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

augmented VISION

ResFPN: Residual Skip Connections in Multi-Resolution Feature Pyramid Networks for Accurate Dense Pixel Matching

Rishav, René Schuster, Ramy Battrawy, Oliver Wasenmüller, Didier Stricker

ICPR 2020

Paper ID 48

Features Extraction in Computer Vision Tasks

- Basic cues for all computer vision tasks
- Robust features → robust results
- Even more important for dense pixel-wise matching tasks
- Multi-scale feature extraction popular for dense problems

Dense Pixel Matching

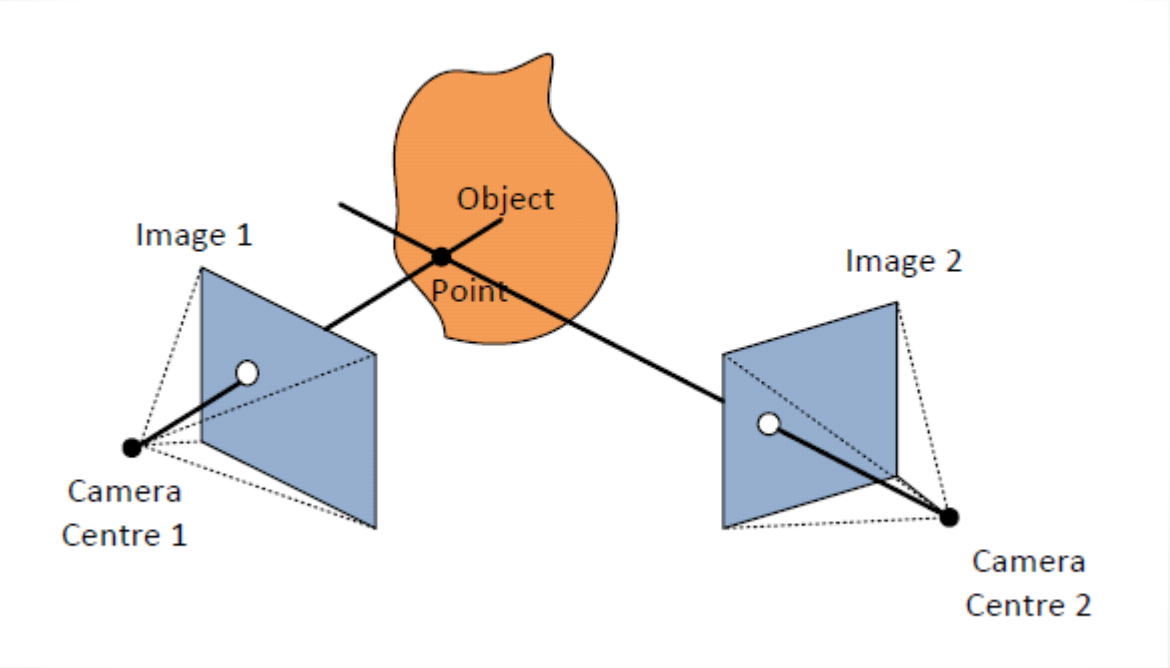
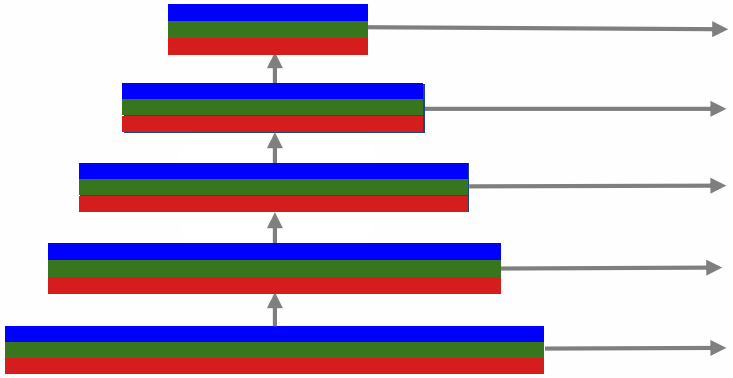
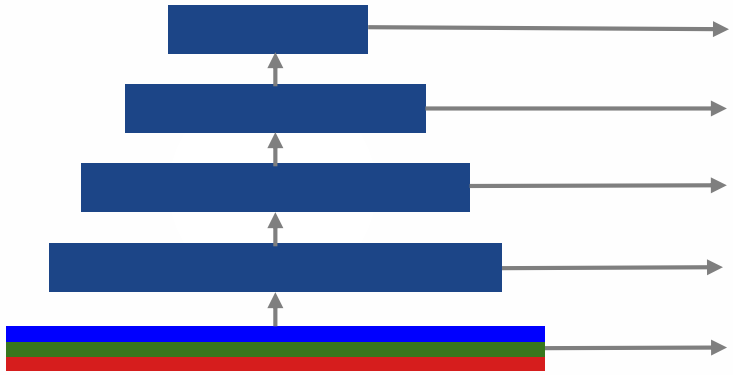


Image Source: Google

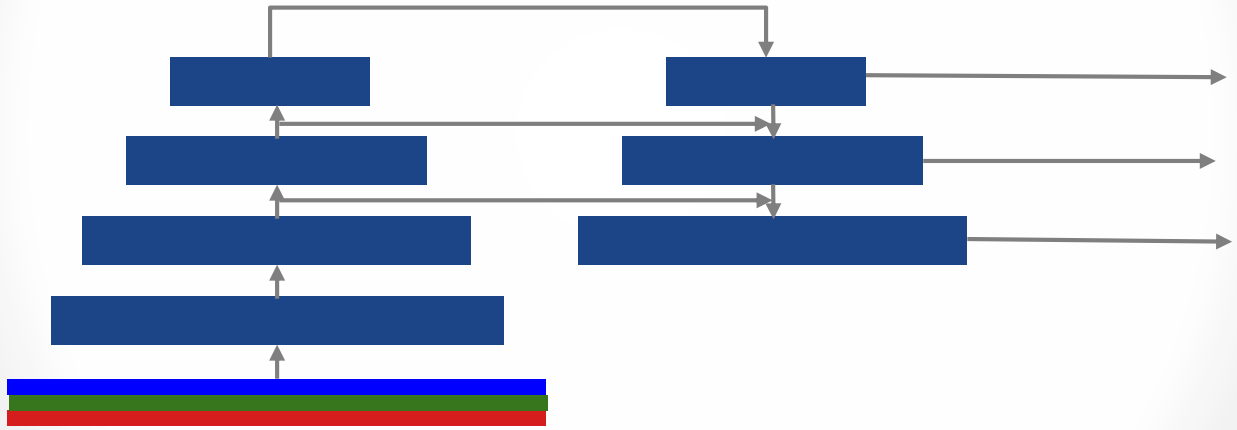
Image Pyramids



Feature Pyramid

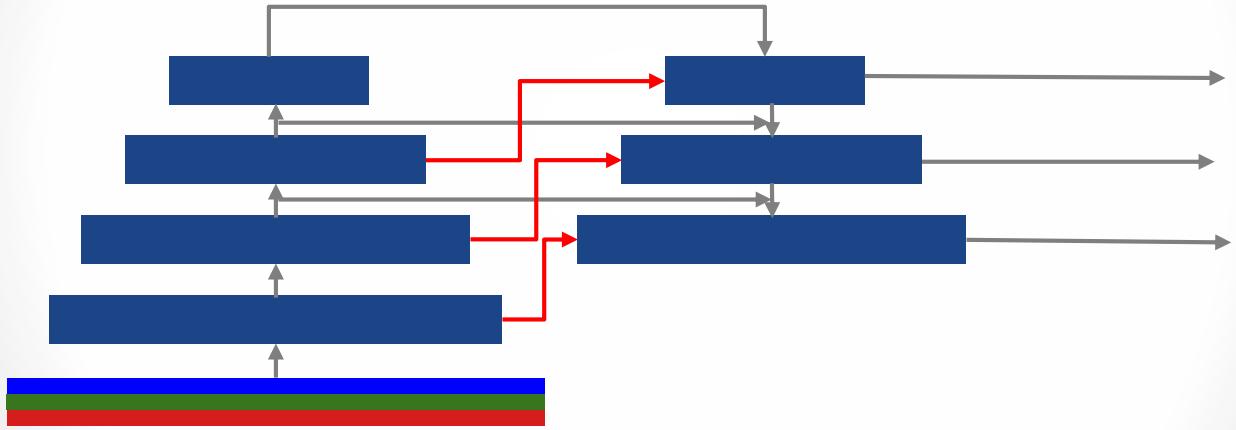


Feature Pyramid Networks [1]



[1] T.-Y. Lin et. al, Feature pyramid networks for object detection. In Conference on Computer Vision and Pattern Recognition (CVPR) 2017.

Residual connections from higher resolution maps

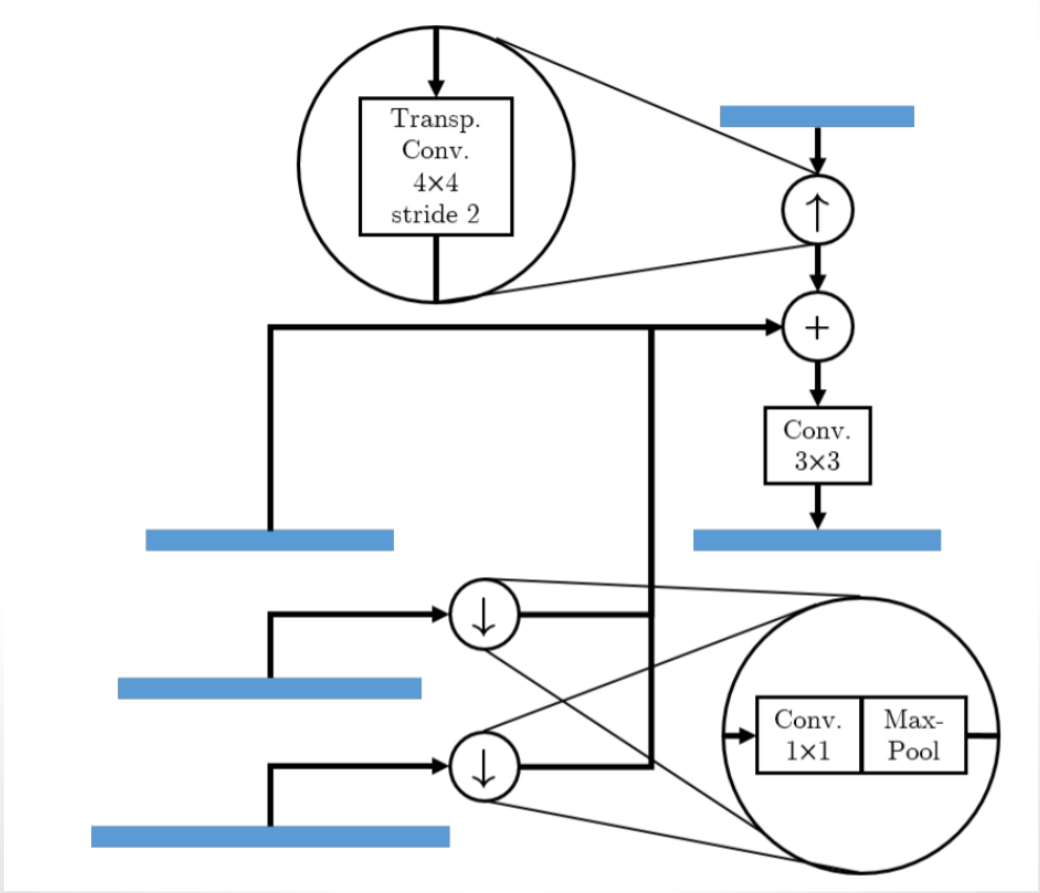


Residual connections from higher resolution maps (Ablation)

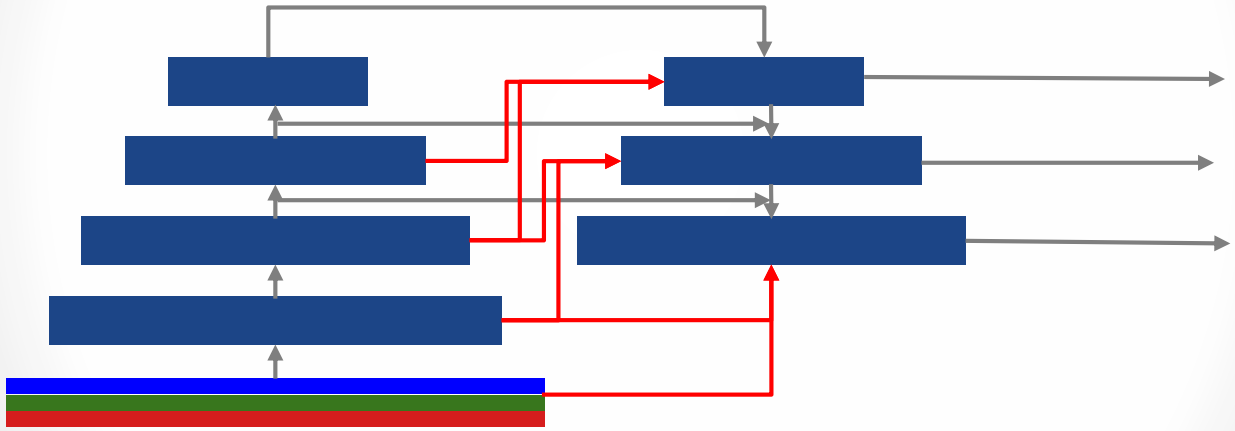
	h	Re-shaping	Merging	FT3D [2] >3px	FT3D [2] EPE	KITTI [3] >3px	KITTI [3] EPE	Parameters $\times 10^6$	FLOPs $\times 10^{12}$
FPN [1]	0	–	addition	21.49	9.15	12.55	3.22	**8.05**	**6.07**
	1	1×1, max-pool	addition	20.95	8.28	11.37	3.09	8.09	6.50
	2	max-pool	concatenation	19.90	7.91	11.21	3.04	8.67	8.94
	2	1×1, max-pool	concatenation	21.16	8.34	11.83	3.02	9.03	12.09
	2	3×3, stride	addition	21.65	8.42	13.67	3.50	8.74	7.43
	2	1×1, bi-linear	addition	20.89	8.09	11.55	3.21	8.12	7.26
	2	max-pool, 1×1	addition	20.28	7.67	12.24	3.06	8.12	6.24
ResFPN	2	1×1, max-pool	addition	**18.91**	**7.19**	**10.63**	**2.98**	8.12	7.30

[1] T.-Y. Lin et. al, Feature pyramid networks for object detection. In Conference on Computer Vision and Pattern Recognition (CVPR) 2017.
[2] Nikolaus Mayer, Eddy Ilg, Philip Hausser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[3] Moritz Menze and Andreas Geiger. Object scene flow for autonomous vehicles. In Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

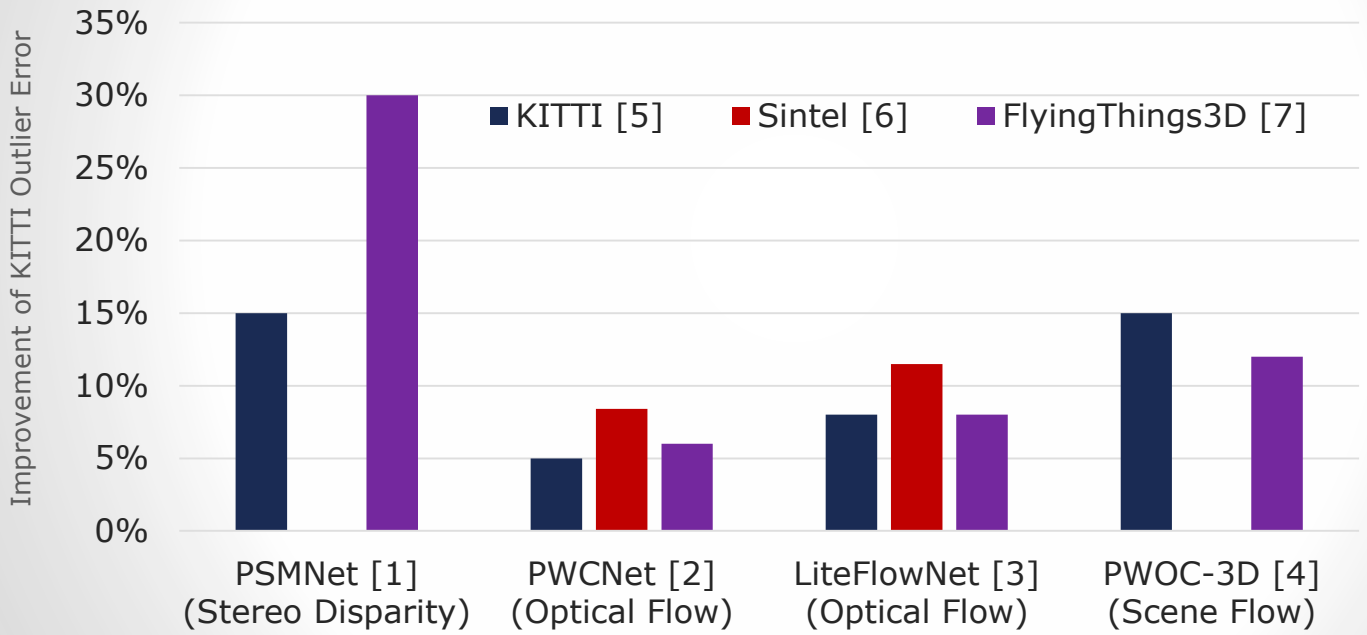
Final Connection Strategy



Our ResFPN

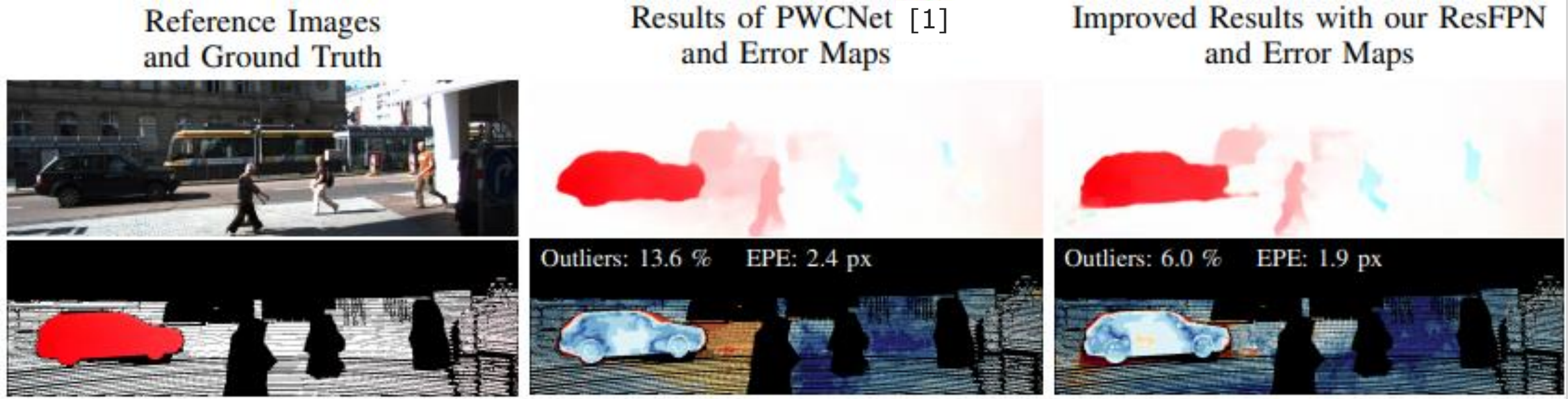


Quantitative Results (Improvement Over Baseline)



[1] Jia-Ren Chang and Yong-Sheng Chen. Pyramid stereo matching network. In Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
[2] Deqing Sun, Xiaodong Yang, Ming-Yu Liu, and Jan Kautz. PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume. In Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
[3] Tak-Wai Hui, Xiaoou Tang, and Chen Change Loy. LiteFlowNet: A lightweight convolutional neural network for optical flow estimation. In Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
[4] Rohan Saxena, René Schuster, Oliver Wasenmüller, and Didier Stricker. PWOC-3D: Deep occlusion-aware end-to-end scene flow estimation. In Intelligent Vehicles Symposium (IV), 2019.
[5] Moritz Menze and Andreas Geiger. Object scene flow for autonomous vehicles. In Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
[6] Daniel J Butler, Jonas Wulff, Garrett B Stanley, and Michael J Black. A naturalistic open source movie for optical flow evaluation. In European Conference on Computer Vision (ECCV), 2012.
[7] Nikolaus Mayer, Eddy Ilg, Philip Hausser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

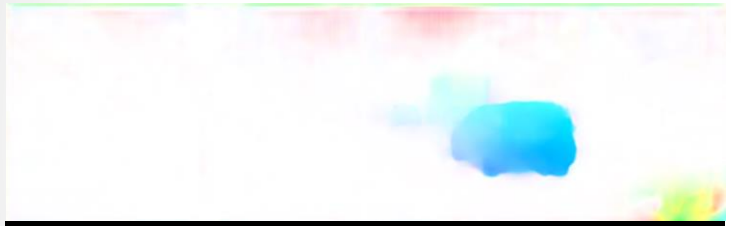
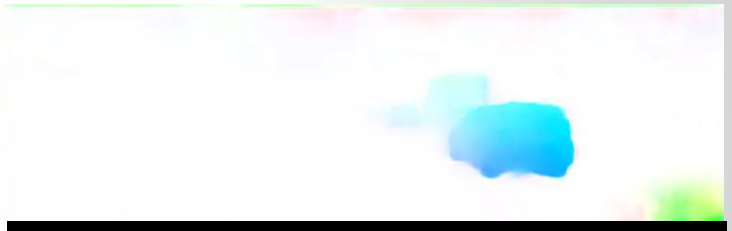
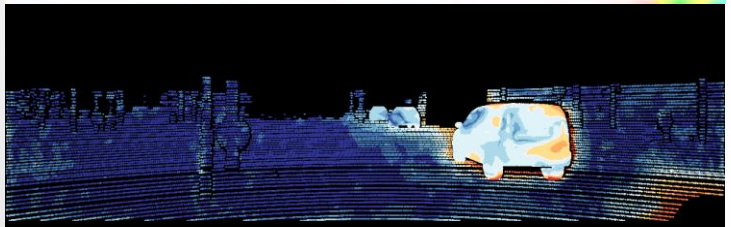
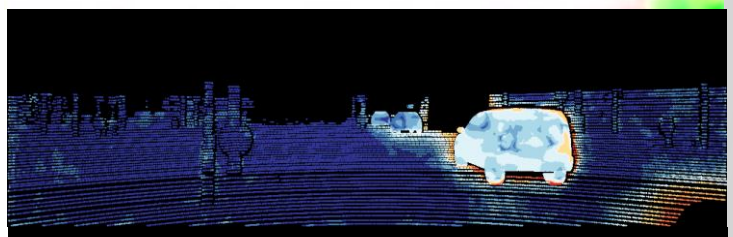
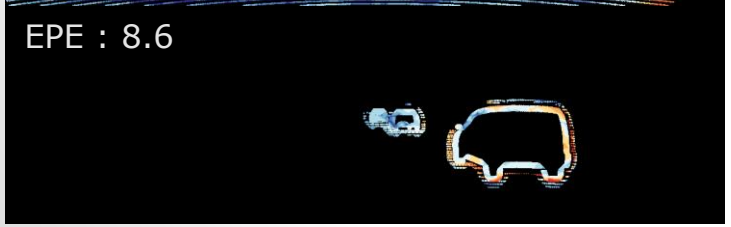
Qualitative Results



[1] Deqing Sun, Xiaodong Yang, Ming-Yu Liu, and Jan Kautz. PWC-Net: CNNs for optical flow using pyramid, warping, and cost volume. In Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

Localization

Some Qualitative Insights (On KITTI [1])

	LiteFlowNet [2] with FPN	LiteFlowNet [2] with **ResFPN**
Optical Flow Output		
Error Map		
20 pixel wide error map		

[1] Moritz Menze and Andreas Geiger. Object scene flow for autonomous vehicles. In Conference on Computer Vision and Pattern Recognition (CVPR), 2015.
[2] Tak-Wai Hui, Xiaoou Tang, and Chen Change Loy. LiteFlowNet: A lightweight convolutional neural network for optical flow estimation. In Conference on Computer Vision and Pattern Recognition (CVPR), 2018

Summary

- ResFPN when used as a feature extractor improves results for several state-of-the-art algorithms on diverse data sets
- Higher resolution maps provide improved localization
- ResFPN is a general concept

Thank You!