Queen Mary
University of London
Science and Engineering

QMUL-BUPT Joint Programme
JP Student Innovation Centre
Annual Showcase 2021/22

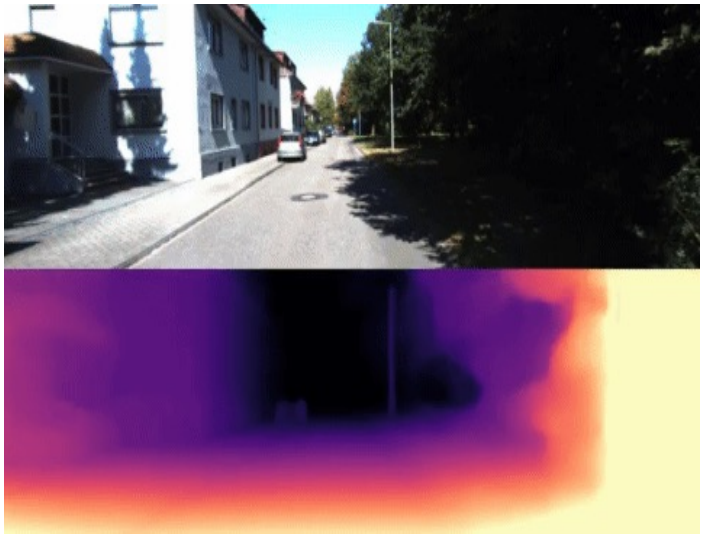
Monocular Depth Estimation using Deep Neural Networks

Student: Yuheng Li

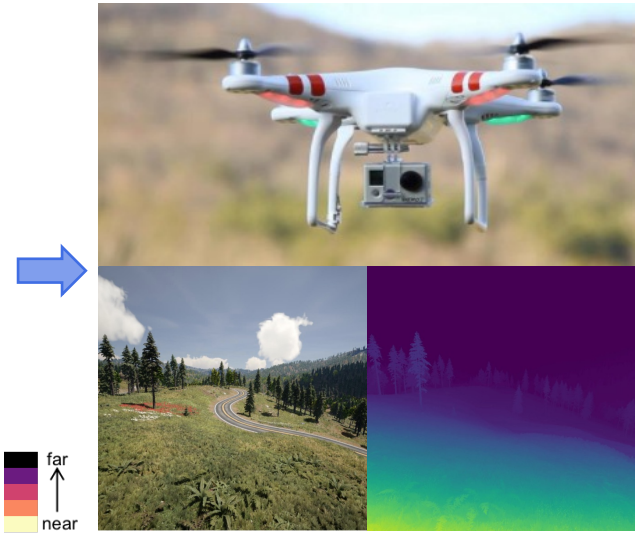
Project Leader: Dr Changjae OH

Introduction of background

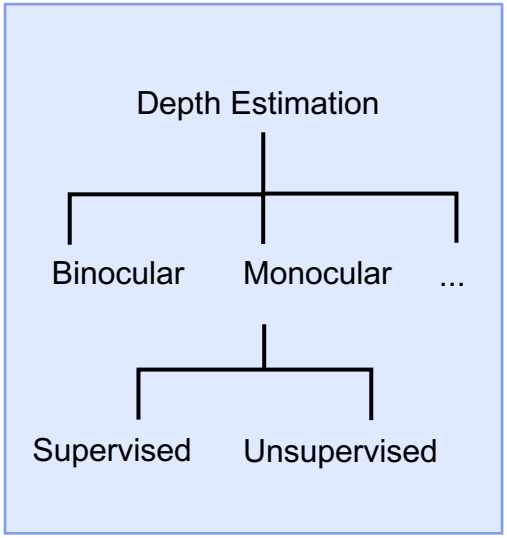
- **Depth estimation:** estimate the distance of each pixel in the image relative to the shooting source
- **Object:** investigate a depth estimation model for **UAV**(unmanned aerial vehicle) based on deep neural networks



Depth estimation in automatic driving



Depth estimation in UAV

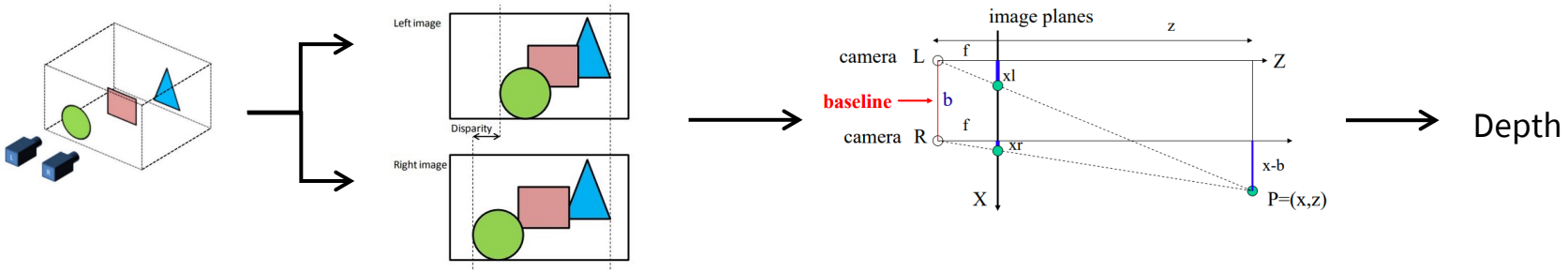


Which is the most suitable for UAV?

Introduction of background

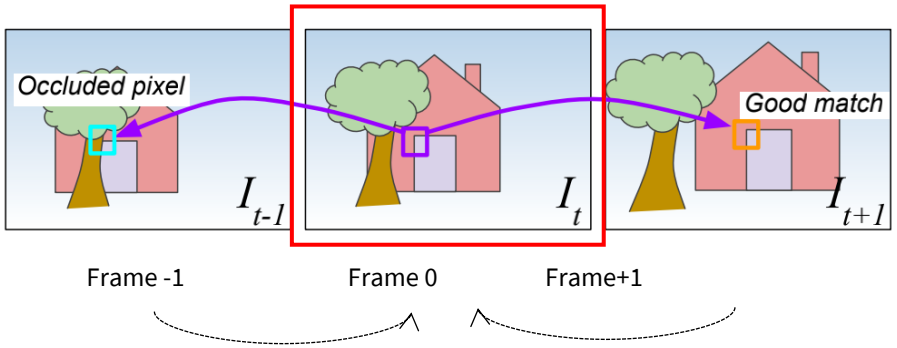
- **Binocular VS. Monocular depth estimation**

- Binocular method —— require more devices, not light！

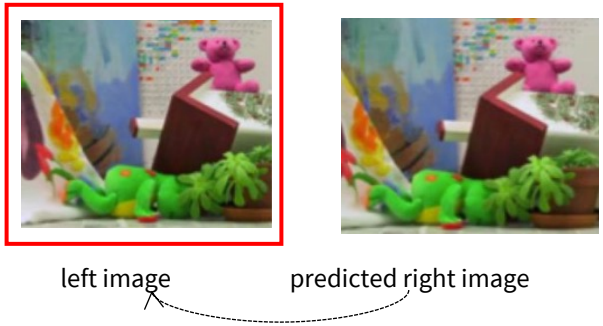


Monocular method

adjacent frames



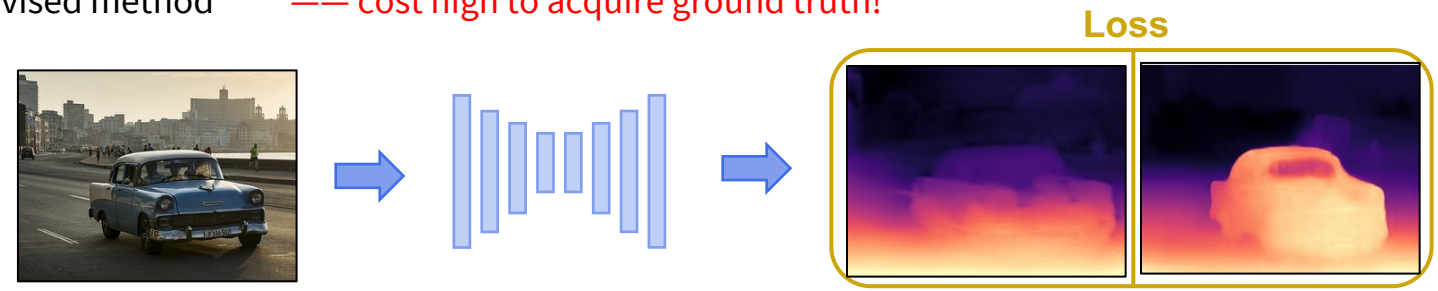
stereo pairs



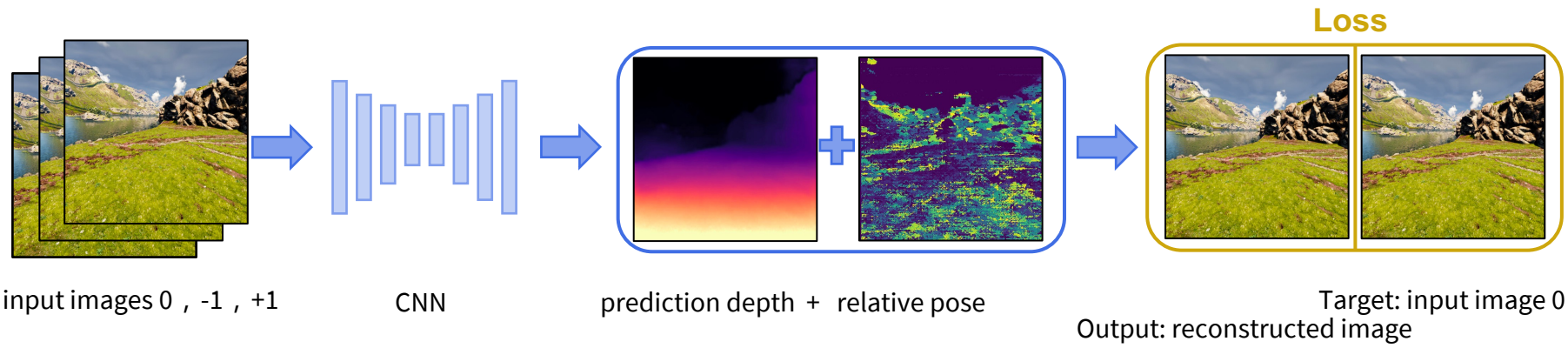
Introduction of background

- **Supervised VS. Unsupervised monocular depth estimation**

- Supervised method —— cost high to acquire ground truth!



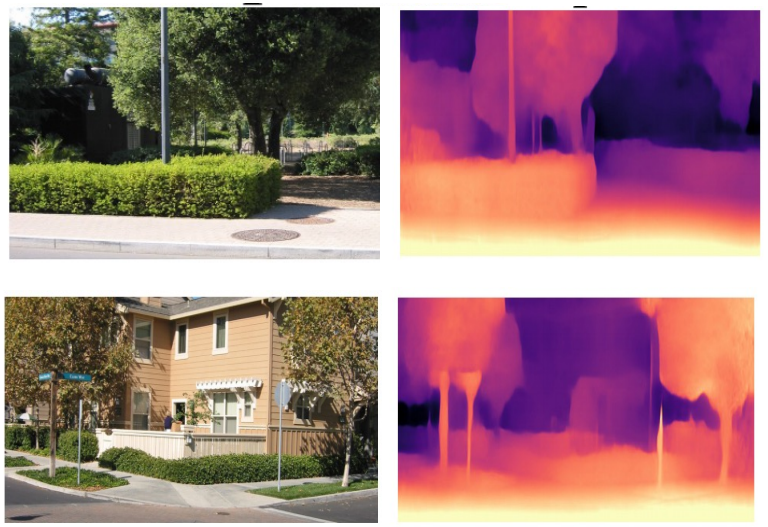
Unsupervised method



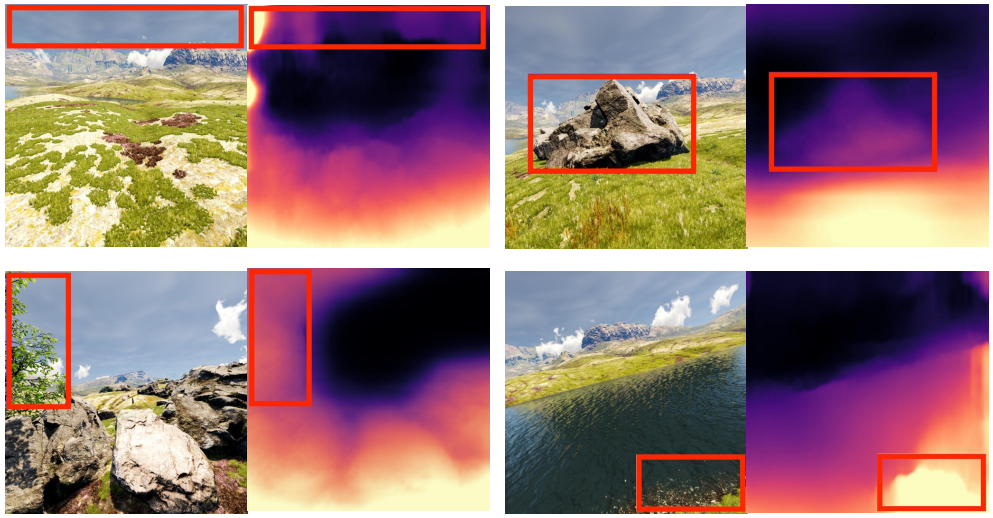
Introduction of background

- **Baseline model—Monodepth2**

Monodepth2 on **KITTI** dataset



Monodepth2 on **MidAir** dataset



Strengths of Monodepth2

- **Unsupervised monocular** depth estimation model
- **State-of-art** model in KITTI dataset

Limitations of Monodepth2

- Almost can't improve the effect of **high-resolution** inputs
- Inaccurate in predicting the depth of **large gradient areas**

Proposed model SS-MDE

SS-MDE: Self-Supervised Monocula

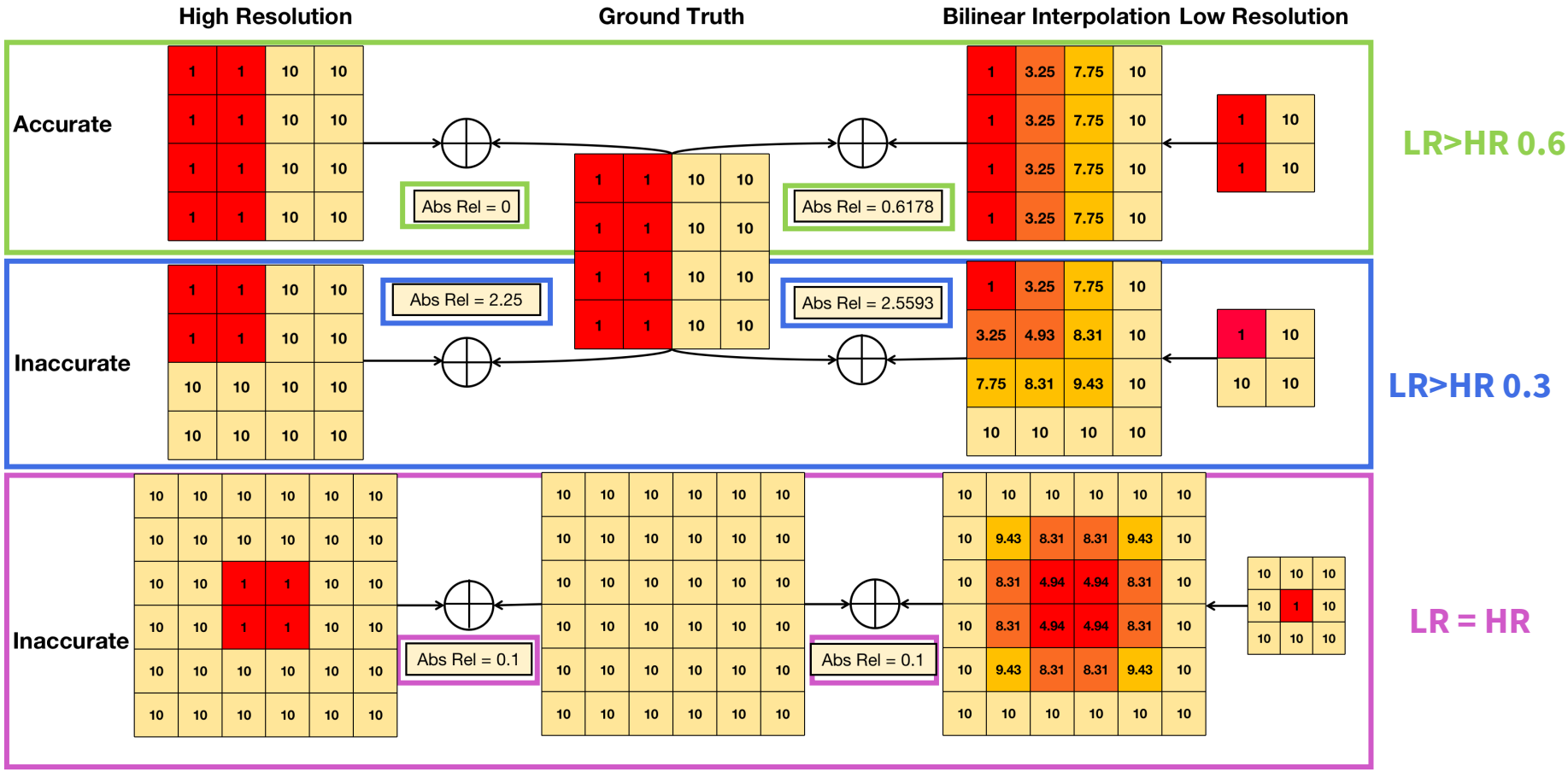
Depth Estimation

HR: High Resolution

LR: Low Resolution

Abs Rel: Absolute relative error

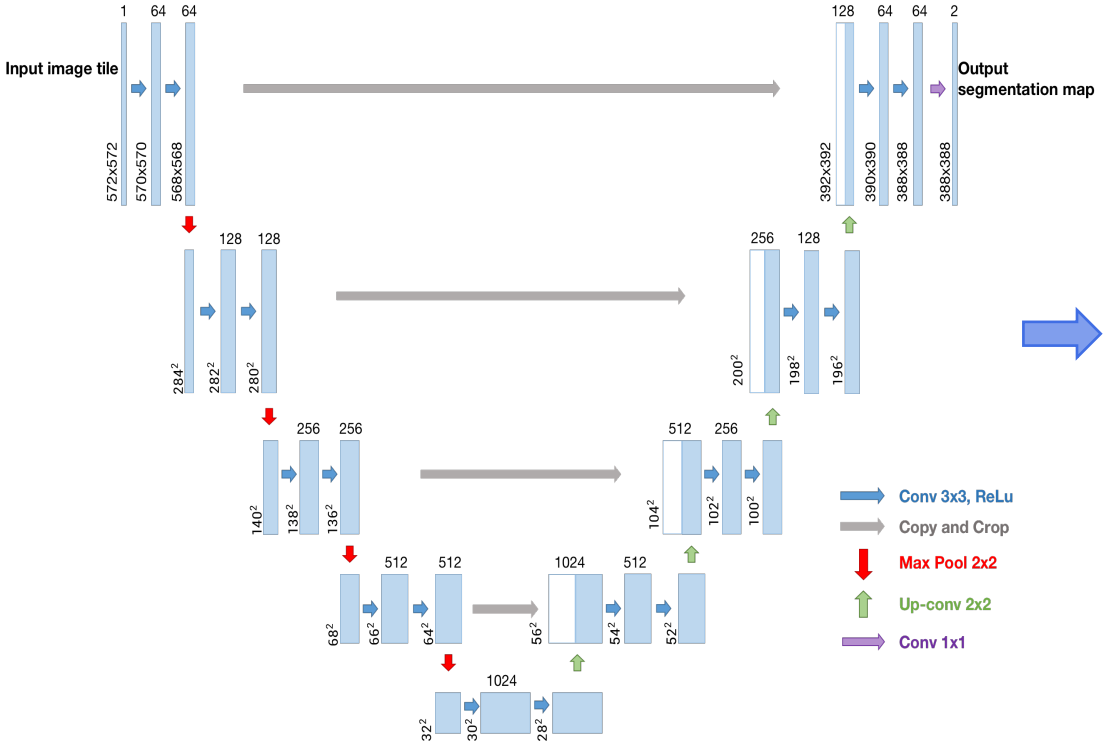
- **Idea for improvement—bilinear interpolation**

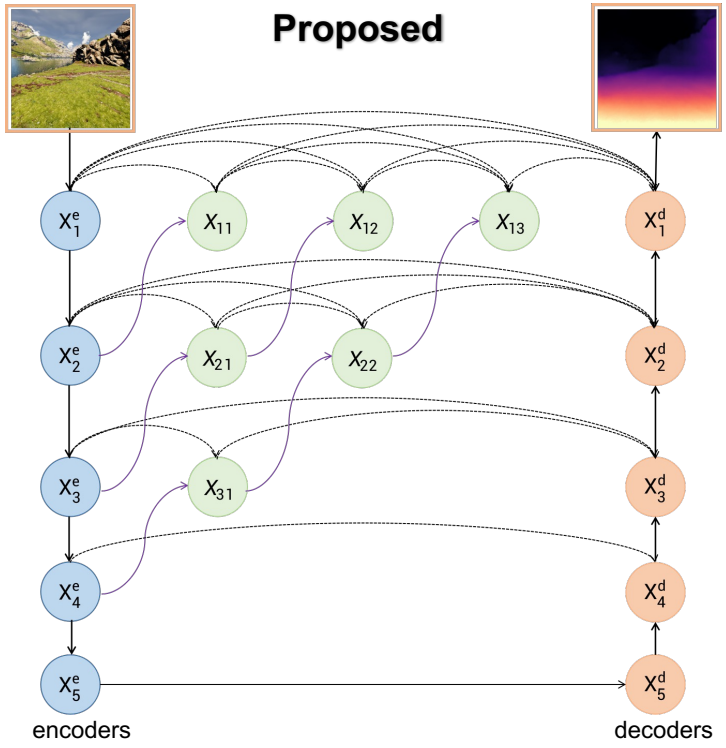


Proposed model SS-MDE

- **Depth estimation network-U-net structure**

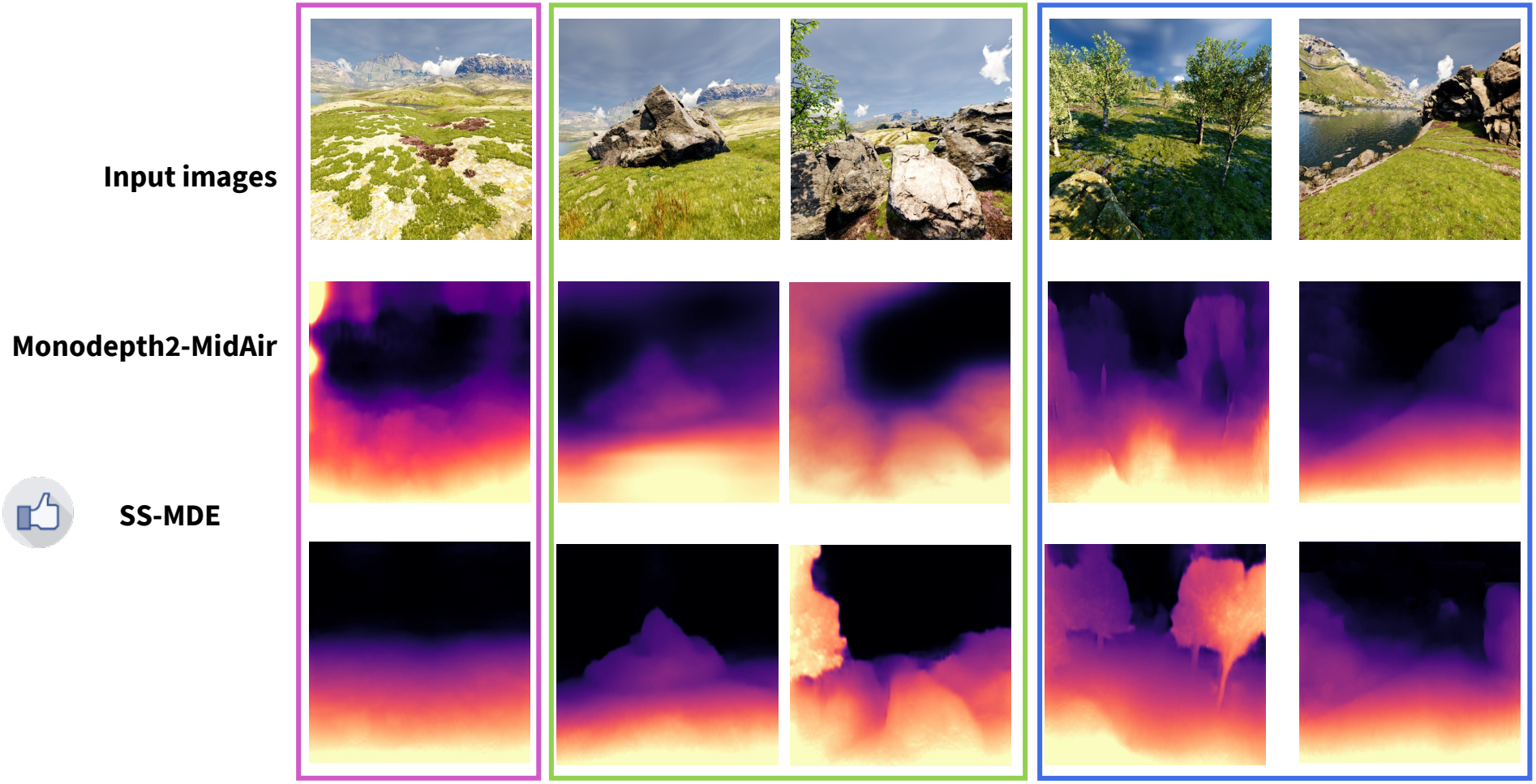
Zhou, Z., & Rahman Siddiquee, & M. , & Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In Deep learning in medical image analysis and Multimodal Learning for Clinical Decision Support (pp. 3-11). Springer, Cham.





Results of the experiment

- **Qualitative results(tested on MidAir)**



Results of the experiment

- **Quantitative results**

Abs Rel：Absolute Relative Error

Sq Rel：Squared Relative Error

RMSE：Root Mean Squared Error

RMSE log：Root Mean Squared Logarithmic Error

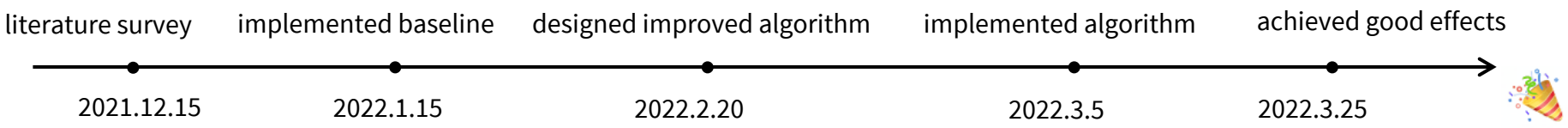
δ：Standard Deviation

Model	Train	Abs Rel↓	Sq Rel↓	RMSE↓	RMSE log↓	$\delta < 1.25 \uparrow$	$\delta < 1.25^2 \uparrow$	$\delta < 1.25^3 \uparrow$
Wang	KITTI	0.241	5.532	12.599	0.368	0.648	0.831	0.911
Monodepth	KITTI	0.314	8.713	13.595	0.438	0.678	0.828	0.895
ST-CLSTM	KITTI	0.404	6.390	13.685	0.438	0.751	0.865	0.911
Monodepth2-KITTI	KITTI	0.717	37.164	74.552	0.882	0.281	0.425	0.521
Monodepth2-MidAir	MidAir	0.135	2.500	13.214	0.222	0.720	0.910	0.996
M4Depth	MidAir	0.143	3.680	**8.864**	0.246	**0.840**	0.924	0.959
 **SS-MDE**	MidAir	**0.114**	**1.742**	10.766	**0.173**	0.791	**0.967**	**0.998**

Summary of my final project

- ✓ Learned deep learning and completed literature survey for monocular depth estimation

- ✓ Implemented the baseline Monodepth2
- ✓ Performded a self-supervised monocular depth estimation in UAV data
- ✓ Designed the algorithm to improve baseline model : dense skip connections in U-Net structure contributing to predicting more accurate depth maps
- ✓ Implemented the algorithm and experiments validated that proposed model SS-MDE achieved state-of-art performance in UAV dataset
- ✓ Realized the application of proposed model SS-MDE with real UAV video and got expected results



Queen Mary
University of London
Science and Engineering

QMUL-BUPT Joint Programme
JP Student Innovation Centre
Annual Showcase 2021/22

Thank you