SPLATNet: Sparse Lattice Networks for Point Cloud Processing

Hang Su[1,2], Varun Jampani[2], Deqing Sun[2], Subhransu Maji[1], Evangelos Kalogerakis[1], Ming-Hsuan Yang[2,3], Jan Kautz[2]

[1]UMass Amherst [2] NVIDIA [3]UC MERCED

video intro

2018 CVPR SALT LAKE CITY • JUNE 18-22

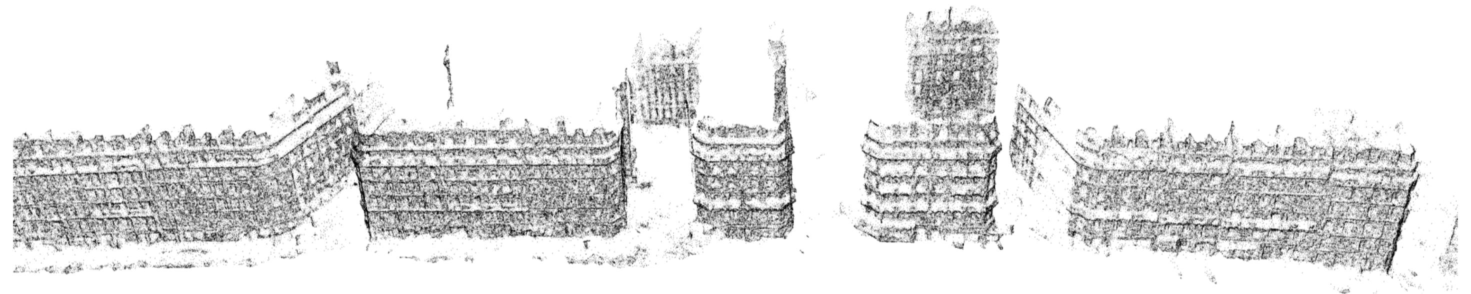
***SPLATNet* operates on point clouds directly and allows joint 2D-3D processing**

Code: `https://github.com/nvlabs/splatnet`

1 Challenges in Point Cloud Processing

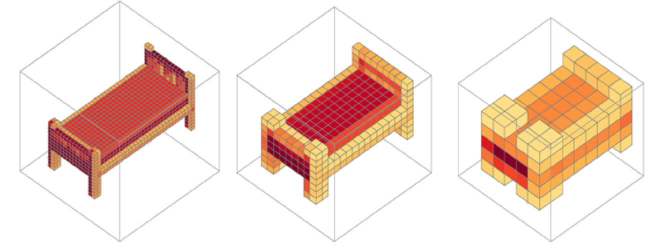
Traditional CNNs are <u>not a good fit</u> for point clouds:

- Point clouds are sets of **sparse**, **unordered** points;
- Traditional CNNs operate on **dense**, **regular** grids.

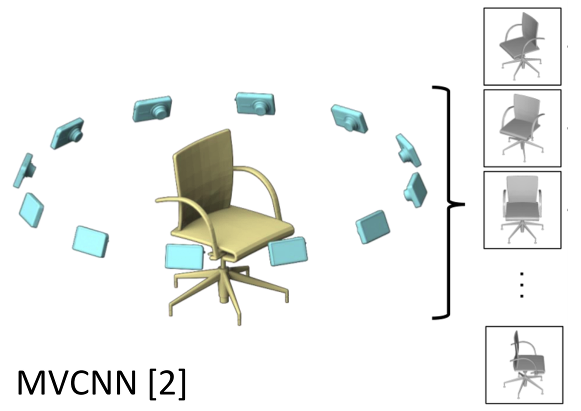


Pre-processing into voxels or 2D projections:

- Loss of details
- Artifacts
- Computational overheads

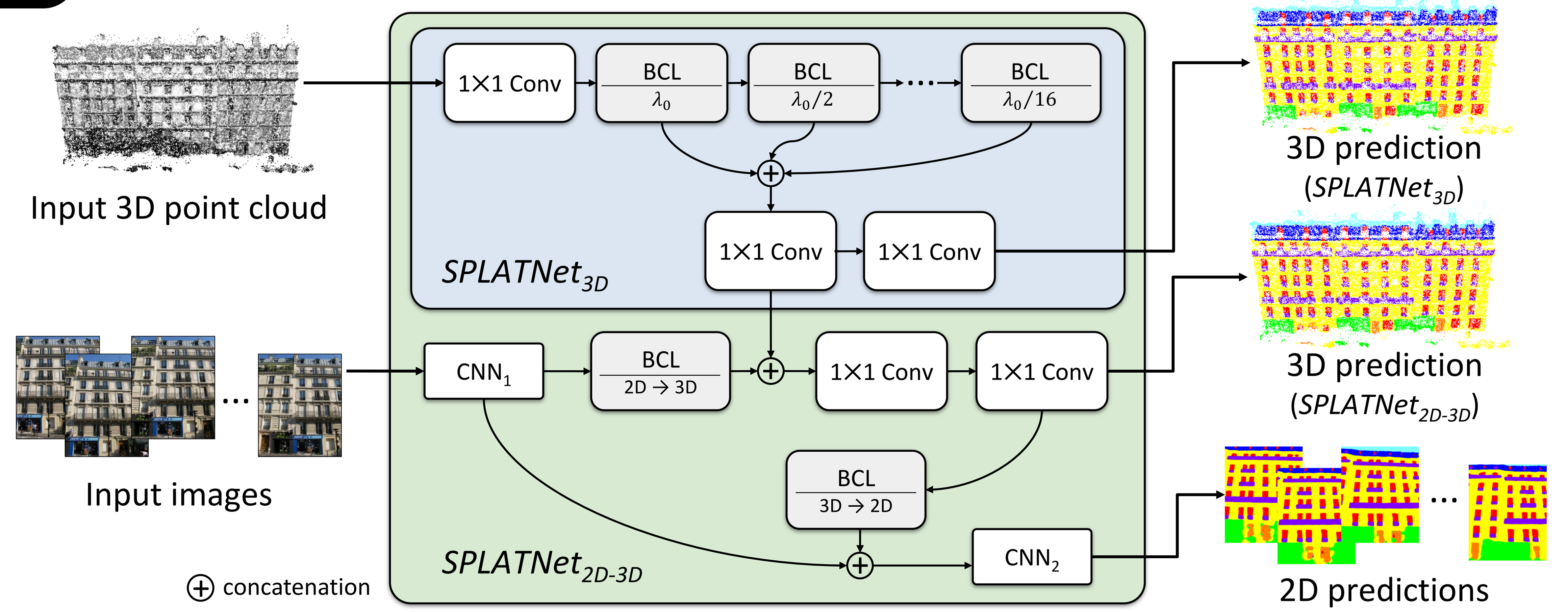


OctNet [1]



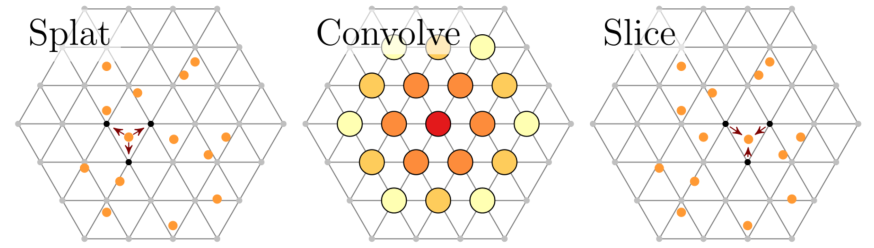
MVCNN [2]

3 Architectures: *SPLATNet*$_{3D}$ and *SPLATNet*$_{2D\text{-}3D}$



2 Bilateral Convolution Layer (BCL) on Point Clouds

Originally introduced in [3], BCL includes:



1. <u>Splat</u>: BCL first interpolates points onto a permutohedral lattice.
2. <u>Convolve</u>: convolution is performed over the sparsely populated lattice vertices.
3. <u>Slice</u>: the filtered signal is interpolated back onto the original point locations.

Efficient sparse high-dim. filtering:

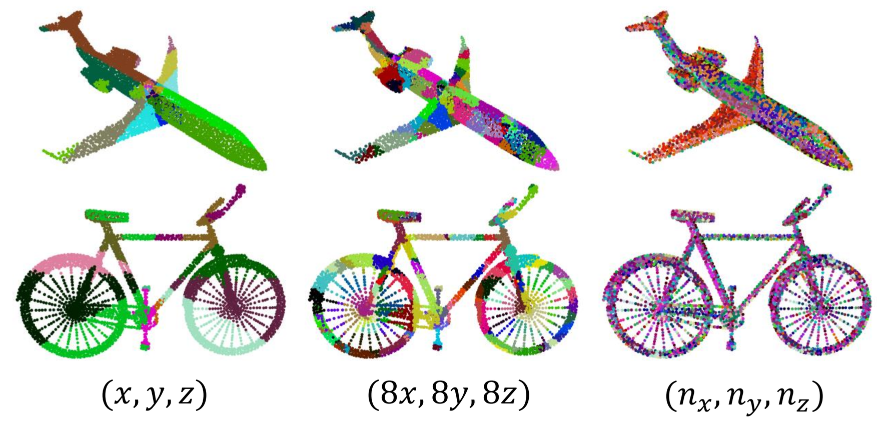
- hash table & permutohedral lattice

<u>Two separate sets of features</u>

Image				
point feature	(r,g,b)	$f(\ldots)$		
lattice feature	(x,y)	(x,y)		
Point cloud				
point feature	1	(x,y,z)	(r,g,b)	$f(\ldots)$
lattice feature	(x,y,z)	(x,y,z)	(n_x,n_y,n_z)	(x,y,z)

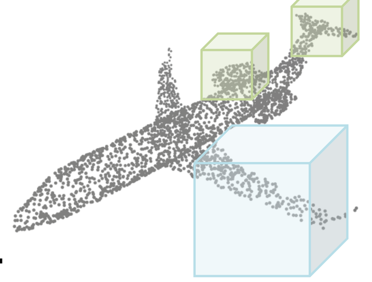
point feature → "what"

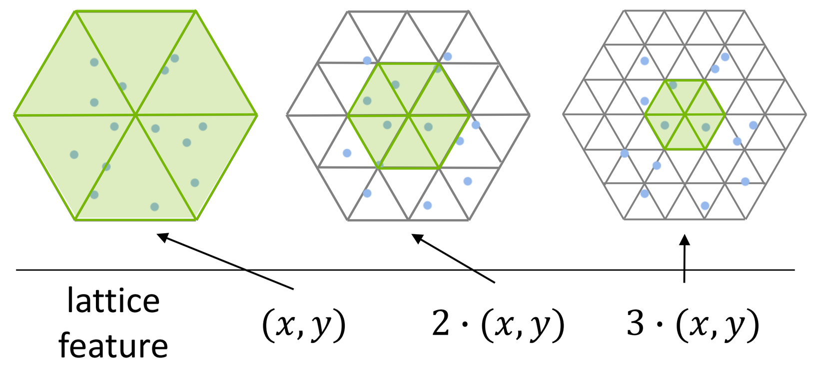
lattice feature → "where"



<u>Control over filter receptive fields</u>

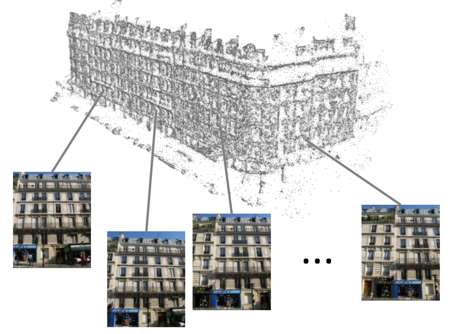
- Varying receptive field sizes can help capture information from multiple scales.
- This can be easily achieved by simply scaling lattice features.

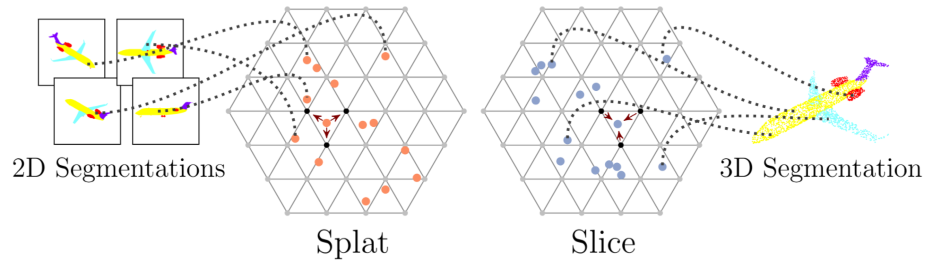




<u>Smooth projection between 2D&3D</u>

- Many sensors output 3D points and 2D images simultaneously.
- Images can provide complementary information to 3D data, and is also often less noisy.
- We achieve this capability by a special type of BCL.





Splatting 2D data (given (l,x,y,z)) into 3D

4 Facade Semantic Segmentation

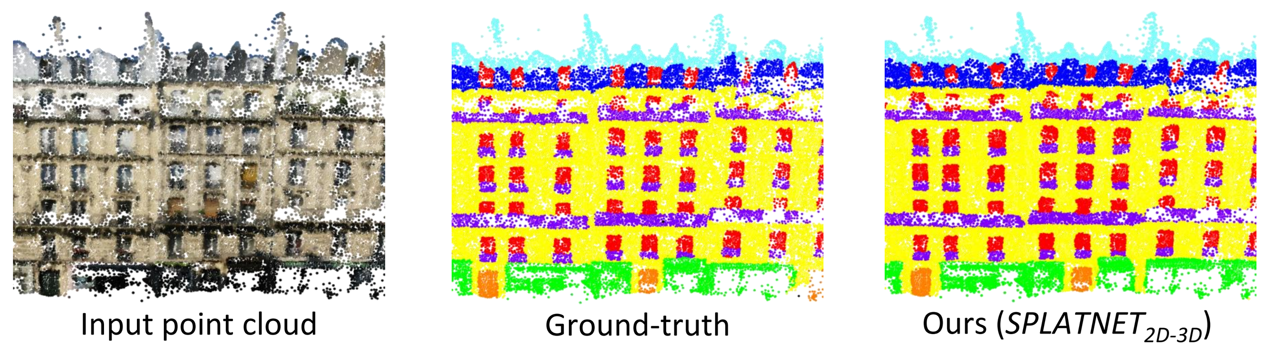
Ruemonge2014 [4]

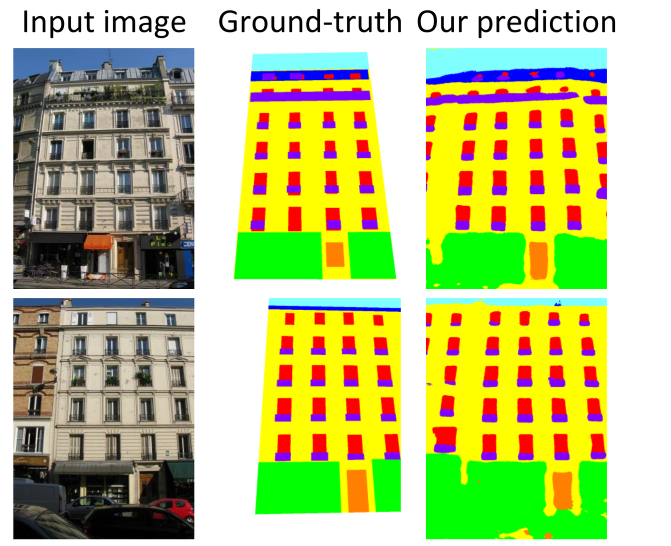
with only 3D data

	IoU	runtime (min)
OctNet [1]	59.2	-
Autocontext$_{3D}$ [5]	54.4	16
SPLATNet$_{3D}$	**65.4**	**0.06**

with both 2D & 3D data

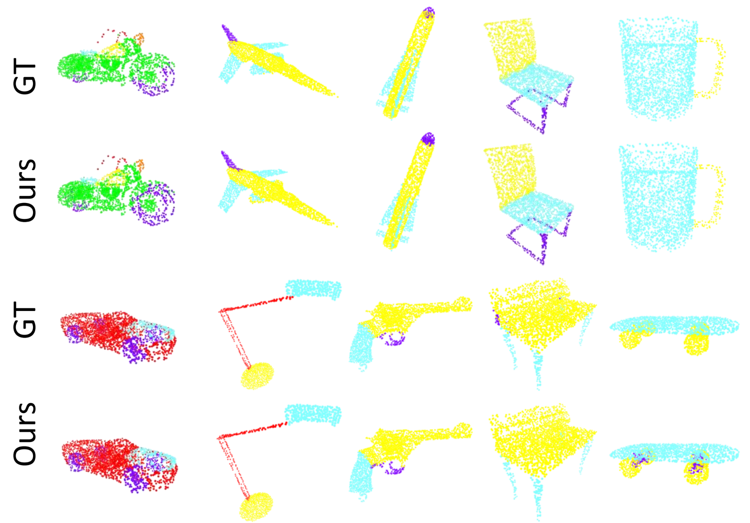
	IoU	runtime (min)
Autocontext$_{2D\text{-}3D}$ [5]	62.9	87
SPLATNet$_{2D\text{-}3D}$	**69.8**	**1.2**



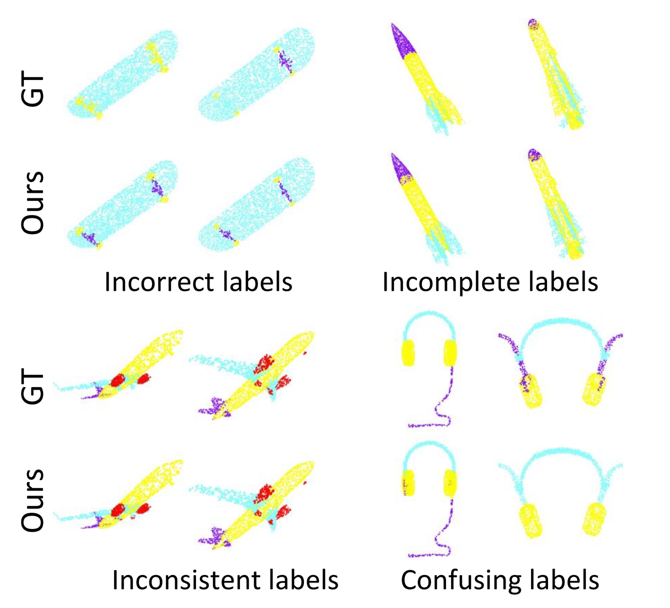


	IoU	runtime (min)
Autocontext$_{2D}$ [5]	60.5	117
Autocontext$_{2D\text{-}3D}$ [5]	62.7	146
2D CNN [6] only	69.3	0.84
SPLATNet$_{2D\text{-}3D}$	**70.6**	**4.34**

5 ShapeNet Part Segmentation



	class avg. IoU	instance avg. IoU
Yi *et al.* [7]	79.0	81.4
3DCNN [8]	74.9	79.4
Kd-network [9]	77.4	82.3
PointNet [8]	80.4	83.7
PointNet++ [10]	81.9	85.1
SyncSpecCNN [11]	82.0	84.7
SPLATNet$_{3D}$	**82.0**	**84.6**
SPLATNet$_{2D\text{-}3D}$	**83.7**	**85.4**



↑ signs of performance saturation on the benchmark!

6 Conclusion

We propose SPLATNet for efficient and flexible processing directly on point clouds. It can also incorporate 2D images for seamless joint 2D-3D learning.

References:

[1] G. Riegler et al. Octnet: Learning deep 3D representations at high resolutions. CVPR'17.

[2] H. Su et al. Multi-view convolutional neural networks for 3D shape recognition. ICCV'15.

[3] V. Jampani et al. Learning sparse high dimensional filters: Image filtering, dense CRFs and bilateral neural networks. CVPR'16.

[4] H. Riemenschneider et al. Learning where to classify in multi-view semantic segmentation. ECCV'14.

[5] R. Gadde et al. Efficient 2D and 3D facade segmentation using auto-context. PAMI'17.

[6] L.-C. Chen et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs. ICLR'15.

[7] L. Yi et al. A scalable active framework for region annotation in 3D shape collections. SIGGRAPH Asia'16.

[8] C. R. Qi et al. PointNet: Deep learning on point sets for 3D classification and segmentation. CVPR'17.

[9] R. Klokov and V. Lempitsky. Escape from cells: Deep Kd-Networks for the recognition of 3D point cloud models. ICCV'17.

[10] C. R. Qi et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space. NIPS'17.

[11] L. Yi et al. SyncSpecCNN: Synchronized spectral CNN for 3D shape segmentation. CVPR'17.

hsu@cs.umass.edu; {vjampani, deqings}@nvidia.com