Mask3D: Mask Transformer for 3D Instance Segmentation

Jonas Schult[1] Francis Engelmann[2,3] Alexander Hermans[1] Or Litany[4] Siyu Tang[3] Bastian Leibe[1]
[1]RWTH Aachen University [2]ETH AI Center [3]ETH Zürich [4]NVIDIA

https://JonasSchult.github.io/Mask3D/

Project & Demo Page

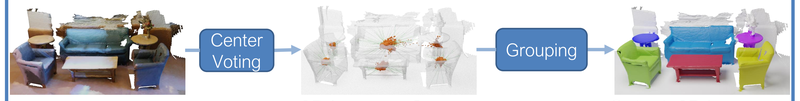
Try on your own scans!

ICRA LONDON 2023

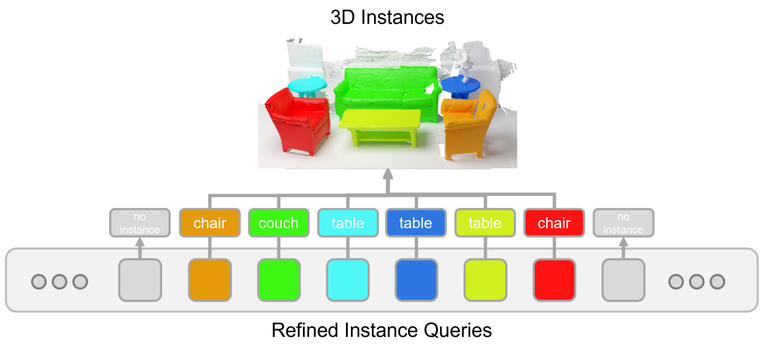
In a nutshell

- ✔ New paradigm for 3D instance segmentation
- ✔ First transformer-based model
- ✔ No need for highly 3D specific components
- ✔ SOTA on ScanNet, ScanNet200, S3DIS and STPLS3D

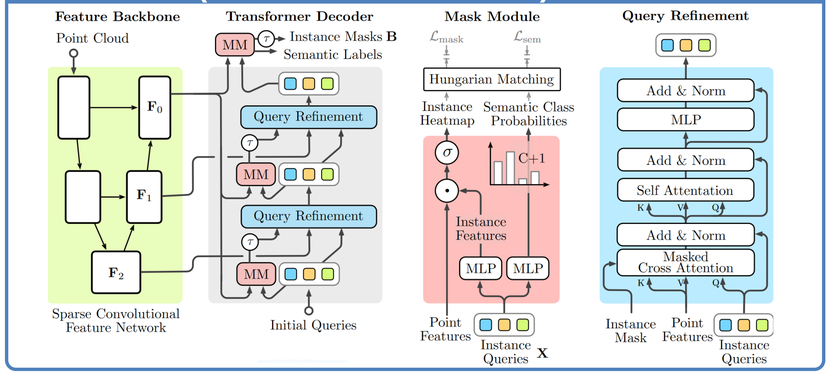
The Dominating Voting & Grouping Paradigm



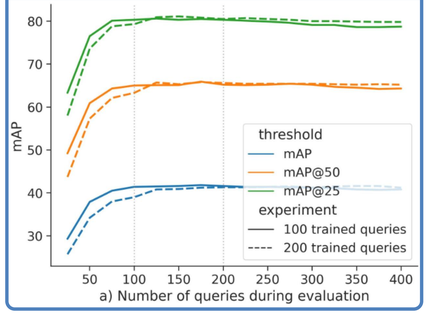
Our Paradigm



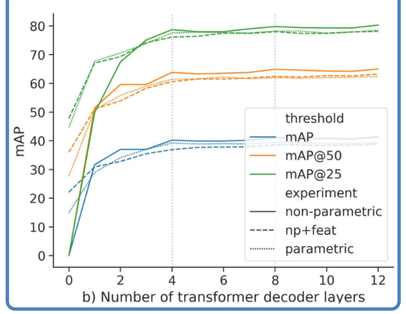
Architecture (based on Mask2Former)



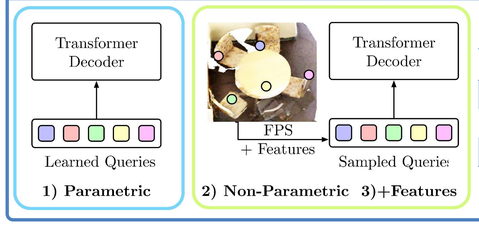
Number Of Queries



Number Of Decoder Steps

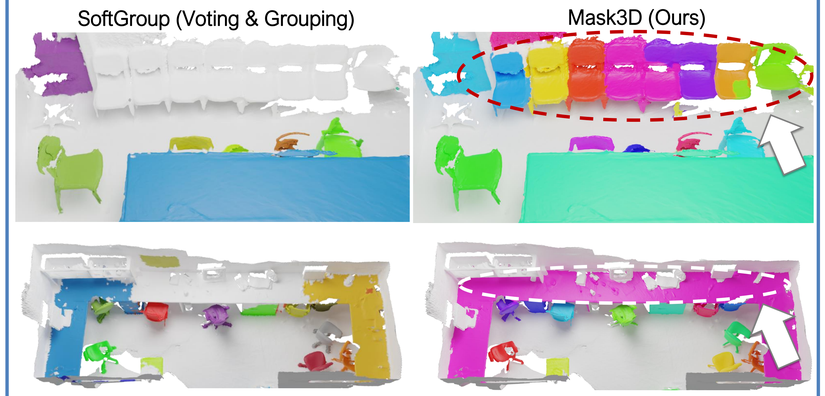


Query Types

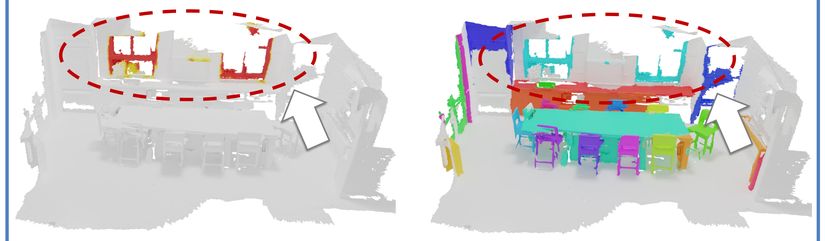


	Positions	Features	mAP
1)	Parametric	Parametric	39.7
2)	Sampled	Zeros	**40.6**
3)	Sampled	Features	38.4

Qualitative Comparison to SoftGroup



Limitations (Merged Instances)



Benchmark Results

Method	ScanNet200 (mAP)
CSC	12.3
Mink34D	13.0
LGround	15.4
Mask3D (**ours**)	27.8

+12.4

Method	STPLS3D (mAP)
PointGroup	23.3
HAIS	35.1
SoftGroup	46.2
Mask3D (**ours**)	57.1

+10.9

Method	ScanNet (mAP)
HAIS	45.7
OccuSeg	48.6
SoftGroup	50.4
SSTNet	50.6
Mask3D (**ours**)	56.6

+6.0