SAMO: A Lightweight Sharpness-Aware Approach for Multi-Task Optimization with Joint Global-Local Perturbation

Supplementary Material

A. Sharpness and task conflicts

For the experiment in Section 4, we use the same model and experimental setup as in Section 6. Following Foret et al. 26, we remove batch normalization, as it affects the interpretation of Hessian. We provide the cosine similarities of task gradients for LS and LS with SAM after the 5th and 10th epoch, corresponding to one-third and two-thirds of the training progression, respectively.

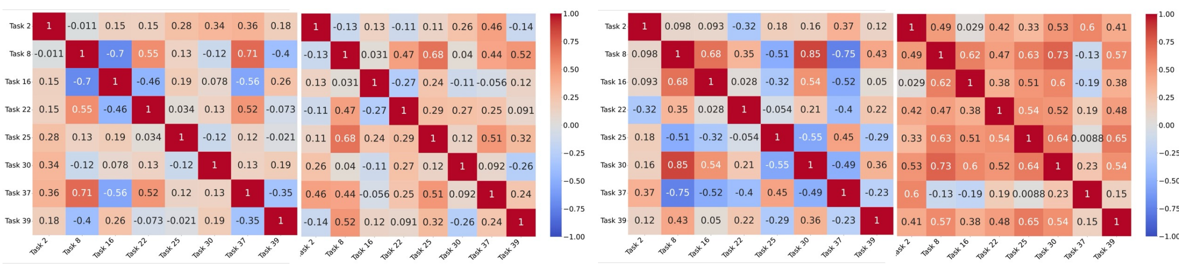

Figure 5. Cosine similarities of task gradients after 5th (left) and 10th (right) epoch. In each figure, LS is on the left and LS with SAM is on the right.

B. Hyperparameter selection and sensitivity analysis

For each dataset, we first search ρ within a wide range, $[0.00001, 0.0001, 0.001, 0.01, 0.1]$, and evaluate the best choice. We then refine the search space to determine an optimal value. For Cityscapes, NYU-v2, CelebA, Office Home, we use $\rho = 0.001$ and $\alpha = 0.5$. For QM9, the loss scales vary across different tasks; therefore, the parameter ρ should be carefully selected. We set $\rho = 0.00001$, and $\alpha = 0.1$. We set $\mu = 0.01$ in Equation (3) to approximate each task gradient using only forward passes.

Furthermore, we conduct the sensitivity analysis of α on Cityscapes, NYU-v2, QM9, and Office Home datasets. The results below indicate that model's performance is not highly sensitive to α. Empirically, we observe that the model remains relatively robust within the range $\alpha \in [0.1, 0.6]$.

In Addition, we also perform a sensitivity analysis of ρ on Cityscapes and NYU-v2 datasets. We explore values $\rho \in \{0.0001, 0.01\}$. We observe from Table 10 that model performance is more sensitive to ρ mainly because it plays a crucial role in the behavior of SAM.

Method	Cityscapes ↓	NYU-v2 ↓	QM9 ↓	Office ↑
LS	22.60	5.59	177.6	77.27
SAMO-LS ($\alpha = 0.1$)	18.84	3.28	141.8	77.14
SAMO-LS ($\alpha = 0.5$)	14.30	2.88	160.0	77.58
SAMO-LS ($\alpha = 0.9$)	24.81	6.38	213.9	76.68
FairGrad	3.90	-4.96	57.9	77.12
SAMO-FairGrad ($\alpha = 0.1$)	1.45	-5.28	53.0	76.65
SAMO-FairGrad ($\alpha = 0.5$)	-0.62	-6.55	55.7	77.66
SAMO-FairGrad ($\alpha = 0.9$)	6.49	-4.56	68.5	76.29

Table 9. Sensitivity of α. The reported results are highlighted.

Method	Cityscapes ↓	NYU-v2 ↓
LS	22.60	5.59
SAMO-LS ($\rho = 0.01$)	24.03	3.59
SAMO-LS ($\rho = 0.001$)	14.30	2.88
SAMO-LS ($\rho = 0.0001$)	18.78	3.67
FairGrad	3.90	-4.96
SAMO-FairGrad ($\rho = 0.01$)	14.03	-4.38
SAMO-FairGrad ($\rho = 0.001$)	-0.62	-6.55
SAMO-FairGrad ($\rho = 0.0001$)	6.11	-4.15

Table 10. Sensitivity of ρ. The reported results are highlighted.

C. Detailed results on multi-task regression

We provide more details about per-task results on the QM9 dataset in Table 11.

Method	μ	α	ϵ_{HOMO}	ϵ_{LUMO}	$\langle R^2 \rangle$	ZPVE	U_0	U	H	G	c_v	$\Delta m\% \downarrow$
	MAE ↓											
STL	0.067	0.181	60.57	53.91	0.502	4.53	58.8	64.2	63.8	66.2	0.072	
LS	0.106	0.325	73.57	89.67	5.19	14.06	143.4	144.2	144.6	140.3	0.128	177.6
SI	0.309	0.345	149.8	135.7	**1.00**	4.50	**55.3**	**55.75**	**55.82**	**55.27**	0.112	77.8
RLW	0.113	0.340	76.95	92.76	5.86	15.46	156.3	157.1	157.6	153.0	0.137	203.8
DWA	0.107	0.325	74.06	90.61	5.09	13.99	142.3	143.0	143.4	139.3	0.125	175.3
UW	0.386	0.425	166.2	155.8	1.06	4.99	66.4	66.78	66.80	66.24	0.122	108.0
MGDA	0.217	0.368	126.8	104.6	3.22	5.69	88.37	89.4	89.32	88.01	0.120	120.5
PCGrad	0.106	0.293	75.85	88.33	3.94	9.15	116.36	116.8	117.2	114.5	0.110	125.7
IMTL-G	0.136	0.287	98.31	93.96	1.75	5.69	101.4	102.4	102.0	100.1	0.096	77.2
CAGrad	0.118	0.321	83.51	94.81	3.21	6.93	113.99	114.3	114.5	112.3	0.116	112.8
Nash-MTL	0.102	**0.248**	82.95	**81.89**	2.42	5.38	74.50	75.02	75.10	74.16	**0.093**	62.0
FAMO	0.150	0.300	94.00	95.20	1.63	4.95	70.82	71.20	71.20	70.30	0.100	58.5
FairGrad	0.117	0.253	87.57	84.00	2.15	5.07	70.89	71.17	71.21	70.88	0.095	57.9
SAMO-LS	**0.096**	0.301	**66.66**	82.81	4.69	11.16	117.49	118.11	118.56	116.86	0.117	141.8
SAMO-MGDA	0.203	0.323	119.58	86.63	2.21	5.35	97.80	98.53	98.26	98.10	0.107	96.8
SAMO-FairGrad	0.137	0.255	99.53	87.31	1.72	**4.30**	70.39	70.88	70.70	70.98	**0.093**	**53.0**

Table 11. Detailed results of on QM9 (11-task) dataset. The best results are highlighted in **bold** with gray background.

D. Ablation study

We provide more detailed per-task results of ablation study on Cityscapes, and NYU-v2 datasets. The results are presented in Table 12 and Table 13.

Method	Segmentation		Depth		$\Delta m\% \downarrow$
	mIoU ↑	Pix Acc ↑	Abs Err ↓	Rel Err ↓	
LS	75.18	93.49	0.0155	46.77	22.60
G-SAM-LS	75.78	93.59	0.0155	41.68	17.85
L-SAM-LS	74.48	93.34	0.0154	40.07	16.71
SAMO-LS	**76.46**	**93.76**	**0.0147**	**39.85**	**14.30**
MGDA	68.84	91.54	0.0309	33.50	44.14
G-SAM-MGDA	72.64	92.97	0.0150	30.01	7.51
L-SAM-MGDA	72.77	92.50	0.0185	**26.97**	11.94
SAMO-MGDA	**73.28**	**93.26**	**0.0133**	30.57	**4.30**
FairGrad	74.10	93.03	0.0135	29.92	3.90
G-SAM-FairGrad	**74.81**	93.12	**0.0126**	28.85	0.93
L-SAM-FairGrad	74.16	93.11	0.0135	26.60	1.01
SAMO-FairGrad	74.37	**93.14**	0.0129	**26.30**	**-0.62**

Table 12. Ablation study of SAMO on Cityscapes (2-task) dataset. The best results of each method are highlighted in **bold** with gray background.

Method	Segmentation		Depth		Surface Normal					$\Delta m\% \downarrow$
					Angle Distance ↓		Within t° ↑			
	mIoU ↑	Pix Acc ↑	Abs Err ↓	Rel Err ↓	Mean	Median	<11.25	<22.5	<30	
LS	39.29	65.33	0.5493	0.2263	28.15	23.96	22.09	47.50	61.08	5.59
G-SAM-LS	39.17	64.84	0.5500	0.2301	27.47	23.02	23.60	49.33	62.74	3.85
L-SAM-LS	38.14	64.76	**0.5324**	**0.2215**	**27.12**	**22.40**	**25.04**	**50.50**	**63.59**	**2.11**
SAMO-LS	**39.59**	**65.72**	0.5514	0.2246	27.38	22.78	24.09	49.82	63.01	2.88
MGDA	**30.47**	59.90	**0.6070**	0.2555	24.88	19.45	29.18	56.88	69.36	1.38
G-SAM-MGDA	29.49	59.60	0.6107	**0.2349**	24.77	18.94	30.77	57.90	69.75	-0.23
L-SAM-MGDA	27.44	59.89	0.6370	0.2356	24.53	18.61	31.42	58.59	70.33	0.01
SAMO-MGDA	29.85	**60.83**	0.6111	0.2388	**24.11**	**18.18**	**32.16**	**59.59**	**71.15**	**-2.19**
FairGrad	38.80	65.29	0.5572	0.2322	24.55	18.97	30.50	57.94	70.14	-4.96
G-SAM-FairGrad	37.90	65.08	**0.5269**	0.2162	24.50	19.03	30.38	57.76	70.04	-5.70
L-SAM-FairGrad	38.72	**65.42**	0.5737	0.2455	**24.16**	**18.43**	**31.67**	**59.06**	**70.92**	-5.42
SAMO-FairGrad	**39.05**	65.06	0.5359	**0.2137**	24.43	18.79	30.98	58.35	70.42	**-6.55**

Table 13. Ablation study of SAMO on NYU-v2 (3-task) dataset. The best results of each method are highlighted in **bold** with gray background.

Moreover, we conduct ablation study on the perturbation normalization and present the results in Table 8. Global normalization scales the local perturbation for task i by the factor $\frac{\|\nabla_\theta l_0(\theta)\|}{\|\hat{\nabla}_\theta l_i(\theta)\|}$, where numerator and denominator are the norms of global and local perturbations across all layers, respectively. It can be observed that our layer-wise normalization consistently outperforms alternative methods.

Method	Cityscapes					NYU-v2
	Segmentation		Depth		$\Delta m\% \downarrow$	$\Delta m\% \downarrow$
	mIoU ↑	Pix Acc ↑	Abs Err ↓	Rel Err ↓		
LS	75.18	93.49	0.0155	46.77	22.60	5.59
SAMO-LS	76.46	93.76	0.0147	39.85	**14.30**	**2.88**
SAMO-LS (glob. norm.)	75.79	93.69	0.0142	44.00	17.21	4.79
SAMO-LS (w/o norm.)	75.82	93.76	0.0143	43.38	16.80	4.91
FairGrad	74.10	93.03	0.0135	29.92	3.90	-4.96
SAMO-FairGrad	74.37	93.14	0.0129	26.30	**-0.62**	**-6.55**
SAMO-FairGrad (glob. norm.)	74.54	93.10	0.0130	27.35	0.42	-5.44
SAMO-FairGrad (w/o norm.)	74.23	93.02	0.0127	26.98	-0.27	-4.72

Table 14. Ablation study of perturbation normalization on Cityscapes and NYU-v2 datasets. glob. norm. refers to global normalization of local perturbations, while w/o norm. means no normalization.

E. Feature visualization

For NYU-v2 dataset, we extract the intermediate features from the task-specific heads and visualize them using t-SNE [59]. SAMO-FairGrad achieves a higher inter-cluster distance (52.35 vs. 50.48), indicating improved task feature separability.

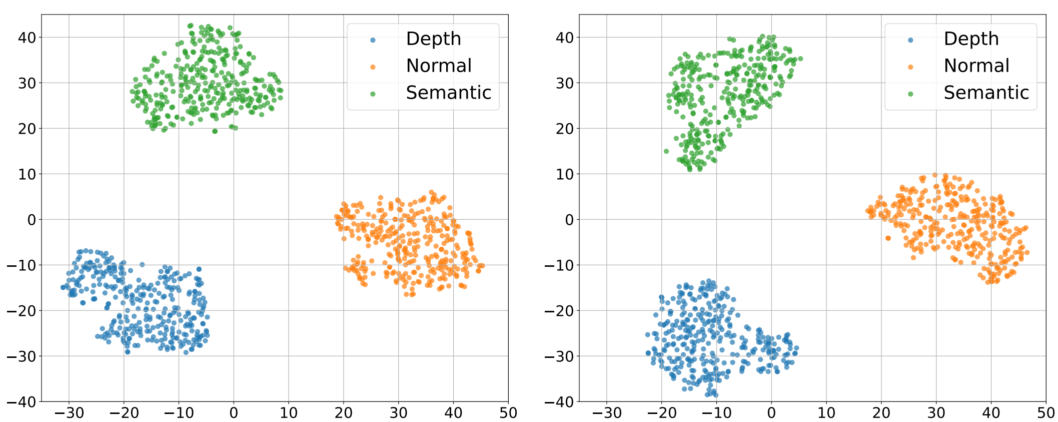

Figure 6. Feature visualization. Left: FairGrad . Right: SAMO-FairGrad.