NEURIPS 2026

Relational Feature Distillation for Lightweight 3D Point Cloud Segmentation

TrimPT → TopoPT: 2.18× fewer parameters, on par with its 1200-epoch teacher

Mohammad Saeid · Amir Salarpour · Pedram MohajerAnsari · Mert D. Pesé

Project Page: m-saeid.github.io/TopoPT

Code and models: github.com/m-saeid/TopoPT



MOTIVATION

Efficient 3D segmentation needs small models that stay accurate

12.71M

LitePT-S teacher

25.42 encoder GFLOPs (analytic)

74.9

Compressed, no distillation

ScanNet val mIoU, 100 epochs

76.5

Official LitePT-S

ScanNet val mIoU, 1200 epochs

Can distillation close the gap with no added inference cost?

ARCHITECTURE

TrimPT narrows the attention stages

LitePT-S (teacher) · 12.71M params · 25.42 GFLOPs

Stage 0 · Conv	Stage 1 · Conv	Stage 2 · Conv	Stage 3 · Attn	Stage 4 · Attn
36 ch	72 ch	144 ch	252 ch	504 ch
2 blocks	2 blocks	2 blocks	6 blocks	2 blocks

▼ channels −25% to −29% in stages 1–4 · stage 3 depth 6 → 4 blocks · window 1024 kept

TrimPT (student) · 5.84M params · 12.95 GFLOPs

Stage 0 · Conv	Stage 1 · Conv	Stage 2 · Conv	Stage 3 · Attn	Stage 4 · Attn
36 ch	54 ch	108 ch	180 ch	360 ch
2 blocks	2 blocks	2 blocks	4 blocks	2 blocks

Brown numbers mark what changed. Stage layout and the 1024-token attention window are kept.

STEP 1 · TRIMPT

Shrink width and stage-3 depth, keep the 1024-token window

Compression axis	GFLOPs saved	Δ mIoU
Depth	−19%	+0.1
Width	−38%	−0.4
Depth + width (TrimPT)	−49%	−0.4
Patch size	−13%	−1.1
All three	−56%	−1.2

Deployed TrimPT

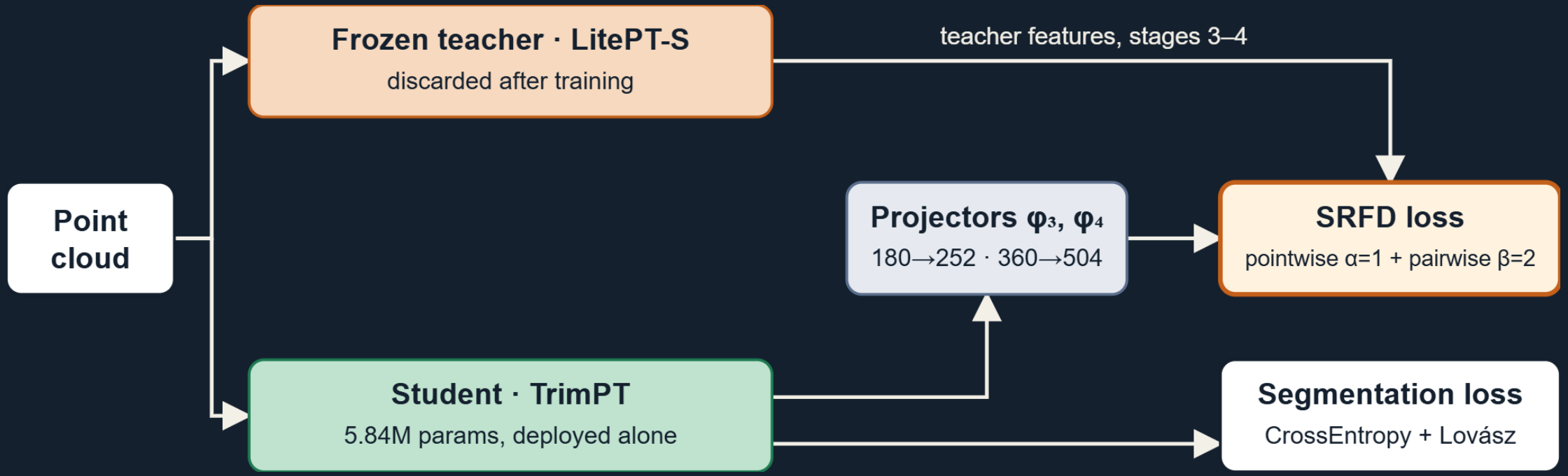
5.84M

2.18× fewer parameters; 12.95 GFLOPs (1.96× fewer, analytic); 30.75 vs 34.08 ms mean latency on an RTX 3090

ScanNet val, 100 epochs, Δ vs. reproduced LitePT-S (75.3 mIoU). GFLOPs: analytic, encoder only.

STEP 2 · SRFD

Stage-wise Relational Feature Distillation



$$L = L_{seg} + \alpha \cdot L_{pw} + \beta \cdot L_{rel}, \text{ with } \alpha = 1 \text{ and } \beta = 2$$

RESULTS · SCANNET VALIDATION

TopoPT is on par with the 1200-epoch teacher, with 2.18× fewer parameters

76.6

TopoPT mIoU

5.84M params, 100 epochs

+1.7

over TrimPT

74.9 without distillation

76.5

Official LitePT-S

12.71M params, 1200 epochs

Single runs. Baseline: batch 12 (TopoPT: 4). Pretrained teacher required.

BEYOND SCANNET SEMANTIC SEGMENTATION

Gains transfer, but vary widely across benchmarks

Benchmark	Metric	TopoPT	Comment
ScanNet instance	mAP50	63.9	7.61M model, 100 ep; teacher: 64.9
ScanNet200	mAP50	33.0	100 ep; teacher: 33.1
nuScenes	mIoU	81.4	+0.2 vs. TrimPT, 0.8 below LitePT-S
Structured3D	mIoU	70.3 (val)	+0.9 from SRFD; well below LitePT-S (83.6)

Single runs. Structured3D: our 50 epochs vs 200 for LitePT, so capacity and schedule cannot be separated.

TAKEAWAYS

Compress, distill relations, know the limits

Compress

Trim width and stage-3 depth: −49% analytic encoder FLOPs, 1024-token window kept

Distill relations

SRFD: +1.7 mIoU over TrimPT on ScanNet, no added inference cost

Know the limits

Single runs, a pretrained teacher is required, and gaps remain on Structured3D

Project Page: m-saeid.github.io/TopoPT · Thank you