Medical Image Seminar

Yicheng Jiang
CUHK, Shenzhen
May 27, 2020

香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

Paper Information

- **TITLE**

  Semi-supervised Breast Lesion Detection In Ultrasound Video Based On Temporal Coherence

- **Authors**

Sihong Chen[1], Weiping Yu*[2], Kai Ma[1], Xinlong Sun[3], Xiaona Lin[4], Desheng Sun[4] and Yefeng Zheng†[1]

[1]Tencent YouTu X-Lab, Shenzhen, China
[2]School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China
[3]Tsinghua University, Shenzhen, China
[4]Peking University Shenzhen Hospital, Shenzhen, China

Problem

- **Target**

Detect breast lesions in ultrasound video

- **Difficulties**

1. The lesion boundary is blurry in some frames
2. High similarity to the background of soft tissues and shadow artifacts result in failure detection of lesions
3. There is only 2D image annotations because sonographers commonly save and label the key frames only. Labeled video data is lack.

Problem

- **Target & Challenges**

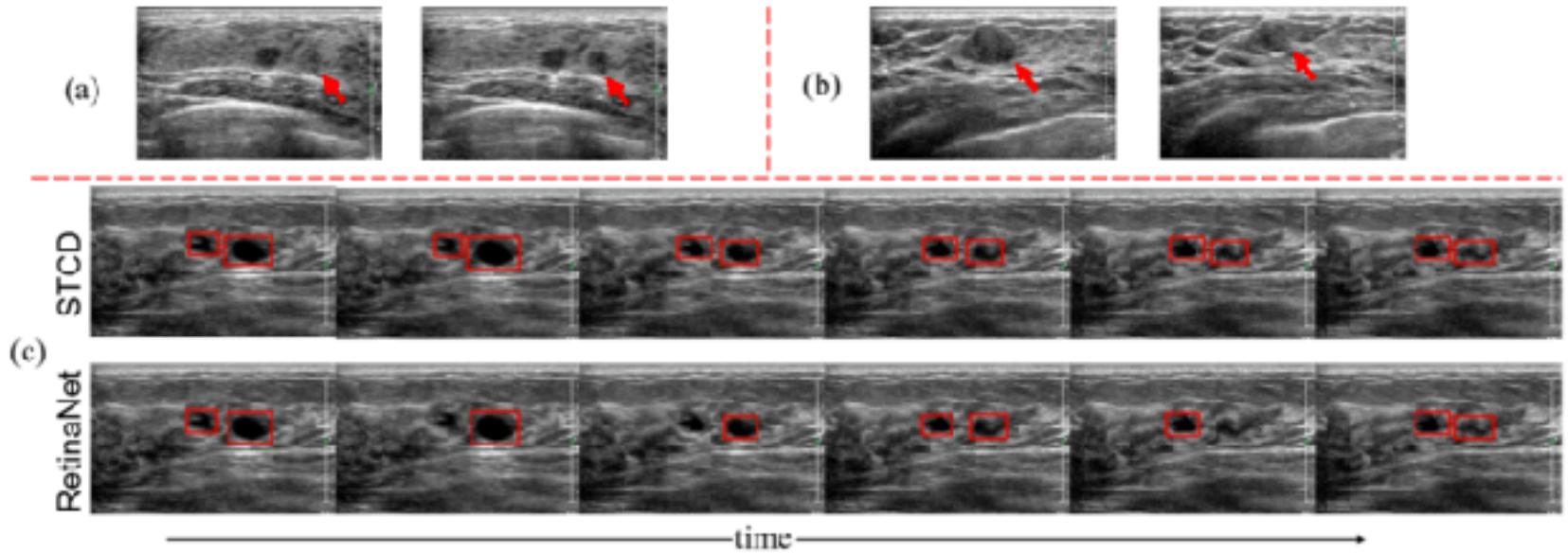


Figure 1: The challenging cases in breast lesion video detection. (a) Blurry boundaries; (b) High similarity to background; (c) Our proposed method on challenging samples

Method

- **STCD (Semi-supervised Temporal Coherent Detection)**
- **Contributions:**

1. STCD offers a semi-supervised solution to train a network with unlabeled video data and a different set of labeled images.
2. STCD improves breast lesion detection accuracy by automatically analyzing the temporal relationship across frames
3. Due to the sparse distribution of key frames in the video, most of the video frames are processed by the proposed efficient MotionNet and WarpNet, which help the network achieve real-time performance.

Method

- **STCD (Semi-supervised Temporal Coherent Detection)**
-

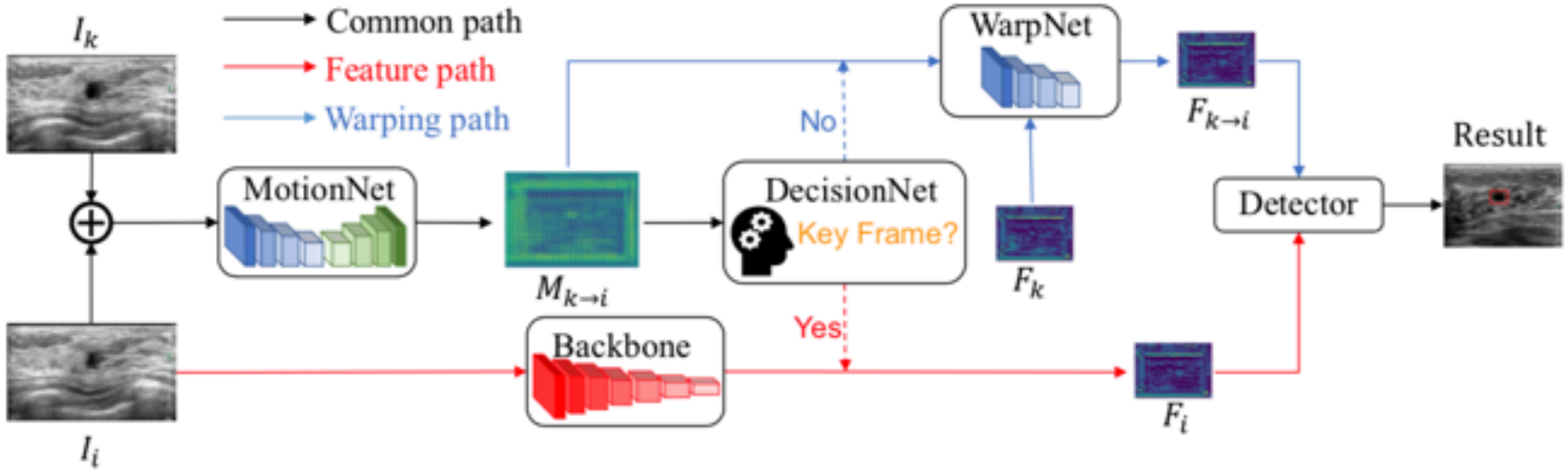


Figure 2: STCD Workflow

Method

- **STCD (Semi-supervised Temporal Coherent Detection)**
- **Training Strategy**

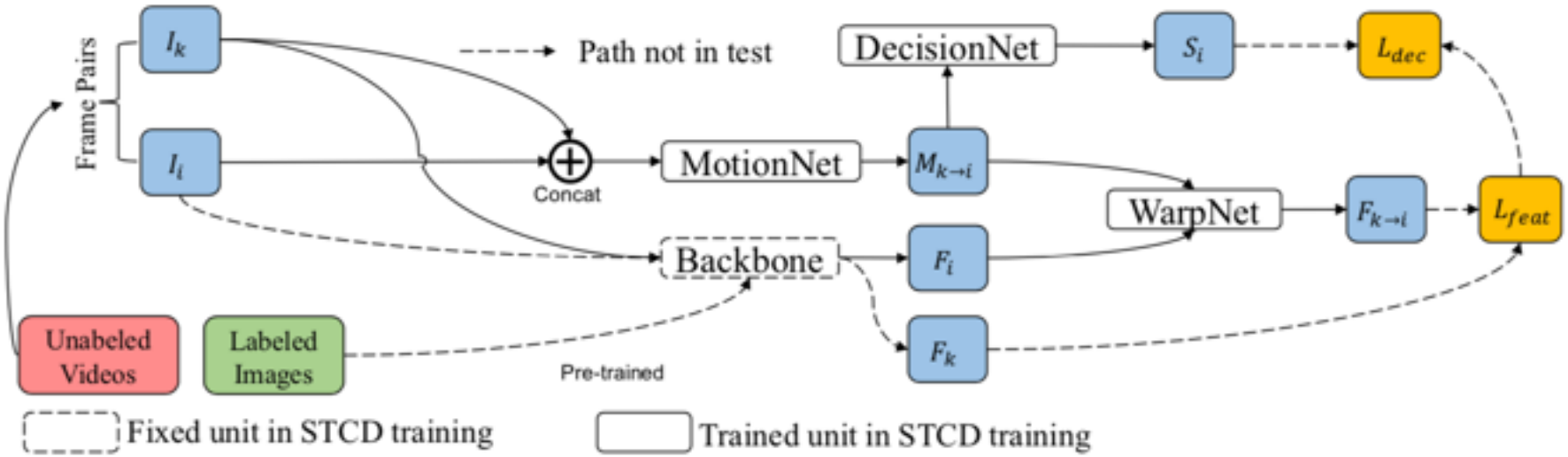


Figure 3: Semi-supervised Training of STCD

Method

- STCD (Semi-supervised Temporal Coherent Detection)
- Training Strategy
- Loss:
- Feature loss L_{feat} to measure correlation between $F_{k\to i}$ and F_i (Not Given)
- L_{dec} to achieve adaptive key-frame scheduling mechanism

$$Q_{k\to i} = \frac{\sum (F_{k\to i}(p) - F_i(p))^2}{N_p}$$

$$L_{dec} = (Q_{k\to i} - S_i)^2.$$

Data

- 5,608 labeled breast lesion ultrasound images for training detection network
- 80 unlabeled sequences and 10 labeled sequences (from 90 patients, where each video sequence contains 105 to 188 frames)
- Randomly Select 1/3 frames of the unlabeled training sequences as the key frames.
- Each of the above key frames is paired with a non-key frame in 20 adjacent frames.
- Form 59,962 training pairs
- 10 labeled sequences are used to be the test set

Result

- **STCD is effective**

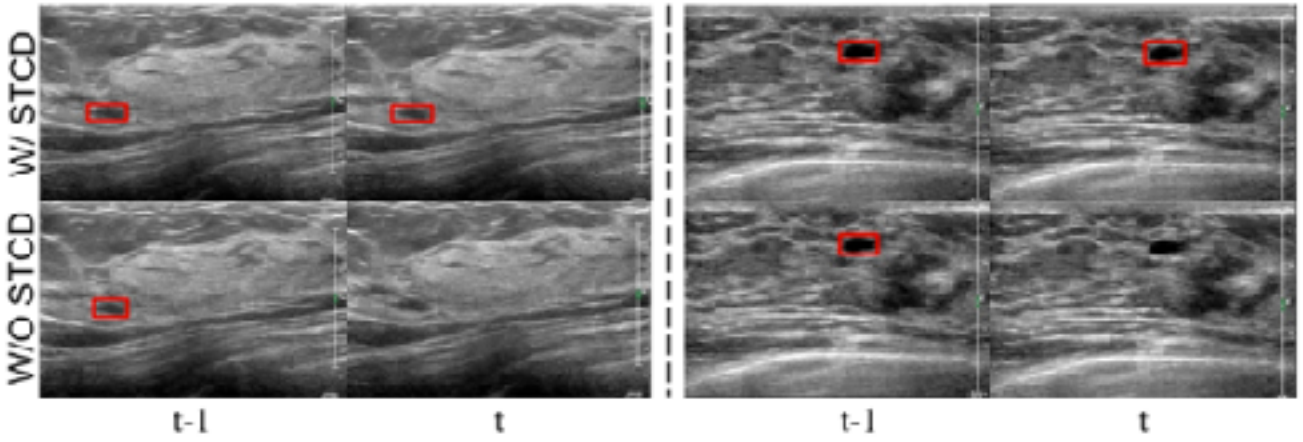


Figure 4: Breast lesion detection results of RetinaNet and the proposed STCD method

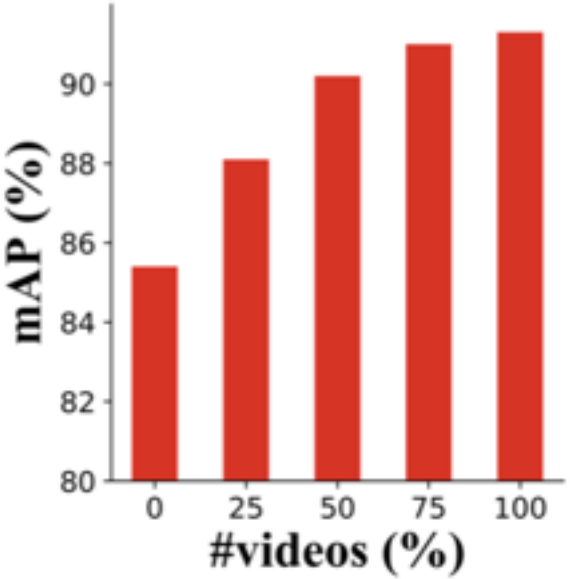


Figure 5: Effectiveness of Unlabeled Data

Result

- **MotionNet and WarpNet is efficient : Faster and More Accurate**

Table 1: Performance of STCT with different backbones for breast lesion detection on the test set of ultrasound videos

Method	Backbone	mAP(%)	GPU Runtime (ms)	CPU Runtime (ms)
YOLOV3 [7]	DarkNet	80.6	25	1071
YOLOV3 (STCD)	DarkNet	**81.8**	**19**	**630**
RetinaNet [6]	ResNet-50	82.0	23	913
RetinaNet (STCD)	ResNet-50	86.0	**16**	**540**
RetinaNet [6]	ResNet-101	85.4	28	1161
RetinaNet (STCD)	ResNet-101	89.9	17	577
RetinaNet [6]	ResNet-152	86.6	32	1548
RetinaNet (STCD)	ResNet-152	**91.3**	19	619

Ablation Study

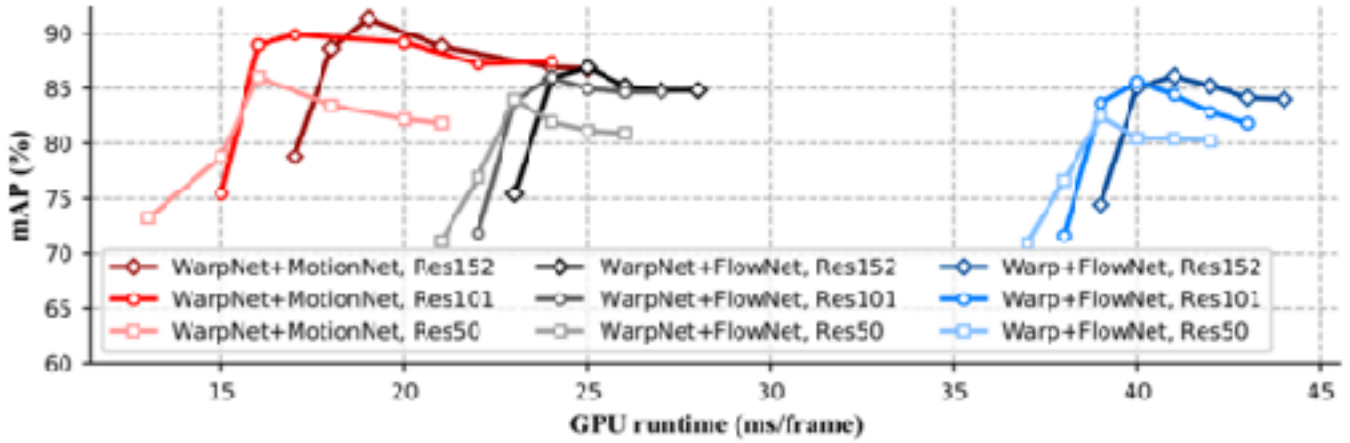


Figure 6: Accuracy (mAP) and runtime of breast lesion detection under different threshold τ for key frame decision

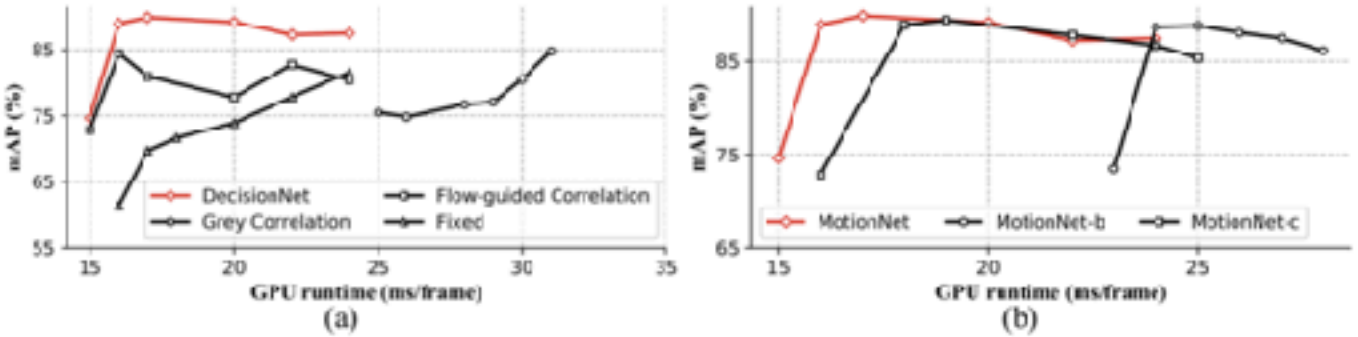


Figure 7: Accuracy (mAP) and runtime of (a) different key-frame scheduling; (b) different MotionNet architectures under different threshold τ for key frame decision

Paper Information

- **TITLE**

  Semi-supervised Segmentation of Lesion from Breast Ultrasound Images with Attentional Generative Adversarial Network

- **Authors**

Luyi Han[a], Yunzhi Huang[a,b,*], Haoran Dou[c], Shuai Wang[d], Sahar Ahamad[d], Honghao Luo[e], Qi Liu[a], Jingfan Fan[f], Jiang Zhang[a]

[a] College of Electrical Engineering, Sichuan University, Chengdu 610065, China
[b] Department of Biomedical Engineering, Sichuan University, Chengdu 610065, China
[c] National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Shenzhen University, Shenzhen 518060, China
[d] Department of Radiology and Biomedical Research Imaging Center, University of North Carolina, Chapel Hill, USA
[e] Department of Ultrasound, West China Hospital of Sichuan University, Chengdu 610041, China
[f] Beijing Engineering Research Center of Mixed Reality and Advanced Display, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

Problem

- **Target**

Automatic segmentation of breast lesion

- **Large-scale breast ultrasound (BUS) images remain unannotated and need to be effectively explored to improve the segmentation quality**

Method

- **BUS-GAN**

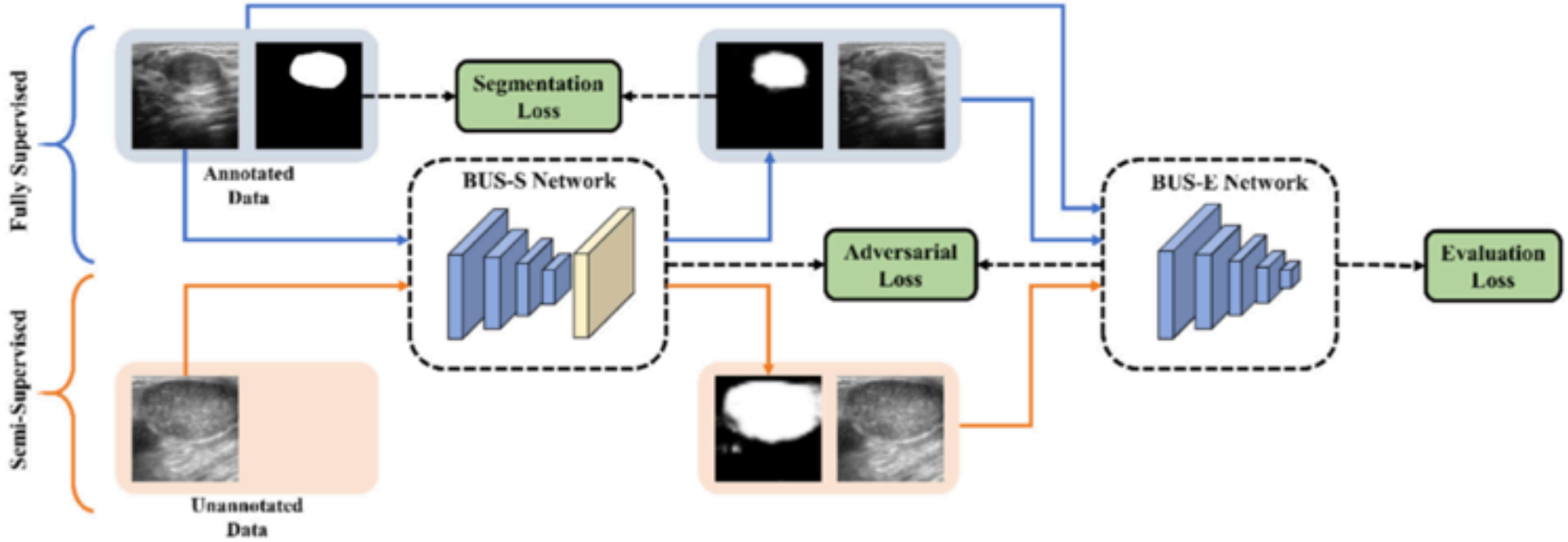


Fig. 2. Overview of the dual-attentive GAN-based semi-supervised segmentation network.

Method

- **BUS-GAN (Detail-Architecture)**

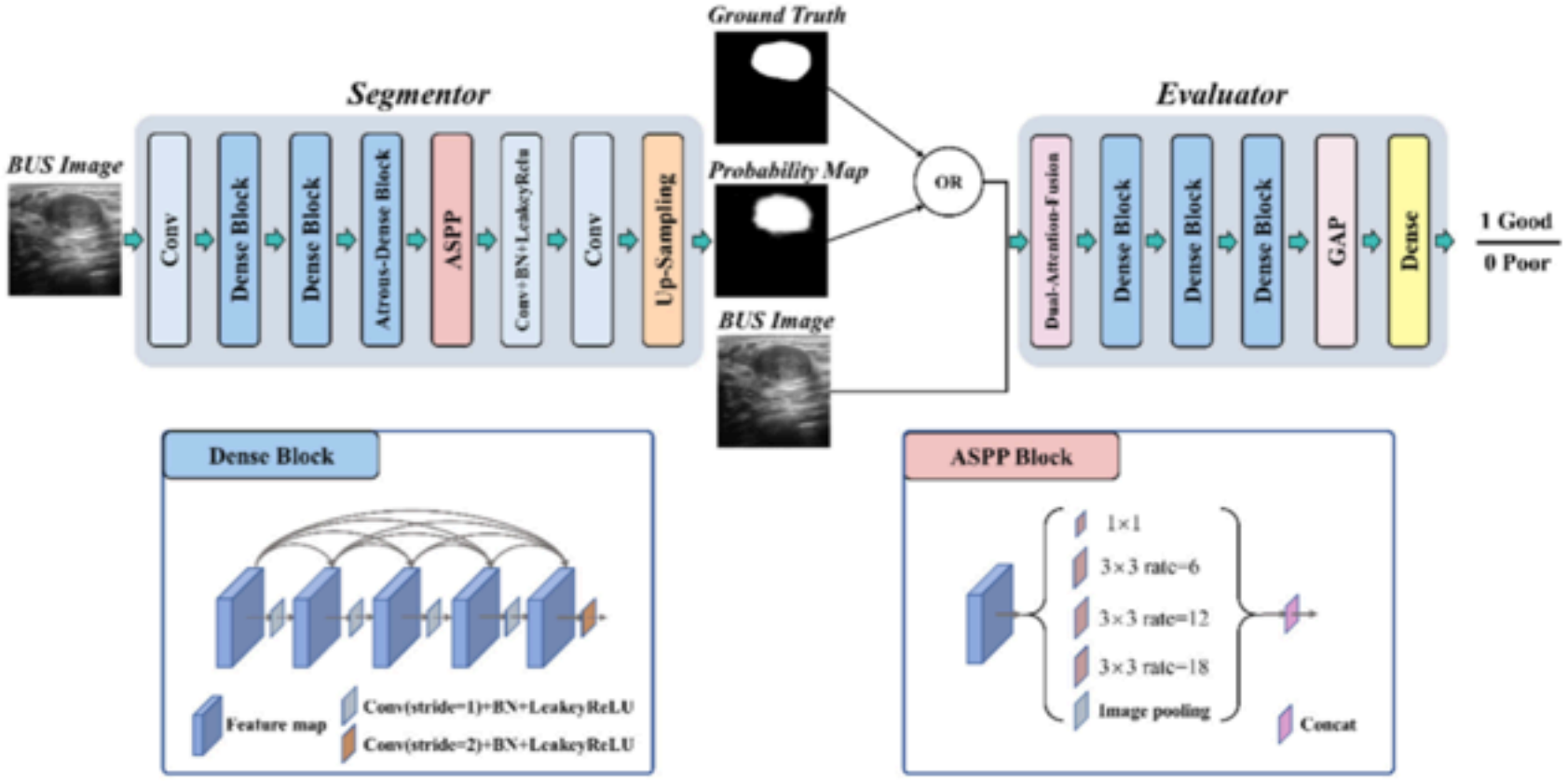


Fig. 3. The architecture of the BUS-S network and BUS-E network.

Method

- **BUS-GAN (Detail-Architecture)**

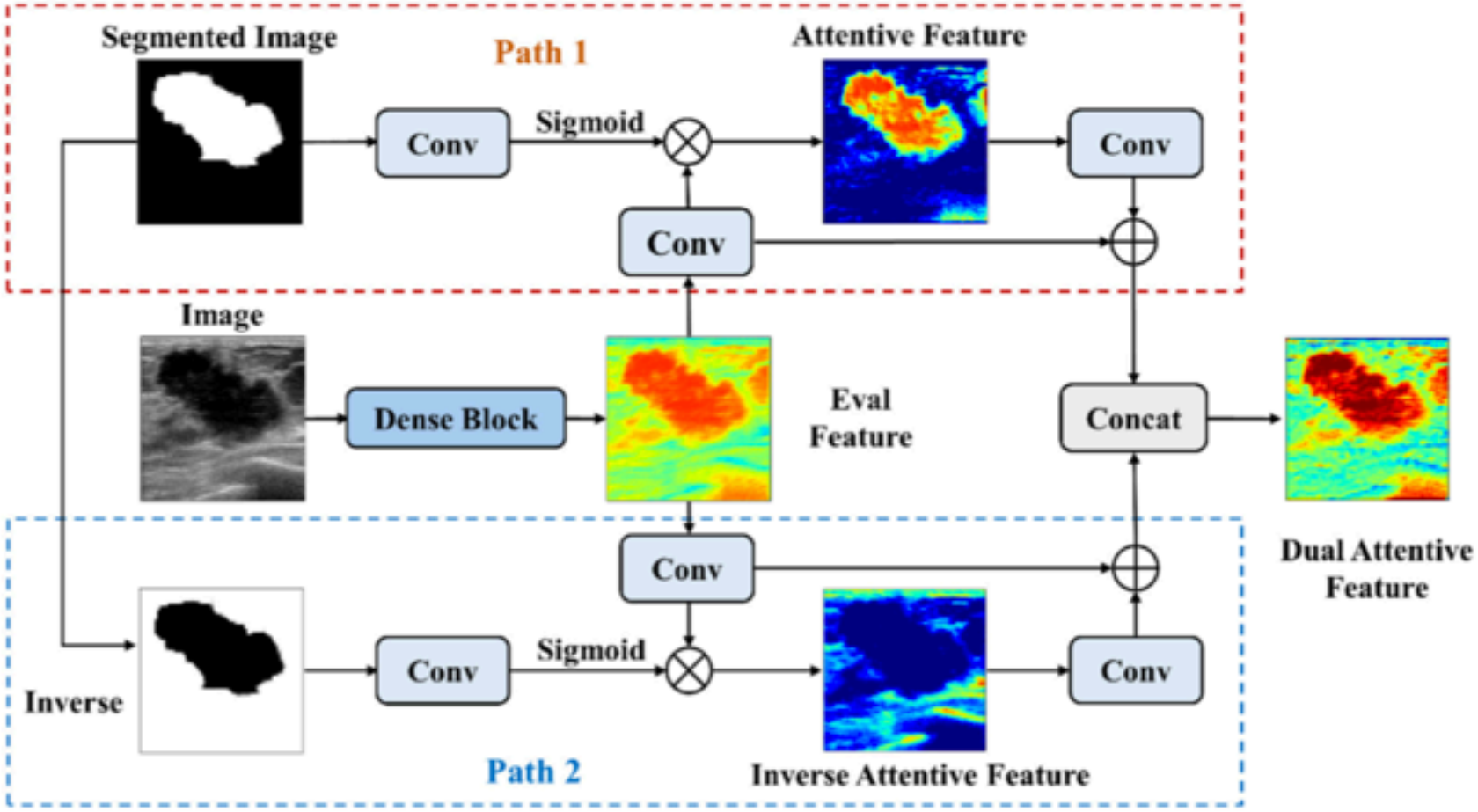


Fig. 4. Illustration of the dual-attentive-fusion block.

Method

- **Loss:**
-

$$\ell_S = \ell_{seg} + \lambda_{adv} \cdot \ell_{adv} \tag{1}$$

$$\ell_{seg}(X_a, Y_a; \theta_S) = \ell_{dce}(Seg(X_a), Y_a) + \ell_{bce}(Seg(X_a), Y_a) \tag{2}$$

$$\begin{aligned} \ell_{adv}(X_a, X_u; \theta_S) = {} & \lambda_a \cdot \ell_{bce}(Eva(X_a, Seg(X_a)), 1) \\ & + \lambda_u \cdot \ell_{bce}(Eva(X_u, Seg(X_u)), 1) \end{aligned} \tag{3}$$

$$\begin{aligned} & \ell_E(X_a, X_u, Y_a; \theta_E) \\ & = \ell_{bce}(Eva(X_a, Y_a), 1) + \lambda_a \cdot \ell_{bce}(Eva(X_a, Seg(X_a)), 0) \\ & \quad + \lambda_u \cdot \ell_{bce}(Eva(X_u, Seg(X_u)), 0) \end{aligned} \tag{4}$$

Data

Table 1
Diagnosis for the in-house dataset.

	Biopsy Results	Training		Testing	Total
		Anno.	Unan.		
Benign	Fibroadenomas	17	340	155	1400
	Adenosis of Breast	14	255	117	
	Adenosis with Fibroadenoma Formation of Breast	16	311	87	
	Sclerosing Adenosis	1	22	10	
	Intraductal Papillary Neoplasms	1	12	6	
	Granulomatous Mastitis	1	24	11	
Malignant	Invasive Ductal Carcinoma	47	867	384	1400
	Ductal Carcinoma in Situ	2	46	21	
	Invasive Lobular Carcinoma	1	23	9	
	Total	100	1900	800	2800

Result

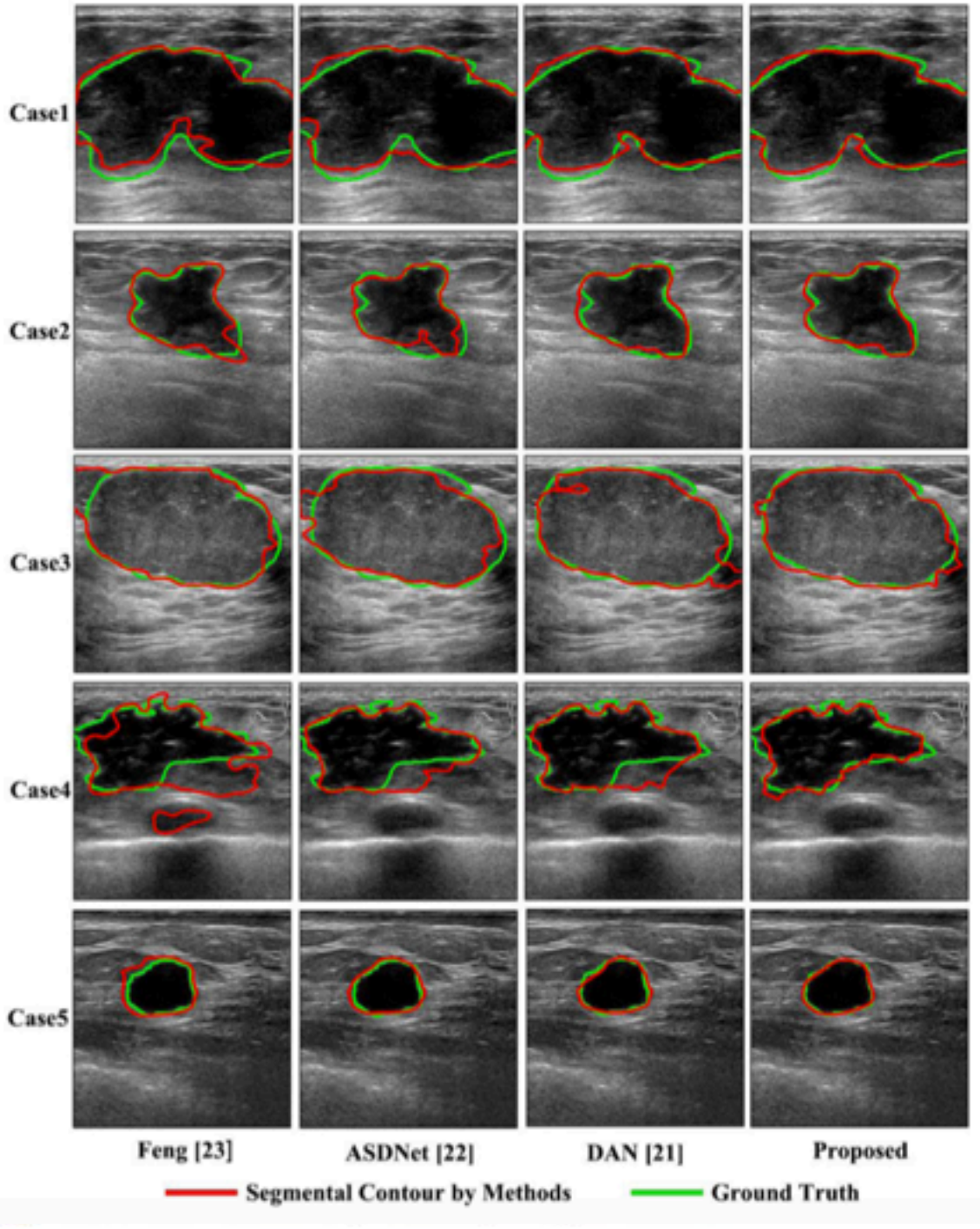
- **Semi- is effective**

Table 2
Testing results of the BUS-GAN model by training with a fixed number of annotated images and a variable number of unannotated images. The segmentation results of the testing data are evaluated in terms of area overlap and boundary error. (Dataset A: in-house testing dataset, Dataset B: public testing dataset).

#images involved		Area Overlap						Boundary Error			
		DSC (%)		JI (%)		Precision (%)		HD (pixels)		AvgD (pixels)	
Anno.	Unanno.	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B
100	100	84.31 (8.482)	79.08 (11.25)	74.01 (11.18)	67.68 (15.17)	83.85 (12.83)	78.44 (16.52)	25.87 (18.49)	33.19 (34.42)	6.105 (5.182)	9.264 (13.28)
100	500	85.37 (7.726)	79.25 (10.86)	75.75 (10.90)	67.57 (15.21)	84.65 (11.33)	78.38 (16.23)	24.50 (17.87)	32.96 (33.82)	5.940 (5.058)	9.201 (13.39)
100	1000	86.50 (6.710)	79.53 (10.80)	76.51 (9.837)	67.89 (15.03)	85.63 (11.46)	78.42 (16.02)	24.43 (16.16)	32.65 (32.53)	5.872 (5.641)	9.126 (12.91)
100	1500	86.75 (6.383)	79.75 (10.35)	76.80 (9.023)	68.00 (14.87)	86.02 (11.24)	78.50 (15.75)	24.37 (16.54)	32.23 (31.74)	5.805 (5.550)	9.098 (12.44)
100	1900	**87.12 (5.650)**	**79.82 (10.38)**	**77.62 (8.642)**	**68.03 (14.98)**	**86.38 (10.79)**	**78.58 (15.95)**	**23.61 (15.76)**	**32.15 (31.85)**	**5.364 (4.902)**	**9.089 (12.55)**

Result

- **Comparison with other methods**



Result

- **Comparison with other methods**

Table 3
Comparison between the proposed BLS-GAN and other semi-supervised learning methods with five evaluation metrics, including DSC, JI, Precision, HD and AvgD. (Dataset A: in-house testing dataset, Dataset B: public testing dataset).

Methods	Area Overlap						Boundary Error			
	DSC (%)		JI (%)		Precision (%)		HD (pixels)		AvgD (pixels)	
	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B
Feng [23]	83.06 (8.620)	76.37 (15.64)	72.89 (11.86)	65.72 (19.94)	82.85 (13.49)	76.88 (17.39)	28.79 (18.99)	34.98 (32.79)	7.150 (6.252)	9.328 (14.68)
ASDNet [22]	85.00 (7.544)	77.50 (14.67)	74.78 (11.71)	66.89 (16.32)	84.08 (12.56)	77.22 (16.88)	26.03 (18.48)	33.68 (33.84)	6.729 (5.519)	9.749 (14.59)
DAN [21]	85.20 (7.331)	77.62 (11.77)	75.42 (10.51)	67.13 (17.62)	84.28 (11.52)	77.54 (16.43)	25.5[illegible] (18.33)	34.05 (33.52)	5.894 (6.123)	9.857 (14.84)
Proposed	**87.12 (5.650)**	**79.82 (10.38)**	**77.62 (8.642)**	**68.03 (14.98)**	**86.38 (10.79)**	**78.58 (15.95)**	**23.61 (15.76)**	**32.15 (31.85)**	**5.364 (4.902)**	**9.089 (12.55)**

Ablation Study

Table 4
Quantitative analysis of segmentation performance from different settings of the proposed BUS-GAN. (Dataset A: in-house testing dataset, Dataset B: public testing dataset).

Backbone	+BUS-F	+Semi-	+Attention	Area Overlap						Boundary Error			
				DSC (%)		Jl (%)		Precision (%)		HD (pixels)		AvgD (pixels)	
				Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B
BUS-S				83.11 (9.094)	77.81 (13.36)	72.44 (11.90)	67.54 (15.26)	79.55 (13.33)	77.94 (16.58)	29.32 (19.52)	34.25 (34.52)	6.894 (5.914)	9.989 (14.56)
	✓			83.45 (9.102)	78.97 (12.02)	72.87 (12.08)	67.67 (15.65)	83.34 (13.48)	77.34 (17.41)	28.92 (19.02)	33.98 (35.36)	6.687 (6.256)	9.613 (15.25)
	✓		✓	83.88 (9.010)	78.94 (11.10)	73.01 (10.80)	67.49 (15.35)	84.55 (13.73)	78.55 (16.39)	28.32 (18.88)	33.32 (34.08)	6.254 (5.953)	9.258 (13.93)
	✓	✓		85.75 (8.255)	79.32 (11.55)	75.87 (11.37)	67.98 (15.17)	82.72 (12.73)	78.32 (16.71)	26.28 (18.65)	33.01 (33.65)	5.977 (5.914)	9.123 (13.94)
	✓	✓	✓	**87.12 (5.650)**	**79.82 (10.38)**	**77.62 (8.642)**	**68.03 (14.98)**	**86.38 (10.79)**	**78.58 (15.95)**	**23.61 (15.26)**	**32.15 (31.85)**	**5.364 (4.902)**	**9.089 (12.55)**

Ablation Study

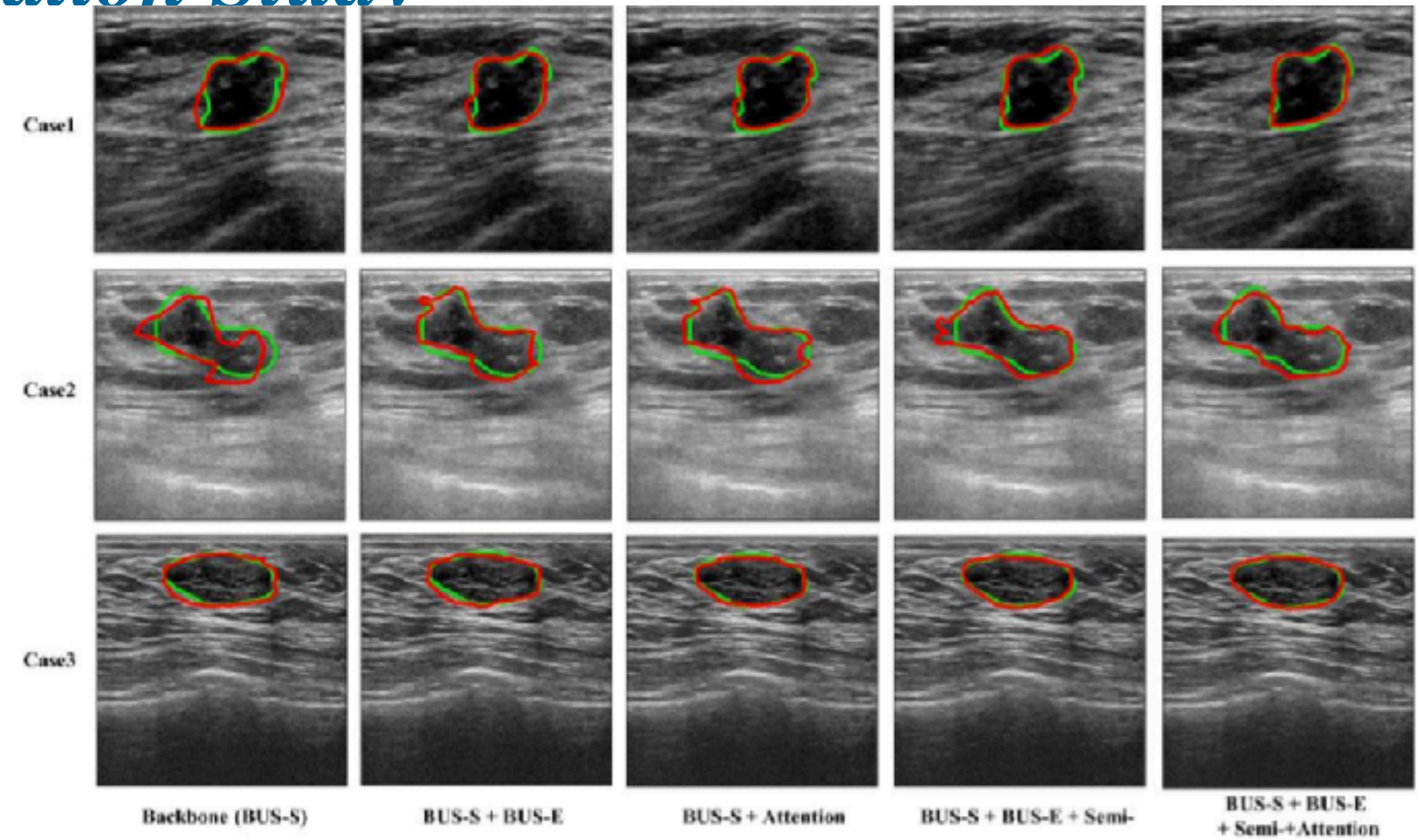


Fig. 6. Sample cases resulted from experiments by adding different elements listed in Table 4. From the first to the fifth column, each column exhibits the predicted contours from experiment "Backbone (BUS-S)", "BUS-S + BUS-E", "BUS-S + Attention", "BUS-S + BUS-E + Semi-", "BUS-S + BUS-E + Semi-+Attention", respectively. Case1 and Case2 derived from the in-house testing dataset and Case3 resulted from the public dataset.

Result

- **Effectiveness of the segmentation backbone**

Table 5
Accuracy comparison between the backbone BUS-S and other fully supervised learning-based methods with five evaluation metrics, including DSC, JI, Precision, HD, and AvgD. (Dataset A: in-house testing dataset. Dataset B: public testing dataset).

Method	Area Overlap						Boundary Error			
	DSC (%)		JI (%)		Precision (%)		HD (pixels)		AvgD (pixels)	
	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B
FCN-16s	79.33 (13.66)	70.86 (16.92)	67.56 (17.24)	60.44 (19.25)	78.90 (15.73)	71.85 (19.71)	33.10 (26.65)	37.43 (35.63)	8.992 (12.06)	11.21 (15.78)
U-Net	80.84 (12.35)	75.40 (15.74)	69.22 (16.11)	63.63 (18.36)	**80.29 (18.04)**	**79.01 (18.04)**	31.85 (25.35)	34.34 (33.05)	8.430 (11.38)	10.93 (14.32)
DeepLab	81.91 (11.21)	77.44 (13.08)	70.98 (12.80)	67.26 (15.80)	79.20 (13.74)	77.58 (16.03)	31.59 (21.05)	35.03 (34.37)	8.034 (9.684)	10.06 (13.84)
BUS-S (backbone)	**83.11 (9.054)**	**77.81 (13.36)**	**72.44 (11.90)**	**67.54 (15.26)**	79.55 (13.33)	77.94 (16.68)	**29.32 (19.52)**	**34.25 (34.52)**	**6.894 (5.914)**	**9.989 (14.56)**

Result

- **Effectiveness of the segmentation backbone**

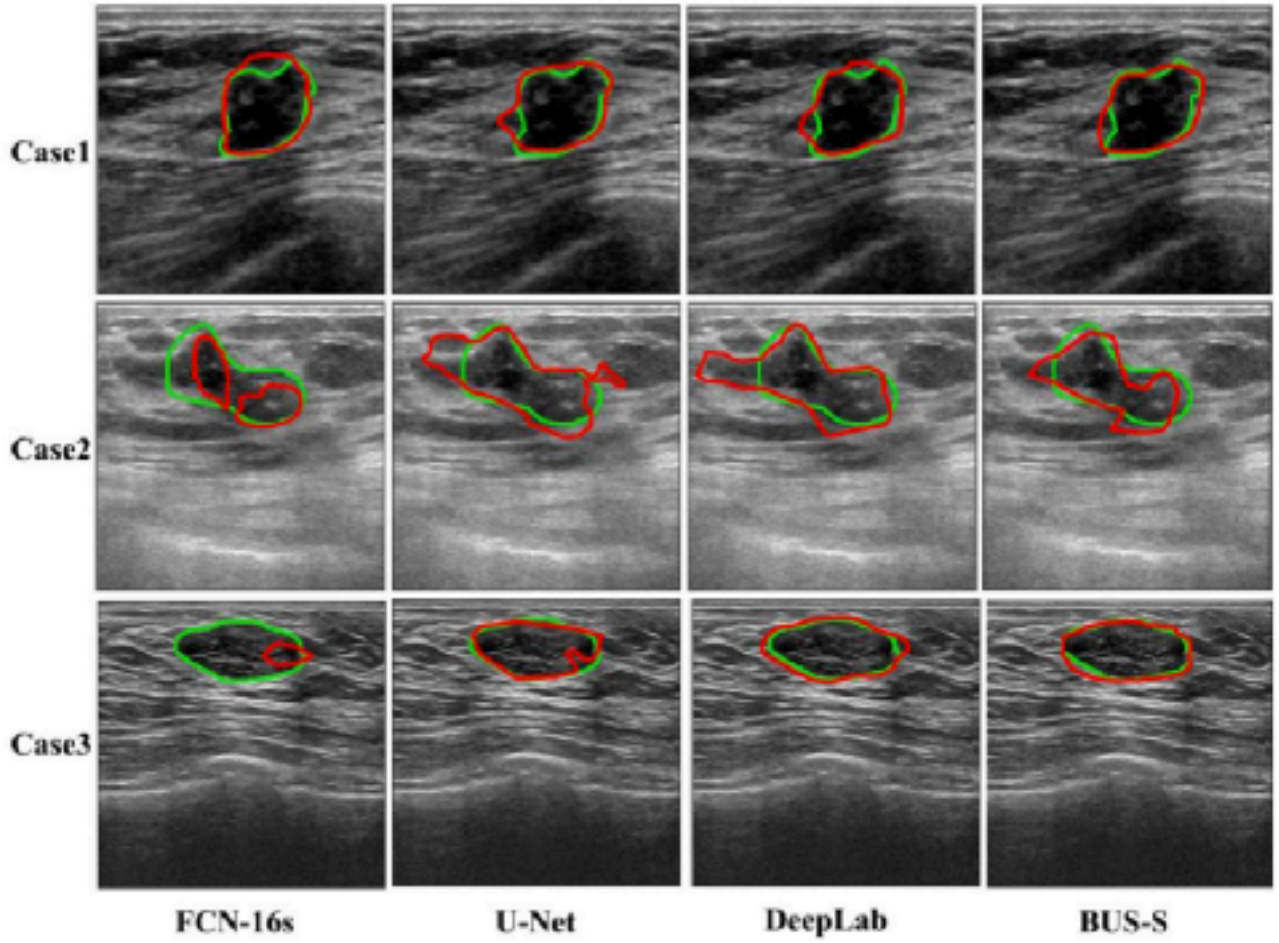


Fig. 7. Sample cases resulted from different fully supervised learning based methods. Each column corresponds to the predicted contours from "FCN", "U Net", "DeepLab", "BUS S", respectively (1st–4th column). Case1 and Case2 are sample cases from in house testing images and Case3 is a sample case from the public dataset.

Result

- **Impact of the Dual-Attentive-Fusion Block**
-

Table 6
Comparisons of the results from different inputs of the proposed DAF block and those from different dual attention methods with five evaluation metrics, including DSC, JI, Precision, HD, and AvgD. (Dataset A: in-house testing dataset, Dataset B: public testing dataset).

Method	Area Overlap						Boundary Error			
	DSC (%)		JI (%)		Precision (%)		HD (pixels)		AvgD (pixels)	
	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B	Dataset A	Dataset B
Seg-only	83.35 (8.758)	78.06 (12.81)	72.33 (11.94)	67.61 (14.85)	82.35 (13.27)	78.00 (16.33)	28.20 (18.17)	33.82 (33.43)	6.634 (5.120)	9.635 (14.08)
Concatenation	85.75 (8.255)	79.32 (11.55)	75.87 (11.37)	67.98 (14.17)	82.72 (12.73)	78.32 (16.71)	26.28 (18.65)	33.01 (33.65)	5.977 (5.914)	9.123 (13.94)
No-inverse	86.32 (8.012)	79.57 (10.99)	76.71 (11.14)	68.01 (15.12)	83.40 (11.90)	78.34 (16.08)	25.66 (18.81)	32.24 (32.31)	5.869 (5.622)	9.201 (13.34)
DANet	86.07 (7.908)	79.43 (10.58)	76.18 (11.10)	68.00 (14.83)	84.31 (12.18)	78.27 (15.73)	24.30 (18.34)	32.42 (32.08)	5.722 (5.233)	9.184 (12.78)
CBAM	86.36 (8.402)	79.69 (10.65)	76.72 (10.16)	68.01 (15.01)	84.02 (11.22)	78.39 (16.32)	25.87 (18.76)	32.59 (31.35)	6.191 (6.486)	9.137 (12.86)
Proposed	**87.12 (5.650)**	**79.82 (10.38)**	**77.62 (8.642)**	**68.03 (14.98)**	**86.38 (10.79)**	**78.58 (15.95)**	**23.61 (15.76)**	**32.15 (31.85)**	**5.364 (4.902)**	**9.089 (12.55)**

Result

- **Impact of the Dual-Attentive-Fusion Block**

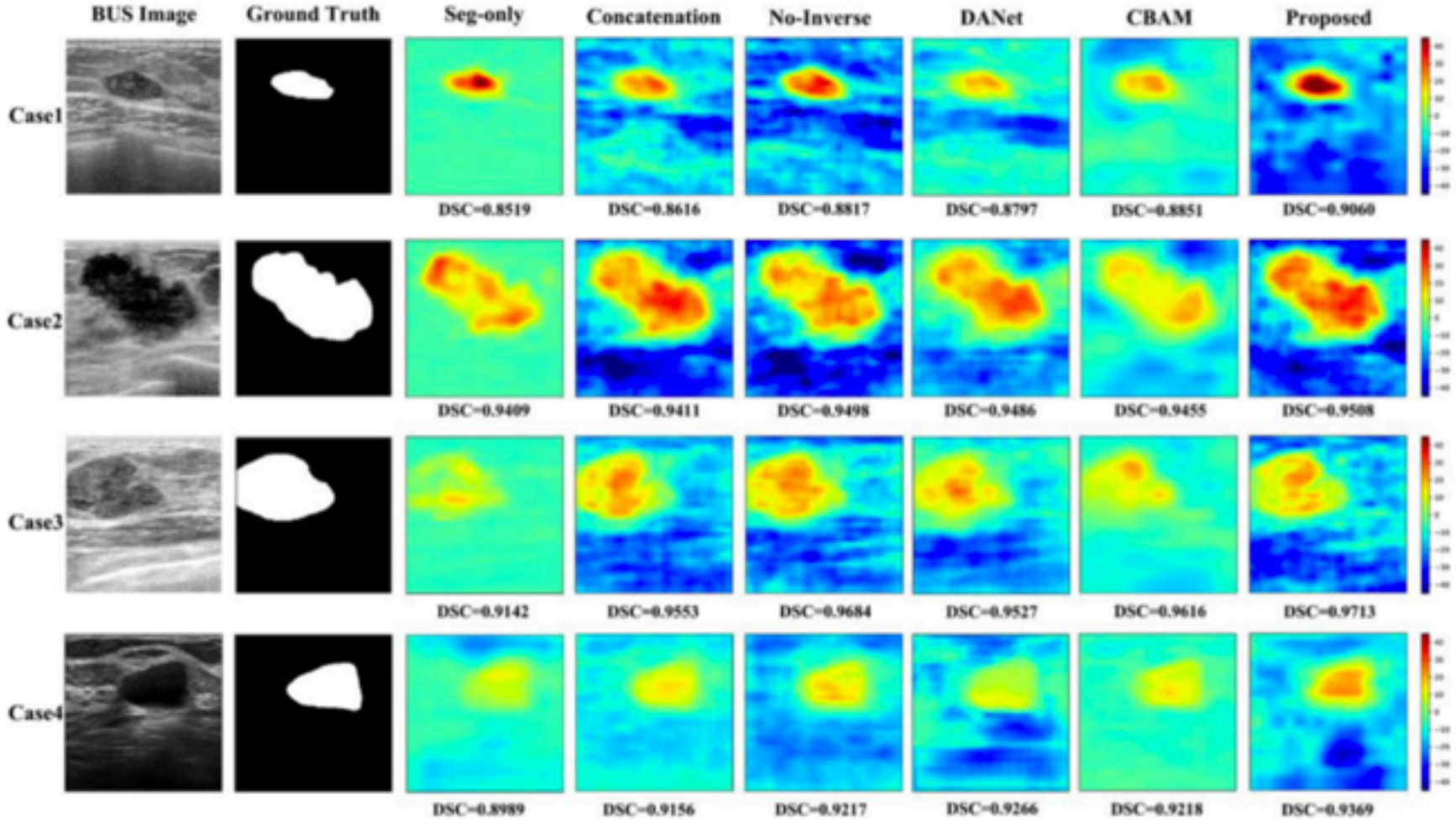


Fig. 8. Samples of the segmentation probability heatmaps resulted from different inputs of the BUS-E network. From the third column to the sixth column, each column represents the segmented probability maps from method "Seg-only", "Concatenation", "No-Inverse", "DANet", "CBAM" and the proposed method, respectively. Case1-3 correspond to the results from the in-house testing dataset and Case 4 refers to the result from the public dataset.

Contribution

- **Make full use of unannotated BUS images to improve segmentation performance**

- **The proposed BUS-GAN model is tested on multi-site breast ultrasound datasets and achieve the best segmentation accuracy compared with the-state-of-art semi- supervised segmentation methods**

- **A developed dual-attentive-fusion (DAF) block is incorporated into the evaluator of the segmentation mode**