Joint Class-Affinity Loss Correction for Robust Medical Image Segmentation with Noisy Labels

Xiaoqing Guo and Yixuan Yuan

Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China

Paper Code

Background & motivation

●Background

With limited budgets and efforts, the resulting dataset would be noisy, and the presence of label noises may mislead the segmentation model to memorize wrong semantic correlations, resulting in severely degraded generalizability. Hence, developing medical image segmentation techniques that are robust to noisy labels in training data is of great importance.

●Motivation

Almost all existing image segmentation methods tackle label noise issues merely in a pixel-wise manner. We found that the pair-wise manner can greatly reduce the noise rate. For example, if one pixel in a pair is mislabeled (e.g. the red rectangle) or even both pixels are mislabeled (e.g. the orange rectangle), the affinity label of this pair might be correct, thereby reducing the noise rate from 44% to 23%.

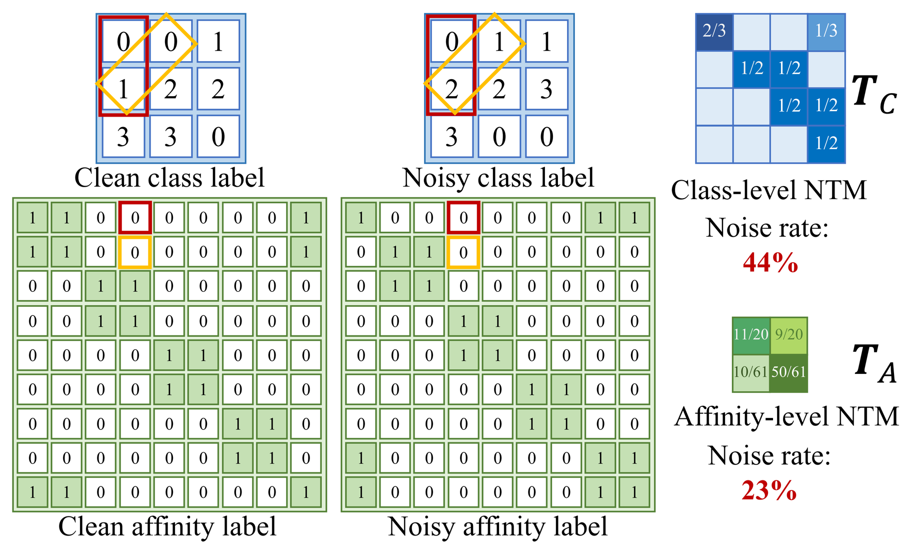

Fig. 1. A toy example to illustrate the comparison between pixel-wise class label and pair-wise affinity label.

●Our contribution

1) Unifying the pixel-wise and pair-wise manners, we propose a robust Joint Class-Affinity Segmentation (JCAS) framework to combat label noise issues in medical image segmentation.

2) We devise a differentiated affinity reasoning (DAR) module to guide the refinement of pixel-wise predictions with differentiated pair-wise affinity relations.

3) We design a class-affinity loss correction (CALC) strategy to further correct both pixel-wise and pair-wise supervision signals, and in the meanwhile, unify the pixel-wise and pair-wise supervisions through the theoretically derived consistency regularization.

5) Extensive experiments on synthetic and real-world noisy labels verify the effectiveness of JCAS.

Methodology

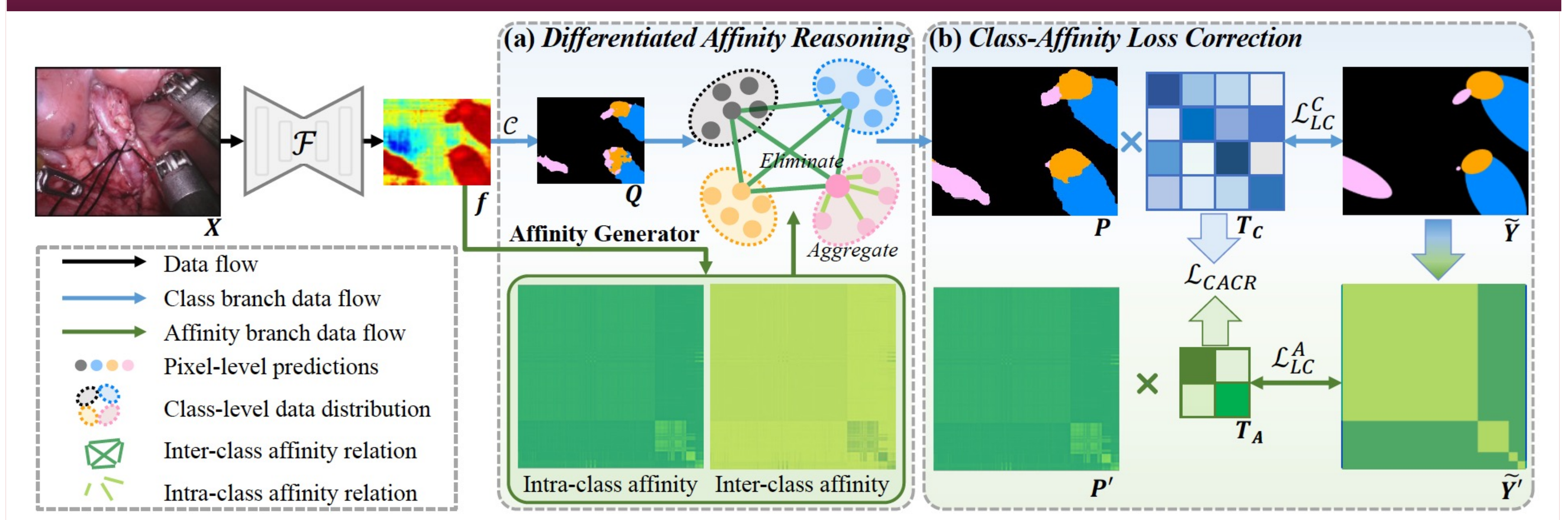

Fig. 2. Illustration of the proposed JCAS framework, composed of (a) DAR module and (b) CALC strategy.

● Joint Class-Affinity Segmentation (JCAS) framework

- JCAS framework has two supervision signals, derived from noisy class labels and noisy affinity labels, for regularizing pixel-wise predictions (the upper branch in Fig. 2) and pair-wise affinity relations (the lower branch in Fig. 2), respectively.
- These two supervision signals are complementary to each other since the pixel-wise one preserves semantics and the pair-wise one reduces noise rate.

● Differentiated Affinity Reasoning (DAR) module

- Pair-wise affinity relations $\boldsymbol{P}'$ derived at the feature level model the contextual dependencies, highlighting these pixel pairs belonging to the same class and revealing the intra-class affinity.
- The reverse affinity map $\boldsymbol{P}'_{re} = norm(1-\boldsymbol{P}')$ measures the dissimilarity between two pixels and reveals the inter-class affinity relations.
- As in Fig. 2.a, DAR differentiates affinity relations to explicitly aggregate intra-class correlated information ($\boldsymbol{P}_{intra}(k_1) = \boldsymbol{P}(k_1) + \sum_{k_2}^{n} \boldsymbol{P}'(k_1,k_2)\boldsymbol{Q}(k_2)$) and eliminate inter-class irrelevant information ($\boldsymbol{P}_{inter}(k_1) = \boldsymbol{P}(k_1) - \sum_{k_2}^{n} \boldsymbol{P}'_{re}(k_1,k_2)\boldsymbol{Q}(k_2)$), guiding the refinement of pixel-wise predictions.

● Class-Affinity Loss Correction (CALC) strategy

- CACL strategy models noise label distributions in class labels and affinity labels as two NTMs ($\boldsymbol{T}_C$ and $\boldsymbol{T}_A$ in Fig. 2.b) for loss correction ($\mathcal{L}_{LC}^C$ and $\mathcal{L}_{LC}^A$ in Fig. 2.b).

$$\mathcal{L}_{LC}^C = -\sum_{k}^{H\times W} \tilde{\boldsymbol{Y}}(k)\log[\boldsymbol{P}(k)\boldsymbol{T}_C]$$

$$\mathcal{L}_{LC}^A = -\sum_{k}^{H\times W} \tilde{\boldsymbol{Y}}(k)\log[\boldsymbol{P}'(k)\boldsymbol{T}_A] + (1-\tilde{\boldsymbol{Y}}(k))\log(1-\boldsymbol{P}'(k)\boldsymbol{T}_A)$$

- CACL strategy unifies pixel-wise and pair-wise supervisions via the theoretically derived class-affinity consistency regularization ($\mathcal{L}_{CACR}$ in Fig. 2.b), thereby facilitating the noise resistance.

$$\mathcal{L}_{CACR} = \|\boldsymbol{T}_{C\to A} - \boldsymbol{T}_A\|_2$$

$$\boldsymbol{T}_{C\to A}(0,0) = 1-\boldsymbol{T}_{C\to A}(0,1),\ \ \boldsymbol{T}_{C\to A}(0,1) = \frac{\sum_m[N_m\sum_n \boldsymbol{T}_C(m,n)]^2 - \sum_m (N_m)^2\|\boldsymbol{T}_C\|_2^2}{\sum_m[N_m(\sum_m N_m - N_m)]},$$

$$\boldsymbol{T}_{C\to A}(1,0) = 1-\boldsymbol{T}_{C\to A}(1,1),\ \ \boldsymbol{T}_{C\to A}(1,1) = \frac{\sum_m (N_m)^2\|\boldsymbol{T}_C\|_2^2}{\sum_m (N_m)^2}.$$

Experiment Results

● Dataset

We validate JCAS on the surgical instrument dataset Endovis18 [1]. It consists of 2384 images (1639 training & 596 test images) annotated with the instrument part labels, including shaft, wrist and clasper classes. Each image is resized into a resolution of 256x320 in preprocessing.

● Noise Patterns

We conduct experiments under both synthetic label noises (i.e., elipse, symmetric and asymmetric noises) and real-world label noise (i.e., noisy pseudo labels in source-free domain adaptation (SFDA)), as illustrated in Fig. 3.

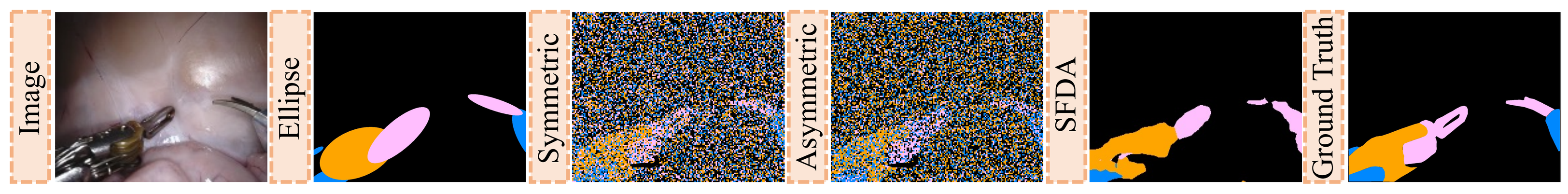

Fig. 3. Illustration of dataset with different kinds of label noises.

● Comparison with State-of-the-art Methods & Ablation Study

Table 1. Comparison under four label noises. Best and second best results are **highlighted** and underlined. 'w/ Affinity' introduces pair-wise supervision $\mathcal{L}_{Bi}^A$ to backbone.

Noises	Method	Shaft Dice (%)	Shaft Jac (%)	Wrist Dice (%)	Wrist Jac (%)	Clasper Dice (%)	Clasper Jac (%)	Average Dice (%)	Average Jac (%)
	Upper bound	88.740	81.699	65.045	52.627	70.531	56.618	74.772	63.648
Ellipse	RAUNet (19') [13]	83.137	74.139	56.941	43.215	61.081	45.883	67.053	54.412
	LWANet (20') [12]	81.945	72.735	53.626	40.886	**64.364**	**49.781**	66.645	54.468
	CSS (21') [14]	<u>84.577</u>	**75.736**	57.597	43.687	63.686	48.347	<u>68.620</u>	<u>55.923</u>
	MTCL (21') [19]	72.719	60.540	39.386	27.474	49.662	35.085	53.922	41.033
	SR (21') [23]	79.966	69.621	53.540	39.747	60.179	44.775	64.561	51.381
	VolMin (21') [11]	81.320	70.758	60.470	46.408	58.203	42.524	66.664	53.230
	Baseline [3]	79.021	68.097	42.069	29.582	55.489	40.175	58.860	45.951
	w/ Affinity	82.158	72.339	49.128	35.455	58.933	43.594	63.406	50.463
	w/ DAR	82.698	72.992	52.207	38.442	61.544	46.027	65.483	52.487
	w/ CALC	82.973	73.126	<u>61.885</u>	<u>47.527</u>	60.416	44.821	68.425	55.158
	Ours (JCAS)	**84.683**	<u>75.378</u>	**65.599**	**51.623**	<u>63.871</u>	<u>48.356</u>	**71.384**	**58.452**
Symmetric	RAUNet (19') [13]	68.044	54.397	31.581	20.676	41.302	27.819	46.976	34.297
	LWANet (20') [12]	0.294	0.150	10.089	5.908	10.228	5.489	6.870	3.849
	CSS (21') [14]	86.555	78.451	32.363	20.767	53.364	37.901	57.427	45.706
	MTCL (21') [19]	78.480	67.855	50.011	38.013	55.515	40.411	61.336	48.760
	SR (21') [23]	86.648	78.823	58.217	46.870	64.643	50.120	69.836	58.604
	VolMin (21') [11]	<u>86.811</u>	<u>78.834</u>	<u>63.712</u>	<u>51.259</u>	<u>66.604</u>	<u>52.096</u>	<u>72.376</u>	<u>60.730</u>
	Baseline [3]	85.021	76.419	57.026	44.563	63.255	48.395	68.434	56.459
	Ours (JCAS)	**88.285**	**80.692**	**65.759**	**53.487**	**68.129**	**53.821**	**74.058**	**62.667**
Asymmetric	RAUNet (19') [13]	87.255	79.983	59.462	46.639	67.347	52.801	71.355	59.808
	LWANet (20') [12]	0.015	0.007	40.548	30.683	9.060	4.825	16.541	11.838
	CSS (21') [14]	**89.825**	**83.543**	43.743	30.569	**69.285**	**54.758**	67.618	56.290
	MTCL (21') [19]	74.544	62.525	41.433	30.533	48.077	33.676	54.685	42.244
	SR (21') [23]	86.360	78.055	62.854	49.651	65.483	50.962	71.566	59.556
	VolMin (21') [11]	86.840	78.796	<u>63.345</u>	<u>51.137</u>	65.220	50.996	<u>71.802</u>	<u>60.310</u>
	Baseline [3]	84.497	75.607	58.717	46.060	61.662	46.770	68.292	56.146
	Ours (JCAS)	<u>88.247</u>	<u>80.730</u>	**67.298**	**54.922**	<u>67.686</u>	<u>53.436</u>	**74.410**	**63.029**
SFDA	RAUNet (19') [13]	73.370	61.568	56.063	42.570	45.979	31.720	58.471	45.286
	LWANet (20') [12]	75.377	64.457	53.203	39.799	<u>48.558</u>	<u>34.191</u>	59.046	46.149
	CSS (21') [14]	74.419	64.261	**61.765**	**47.880**	45.749	31.709	<u>60.644</u>	<u>47.950</u>
	MTCL (21') [19]	72.289	60.346	51.095	37.972	38.762	25.567	54.048	41.295
	SR (21') [23]	75.992	64.835	57.370	43.863	40.471	27.388	57.944	45.362
	VolMin (21') [11]	**76.641**	<u>65.063</u>	58.285	44.389	41.780	28.324	58.902	45.925
	Baseline [3]	76.107	64.858	56.259	42.740	41.364	28.091	57.910	45.230
	Ours (JCAS)	<u>76.540</u>	**65.300**	<u>59.904</u>	<u>46.104</u>	**48.725**	**34.283**	**61.723**	**48.562**

● Visualization of Segmentation Results

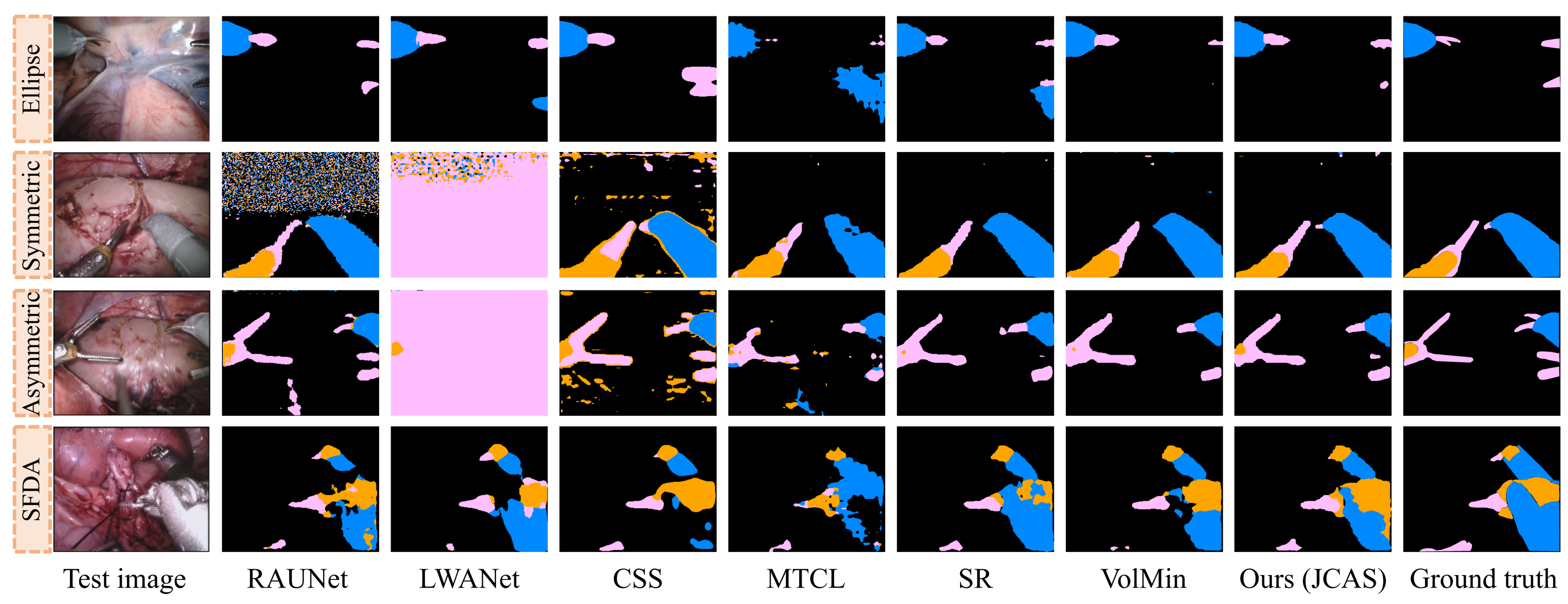

Fig. 4. Comparison of segmentation results.

● Test *Jac* Curve in Training Stage

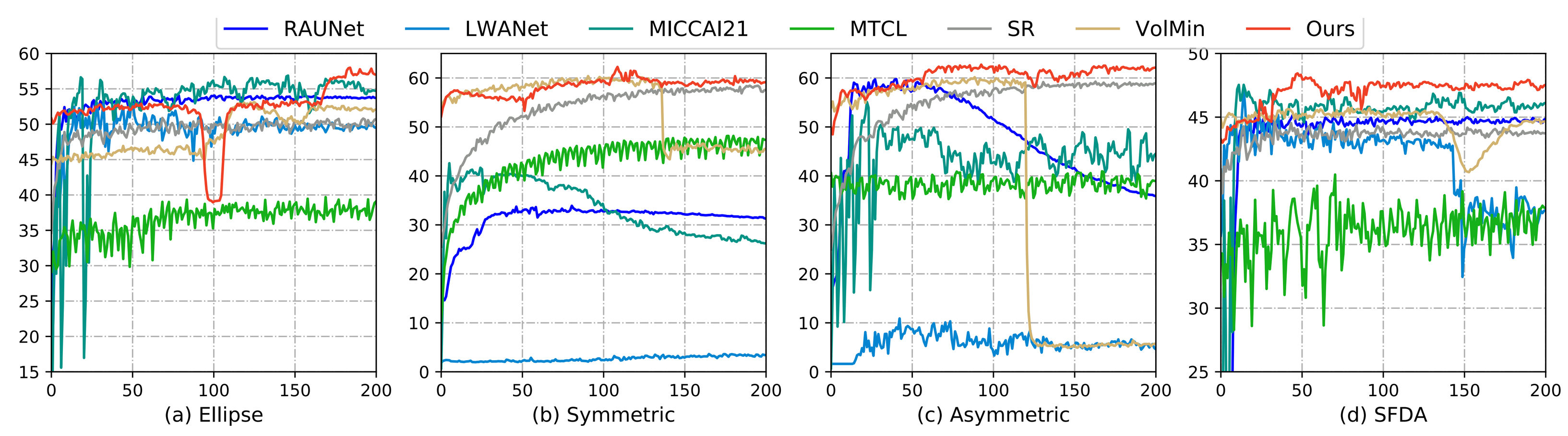

Fig. 5. Curve of test *Jac* vs. epoch with four different types of noise labels.

Conclusion

We propose a robust JCAS framework to combat label noise issues in medical image segmentation. Complementing the widely used pixel-wise manner, we introduce the pair-wise manner by capturing affinity relations among pixels to reduce noise rate. Then a DAR module is devised to rectify pixel-wise segmentation predictions by reasoning about intra-class and inter-class affinity relations. We further design a CALC strategy to unify pixel-wise and pair-wise supervisions, and facilitate noise tolerances of both supervisions. Extensive experiments under four noisy labels corroborate the noise immunity of JCAS.

Reference

[1] Allan,M.,Kondo,S.,Bodenstedt,S.,Leger,S.,Kadkhodamohammadi,R.,Luengo, I., Fuentes, F., Flouty, E., Mohammed, A., Pedersen, M., et al.: 2018 robotic scene segmentation challenge. arXiv preprint arXiv:2001.11190 (2020)

[2] Allan, M., Shvets, A., Kurmann, T., Zhang, Z., Duggal, R., Su, Y.H., Rieke, N., Laina, I., Kalavakonda, N., Bodenstedt, S., et al.: 2017 robotic instrument segmen- tation challenge. arXiv preprint arXiv:1902.06426 (2019)