

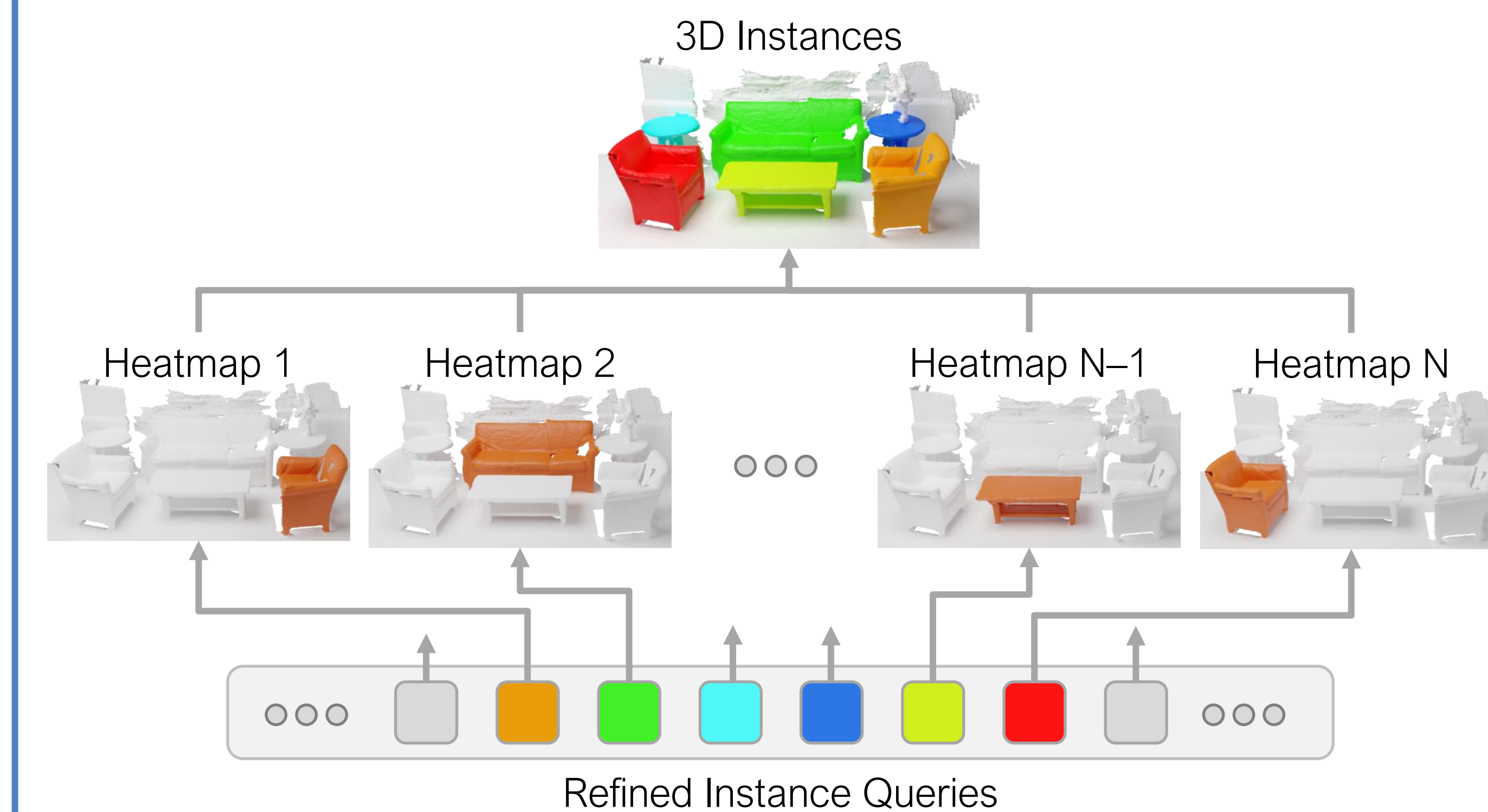
In a nutshell

- ✓ New paradigm for 3D instance segmentation
- ✓ First transformer-based model
- ✓ No need for highly 3D specific components
- ✓ SOTA on ScanNet
- ✓ Fast Method

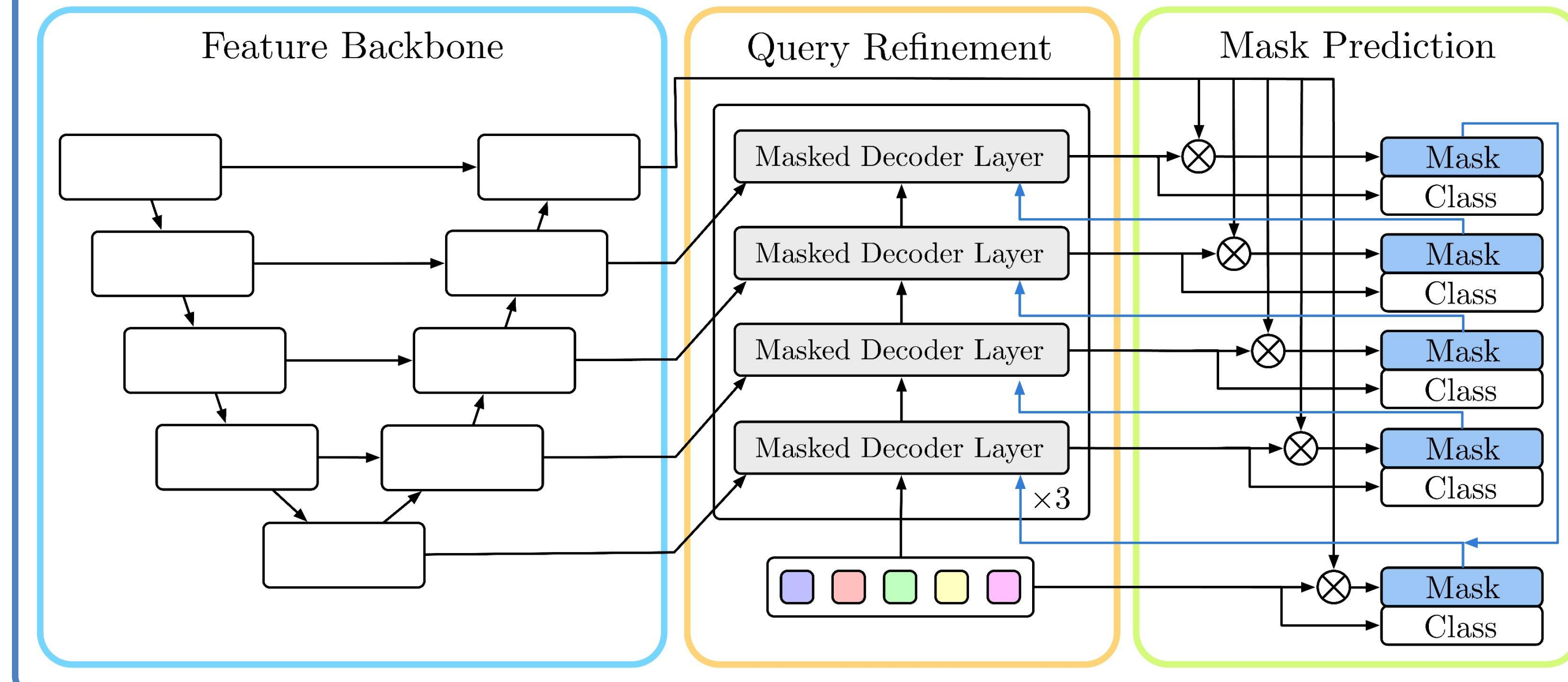
The Dominating Voting & Grouping Paradigm



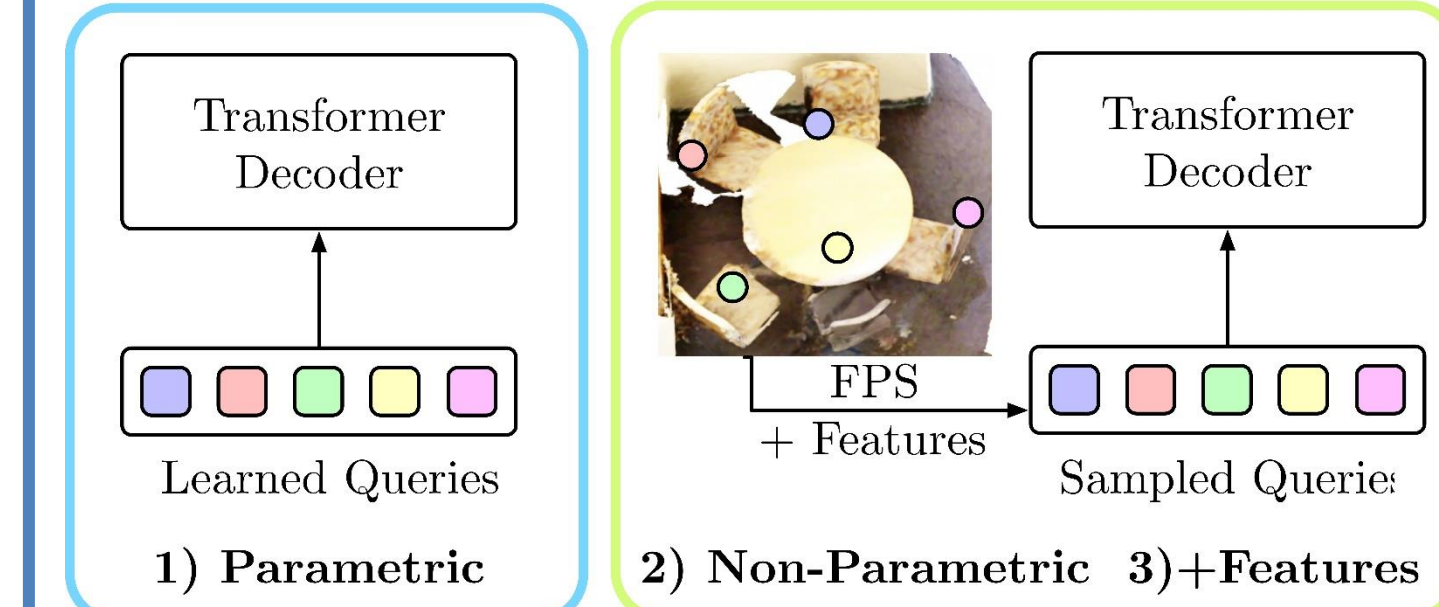
Our Paradigm



Architecture (based on Mask2Former)

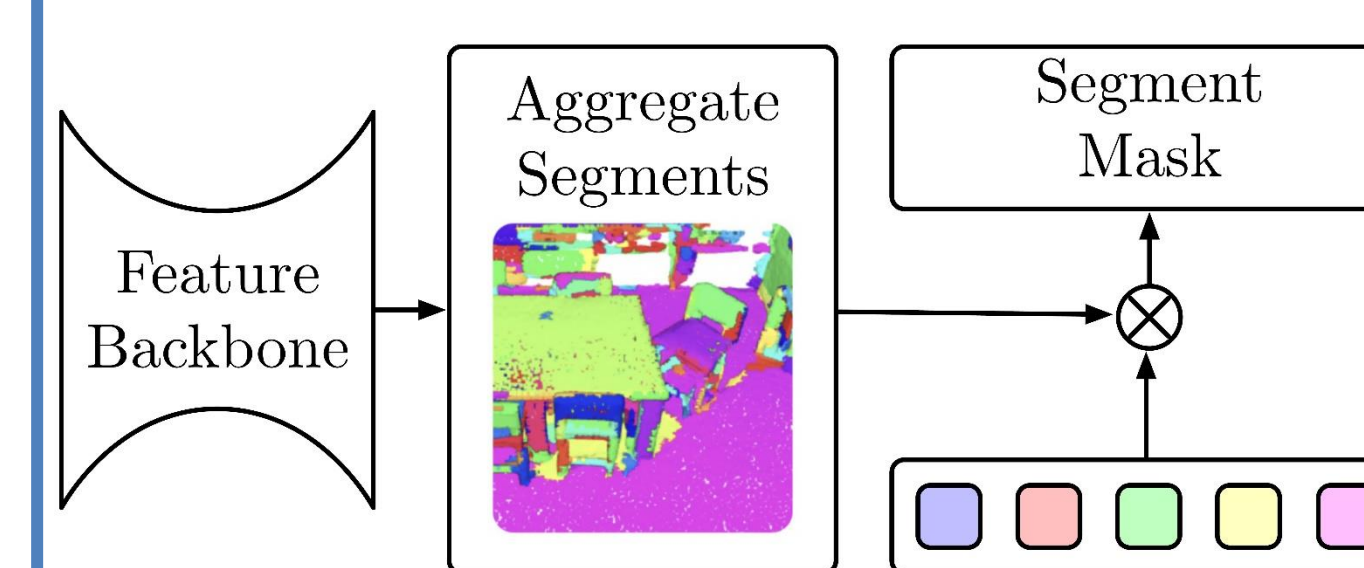


Query Types



Positions	Features	mAP	mAP ₅₀
1) Parametric	Parametric	45.1	63.3
2) Sampled	Zeros	46.8	65.1
3) Sampled	Features	45.5	65.6

Segment Aggregation



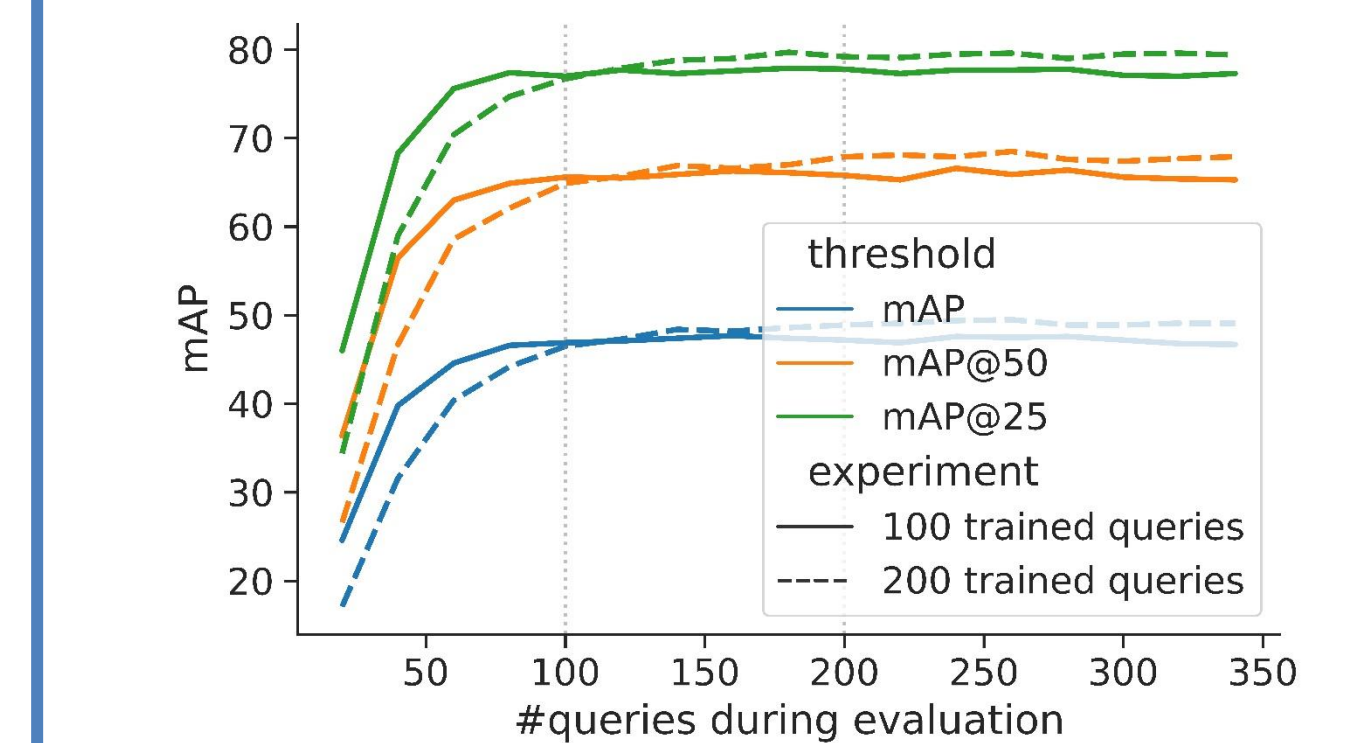
Aggreg.	Memory	mAP	mAP ₅₀
--	29.1 GB	47.7	67.2
Average	26.5 GB -2.6	46.8	65.1

Number Of Queries (Train)

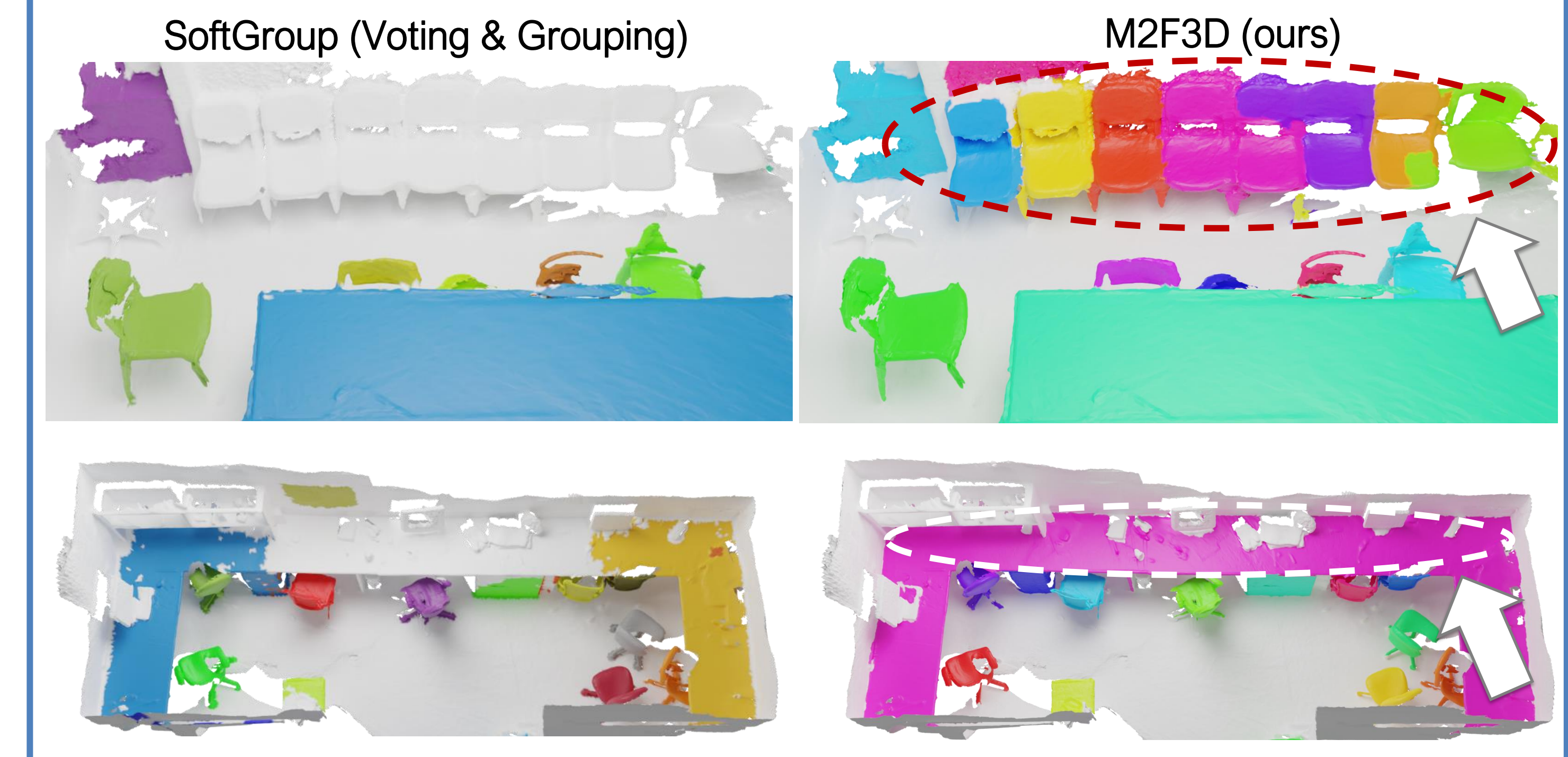
#Queries	mAP	mAP ₅₀	mAP ₂₅
100	46.8	65.1	76.8
200	48.2	66.5	79.1
300	???	???	???

OOM X More improvements expected

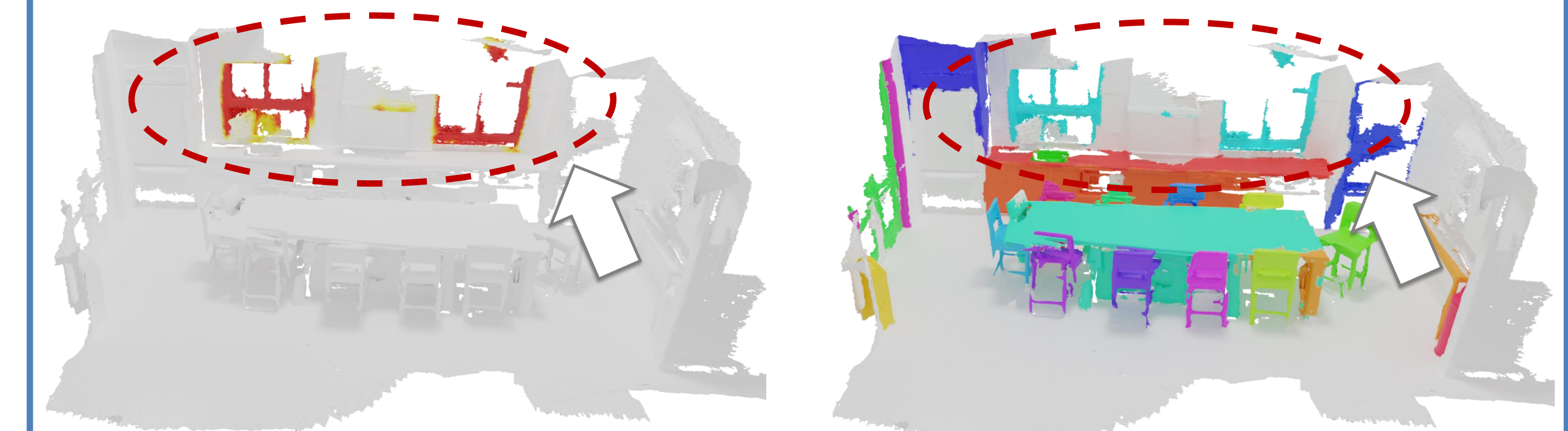
Number Of Queries (Test)



Qualitative Comparison to SoftGroup



Limitations (Merged Instances)



ScanNet Benchmark

Method	mAP	mAP ₅₀	mAP ₂₅	Runtime (in ms)
OccuSeg	48.6	67.2	74.2	1904
SSTNet	50.6	69.8	78.9	428
HAIS	45.7	69.9	80.3	339
SoftGroup	50.4 +5.7	76.1 +0.4	86.5	345
M2F3D (ours)	56.1	76.5	85.3	339