TransforSeg: 3D Indoor Scene Long Tail Segmentation

李奕勳（R10943019）陳昱仁（R10943O117）陳欣妤（R10943017）謝郁楷（R11943025）

NTU DLCV-FALL-2022 INTRAINYEEFA

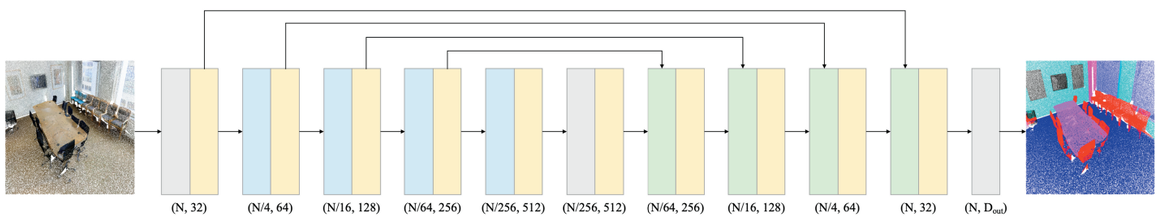
INTRODUCTION

Transformer-based method is often used to solve 3D point cloud segmentation problem since both of them are invariant to permutation and order of the input. To deal with imbalanced data distribution, weighted cross-entropy loss and ensemble leaning are adopted while training and testing respectively.

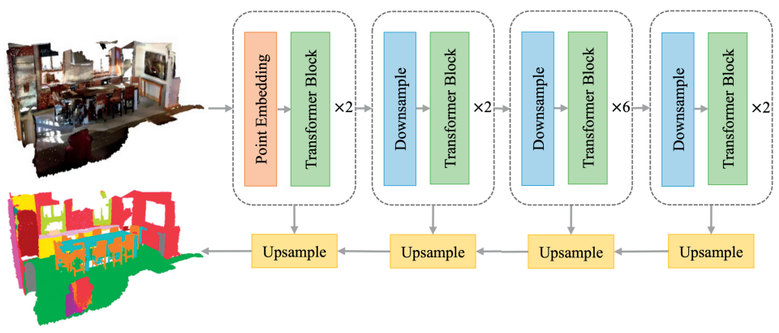
METHODOLOGY

I. MODEL

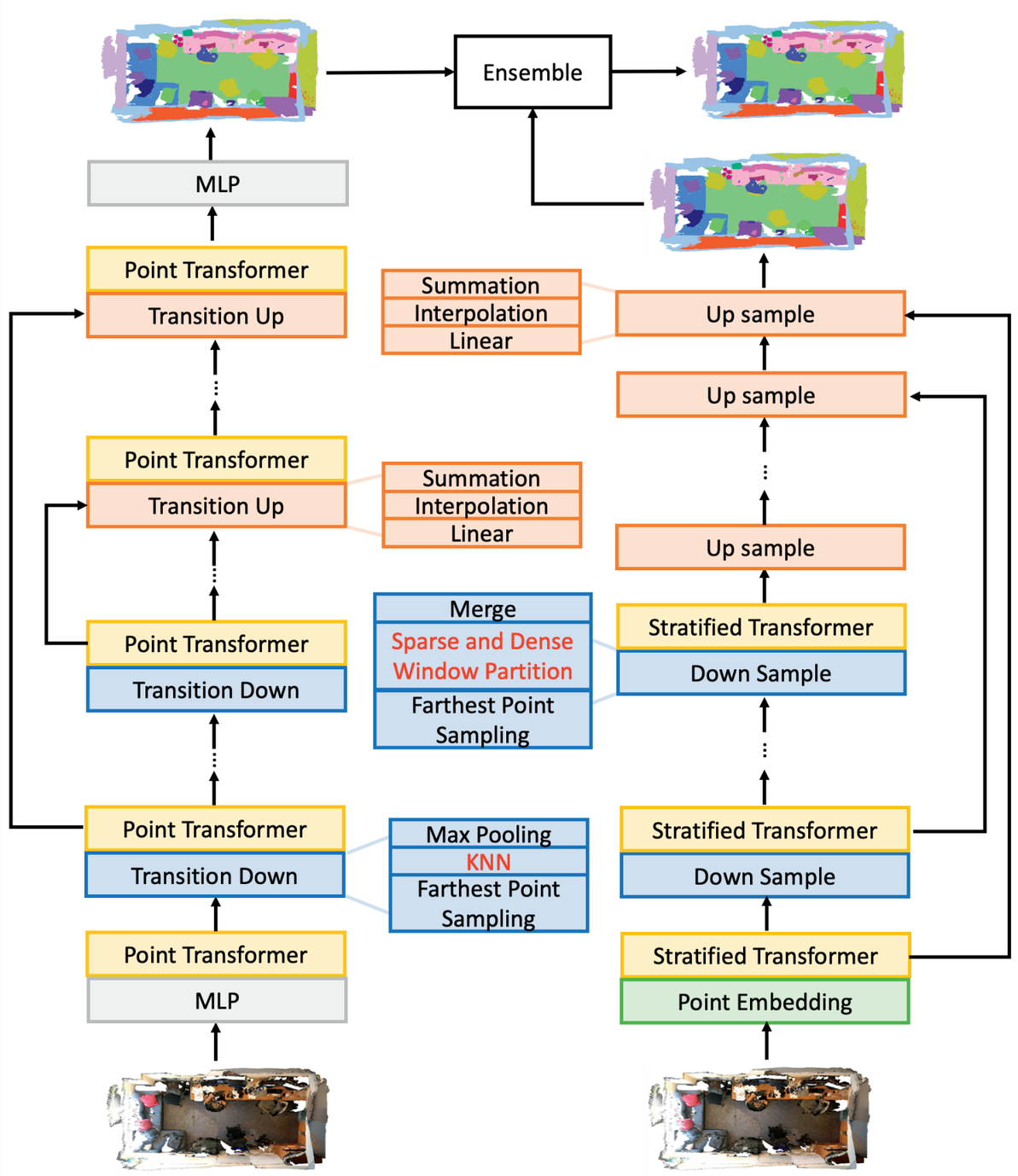
1. Point Transformer



2. Stratified Transformer



3. Our Model : TransforSeg

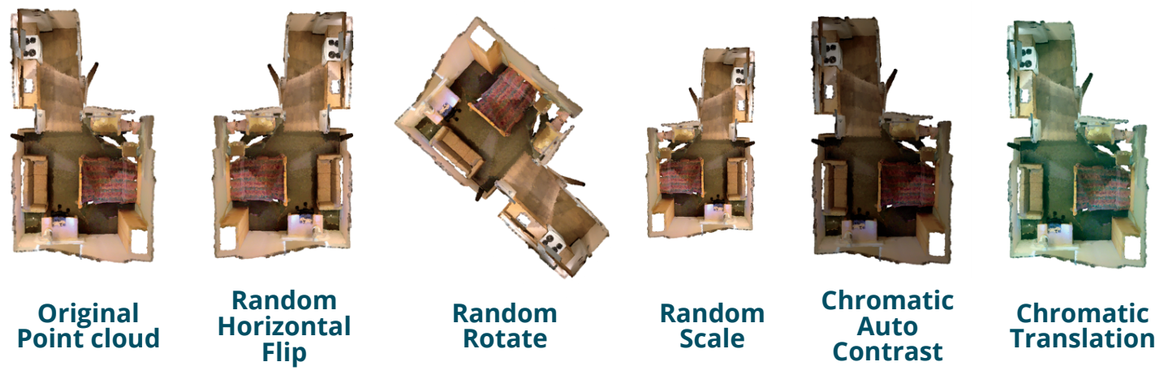


II. LOSS FUNCTION

Weighted (Class-balanced) Loss

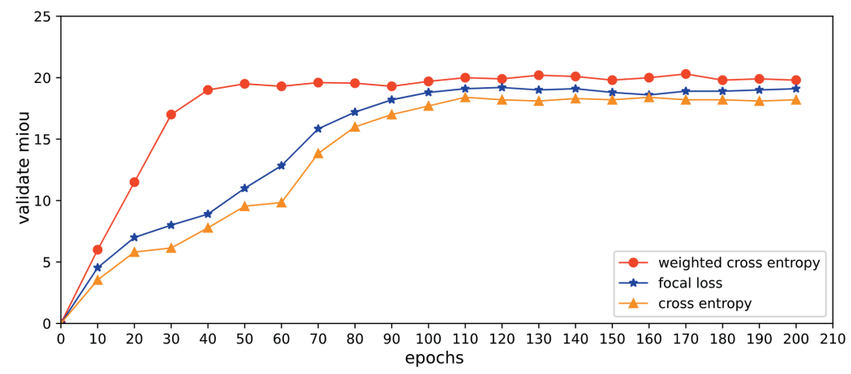
$$CB(\mathbf{P}, y) = \frac{1}{E^{n_y}}\mathcal{L}(\mathbf{P}, y) = \frac{1-\beta}{1-\beta^{\sqrt[3]{n_y}}}\mathcal{L}(\mathbf{P}, y)$$

III. DATA AUGMENTATION



RESULT

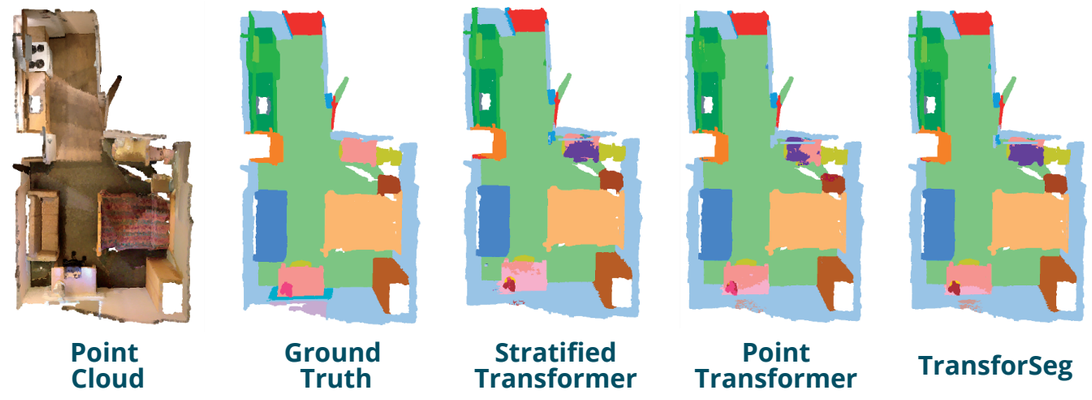
I. LOSS FUNCTION



II. EXPERIMENT (ABLATION STUDY)

Archit.	Loss	mIoU	head	common	tail
PointNet	cross entropy	4.76	11.42	2.15	0.78
PointNet + data aug.	cross entropy	8.12	18.78	3.65	2.04
PointNet++	cross entropy	14.38	30.58	8.39	4.35
Language Grounded	cross entropy	19.41	38.6	12.64	7.18
Language Grounded	focal	19.81	39.65	12.70	7.08
Language Grounded	weighted cross entropy	20.01	40.58	11.78	7.94
Point Transformer	cross entropy	20.03	40.17	12.7	8.32
Point Transformer	weighted cross entropy	**22.36**	41.90	**16.42**	**8.96**
Stratified Transformer	cross entropy	**22.41**	**43.56**	15.79	8.08
Ensemble (PT + ST)					

III. OUTPUT VISUALIZATION



CONCLUSION

Under transformer-based architecture, stratified transformer is conducive to the result of head, while point transformer preforms better on common and tail. We take both of their advantages via ensemble.

Also, we used class-balanced loss to accelerate training time and obtain a more balanced result.

BIBLIOGRAPHY

Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, Serge Belongie; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 9268-9277

Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip H.S. Torr, Vladlen Koltun; Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 16259-16268

Xin Lai, Jianhui Liu, Li Jiang, Liwei Wang, Hengshuang Zhao, Shu Liu, Xiaojuan Qi, Jiaya Jia; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 8500-8509