3D Instances as 1D Kernels

Yizheng Wu[1*], Min Shi[1*], Shuaiyuan Du[1], Hao Lu[1], Zhiguo Cao[1✉], and Weicai Zhong[2]
[1]School of AIA, Huazhong University of Science and Technology
[2]Huawei CBG Consumer Cloud Service Search & Maps BU

ECCV TEL AVIV 2022

Motivation and Insights

Key Idea

Represent 3D instances as 1D vectors, termed ***instance kernel***. Instance masks are recovered by scanning these kernels on the whole scene.

Challenges

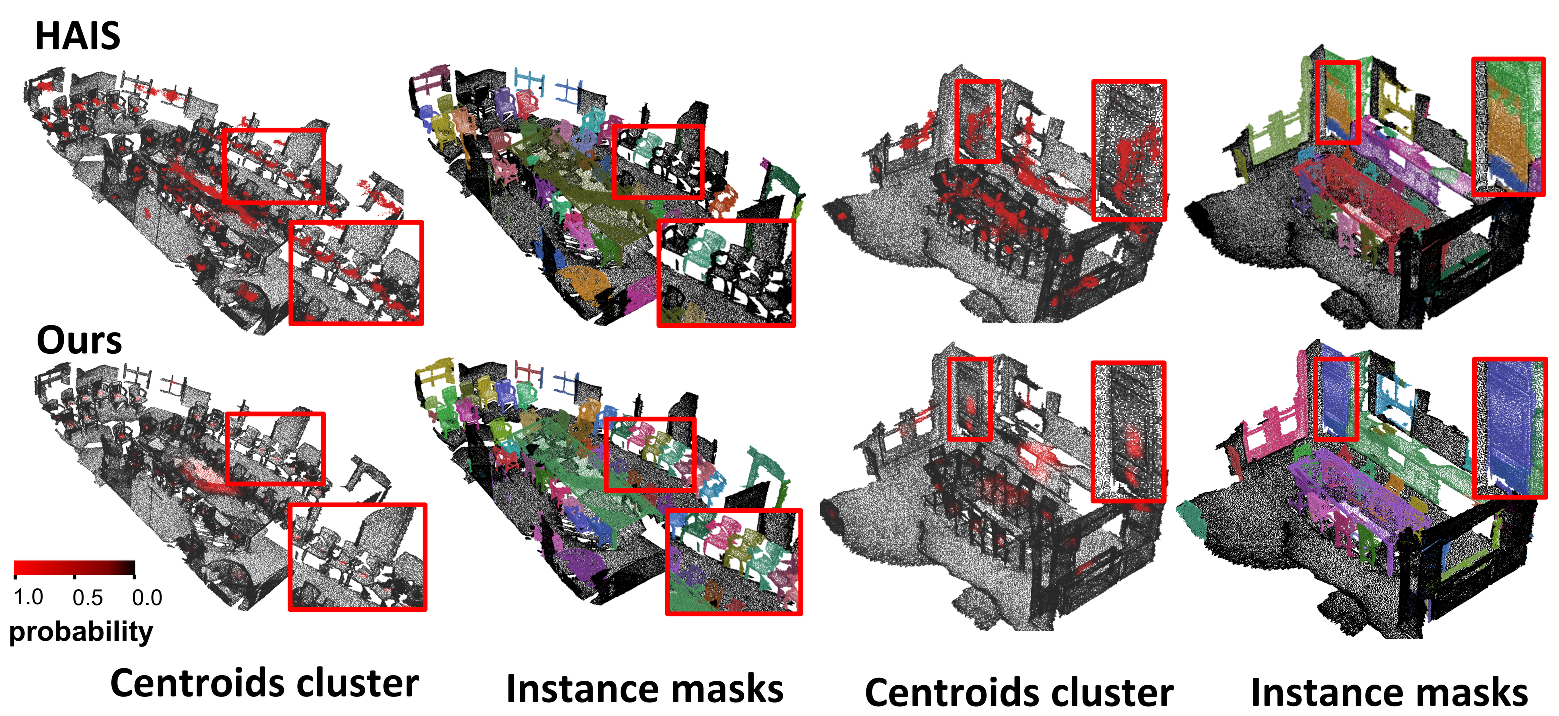
- Precise localization of instances among disordered point cloud data.
- Effective feature aggregation for discriminative instance representation.

Contribution

- We extend the idea of dynamic convolution into ***instance kernel***, a comprehensive representation for 3D instances in point clouds;
- We propose a dynamic kernel network for 3D instance segmentation, with a novel instance kernel encoding paradigm.

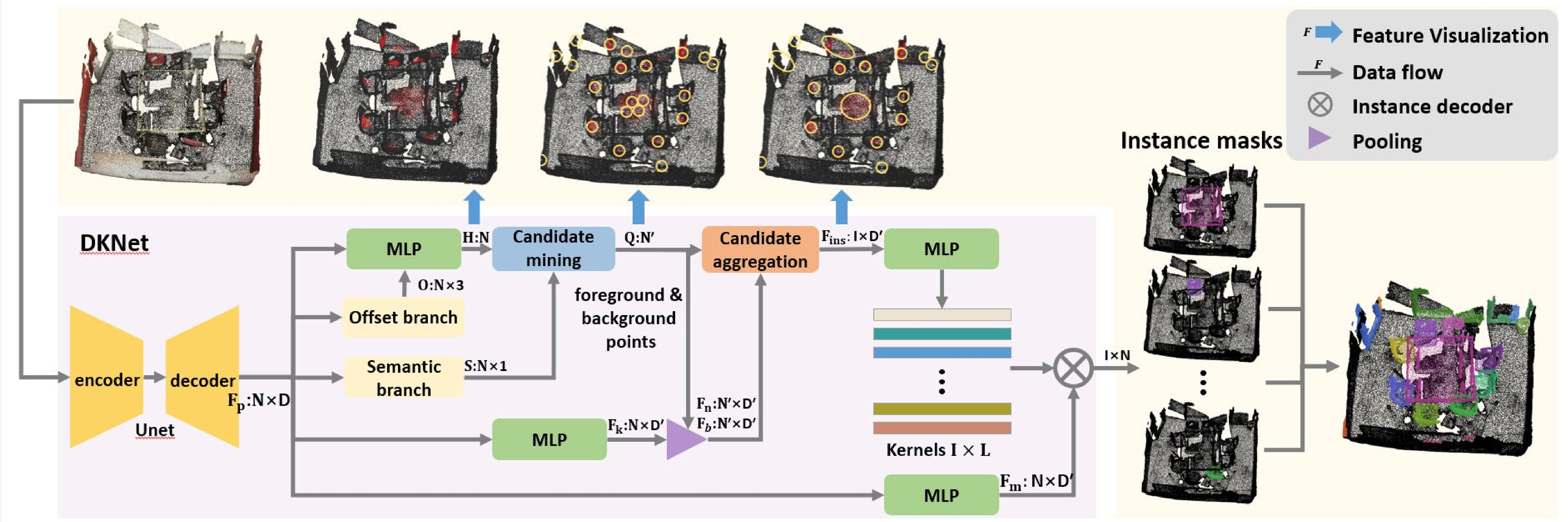
Advantages

Better objectness compared with the prevailing bottom-up methods.



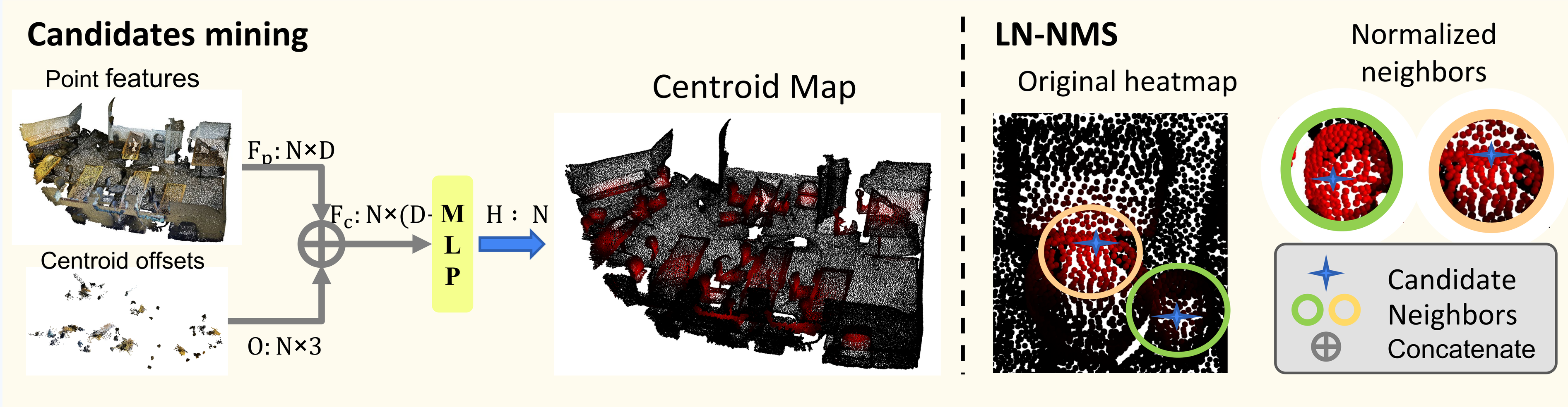
Dynamic Kernel Network

Overall Pipeline



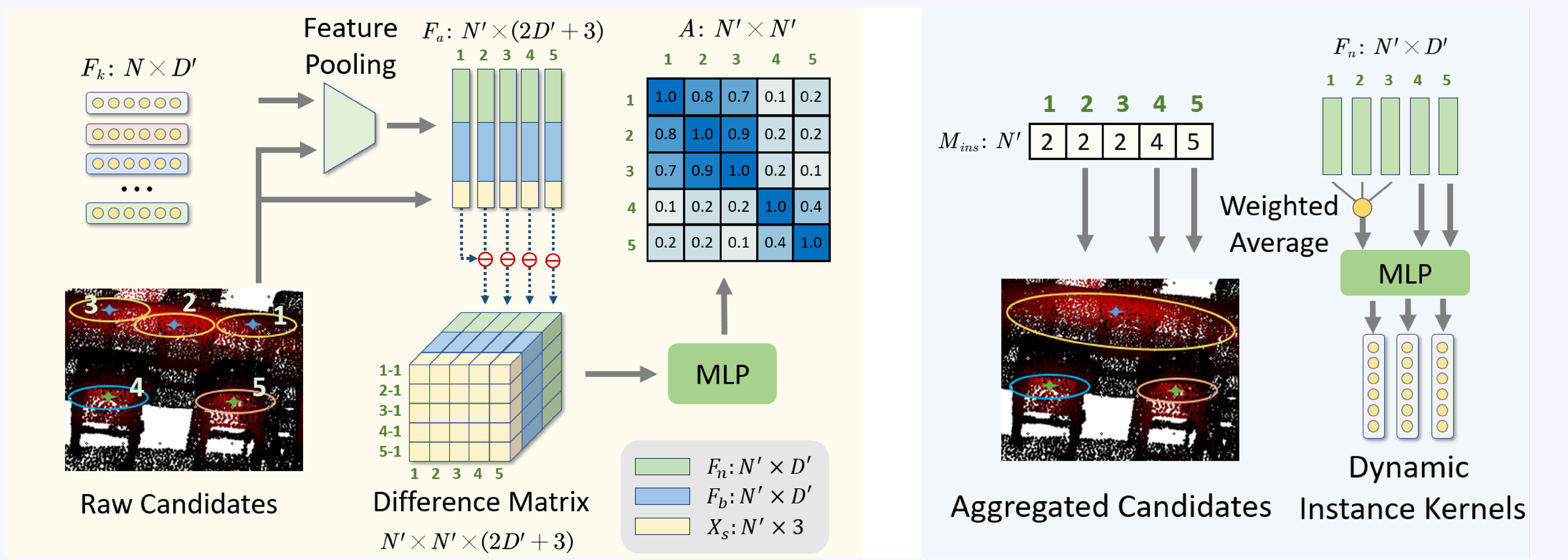
A localize-then-aggregate pipeline for 3D instance segmentation.

Candidate Mining Branch



Localize instance centroids by applying the proposed LN-NMS algorithm on centroids (objectness) map.

Candidate aggregation branch



Merge duplicated candidates and collect context around instance centroids to form instance kernels.

Experiments

Qualitive results on ScanNetV2 test set.

approaches	mAP	$AP@50$	bed	booksh.	cabinet	chair	curtain	desk	door	otherfu.	picture	refrige.	sofa	table	window
3D-BoNet[28]	25.3	48.8	67.2	59.0	30.1	48.4	62.0	30.6	34.1	25.9	12.5	43.4	49.9	51.3	43.9
3D-SIS[10]	16.1	38.2	43.2	24.5	19.0	57.7	26.3	3.3	32.0	24.0	7.5	42.2	69.9	27.1	23.5
MTML[14]	28.2	54.9	80.7	58.8	32.7	64.7	81.5	18.0	41.8	36.4	18.2	44.5	68.8	57.1	39.6
3D-MPA[4]	35.5	61.1	83.3	76.5	52.6	75.6	58.8	47.0	43.8	43.2	35.8	65.0	76.5	55.7	43.0
PointGroup[12]	40.7	63.6	76.5	62.4	50.5	79.7	69.6	38.4	44.1	55.9	47.6	59.6	75.6	55.6	51.3
GICN[17]	34.1	63.8	**89.5**	80.0	48.0	67.6	73.7	35.4	44.7	40.0	36.5	70.0	83.6	59.9	47.3
DyCo3D[9]	39.5	64.1	84.1	**89.3**	53.1	80.2	58.8	44.8	43.8	53.7	43.0	55.0	76.4	65.7	56.8
Occuseg[6]	48.6	67.2	75.8	68.2	57.6	84.2	50.4	52.4	56.7	58.5	45.1	55.7	79.7	56.3	46.7
PE[30]	39.6	64.5	77.3	79.8	53.8	78.6	79.9	35.0	43.5	54.7	54.5	64.6	76.1	55.6	50.1
SSTNet[16]	50.6	69.8	69.7	88.8	55.6	80.3	62.6	41.7	55.6	58.5	70.2	60.0	72.0	69.2	50.9
HAIS[2]	45.7	69.9	84.9	82.0	67.5	80.8	75.7	46.5	51.7	59.6	55.9	60.0	76.7	67.6	56.0
SoftGroup[25]	50.4	**76.1**	80.8	84.5	**71.6**	86.2	**82.4**	**65.5**	62.0	**73.4**	69.9	**79.1**	**84.4**	**76.9**	59.4
Ours	**53.2**	71.8	81.4	78.2	61.9	**87.2**	75.1	56.9	**67.7**	58.5	**72.4**	63.3	81.9	73.6	**61.7**

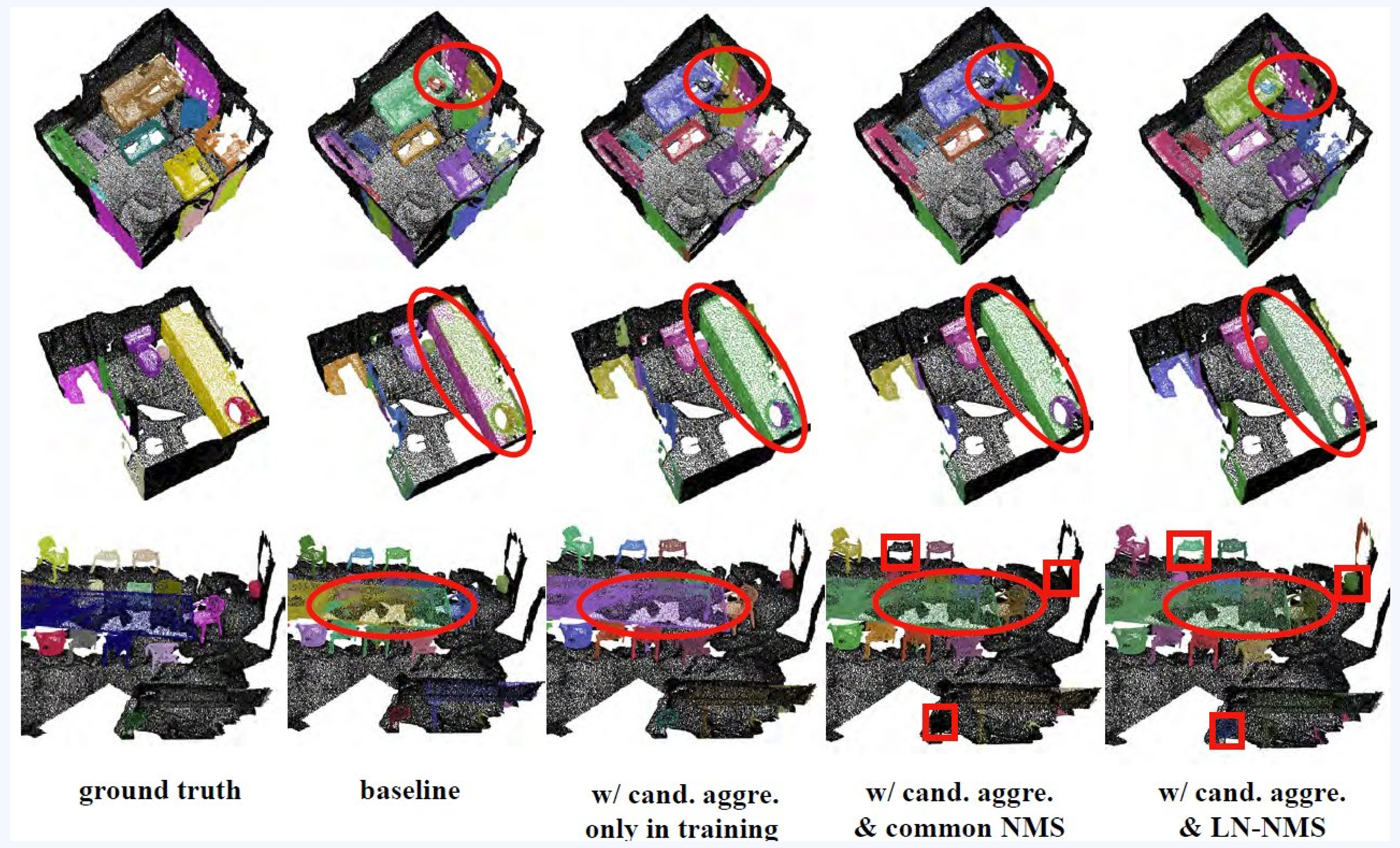
S3DIS dataset

approach	mCov	mWCov	mPre	mRec
PointGroup‡ [12]	-	-	61.9	62.1
DyCo3D‡ [9]	63.5	64.6	64.3	64.2
SSTNet‡ [16]	-	-	65.5	64.2
HAIS‡ [2]	64.3	**66.0**	**71.1**	65.0
Ours‡	**64.7**	65.6	70.8	**65.3**
OccuSeg [6]	-	-	72.8	60.3
GICN [17]	-	-	68.5	50.8
PointGroup [12]	-	-	69.6	69.2
SSTNet [16]	-	-	73.5	**73.4**
HAIS [2]	67.0	70.4	73.2	69.4
SoftGroup [25]	69.3	71.7	**75.3**	69.8
Ours	**70.3**	**72.8**	**75.3**	71.1

Object Detection (ScanNetV2)

approach	$AP@25$	$AP@50$
VoteNet [22]	58.6	33.5
3DSIS [10]	40.2	22.5
3D-MPA [4]	64.2	49.2
DyCo3D [9]	58.9	45.3
HAIS [2]	66.0	54.2
Ours^{-}	65.4	57.9
Ours	**67.4**	**59.0**

Visualization



Email:
✉: yzwu21@hust.edu.cn ✉: zgcao@hust.edu.cn

Code:
https://github.com/W1zheng/DKNet

Acknowledgement

This work was supported in part by the National Key R&D Program of China (No.2018YFB1305504) and the DigiX Joint Innovation Center of Huawei-HUST.