ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Delivering Meaningful Representation for Monocular Depth Estimation

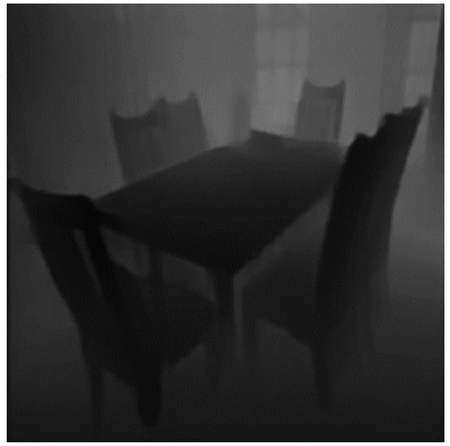
Doyeon Kim, Donggyu Joo, and Junmo Kim
KAIST

Goal

Monocular Depth Estimation

Input image

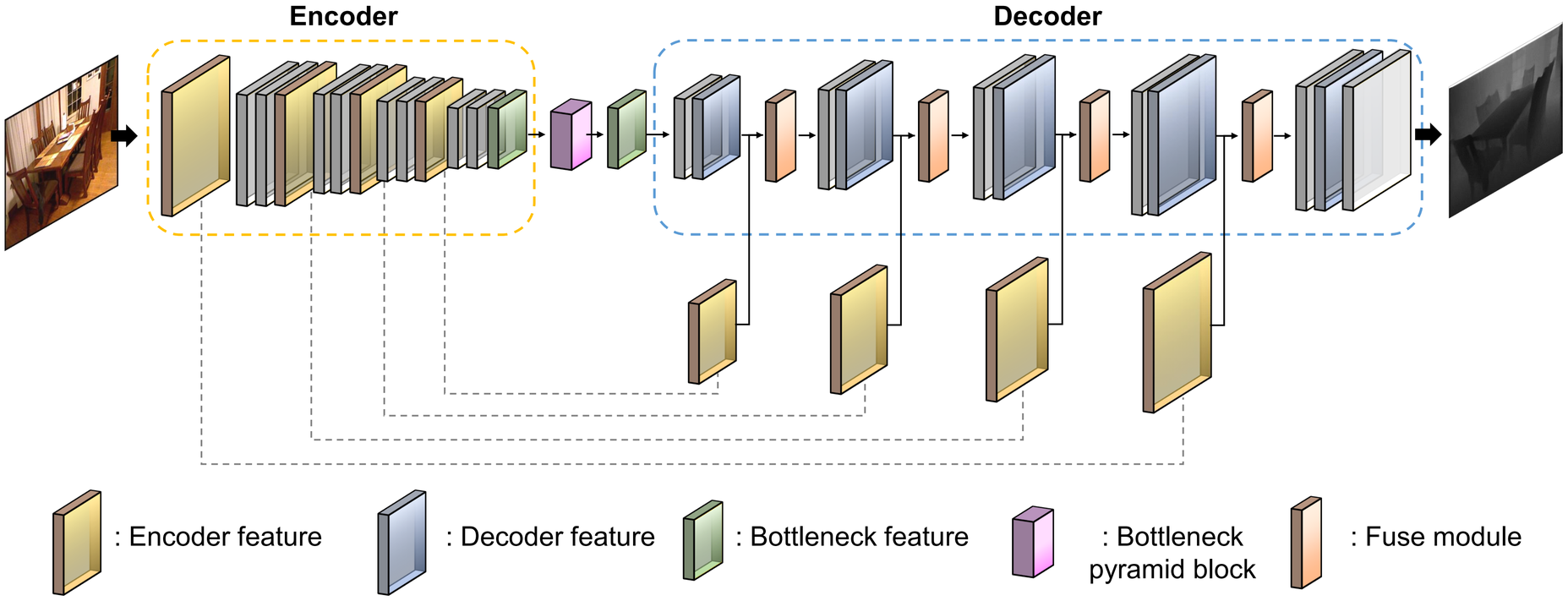
Model



Depth image

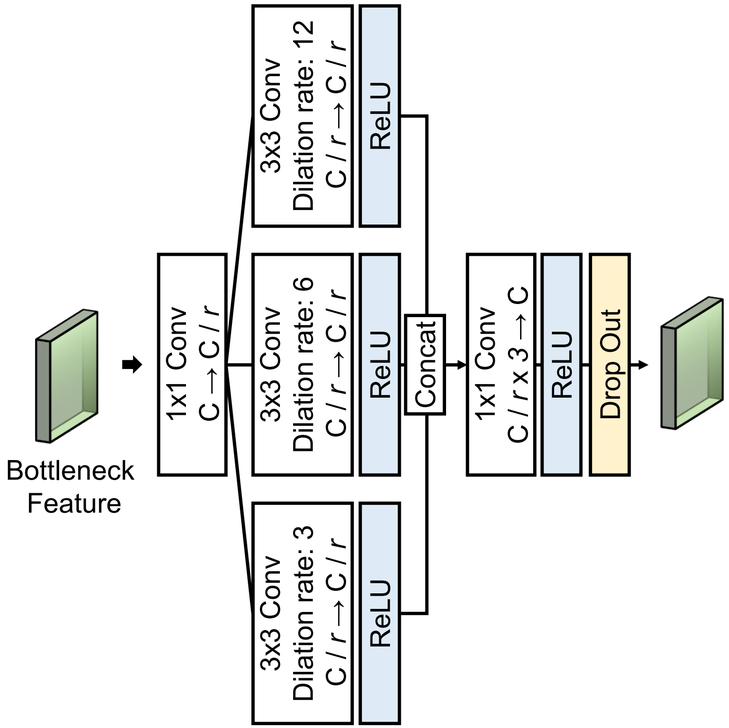
Proposed Method

Overall Architecture

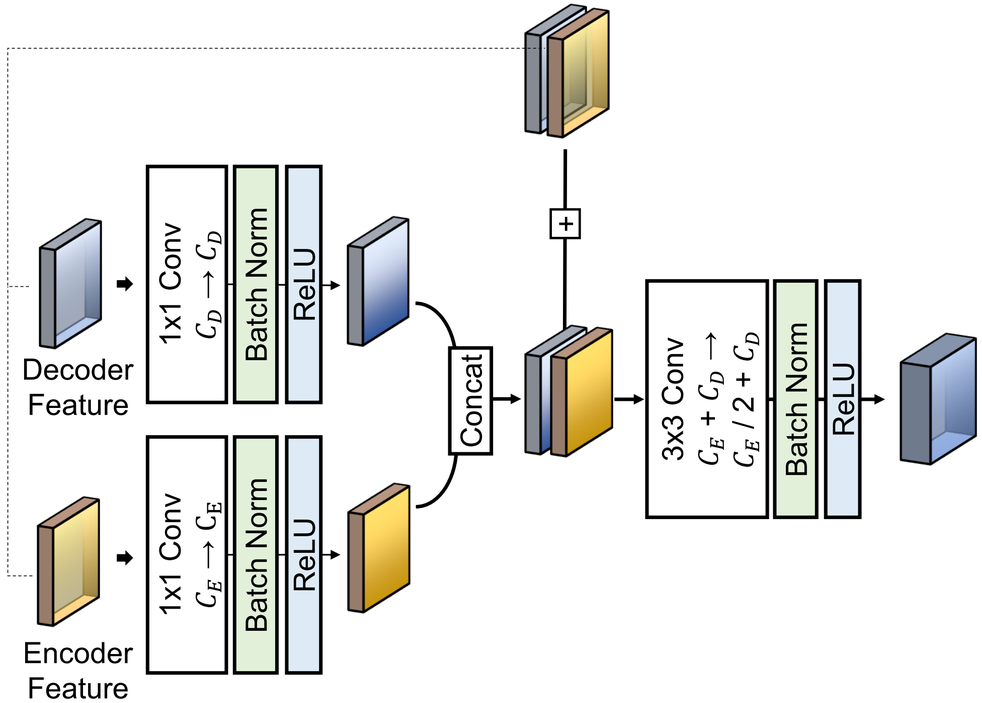


Proposed Method

Detailed structure of proposed modules



Bottleneck pyramid block



Fuse module

Results

Tables

Method	Base Network	*higher is better*			*lower is better*		
		δ_1	δ_2	δ_3	AbsRel	log_{10}	RMSE
Wang *et al.* [10]	VGG-16	0.605	0.890	0.970	0.220	-	0.745
Eigen *et al.* [4]	VGG-16	0.769	0.950	0.988	0.158	-	0.641
Laina *et al.* [11]	ResNet-50	0.811	0.953	0.988	0.127	0.055	0.573
Li *et al.* [22]	ResNet-50	0.789	0.955	0.988	0.152	0.064	0.611
Fu *et al.* [6]	ResNet-101	0.828	0.965	0.992	0.115	0.051	0.509
Qi *et al.* [12]	ResNet-50	0.834	0.960	0.990	0.128	0.057	0.569
Hao *et al.* [23]	ResNet-101	0.841	0.966	0.991	0.127	0.053	0.555
Lee *et al.* [15]	DenseNet-161	0.837	0.971	0.994	0.131	-	0.538
Hu *et al.* [24]	ResNet-101	0.866	0.975	0.993	0.115	0.050	0.530
Lin *et al.* [21]	SENet-154	0.866	0.975	0.993	0.115	0.050	0.523
Yin *et al.* [7]	ResNeXt-101	0.875	0.976	0.994	**0.108**	0.048	0.416
Ours	ResNet-50	0.858	0.977	0.995	0.121	0.051	0.414
Ours	ResNet-101	0.863	0.978	0.995	0.118	0.050	0.412
Ours	ResNeXt-101	**0.876**	**0.982**	**0.996**	0.112	**0.048**	**0.395**

NYU Depth V2

Method	cap	*higher is better*			*lower is better*		
		δ_1	δ_2	δ_3	AbsRel	SqRel	RMSE
Saxena *et al.* [8]	0 - 80m	0.601	0.820	0.926	0.280	3.012	8.734
Liu *et al.* [29]	0 - 80m	0.647	0.882	0.961	0.217	1.841	6.986
Eigen *et al.* [4]	0 - 80m	0.692	0.899	0.967	0.190	1.515	7.156
Nath *et al.* [16]	0 - 80m	0.771	0.922	0.971	0.167	1.257	5.578
Godard *et al.* [18]	0 - 80m	0.861	0.949	0.976	0.114	0.898	4.935
Kuznietsov *et al.* [17]	0 - 80m	0.862	0.960	0.986	0.113	0.741	4.621
Gan *et al.* [13]	0 - 80m	0.890	0.964	0.985	0.098	0.666	3.933
Guo *et al.* [14]	0 - 80m	0.892	0.967	0.986	0.096	0.641	4.095
Fu *et al.* [6]	0 - 80m	0.932	0.984	0.994	0.072	0.307	2.727
Yin *et al.* [7]	0 - 80m	0.938	0.990	0.998	0.072	-	3.258
Ours	0 - 80m	**0.957**	**0.994**	**0.999**	**0.061**	**0.231**	**2.680**
Garg *et al.* [5]	0 - 50m	0.740	0.904	0.962	0.169	1.080	5.104
Godard *et al.* [18] (CS+K)	0 - 50m	0.873	0.954	0.979	0.108	0.657	3.729
Kuznietsov *et al.* [17]	0 - 50m	0.875	0.964	0.988	0.108	0.595	3.518
Gan *et al.* [13]	0 - 50m	0.898	0.967	0.986	0.094	0.552	3.133
Guo *et al.* [14]	0 - 50m	0.901	0.971	0.988	0.092	0.515	3.163
Fu *et al.* [6]	0 - 50m	0.936	0.985	0.995	0.071	0.268	2.271
Ours	0 - 50m	**0.964**	**0.994**	**0.999**	**0.058**	**0.167**	**1.899**

KITTI

Results

Ablation Study

Input | Ground Truth | Baseline | Baseline + Fuse | Baseline + Fuse + BPB

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Thank you