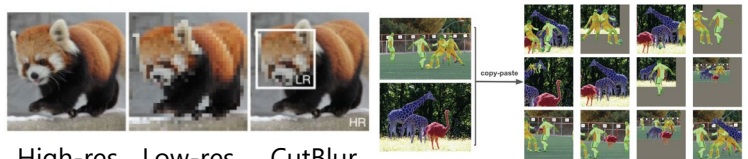
CutDepth:Edge-aware Data Augmentation in Depth Estimation : Ishii (Panasonic), Yamashita (Chubu Univ.)

Data augmentation

It is used for Classification etc., but there are few cases of Enc-Dec task.

Related work

Super resolution CutBlur [Yoo+ CVPR2020]
Semantic segmentation [Ghiasi+ CVPR2021]



High-res image | Low-res image | CutBlur

We propose a data augmentation suitable for depth estimation.

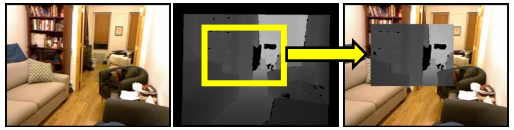
Proposed method

Task: Monocular Depth Estimation

- Edge positions are almost the same before and after data augmentation.
→ Increase the variation of appearance without destroying the features of the contour

- Add depth information as prior knowledge
→ Highly accurate depth estimation with latent variables including depth context

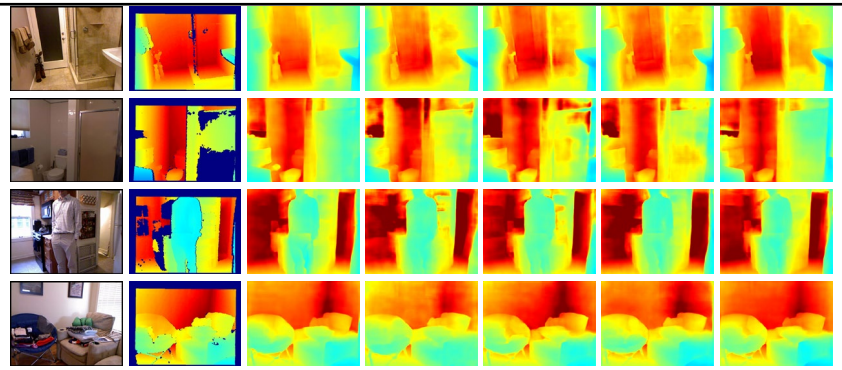
Replace image to depth



RGB image | depth | Proposed

Qualitative evaluation

The accuracy of contours and long distances has been improved.



input | GT | Baseline | Cutmix | Cutout | Random Erasing | Proposed

Quantitative evaluation

BTS [Lee+ arxiv:1907], **Laplacian depth** [Song+ IEEE Trans. CSVT2021]

Comparison with conventional data augmentatino methods

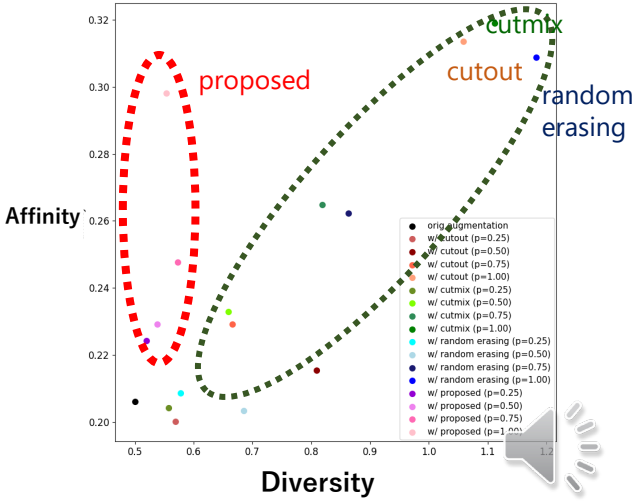
Method	p	BTS Abs Rel ↓	BTS log10 ↓	BTS RMSE ↓	BTS RMSE log ↓	BTS d1 ↑	BTS d2 ↑	BTS d3 ↑	Laplacian Depth Abs Rel ↓	Laplacian Depth log10 ↓	Laplacian Depth RMSE ↓	Laplacian Depth RMSE log ↓	Laplacian Depth d1 ↑	Laplacian Depth d2 ↑	Laplacian Depth d3 ↑
Baseline		0.1122	0.048	0.406	0.145	0.878	0.979	0.995	0.11	0.047	0.39	0.139	0.884	0.983	0.996
CutOut	0.25	0.1122	0.048	0.405	0.144	0.878	0.98	0.996	0.106	0.046	0.384	0.136	0.891	0.984	0.996
	0.50	0.1118	0.048	0.402	0.144	0.879	0.981	0.996	0.109	0.046	0.382	0.137	0.889	0.983	**0.997**
	0.75	0.1146	0.05	0.414	0.148	0.871	0.979	0.996	0.106	0.045	0.382	0.135	0.893	**0.985**	**0.997**
	1.00	0.1194	0.051	0.427	0.152	0.864	0.977	0.996	0.11	0.047	0.394	0.14	0.884	0.984	**0.997**
Random Erasing	0.25	0.1106	0.048	0.4	0.143	0.88	0.981	0.996	0.109	0.046	0.384	0.137	0.89	0.982	0.996
	0.50	0.1116	0.048	0.4	0.143	0.881	0.981	0.996	0.106	0.045	0.378	0.134	0.892	**0.985**	**0.997**
	0.75	0.1132	0.049	0.415	0.147	0.871	0.979	0.996	0.106	0.045	0.379	0.134	0.893	**0.985**	**0.997**
	1.00	0.1186	0.051	0.429	0.152	0.863	0.977	0.996	0.111	0.047	0.394	0.14	0.884	0.983	**0.997**
CutMix	0.25	0.1105	0.047	0.397	0.142	0.882	0.981	0.996	0.107	0.046	0.388	0.137	0.889	0.983	0.996
	0.50	0.1132	0.049	0.406	0.146	0.874	0.979	0.996	0.107	0.046	0.386	0.136	0.891	0.983	0.996
	0.75	0.1231	0.054	0.438	0.158	0.848	0.976	0.996	0.107	0.046	0.386	0.136	0.891	0.983	0.996
	1.00	0.1851	0.086	0.674	0.241	0.659	0.918	0.982	0.11	0.047	0.391	0.139	0.886	0.982	0.996
Proposed	0.25	0.1083	0.047	0.398	0.141	0.884	0.981	0.996	0.106	0.045	0.38	0.135	0.895	0.984	0.996
	0.50	0.1077	**0.046**	**0.391**	**0.14**	0.884	**0.982**	**0.997**	**0.104**	**0.044**	**0.375**	**0.132**	**0.899**	**0.985**	**0.997**
	0.75	**0.1074**	0.047	0.392	**0.14**	**0.885**	**0.982**	0.996	0.106	0.045	0.379	0.135	0.894	0.984	**0.997**
	1.00	0.1127	0.047	0.392	0.142	0.88	0.981	0.996	**0.104**	0.045	0.376	**0.132**	0.898	**0.985**	0.996

Our method has higher performance than other methods.
Our method is effective even when the number of data is small.

Performance comparison when reducing the number of data

ratio	methods	Abs Rel ↓	log10 ↓	RMSE ↓	RMSE log ↓	d1 ↑	d2 ↑	d3 ↑
25%	Baseline	0.1226	**0.052**	0.428	0.154	**0.859**	0.977	0.995
	CutOut	0.1242	0.053	0.432	0.156	0.854	0.976	**0.996**
	RE	0.1268	0.054	0.440	0.158	0.848	0.976	0.995
	CutMix	0.1467	0.064	0.520	0.188	0.782	0.956	0.993
	Proposed	**0.1225**	**0.052**	**0.424**	**0.153**	0.858	**0.978**	0.995
50%	Baseline	0.1174	0.050	0.414	0.150	0.867	0.978	0.995
	CutOut	0.1168	0.050	0.418	0.150	0.867	0.979	**0.996**
	RE	0.1184	0.051	0.422	0.151	0.862	0.978	**0.996**
	CutMix	0.1307	0.056	0.460	0.168	0.832	0.970	0.994
	Proposed	**0.1155**	**0.049**	**0.411**	**0.148**	**0.870**	**0.981**	**0.996**
75%	Baseline	0.1154	0.049	0.410	0.147	0.871	0.979	0.996
	CutOut	0.1148	0.050	0.413	0.147	0.870	0.980	**0.997**
	RE	0.1179	0.051	0.424	0.151	0.863	0.977	0.996
	CutMix	0.1353	0.058	0.465	0.172	0.826	0.967	0.993
	Proposed	**0.1142**	**0.048**	**0.401**	**0.144**	**0.876**	**0.981**	0.996

Analysis by diversity and affinity [1]



Our data augmentation increases the variation of data, but does not change it excessively.
This property seems to be suitable for image to image translation tasks.
※Affinity means that there is little deviation in the data distribution.
Diversity means the size of the distribution of data.

[1] Affinity and diversity: Quantifying mechanisms of data augmentation,arixv