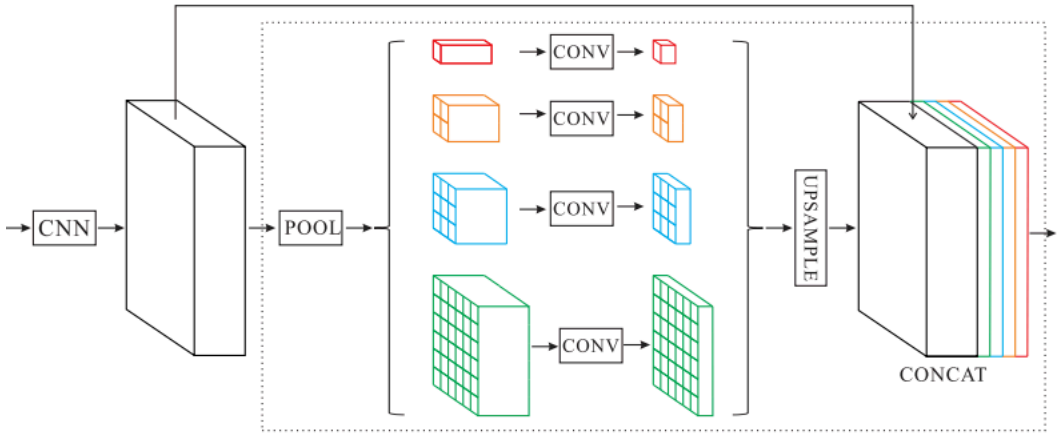
The 1st Solution for AutoNUE21 Challenge Semantic Segmentation Track

Liulei Li, Tianfei Zhou, Yi Liu, Lu Yang, Zeyu Chen, Wenguan Wang

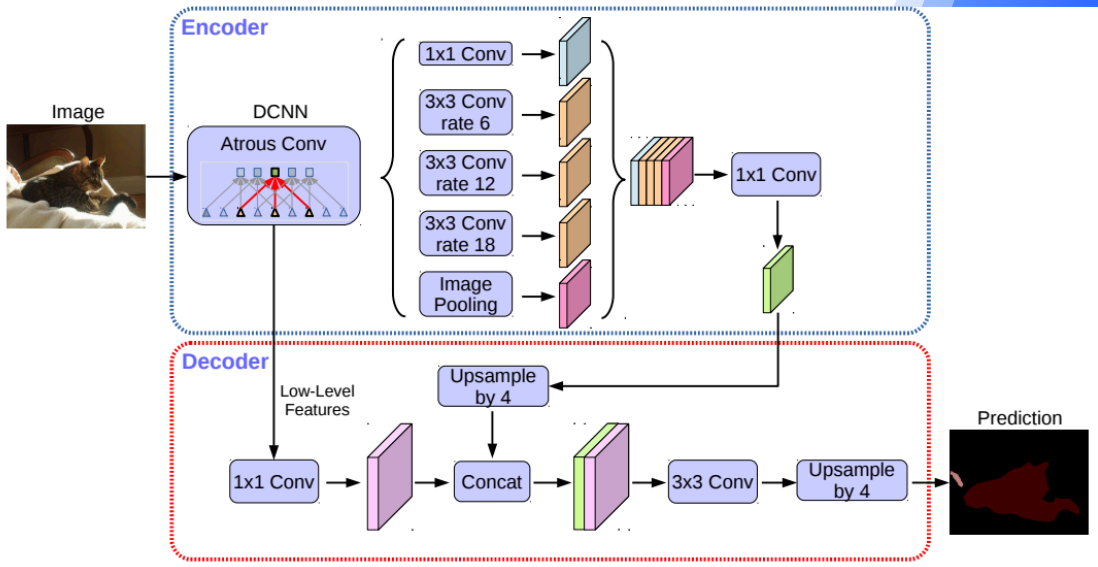
Work done while Liulei Li was interning at PaddleSeg, Baidu, Inc.

Preliminary

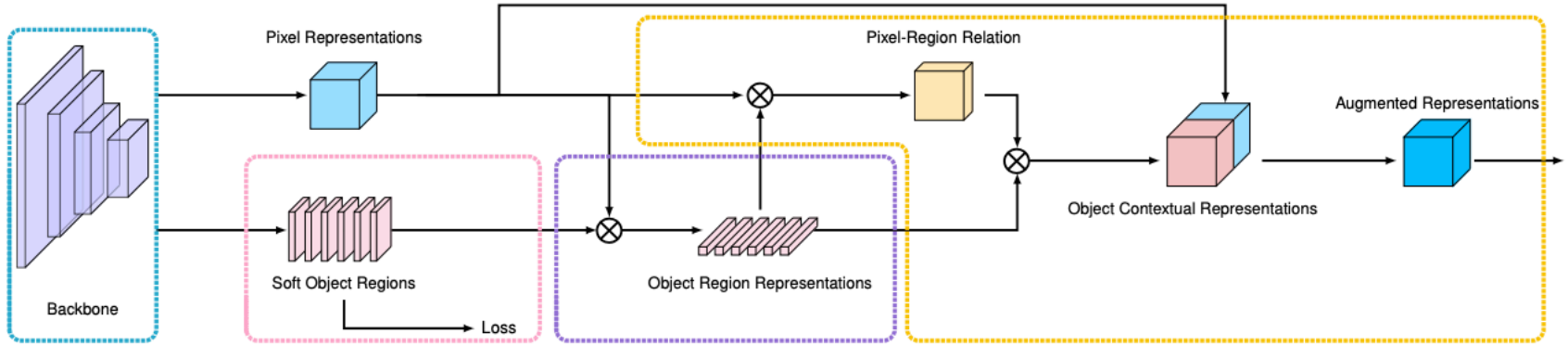
Encodr-Decoder Based CNNs for Semantic Segmentation



a) PSPNet[1]



b) DeepLabV3+[2]

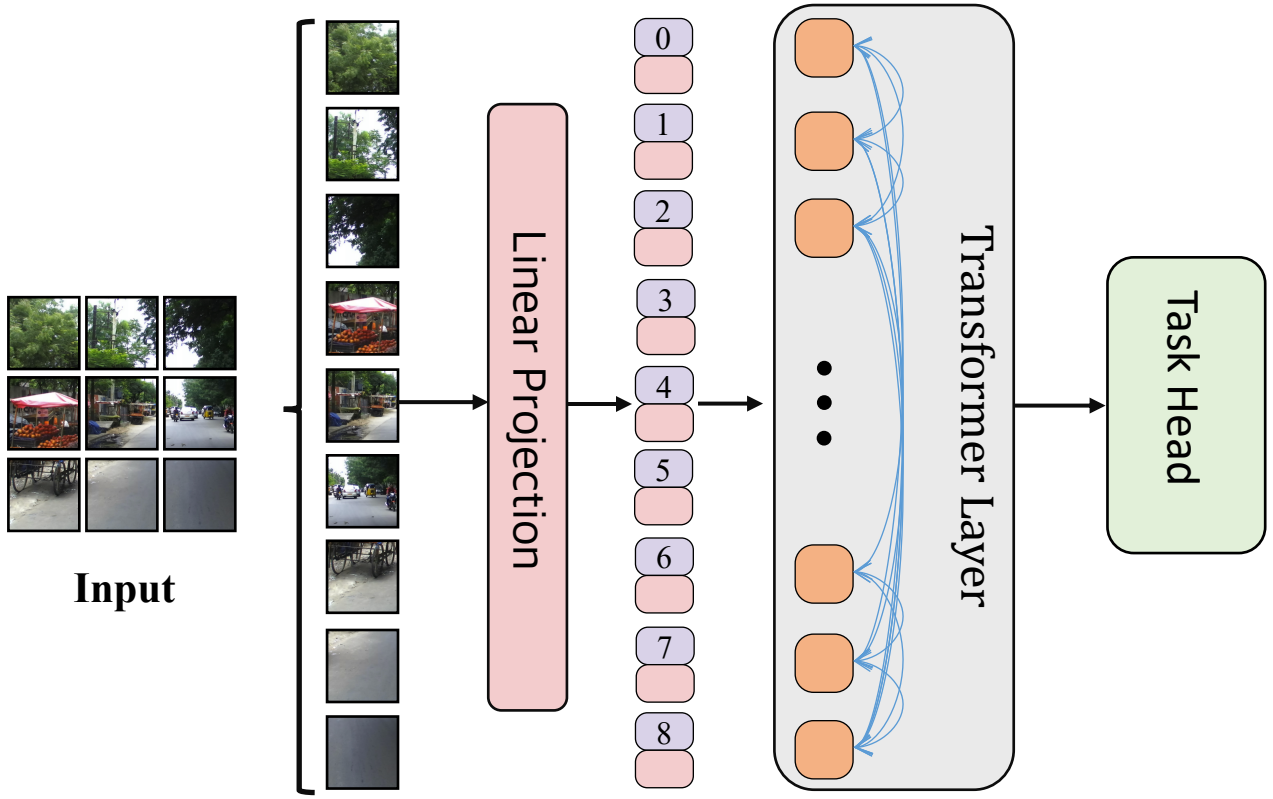


c) OCRNet[3]

1. Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2881-2890.
2. Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 801-818.
3. Yuan Y, Chen X, Wang J. Object-contextual representations for semantic segmentation[J]. arXiv preprint arXiv:1909.11065, 2019.

Motivation

Vision Transformer[1]

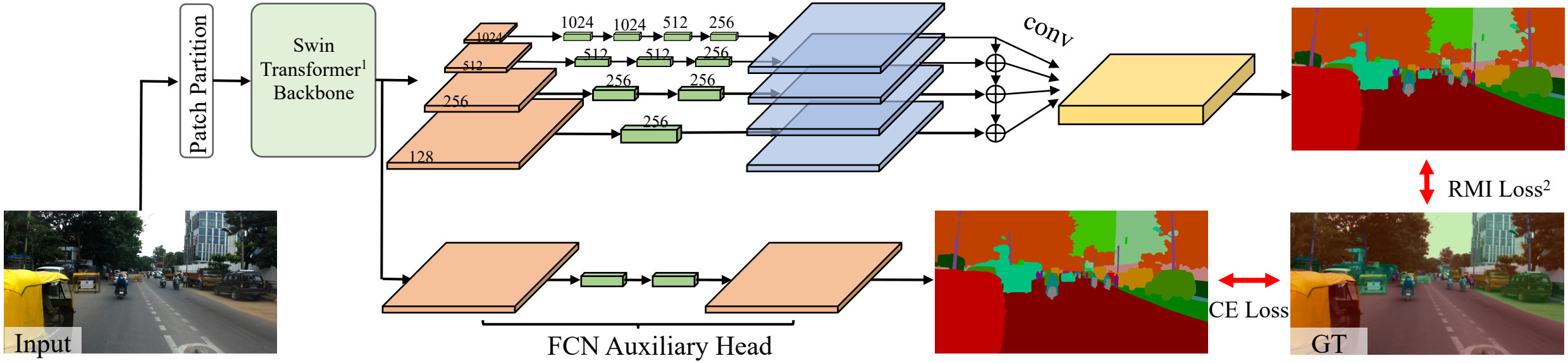


Interaction happens between every two patches

1. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.

Approach

Transformer-based Semantic Segmentation



1. Liu Z, Lin Y, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[J]. arXiv preprint arXiv:2103.14030, 2021.
2. Zhao S, Wang Y, Yang Z, et al. Region mutual information loss for semantic segmentation[J]. arXiv preprint arXiv:1910.12037, 2019.

Approach

Regional Progressive Segmentation

Densely Laid Objects

Images from the IDD Segmentation dataset

Examplar samples of CoCo-Stuff dataset

Approach

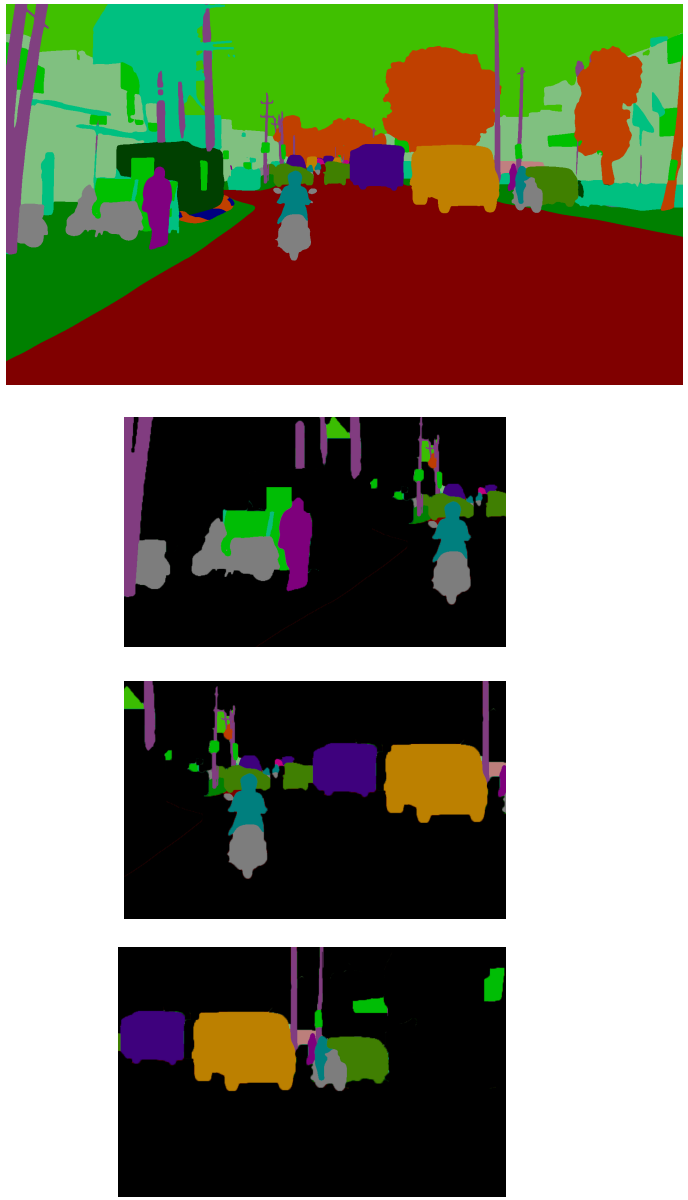
Regional Progressive Segmentation

Approach

Regional Progressive Segmentation

Thing: person, animal, rider, motorcycle, bicycle, auto rickshaw, car, truck, bus, caravan, trailer, vehicle fallback, billboard, traffic sign, traffic light, pole, pole group.

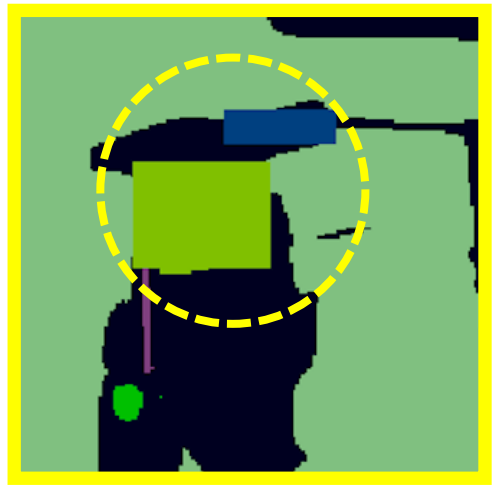
Stuff: road, parking, drivable fallback, sidewalk, non-drivable fallback, rail track, curb, wall, fence, guard rail, building, bridge, tunnel, vegetation, sky, fallback background.



Approach

Pseudo-labeling

Bounding-Box annotation of *IDD Detection Dataset*



Experiments

https://github.com/PaddlePaddle/PaddleSeg

Model	Method	mIoU (%)
Swin Transformer	UperNet	75.73
	Multi-Scale feature Aggre.	76.78(+1.05)
	MSD + pseudo-labeling	77.32(+0.54)
OCRNet	baseline	76.12
	Regional Progressive Seg	76.69(+0.57)
	RPS + pseudo-labeling	77.08(+0.39)
Ensemble	Swin + OCRNet	**78.28** (+2.16)

Table 1: Ablation studies on each component of our method. All of them are evaluated on the IDD Segmentation validation set.

Results

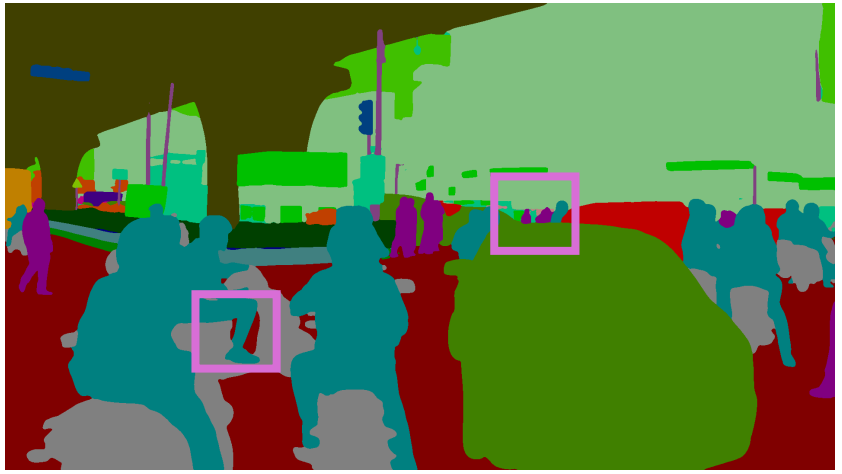
Team/Uploader Name	Method Name	mIoU, L3	mIoU, L2	mIoU, L1
PaddleSeg	PaddleSeg	0.7862	0.8046	0.9099
Anonymous	Anonymous	0.7845	0.8028	0.9081
Anonymous	Anonymous	0.7839	0.8022	0.9077
Anonymous	Anonymous	0.7748	0.7954	0.9022
Prabahkar	HR	0.769	0.7929	0.9044
Александр	SENet	0.767	0.7914	0.9035
Tsubasa	infomer_40	0.7655	0.7904	0.9022
OCRNet	final ocr	0.7649	0.7887	0.9006
lovasz loss	Lovasz	0.7637	0.7857	0.899
SKK.AL	HRNet	0.7621	0.786	0.898

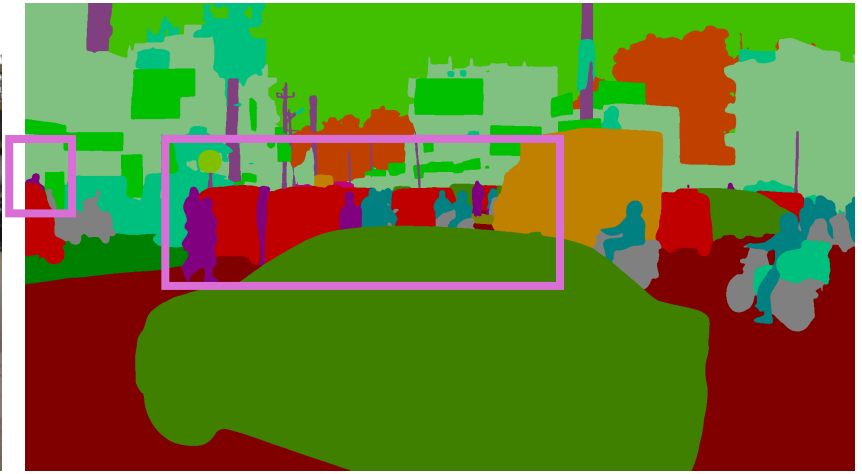
Table 2: Final leaderboard of Semantic Segmentation Track.

Qualitative Results

Prediction

Prediction + Image


Qualitative Results

Prediction

Prediction + Image






Thanks for your attention !

https://github.com/PaddlePaddle/PaddleSeg