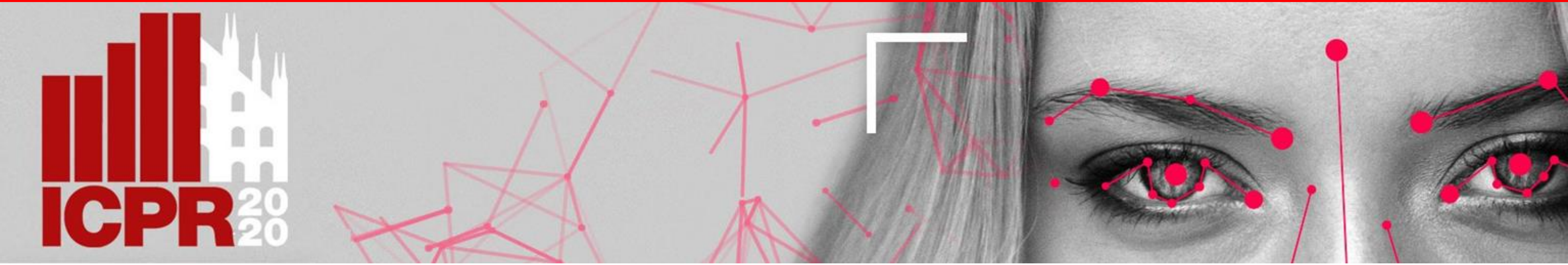
[2]: J. R. Chang and Y. S Chen, CVPR 2018 – «Pyramid Stereo Matching Network »

[3]: F. Tosi, M. Poggi, A. Benincasa and S. Mattoccia, ECCV 2018 – «Beyond local reasoning for stereo confidence estimation with deep learning»

[4]: A. Shaked and L. Wolf, CVPR 2017 – «Improved stereo matching with constant highway networks and reflective confidence learning»

[5]: A. Kendall and Y. Gal, NIPS 2017– «What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision»

[6]: M. Poggi and S. Mattoccia, BMVC 2016 – «Learning from scratch a confidence measure»



Thank you for your attention