Multi Image Depth from Defocus Network with Boundary Cue for Dual Aperture Camera

Gwangmo Song, Yumee Kim, Kukjin Chun, Kyoung Mu Lee

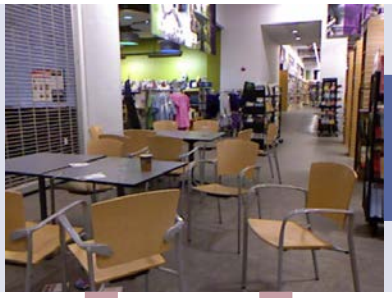
Seoul National University

Presenter : Gwangmo Song

Introduction

- Depth estimation problem
- Double defocused images (Different *Depth-of-Field*)

Single All-in-Focus



Depth Estimation

Shallow *DoF* ($F_{\#} = 2$)

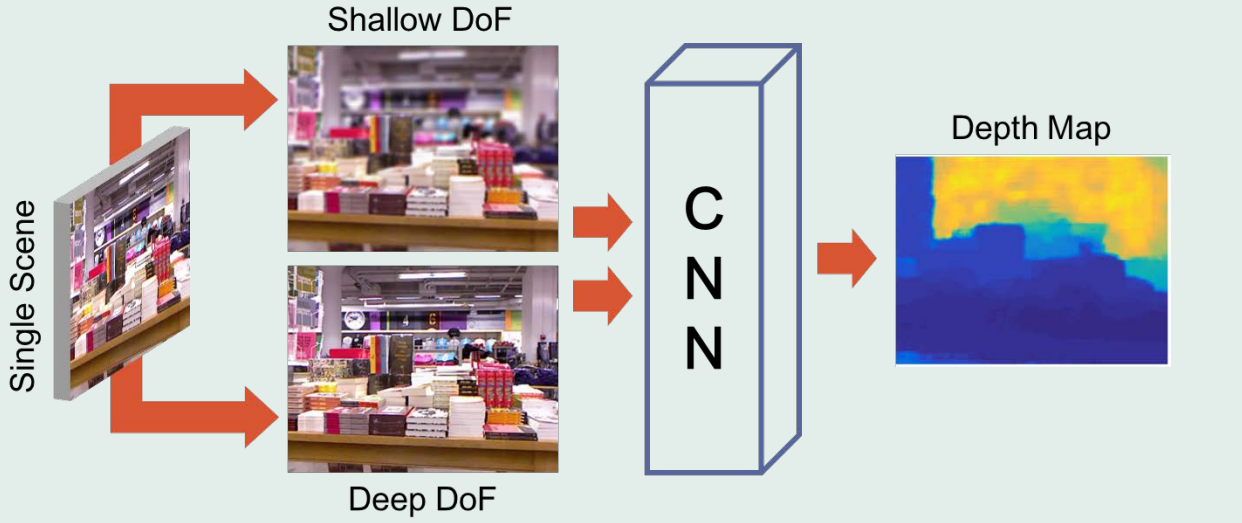
Deep *DoF* ($F_{\#} = 14$)

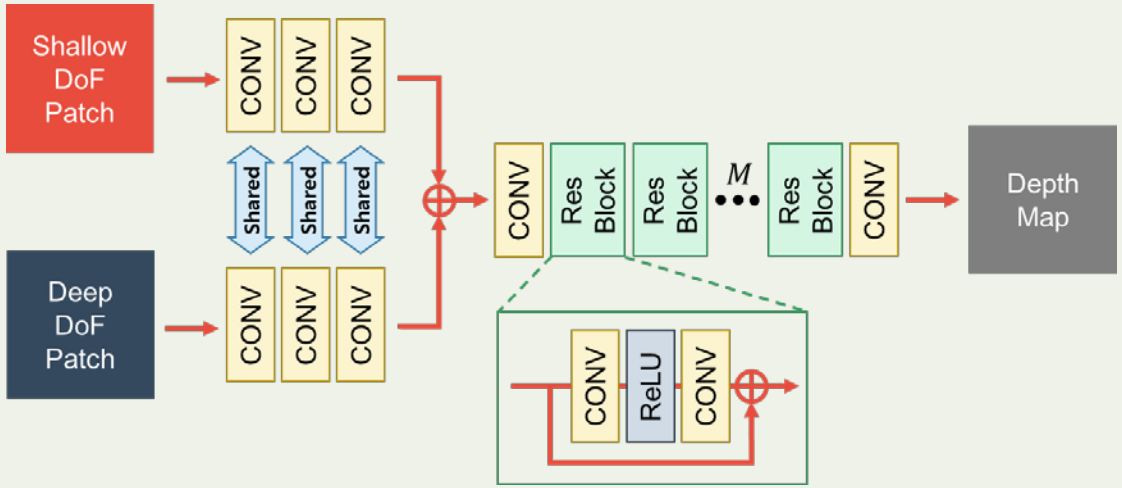
Double Out-of-Focus

Depth Estimation

Previous Works

- *MRF*-based vs. *DNN*-based
- Depth from Defocus(DFD) problem [1,2]
- Dual DFD [3]



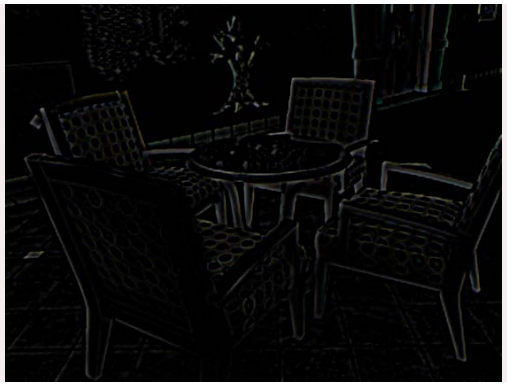


[1] Saeed Anwar, Zeeshan Hayder, and Fatih Porikli, “Depth estimation and blur removal from a single out-of-focus image,” in BMVC, 2017.
[2] Marcela Carvalho, Bertrand Le Saux, Pauline Trouvé Peloux, Andrés Almansa, and Frédéric Champagnat, “Deep depth from defocus: how can defocus blur improve 3d estimation using dense neural networks?,” in ECCV, 2018.
[3] Gwangmo Song and Kyoung Mu Lee, “Depth estimation network for dual defocused images with different depth-of-field,” in ICIP, 2018.

Introduction

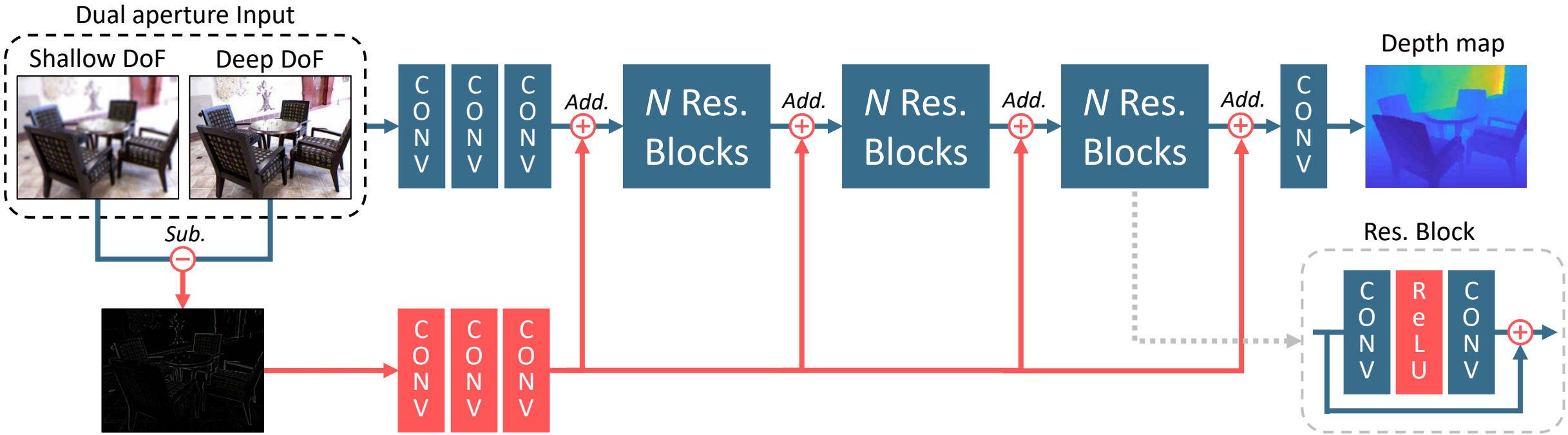
- Contributions

Boundary Cue
- Edge information
- Improve accuracy



Real dataset
- Collect new dataset
- Limited circumstance

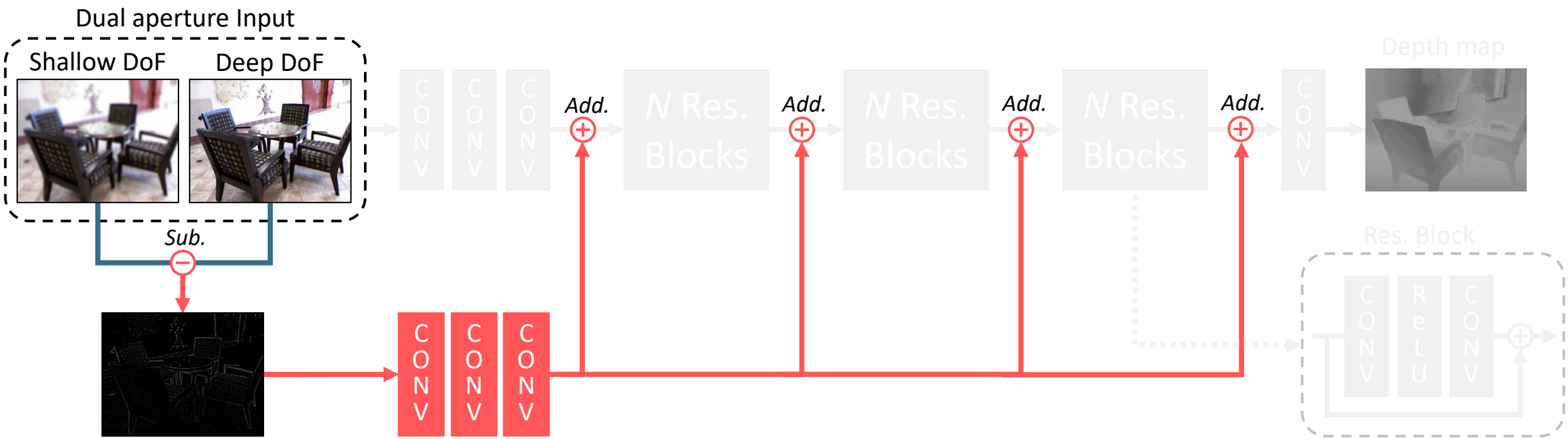
Dual DFD Network



Dual DFD Network

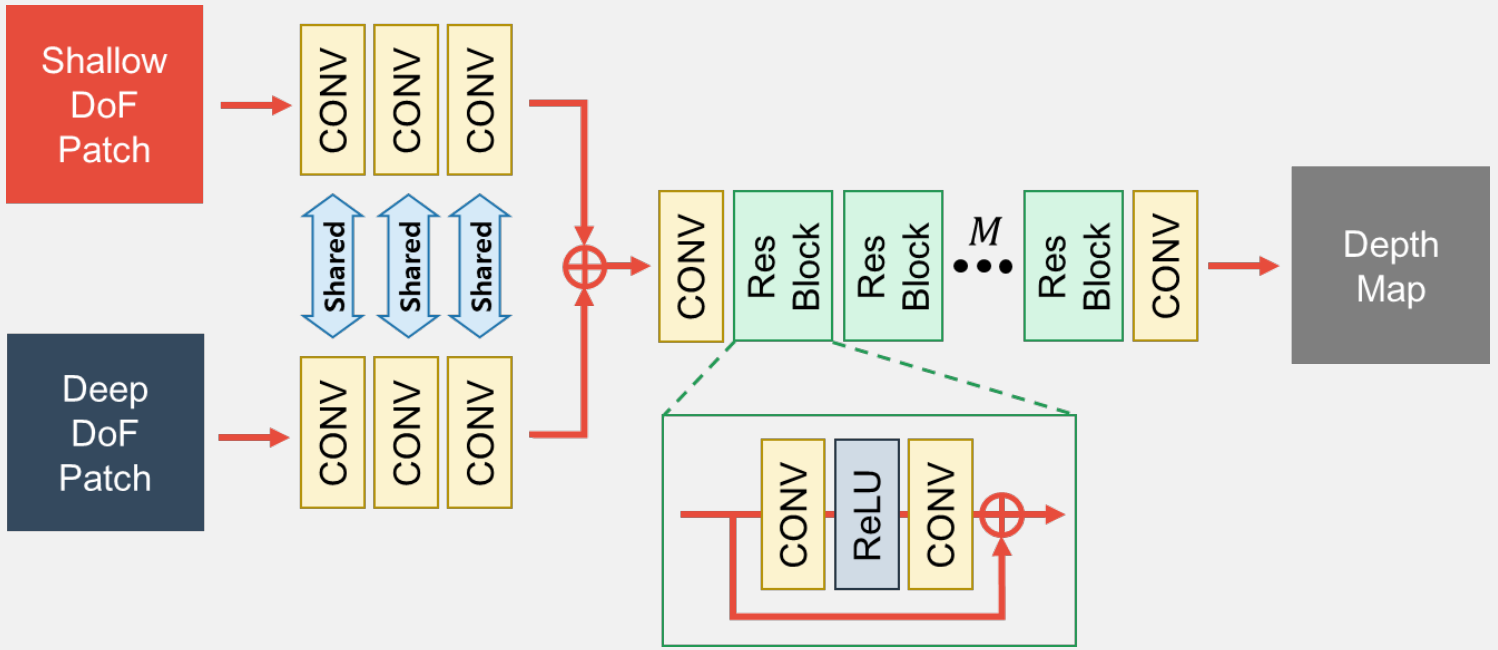
- Boundary Cue
  - Homogeneous region has **less** effect of blur
  - **Subtraction** highlights around the **edge** of the object

Dual DFD Network



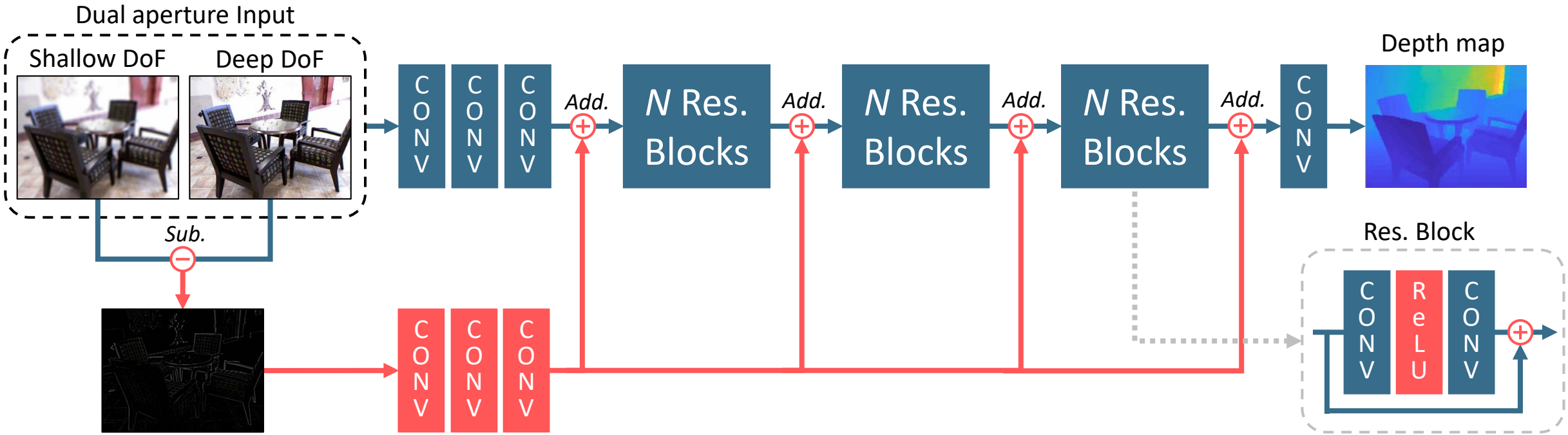
Dual DFD Network

- Main Network
  - EDSR [1] based
  - Baseline network [2]



[1] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee, "Enhanced deep residual networks for single image super-resolution," in CVPRW, 2017.
[2] Gwangmo Song and Kyoung Mu Lee, "Depth estimation network for dual defocused images with different depth-of-field," in ICIP, 2018.

Dual DFD Network

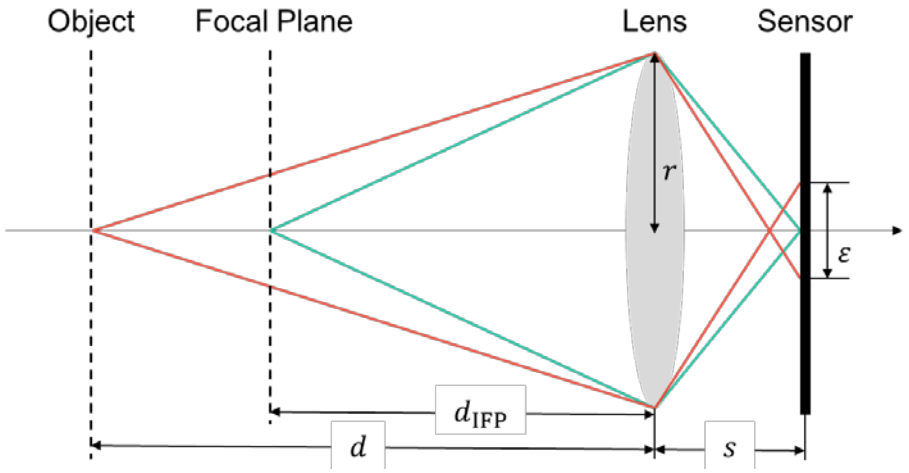


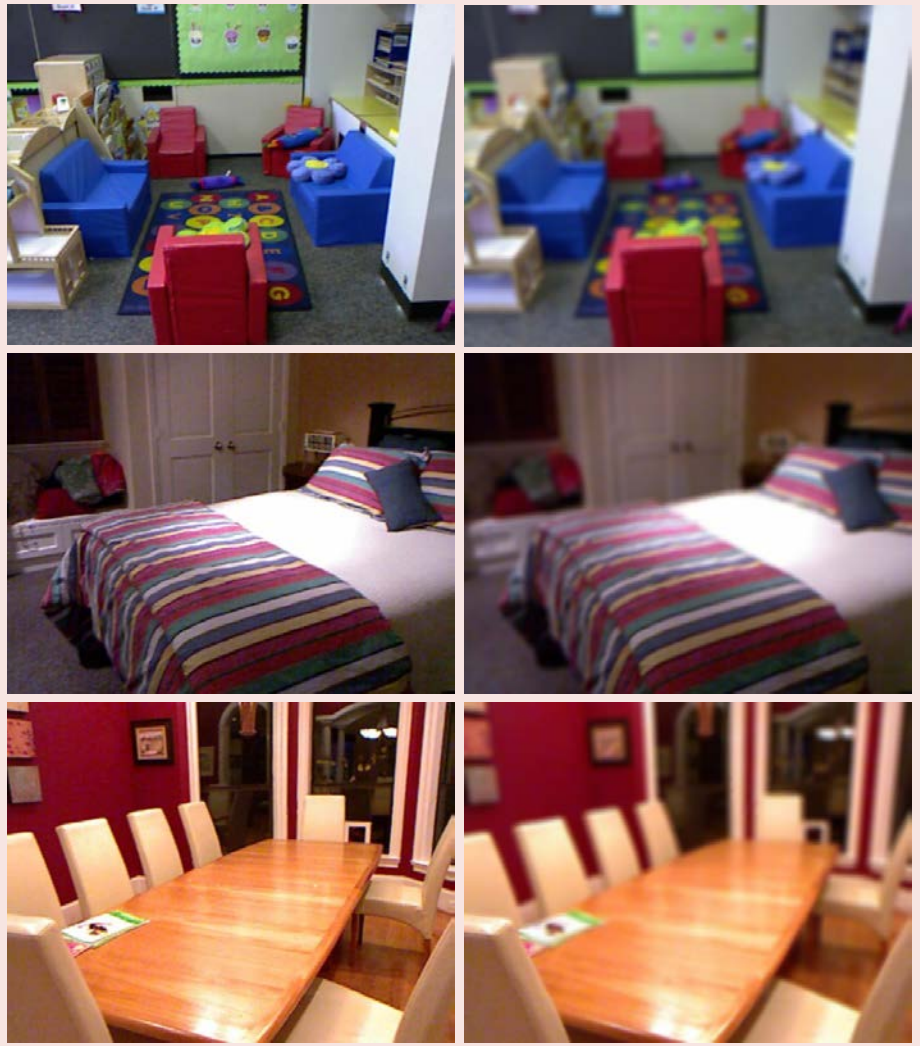
Datasets

- Synthetic dataset [1]
  - Dual defocused dataset
  - *NYU-v2* dataset
  - Using thin lens model ($F_{\#}$ = 2, 14)

$$\sigma = \frac{1}{\sqrt{2}} \frac{1}{p} \frac{f^2}{F_{\#}} \frac{1}{d_{IFP} - f} \left(1 - \frac{d_{IFP}}{d}\right)$$

  - *795* training data and *654* test data





[1] Gwangmo Song and Kyoung Mu Lee, "Depth estimation network for dual defocused images with different depth-of-field," in ICIP, 2018.

Datasets

- Real dataset
  - Tunable aperture camera
  - $F_{\#}$ = 1.8, 4.0
  - Static scene
  - LIDAR align
  - *199* training data and *100* test data

Datasets

- Real dataset
  - Parking lot
  - Maximum distance : 70m
  - 3 types of car, 3 types of pedestrian

Experiments

- Synthetic dataset
  - *NYU v2*-based
- *rel*
  - Average relative error
- *log10*
  - Average $\log_{10}$ error
- *rms*
  - Root mean square error

Method	*rel*	*log10*	*rms*
Single DFD			
Anwar *et al.* [1]	0.094	0.039	0.347
D3-Net [2]	0.068	0.028	0.274
D3-Net* [2]	0.036	0.016	0.144
D3-Net** [2]	0.056	0.024	0.244
Dual DFD			
D3-Net** [2]	0.030	0.013	0.164
Song *et al.* [3]	0.028	0.012	0.154
Our	**0.026**	**0.011**	**0.139**

* Using dataset of [1]
** Using dataset of [3]

[1] Saeed Anwar, Zeeshan Hayder, and Fatih Porikli, "Depth estimation and blur removal from a single out-of-focus image," in BMVC, 2017.
[2] Marcela Carvalho, Bertrand Le Saux, Pauline Trouvé Peloux, Andrés Almansa, and Frédéric Champagnat, "Deep depth from defocus: how can defocus blur improve 3d estimation using dense neural networks?," in ECCV, 2018.
[3] Gwangmo Song and Kyoung Mu Lee, "Depth estimation network for dual defocused images with different depth-of-field," in ICIP, 2018.

Experiments

- Ablation Study
- Effect of boundary cue

Method	*rel*	*log10*	*rms*
Baseline	0.031	0.013	0.162
Boundary Cue	0.028	0.012	0.146
Boundary Cue + Skip Conn.	**0.026**	**0.011**	**0.139**

Experiment

- Qualitative

Image ($F_{\#}$ = 2) | Image ($F_{\#}$ = 14) | GT depth map | [1] result | Our result

[1] Gwangmo Song and Kyoung Mu Lee, “Depth estimation network for dual defocused images with different depth-of-field,” in ICIP, 2018.

Experiments

- Real dataset
- Patch-size
  - 224 vs. 48
  - Homogeneous region

- GT depth map
  - **Sparse** LIDAR point
  - Boundary is **not clean**
  - Boundary cue ↓

Method	*rel*	*log10*	*rms*
Single DFD			
D3-Net [1]	0.027	0.012	2.070
Dual DFD			
D3-Net [1]	0.027	0.012	**1.202**
Song *et al*. [2]	0.019	**0.008**	1.400
Our	**0.018**	**0.008**	1.320

[1] Marcela Carvalho, Bertrand Le Saux, Pauline Trouvé Peloux, Andrés Almansa, and Frédéric Champagnat, "Deep depth from defocus: how can defocus blur improve 3d estimation using dense neural networks?," in ECCV, 2018.
[2] Gwangmo Song and Kyoung Mu Lee, "Depth estimation network for dual defocused images with different depth-of-field," in ICIP, 2018.

Experiment

- Qualitative

Image ($F_{\#}$ = 1.8) Image ($F_{\#}$ = 4) GT depth map Our result

Experiments

- Object-based measure
  - Ignore *sky*, *road*
  - Center of object
  - Mean value of bounding box

Method	*rel*	*log10*	*rms*
Single DFD			
D3-Net [1]	0.045	0.020	1.064
Dual DFD			
D3-Net [1]	0.041	0.018	0.943
Song *et al*. [2]	0.040	0.016	0.887
Our	**0.031**	**0.013**	**0.718**



[1] Marcela Carvalho, Bertrand Le Saux, Pauline Trouvé Peloux, Andrés Almansa, and Frédéric Champagnat, "Deep depth from defocus: how can defocus blur improve 3d estimation using dense neural networks?," in ECCV, 2018.
[2] Gwangmo Song and Kyoung Mu Lee, "Depth estimation network for dual defocused images with different depth-of-field," in ICIP, 2018.

Summary

- Create **boundary cue** through dual defocused images
- Proposal of **DFD network** structure using boundary cue
- **Dataset collection** using tunable aperture camera
- Record **SOTA** in synthetic and real dataset