MULTI-CLASS PART PARSING BASED ON MULTI-CLASS BOUNDARIES SUPPLEMENTARY MATERIAL

Anonymous ICIP submission

1. IMPLEMENTATION DETAILS

1.1. Dataset

The widely used PASCAL-Part [1] and the large-scale ADE20K-Part [2] datasets are used to train and evaluate the proposed method. PASCAL-Part includes PASCAL-Part-58, PASCAL-Part-108, and PASCAL-Person-Part. Both PASCAL-Part-58 and PASCAL-Part-108 contain 10103 images of varying sizes, along with 58 (PASCAL-Part-58) or 108 (PASCAL-Part-108) part-level annotations of 21 semantic object classes, including the background class. 4998 images are used for training and 5105 images for testing, following the original split in [1]. PASCAL-Person-Part contains 3533 images of multi-person on various scales and with 7 part-level annotations, including the background class. 1716 images are used for training and 1817 images for testing, following the original split in [3, 4, 5]. ADE20K-Part dataset contains 22210 images of different sizes, along with 544 part-level annotations of 150 object- and stuff-level classes as in [6]. 20210 images are used for training and 2000 images are used for testing, following the original split in [2]. Also, we follow the same evaluation metrics of the state-of-the-art and other well-known part parsing methods [5, 7, 8, 9] by using the mean Intersection over Union (mIoU); and applying the same evaluation strategy.

1.2. Training details

During training, the input images are cropped to 513×513 and randomly left-right flipped and scaled with a factor ranging from 0.5 to 2.0 times the original resolution. In the experiments, the learning rate is set to 0.05. The Stochastic gradient descent (SGD) optimizer with weight decay regularization $1e{-4}$ and momentum 0.9 is used. In all the models, the atrous rates of the ASPP are set to $(6, 12, 18)$ and the down-sampling stride is set to 16 as in prior works [10, 5, 7].

2. ADDITIONAL RESULTS ON PASCAL-PART AND ADE20K-PART DATASETS

Herein the performance of the proposed approach is further evaluated on PASCAL-Part-58, PASCAL-Part-108, PASCAL-Person-Part and ADE20K-Part benchmark datasets alongside DeepLab v3+ [10] and the published results from other multi-class part parsing methods [5, 7, 6, 9, 11].

2.1. PASCAL-Part-58

The segmentation performance of these methods is first compared based on the PASCAL-Part-58 benchmark. Despite the modest overall mIoU improvement achieved by incorporating multi-class boundaries, its impact becomes evident when analyzing performance at a more granular level, *i.e.*, per-class and per-part. Closer examination of the class-level segmentation in Table 1 further shows that the proposed model achieves the highest mIoU for 6 out of 21 categories (including background), matching the performance of GRPSNet, which achieved the highest overall mIoU. Additionally, AFPSNet+MCB improves the performance of the baseline, AFPSNet, in 11 categories, demonstrating its effectiveness across a wide range of object classes.

The segmentation performance of AFPSNet+MCB on the dataset, as shown in Table 1, demonstrates varying results across different object classes, highlighting its strengths and limitations. AFPSNet+MCB achieves the highest performance for birds, buses, cows, horses, persons, and sheep. These classes likely exhibit smaller, detailed part structures such as cow and horse tails, bird legs, and bus wheels, where the ability of AFPSNet+MCB to delineate boundaries separately for each part class enhances segmentation accuracy effectively. However, AFPSNet+MCB performs comparably to the baseline for classes such as cats, and shows lower performance for dogs, bikes, bottles, plants, and TVs. The variability in part appearance and positioning within these classes can limit the ability of the model to leverage multi-class boundaries effectively. Additionally, classes such as boats, chairs, sofas, and trains are presented in the dataset as single entities without distinct internal parts. The weighted loss function might not prioritise these large, less segmented classes, leading to reduced performance in these categories. This indicates that while AFPSNet+MCB excels in classes with small, detailed parts, its effectiveness decreases with variability in part appearance and larger classes presented as single entities.

In Table 2, a comparison is made of the per-part IoU on the 58 part classes achieved by AFPSNet+MCB, DeepLab v3+ and the published results from state-of-the-art multi-class methods [5, 7, 9, 11]. As can be seen, AFPSNet+MCB achieves the highest IoU for 19 out of 58 part classes, placing it second among all compared methods. AFPSNet+MCB shows especially superior performance in segmenting aeroplane body, bird leg, bus wheel, car light, cow head/tail, horse tail, etc., achieving more than 1.0% better than the first best method GRPSNet. AFPSNet+MCB demonstrates better performance in accurately delineating small parts across various classes, highlighting its efficacy in handling detailed structures.

2.2. PASCAL-Part-108

The performance of AFPSNet+MCB is further compared on the PASCAL-Part-108 benchmark. The proposed method is compared with the DeepLab v3+ [10] and 4 of the multi-class part parsing methods [5, 7, 9, 13] with the reported performances on the mean per-part IoU, as shown in Table 3. As can be seen, AFPSNet+MCB is 1.1% better than the baseline method. Further examination of the class-level segmentation results shows that the proposed model achieves the highest mIoU for 4 out of 21 categories and improves the performance of the baseline, AFPSNet, in 15 categories, demonstrating its effectiveness across a wide range of object classes.

The segmentation performance of AFPSNet+MCB on the dataset, as shown in Table 3, confirms the observations reported in Table 1 compared to the baseline. AFPSNet+MCB continues to demonstrate strong performance in accurately delineating small and detailed parts of objects, such as birds, buses, and horses. However, AFPSNet+MCB shows lower performance on classes such as sofas and TVs, which consist of single parts. This may be due to the weighted loss function, which may prioritise smaller, more segmented classes. Overall, this consistency across different datasets underscores the robustness and reliability of the proposed approach.

Fig. 1 shows qualitative results comparing AFPSNet+MCB with DeepLab v3+, GMNet, AFPSNet and GRPSNet. the proposed method shows overall better segmentation results with more details of object parts and more accurate boundaries. As can be seen, AFPSNet+MCB can better detect and segment the dog neck in the first column, the plant in the second column, the train in the fourth column and the tail of the small horse in the last column. Additionally, AFPSNet+MCB can better predict the boundaries of the horse legs in the third column and the dog ears in the fifth column.

In Table 5, a comparison is made of the per-part IoU on the 108 part classes achieved by AFPSNet+MCB and these methods. The results show that the proposed model achieves the highest mIoU for 44 out of 108 part classes (including background), placing it second among all compared methods. AFPSNet+MCB shows especially superior performance in segmenting bird beak/wing, bus mirror/light, cow tail, dog ear/neck, horse torso/neck, etc., achieving more than 2.0% better than the first best method GRPSNet. AFPSNet+MCB demonstrates better performance in accurately delineating small parts across various classes, highlighting its efficacy in handling detailed structures.

2.3. ADE20K-Part

Fig. 2, additionally compares the segmentation results of the proposed method, AFPSNet+MCB, with DeepLab v3+, AFPSNet and GRPSNet on the ADE20K-Part dataset. As can be seen, AFPSNet+MCB can better predict the boundaries of the house roof in the second row, the TV and the legs of the pool table in the fourth row, the aeroplane in the fifth row, the curtain in the sixth row, the building dome in the seventh row and the wardrobe door in the last row. Moreover, AFPSNet+MCB shows superior performance in localising parts. For example, the car door in the first row, the human head/legs in the third row and the small dome in the seventh row. The segmentation results of AFPSNet+MCB on this dataset validate the observations reported earlier in Table 1 and Table 3. AFPSNet+MCB demonstrate strong performance in accurately delineating small and detailed parts of objects.

2.4. PASCAL-Person-Part

The segmentation performance of AFPSNet+MCB is further compared with the DeepLab v3+ and the reported performances of the state-of-the-art multi-class part parsing methods [5, 7, 6, 9, 11], on PASCAL-Person-Part benchmark. The proposed approach, AFPSNet+MCB, achieved the second-highest per-part mIoU. Further analysis of the segmentation results for human parts in Table 4 shows that AFPSNet+MCB improves upon the baseline AFPSNet in 5 out of 7 categories, demonstrating its effectiveness across various part classes.

3. REFERENCES

[1] Xianjie Chen, Roozbeh Mottaghi, Xiaobai Liu, Sanja Fidler, Raquel Urtasun, and Alan Yuille, "Detect what you can: Detecting and representing objects using holistic models and body parts," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1971–1978.

[2] Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba, "Scene parsing through ade20k dataset," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 633–641.

[3] Liang-Chieh Chen, Yi Yang, Jiang Wang, Wei Xu, and Alan L Yuille, "Attention to scale: Scale-aware semantic image segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3640–3649.

[4] Fangting Xia, Peng Wang, Xianjie Chen, and Alan L Yuille, "Joint multi-person pose estimation and seman-

Table 1. Segmentation performance of mIoU on PASCAL-Part-58 benchmark. mIoU: per-part-class mIoU. Avg.: average per-object-class mIoU.

Method	backg	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	potted	sheep	sofa	train	tv	mIoU	Avg.
DeepLab v3+ [10]	94.5	46.4	65.2	53.6	63.7	51.5	67.1	51.6	62.6	38.5	52.6	45.2	58.6	66.5	72.5	56.5	55.4	52.1	46.0	80.2	61.0	57.6	59.1
BSANet [5]	91.6	50.0	65.7	54.8	60.2	49.2	70.1	53.5	63.8	36.5	52.8	43.7	58.3	66.0	71.6	58.4	55.0	49.6	43.1	82.2	61.4	58.2	58.9
GMNet [7]	92.7	46.7	66.4	52.0	70.0	55.7	71.1	52.2	63.2	51.4	54.8	51.3	59.6	64.4	**73.9**	56.2	56.2	53.6	**56.1**	85.0	65.6	59.0	61.8
GMENet [6]	92.6	46.5	66.6	52.2	**70.7**	55.8	71.6	52.7	63.8	**51.6**	55.5	**51.5**	59.9	64.8	73.7	57.2	56.5	54.2	55.8	**85.8**	**66.4**	59.6	*62.2*
CSR [8]	91.9	52.0	64.9	56.0	61.7	56.9	72.0	**56.9**	64.0	36.3	59.2	45.1	**62.3**	68.6	72.9	55.2	56.9	53.6	43.5	79.8	63.5	60.7	60.6
AFPSNet [9]	94.8	50.9	68.1	55.7	64.0	57.7	72.0	55.7	65.1	39.8	60.7	44.6	61.9	70.4	72.8	61.4	58.3	57.0	46.4	81.6	63.1	61.3	62.0
GRPSNet [11]	**94.9**	**52.4**	**68.3**	55.7	61.5	**58.9**	71.8	55.8	**65.3**	37.6	60.3	45.5	62.2	71.1	73.7	61.7	**60.2**	57.5	48.6	79.3	63.1	**61.6**	**62.2**
AFPSNet+MCB	94.7	51.2	67.8	**56.1**	60.4	56.6	**72.1**	56.2	65.1	38.2	**62.1**	45.1	61.3	**72.2**	73.2	**61.9**	57.2	**57.8**	44.8	81.3	62.0	61.4	61.8

Table 2. Segmentation performance per-part IoU on the 58 part classes of PASCAL-Part-58 dataset.

Parts name	DeepLab v3	DeepLab v3+	BSANet	GMNet	AFPSNet	GRPSNet	AFPSNet+MCB	Parts name	DeepLab v3	DeepLab v3+	BSANet	GMNet	AFPSNet	GRPSNet	AFPSNet+MCB
	IoU	IoU	IoU	IoU	IoU	IoU	IoU		IoU	IoU	IoU	IoU	IoU	IoU	IoU
background	91.1	90.2	91.6	92.7	94.8	**94.9**	94.7	cow tail	0.0	1.0	7.9	8.1	21.4	20.9	**25.7**
aeroplane body	66.6	68.4	70.0	69.6	69.7	70.7	**72.4**	cow leg	46.1	55.1	53.4	53.3	59.2	59.3	**59.9**
aeroplane engine	25.7	27.8	29.1	25.7	31.0	**32.0**	31.3	cow torso	69.9	74.0	73.5	77.1	78.3	78.0	**78.5**
aeroplane wing	33.5	38.3	38.3	34.2	**42.3**	**42.3**	40.1	dining table	43.0	43.1	43.7	**51.3**	44.6	45.5	45.1
aeroplane stern	57.1	52.6	59.2	57.2	60.5	**61.0**	60.1	dog head	78.7	81.7	82.5	**85.0**	84.5	84.7	84.1
aeroplane wheel	45.4	50.5	53.2	46.8	51.1	**56.2**	52.0	dog leg	48.1	50.8	53.8	53.8	56.2	**56.3**	56.1
bike wheel	78.0	75.7	78.0	**81.3**	79.8	80.3	80.0	dog tail	27.1	32.6	31.3	31.4	39.3	**39.9**	37.3
bike body	48.4	52.2	53.4	51.5	**56.3**	56.2	55.6	dog torso	63.7	62.9	65.7	**68.0**	67.5	**68.0**	67.6
bird head	64.6	71.8	74.0	71.1	72.5	74.1	**74.7**	horse head	74.7	75.4	76.6	73.9	82.1	83.4	**84.2**
bird wing	35.1	38.3	39.7	38.6	**44.5**	**44.5**	42.6	horse tail	47.0	47.2	51.0	50.4	57.2	57.1	**60.7**
bird leg	29.3	34.1	34.8	28.7	35.9	34.0	**36.8**	horse leg	55.9	62.3	61.6	59.3	63.9	**65.3**	64.8
bird torso	66.9	66.8	**70.9**	69.5	70.2	70.0	70.4	horse torso	70.3	72.8	74.9	73.9	78.4	78.4	**78.9**
boat	54.4	64.0	60.2	**70.0**	64.0	61.5	60.4	mbike wheel	70.9	69.9	71.6	73.5	73.0	**74.4**	73.5
bottle cap	30.7	28.9	29.8	33.9	39.6	**42.2**	40.3	mbike body	65.1	71.5	71.5	**74.3**	72.6	73.0	72.8
bottle body	68.8	70.5	68.6	**77.6**	75.8	75.6	72.8	person head	83.5	84.8	85.0	84.7	86.2	86.5	**86.8**
bus window	72.7	74.5	74.8	75.4	**78.5**	77.9	77.5	person torso	65.9	65.9	68.2	67.0	71.3	71.5	**71.6**
bus wheel	55.3	55.5	57.1	58.1	58.2	57.0	**58.3**	person larm	46.9	48.7	52.0	48.6	55.7	**56.8**	56.4
bus body	74.8	77.6	78.3	79.9	79.6	80.4	**80.5**	person uarm	51.5	48.6	54.4	52.4	58.9	**59.7**	59.6
car window	62.6	66.7	68.1	64.8	**71.2**	**71.2**	**71.2**	person lleg	38.6	39.4	43.5	40.2	46.0	46.2	**46.9**
car wheel	64.8	**72.1**	68.5	70.3	71.9	70.7	71.8	person uleg	43.8	44.5	47.4	44.5	**50.3**	49.2	50.0
car light	46.2	53.5	53.7	48.4	57.6	58.6	**59.8**	pplant pot	45.3	50.0	53.5	56.0	57.3	**59.9**	55.4
car plate	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	pplant plant	52.4	59.9	56.6	56.4	59.3	**60.5**	58.9
car body	72.1	76.2	77.0	77.6	78.0	**78.3**	78.0	sheep head	60.9	70.8	65.4	70.8	72.1	73.6	**73.8**
cat head	80.2	82.3	83.7	83.8	**85.0**	84.9	84.1	sheep leg	8.6	19.3	11.7	14.3	25.4	**25.6**	25.2
cat leg	48.6	47.3	50.1	49.4	**53.2**	52.2	52.6	sheep torso	68.3	73.0	71.6	**75.6**	73.5	73.2	74.3
cat tail	40.2	45.9	48.8	46.0	49.1	**50.2**	50.1	sofa	43.2	42.4	43.1	**56.1**	46.4	48.6	44.8
cat torso	70.3	69.6	72.6	73.8	73.0	**73.9**	73.8	train	79.6	82.6	82.2	**85.0**	81.6	79.3	81.3
chair	35.4	38.2	36.5	**51.4**	39.8	37.6	38.2	tv screen	69.5	69.1	73.1	**77.0**	74.0	73.9	71.7
cow head	74.3	77.2	76.4	80.7	83.7	83.0	**84.2**	tv frame	45.9	46.3	49.8	**54.1**	52.1	52.2	52.3

Table 3. Segmentation performance of mIoU on PASCAL-Part-108 benchmark. mIoU: per-part-class mIoU. Avg.: average per-object-class mIoU.

Method	backg	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	potted	sheep	sofa	train	tv	mIoU	Avg.
DeepLab v3 [12]	90.9	41.9	44.5	35.3	53.7	47.0	34.1	42.3	49.2	35.4	39.8	33.0	48.2	48.8	23.2	50.4	43.6	35.4	39.2	20.7	60.8	41.3	43.7
DeepLab v3+ [10]	94.5	48.8	45.4	41.6	59.5	49.5	36.5	45.3	51.3	37.3	50.9	44.1	52.0	54.5	23.9	55.8	54.0	42.6	47.4	23.3	69.7	46.5	48.9
BSANet [5]	91.6	45.3	40.9	41.0	61.4	48.9	32.2	43.3	50.7	34.1	39.4	45.9	52.1	50.0	23.1	52.4	50.6	37.8	44.5	20.7	66.3	42.9	46.3
GMNet [7]	92.7	48.0	46.2	39.3	69.2	56.0	37.0	45.3	52.6	**49.1**	50.6	**60.6**	52.0	51.5	24.8	52.6	56.0	40.1	**53.9**	21.6	70.7	45.8	50.5
GMENet [6]	92.9	48.9	47.3	40.2	**69.6**	55.3	37.8	46.7	53.3	48.4	51.8	50.1	52.3	51.1	27.4	54.2	57.8	41.5	53.4	24.3	70.3	46.3	51.2
AFPSNet [9]	94.9	50.4	**52.0**	43.8	61.1	52.1	41.1	48.9	54.0	38.0	54.5	43.0	55.0	57.7	**25.4**	58.5	57.2	44.5	47.2	23.1	**73.1**	49.2	51.2
GRPSNet [11]	**95.0**	**51.5**	51.5	**46.3**	61.6	**57.2**	**44.2**	**50.0**	**55.3**	40.0	**56.4**	46.4	55.7	58.9	**25.4**	**59.9**	56.8	**45.7**	47.0	23.7	70.9	**50.5**	**52.4**
AFPSNet+MCB	94.9	51.3	50.6	45.8	61.7	54.0	43.1	49.7	54.9	37.3	55.6	46.3	**56.0**	**59.9**	24.8	59.5	**60.7**	45.2	46.8	**25.5**	72.2	50.3	52.2

Table 4. Segmentation performance of mIoU on Pascal-Person-Part benchmark. mIoU: per-part-class mIoU.

Method	backg	head	torso	u-arms	l-arms	u-legs	l-legs	mIoU
DeepLab v3 [12]	94.79	84.06	66.69	54.26	52.80	48.08	43.59	63.5
DeepLab v3+ [10]	97.12	87.00	70.91	59.69	59.54	52.96	49.42	68.1
BSANet-101 [5]	95.62	86.49	70.20	59.31	58.72	51.91	49.32	67.4
BSANet-152 [5]	95.79	86.98	71.35	61.36	60.26	53.28	49.95	68.4
GMNet [7]	-	-	-	-	-	-	-	67.5
GMENet [6]	-	-	-	-	-	-	-	68.4
AFPSNet [9]	97.28	87.60	72.68	62.07	61.48	54.59	51.22	69.6
GRPSNet [11]	**97.30**	**87.83**	**72.72**	**63.40**	**62.84**	**55.23**	**51.43**	**70.1**
AFPSNet+MCB	97.20	87.63	72.58	63.00	62.18	55.02	**51.43**	69.9

tic part segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 6769–6778.

[5] Yifan Zhao, Jia Li, Yu Zhang, and Yonghong Tian, "Multi-class part parsing with joint boundary-semantic awareness," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 9177–9186.

[6] Umberto Michieli and Pietro Zanuttigh, "Edge-aware graph matching network for part-based semantic segmentation," *International Journal of Computer Vision*, vol. 130, no. 11, pp. 2797–2821, 2022.

[7] Umberto Michieli, Edoardo Borsato, Luca Rossi, and Pietro Zanuttigh, "Gmnet: Graph matching network for large scale part semantic segmentation in the wild," in *European Conference on Computer Vision*. Springer, 2020, pp. 397–414.

[8] Xin Tan, Jiachen Xu, Zhou Ye, Jinkun Hao, and Lizhuang Ma, "Confident semantic ranking loss for part parsing," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2021, pp. 1–6.

[9] Njuod Alsudays, Jing Wu, Yu-Kun Lai, and Ze Ji, "Afpsnet: Multi-class part parsing based on scaled attention and feature fusion," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 4033–4042.

[10] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818.

[11] Njuod Alsudays, Jing Wu, Yu-Kun Lai, and Ze Ji, "Grpsnet: Multi-class part parsing based on graph reasoning," in *2024 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2024.

[12] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587*, 2017.

[13] Hossein Azizpour and Ivan Laptev, "Object detection using strongly-supervised deformable part models," in *European Conference on Computer Vision*. Springer, 2012, pp. 836–849.

Fig. 1. Segmentation results on PASCAL-Part-108 dataset. AFPSNet+MCB generates notable results by achieving better part localisation and more accurate prediction of small parts compared to the other models.

Table 5. Segmentation performance per-part IoU on the 108 part classes of PASCAL-Part-108 dataset.

Parts name	DeepLab v3	DeepLab v3+	BSANet	GMNet	AFPSNet	GRPSNet	AFPSNet+MCB	Parts name	DeepLab v3	DeepLab v3+	BSANet	GMNet	AFPSNet	GRPSNet	AFPSNet+MCB
	IoU	IoU	IoU	IoU	IoU	IoU	IoU		IoU	IoU	IoU	IoU	IoU	IoU	IoU
background	90.0	94.5	91.6	92.7	94.9	**95.0**	94.9	dining table	33.0	44.1	45.9	**50.6**	43.0	46.4	46.3
aero body	61.9	68.1	68.2	61.9	69.5	71.3	**71.9**	dog head	60.5	63.1	63.8	64.0	64.9	65.7	**65.8**
aero stern	53.2	59.5	54.2	57.4	**61.2**	59.8	60.4	dog reye	50.1	50.2	54.1	54.7	58.5	**61.4**	**61.4**
aero rwing	28.9	38.3	33.1	34.3	40.6	**42.9**	40.0	dog rear	54.0	58.0	57.2	56.8	60.6	60.3	**62.8**
aero engine	24.7	27.0	26.5	27.2	29.3	29.3	**30.9**	dog nose	63.5	68.2	66.3	66.0	70.1	71.4	**72.4**
aero wheel	40.9	51.3	44.5	51.5	51.3	**54.3**	53.4	dog torso	58.4	61.0	62.3	63.2	62.2	**64.1**	64.0
bike fwheel	78.4	79.1	75.3	80.2	80.5	80.0	**81.4**	dog neck	27.1	27.8	26.2	28.1	30.7	28.5	**32.4**
bike saddle	34.1	36.0	31.0	38.0	**42.6**	42.2	40.9	dog rfleg	39.2	43.1	42.4	43.7	44.9	45.0	**45.4**
bike handlebar	23.3	22.1	20.6	22.4	**33.6**	32.9	30.8	dog rfpaw	39.4	44.1	44.2	43.7	46.4	**48.8**	46.2
bike chainwhell	42.3	44.5	36.5	44.1	**51.1**	50.9	49.3	dog tail	24.7	35.8	34.9	30.8	40.1	**40.3**	39.2
birds head	51.5	67.9	66.4	65.3	68.8	**69.3**	68.8	dog muzzle	65.1	68.5	69.4	68.9	71.1	**71.5**	70.6
birds beak	40.4	51.9	47.1	44.3	58.3	60.8	**64.3**	horse head	54.4	64.6	57.1	55.9	**68.5**	67.4	68.1
birds torso	61.7	62.7	65.2	64.8	65.3	65.0	**65.7**	horse rear	49.7	56.1	51.1	52.2	60.3	62.2	**62.8**
birds neck	27.5	38.1	39.1	28.4	36.1	**39.6**	37.6	horse muzzle	61.3	69.4	65.2	62.9	72.3	72.2	**72.8**
birds rwing	35.9	40.1	39.3	37.2	41.3	41.2	**43.3**	horse torso	56.7	62.2	59.5	60.7	65.1	64.9	**66.6**
birds rleg	23.5	26.0	26.5	23.8	27.8	**33.3**	31.8	horse neck	42.1	53.3	49.6	47.2	55.2	53.9	**57.9**
birds rfoot	13.9	13.8	11.6	17.7	18.3	**21.2**	20.0	horse rfuleg	54.1	60.1	57.0	56.4	62.0	63.2	**63.4**
birds tail	28.1	32.2	33.0	32.5	34.7	**39.7**	34.7	horse tail	48.1	53.4	47.6	51.4	56.6	**59.4**	59.0
boat	53.7	59.5	61.4	**69.2**	61.1	61.6	61.7	horse rfho	24.1	17.2	12.9	25.3	21.9	28.2	**28.9**
bottle cap	30.4	31.9	26.2	33.4	35.8	**39.8**	38.4	mbike fwheel	69.6	72.0	69.3	73.6	73.3	**75.0**	73.7
bottle body	63.7	67.1	71.5	**78.7**	68.3	74.6	69.5	mbike hbar	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
bus rightside	70.8	74.8	73.0	75.7	**77.6**	77.4	76.4	mbike saddle	0.0	0.0	0.0	**0.8**	0.0	0.0	0.0
bus roofside	7.5	13.9	0.3	13.5	15.4	**22.9**	18.8	mbike hlight	25.8	23.7	10.6	**28.5**	28.1	26.4	25.5
bus mirror	2.1	8.6	0.3	6.6	15.4	19.2	**21.3**	person head	68.2	72.8	69.7	69.3	74.1	73.8	**74.3**
bus fliplate	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	person reye	35.1	45.2	41.3	38.7	47.9	50.0	**51.2**
bus door	40.1	43.2	37.2	38.1	49.1	**52.0**	44.0	person rear	37.4	48.8	41.9	41.4	52.9	**55.3**	54.6
bus wheel	54.8	49.1	53.1	56.7	57.6	**59.5**	58.5	person nose	53.0	57.8	54.3	56.7	62.2	**66.0**	65.7
bus headlight	25.6	27.2	19.9	30.4	35.7	44.9	**48.6**	person mouth	48.9	54.1	49.5	51.3	56.3	**60.1**	57.9
bus window	71.8	75.2	73.5	74.6	**78.2**	77.6	76.8	person hair	70.8	73.2	72.3	71.8	74.9	75.2	**75.3**
car rightside	64.0	68.8	67.9	70.5	72.4	**72.5**	71.7	person torso	63.4	67.6	64.3	65.2	69.9	**70.2**	69.8
car roofside	21.0	15.8	16.1	22.3	19.3	**25.5**	23.7	person neck	49.7	53.2	50.9	51.2	55.1	55.0	**55.3**
car fliplate	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	person ruarm	54.7	61.1	55.7	57.4	63.8	**63.9**	**63.9**
car door	41.4	45.1	39.6	42.3	49.6	49.7	**51.0**	person rhand	43.0	48.9	47.4	44.1	52.2	53.2	**53.7**
car wheel	65.8	67.8	64.0	70.2	71.7	**73.4**	71.3	person ruleg	50.8	55.1	52.3	53.0	57.0	**57.5**	57.0
car headlight	42.9	51.1	49.4	46.4	57.7	57.4	**58.5**	person rfoot	29.8	31.8	28.9	31.3	35.2	**38.1**	34.8
car window	61.0	68.8	66.5	65.0	**71.9**	71.6	71.4	pplant pot	43.6	52.8	50.6	56.0	56.3	56.1	**60.5**
cat head	73.9	76.7	75.6	77.5	77.7	**78.5**	77.7	pplant plant	42.9	55.2	55.5	56.6	58.1	57.4	**60.8**
cat reye	58.8	57.1	62.0	62.8	67.3	**68.7**	68.6	sheep head	45.6	51.2	47.0	**54.0**	52.9	52.9	52.3
cat rear	65.5	67.7	66.8	67.1	70.7	**71.5**	71.3	sheep rear	43.2	50.6	47.7	45.3	54.1	56.3	**56.5**
cat nose	40.3	39.2	41.2	46.3	46.9	**52.0**	49.4	sheep muzzle	58.2	62.6	61.1	64.9	65.1	**66.9**	63.7
cat torso	64.2	67.0	66.8	68.7	67.9	**68.8**	68.6	sheep rhorn	3.0	46.9	0.0	5.4	44.4	**54.1**	44.3
cat neck	22.8	23.9	19.8	24.4	24.0	23.0	**25.1**	sheep torso	62.6	65.0	66.4	68.8	68.5	**69.0**	67.1
cat rfleg	36.5	39.6	38.5	39.1	41.3	**41.7**	41.1	sheep neck	26.9	**34.5**	25.3	30.3	33.6	31.7	33.3
cat rfpaw	40.6	42.5	43.4	41.7	43.0	**44.9**	43.9	sheep rfuleg	8.6	20.6	17.4	11.7	**21.1**	16.3	19.7
cat tail	40.2	47.9	42.6	45.8	47.0	**48.9**	48.6	sheep tail	6.7	9.5	1.1	9.1	15.9	18.4	**24.4**
chair	35.4	37.3	34.1	**49.1**	38.0	40.0	37.3	sofa	39.2	**47.4**	44.5	53.9	47.2	47.0	46.8
cow head	51.2	66.1	58.2	63.8	66.0	**68.3**	66.0	train head	5.3	4.7	5.6	4.5	5.6	**8.1**	6.7
cow rear	51.2	63.9	53.0	60.0	61.7	64.1	**65.4**	train hrightside	61.9	63.9	63.5	60.8	64.0	62.5	**65.9**
cow muzzle	61.2	71.9	67.2	**74.9**	73.9	74.3	72.5	train hroofside	23.0	22.6	13.7	21.1	22.0	22.9	**27.4**
cow rhorn	28.8	44.7	10.1	44.0	57.6	**59.0**	53.8	train headlight	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
cow torso	63.4	72.9	69.9	73.2	75.1	75.8	**76.2**	train coach	28.6	35.2	**42.0**	31.4	36.9	35.2	38.7
cow neck	9.5	19.9	7.3	20.3	26.1	**27.7**	26.4	train crightside	15.6	16.2	19.0	14.9	18.1	**20.9**	17.4
cow rfuleg	46.5	53.8	49.7	54.8	57.8	58.4	**58.5**	train croofside	10.8	20.2	1.0	18.1	15.1	16.4	**22.7**
cow tail	6.5	13.6	0.1	13.6	17.6	23.6	**26.0**	tv screen	60.8	69.7	66.3	70.7	**73.1**	70.9	72.2

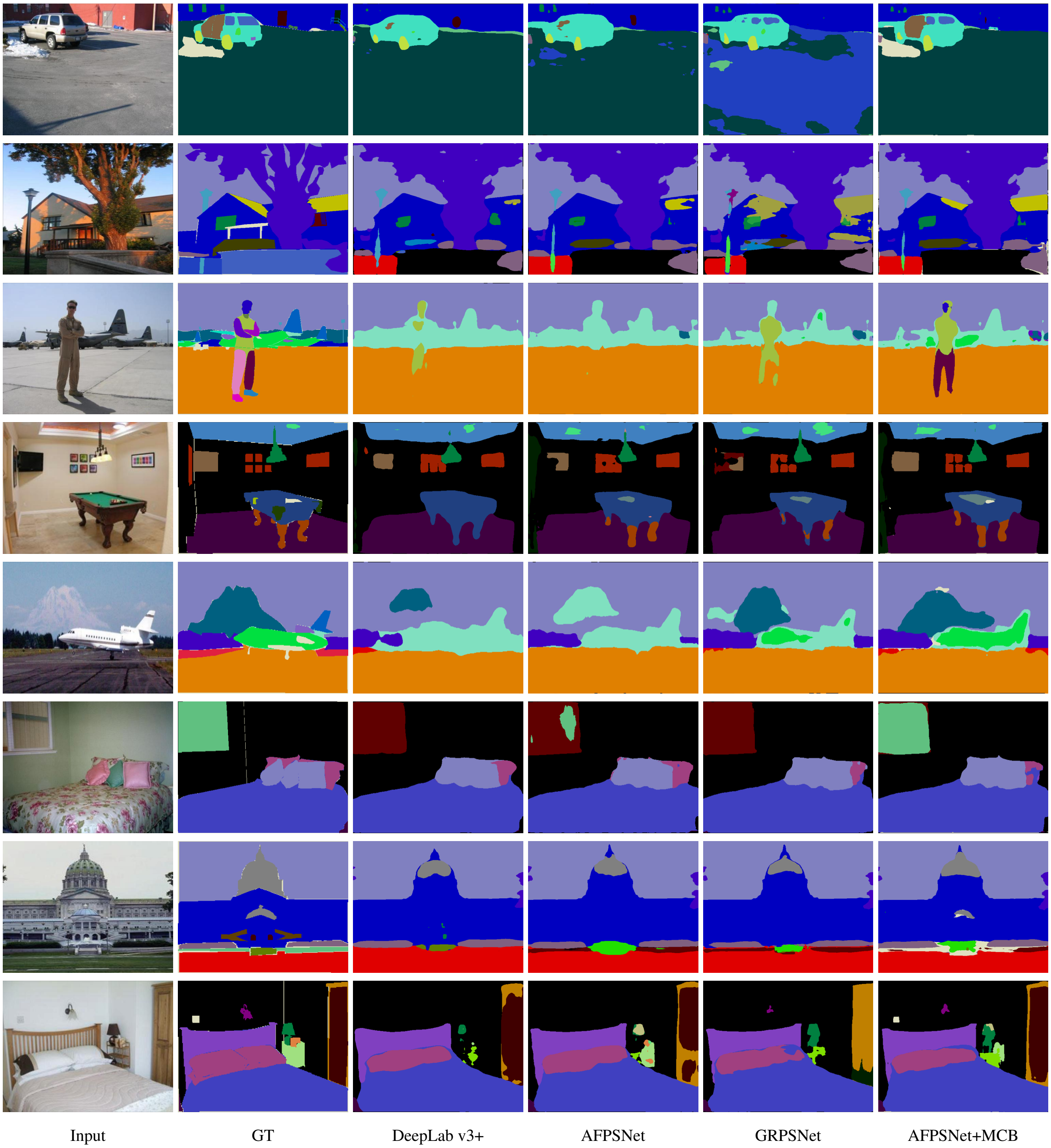

Fig. 2. Segmentation results on ADE20K-Part dataset. The proposed AFPSNet+MCB model shows overall better segmentation results with better part localisation and more accurate boundaries compared to other methods.