7.26 汇报

1、基于U-Net的医学图像多分辨率分割方法

U-Net Based Architecture for an Improved Multiresolution Segmentation in Medical Images

2、基于多源弱监督的显著性检测

Multi-source weak supervision for saliency detection

基于U-Net的医学图像多分辨率分割方法

目标改善U-Net架构下的多分辨率图像分割性能。

mrU-Net与U-Net相比，该结构具有更高的提取图像特征的能力。与U-Net相比，mrU-Net具有更快的训练速度和更精确的图像分割。

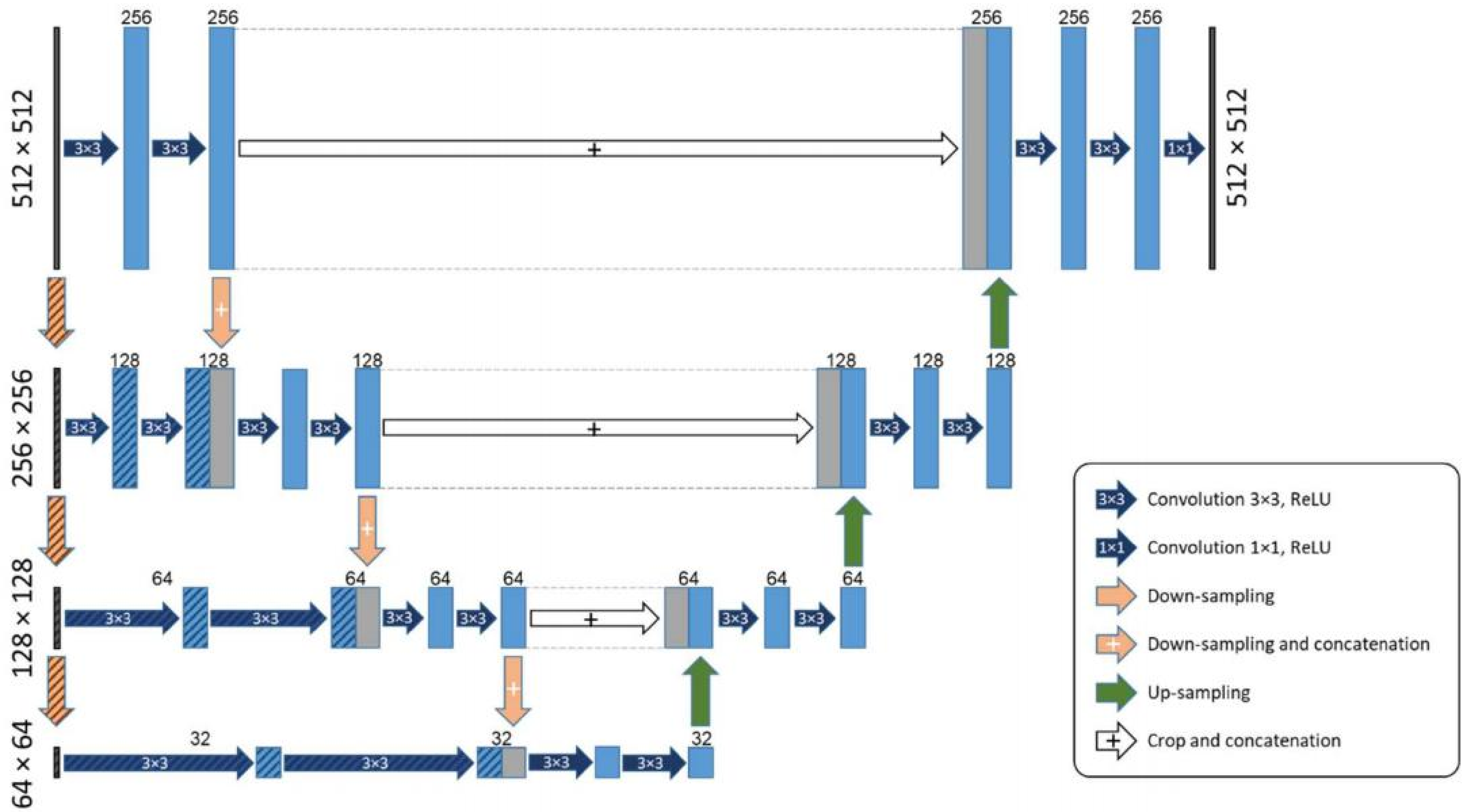


Fig. 2 Four-level mrU-Net architecture. The hashed layers are added layers to the U-Net architecture. The numbers above the feature maps indicate the number of feature channels. The size of the original input image is 512×512 and it has either one channel for grayscale images or three channels for RGB images.

基于U-Net的医学图像多分辨率分割方法

皮肤损伤照片

肺部CT图像（LUNA数据集）
视网膜图像（DRIVE数据集）
前列腺磁共振（MR）图像（PROMISE12数据集）

Table 2 The number of iterations for training the networks.

Dataset	**Network Model**	**Number of iterations (epochs)**
Skin lesion	U-Net	1700
	mrU-Net	1450
LUNA	U-Net	4950
	mrU-Net	4850
DRIVE	U-Net	4800
	mrU-Net	4950
PROMISE12	U-Net	4750
	mrU-Net	4700

基于U-Net的医学图像多分辨率分割方法

皮肤损伤照片

肺部CT图像（LUNA数据集）
视网膜图像（DRIVE数据集）
前列腺磁共振（MR）图像（PROMISE12数据集）

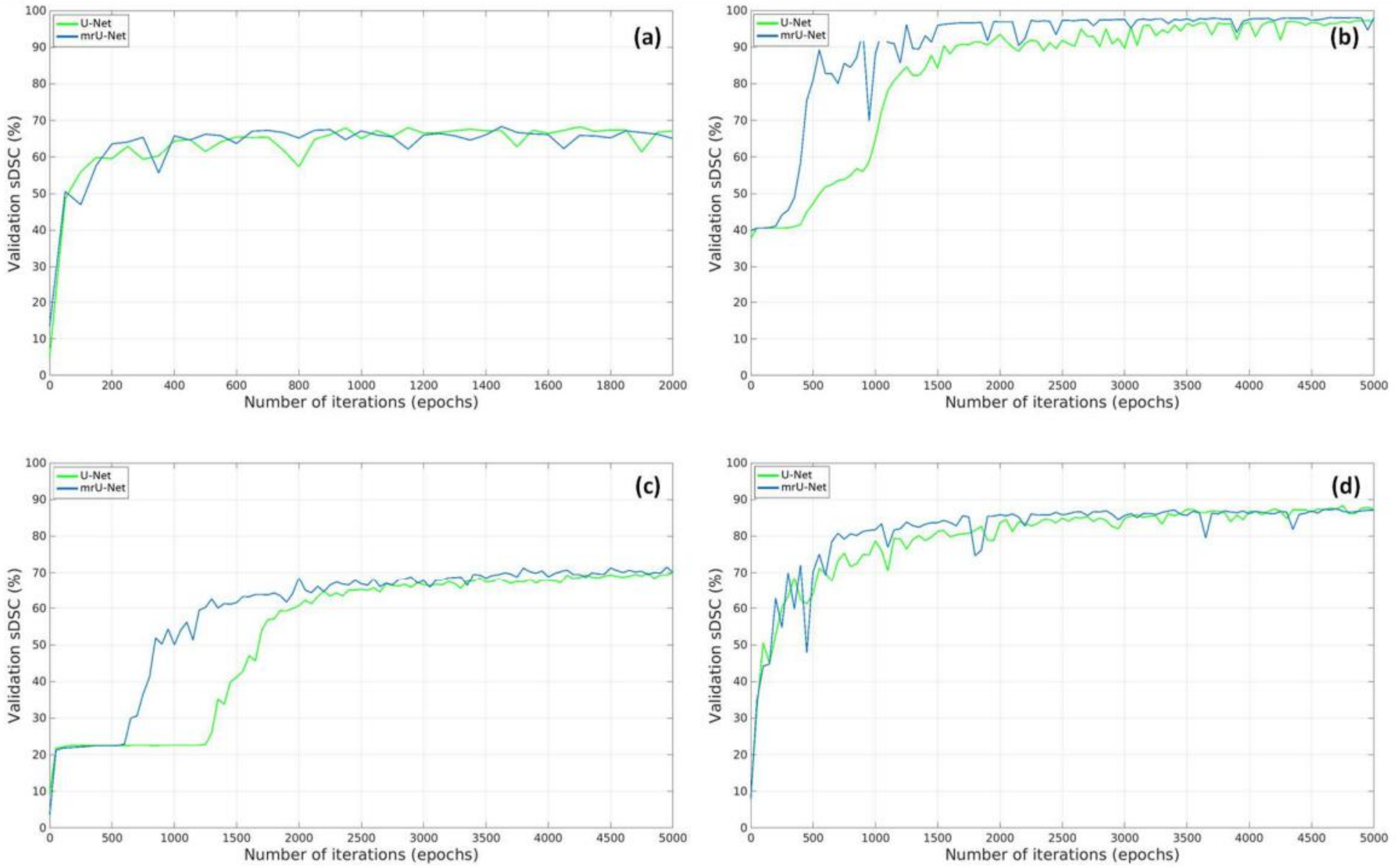


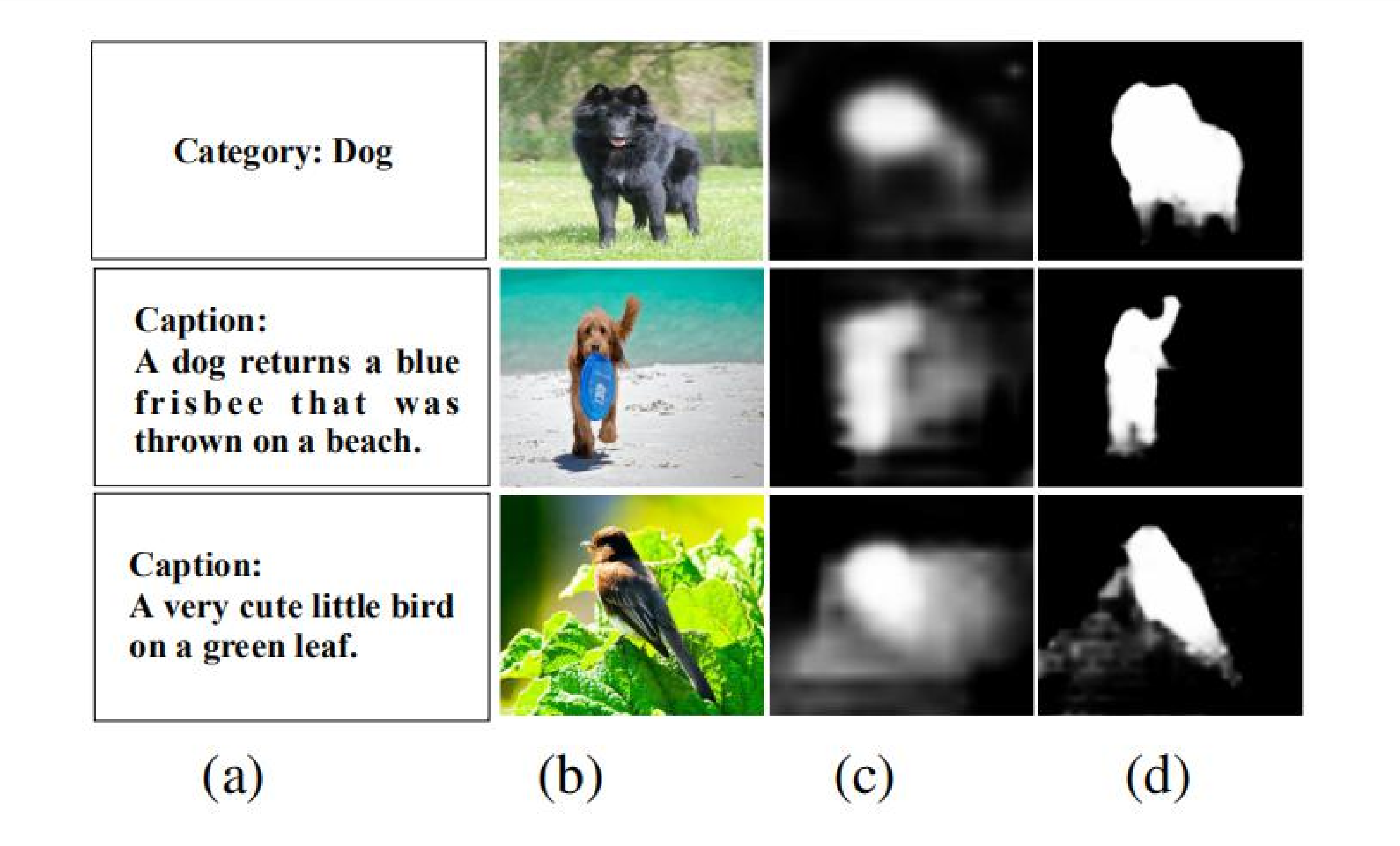
Fig. 3 Validation sDSC for (a) skin, (b) LUNA, (c) DRIVE, and (d) PROMISE12 datasets during training.

基于多源弱监督的显著性检测 CVPR2019

类别标签

标题文本

未标记的数据



基于多源弱监督的显著性检测 CVPR2019

分类网络

CNET

字幕生成网络

PNET

显著性预测网络

SNET

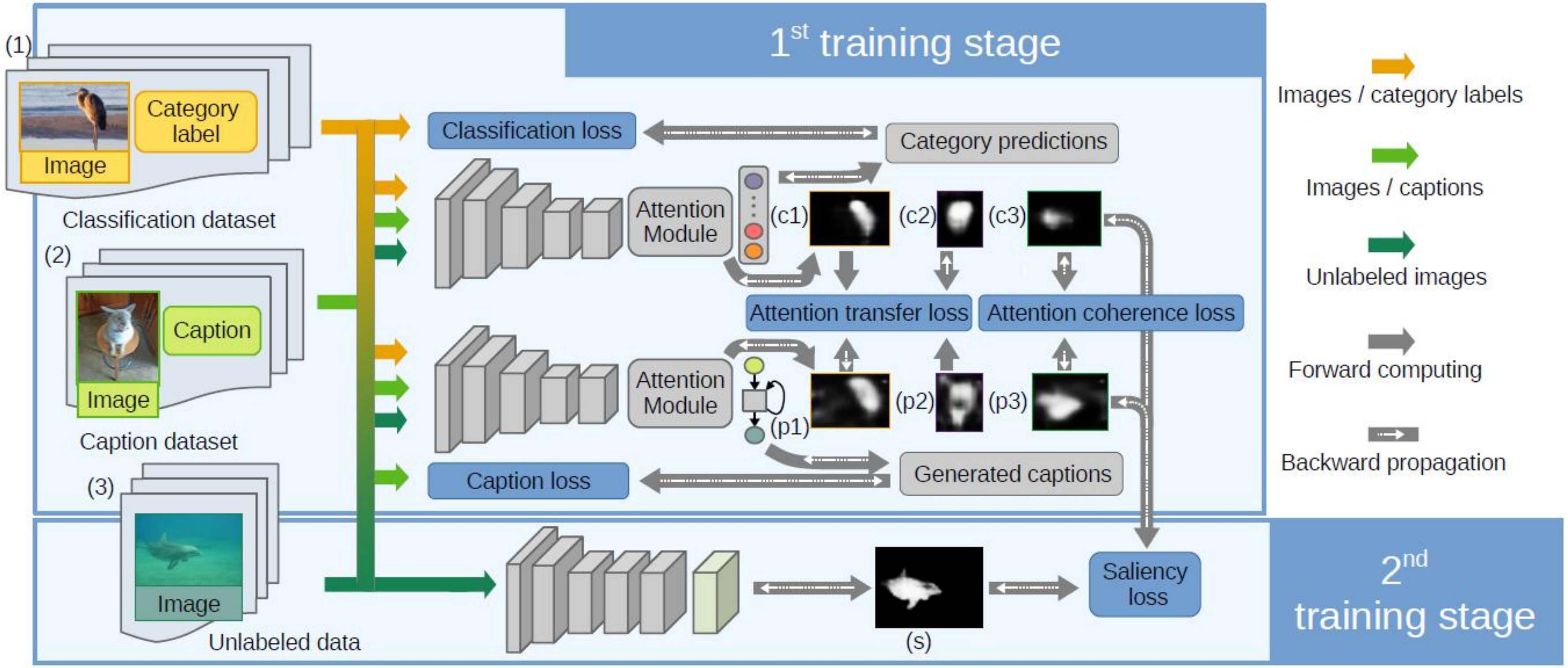


Figure 2. An overview of the proposed multi-source weak supervision framework. (1, 2, 3) images annotated with category labels, caption annotations, and unlabelled images. (c1, c2, c3) saliency maps of images (1, 2, 3) generated by the classification network (CNet). (p1, p2, p3) saliency maps of images (1, 2, 3) generated by the caption generation network (PNet). (s) Final output saliency maps.

基于多源弱监督的显著性检测 CVPR2019

Feature extractors

特征提取

DenseNet-169

Gao Huang, Zhuang Liu, Laurens V an Der Maaten, and Kil-ian Q Weinberger. Densely connected convolutional net-works. In IEEE conference on computer vision and pattern recognition, 2017.

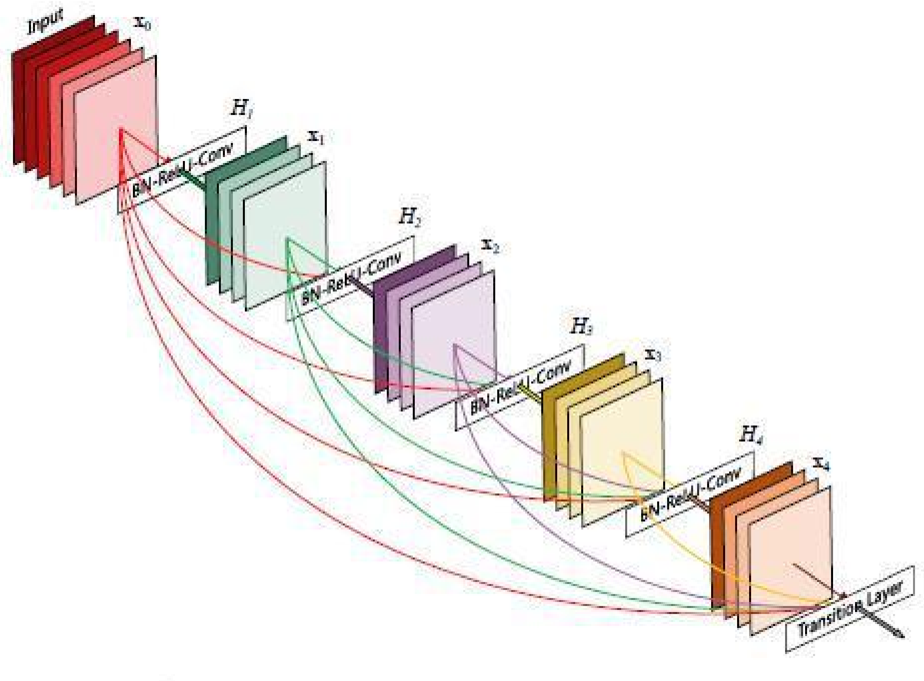


Figure 1. A 5-layer dense block with a growth rate of $k = 4$. Each layer takes all preceding feature-maps as input.

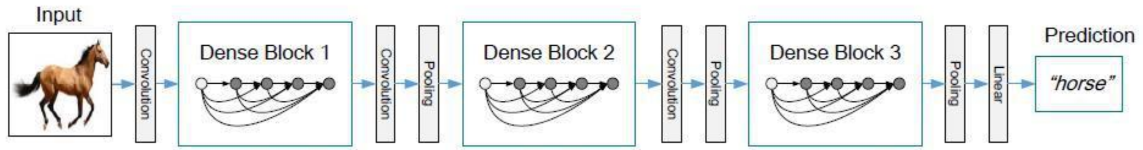


Figure 2. A deep DenseNet with three dense blocks. The layers between two adjacent blocks are referred to as transition layers and change feature map sizes via convolution and pooling.

Layers	Output Size	DenseNet-121($k = 32$)	DenseNet-169($k = 32$)	DenseNet-201($k = 32$)	DenseNet-161($k = 48$)
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 36$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

基于多源弱监督的显著性检测
CVPR2019

Attention module

注意力模块

$$s_i = \sigma(\boldsymbol{w}_s^\intercal \boldsymbol{v}_i + b_s),$$

$$\boldsymbol{f}_i = s_i \cdot (\boldsymbol{w}_f^\intercal \boldsymbol{v}_i + b_f),$$

$$a_i = \boldsymbol{w}_a^\intercal \boldsymbol{f}_i + b_a$$
$$\boldsymbol{\alpha} = \text{softmax}(\boldsymbol{a}),$$

$$\boldsymbol{g} = \sum_{i=1}^{K} \alpha_i \cdot \boldsymbol{f}_i.$$

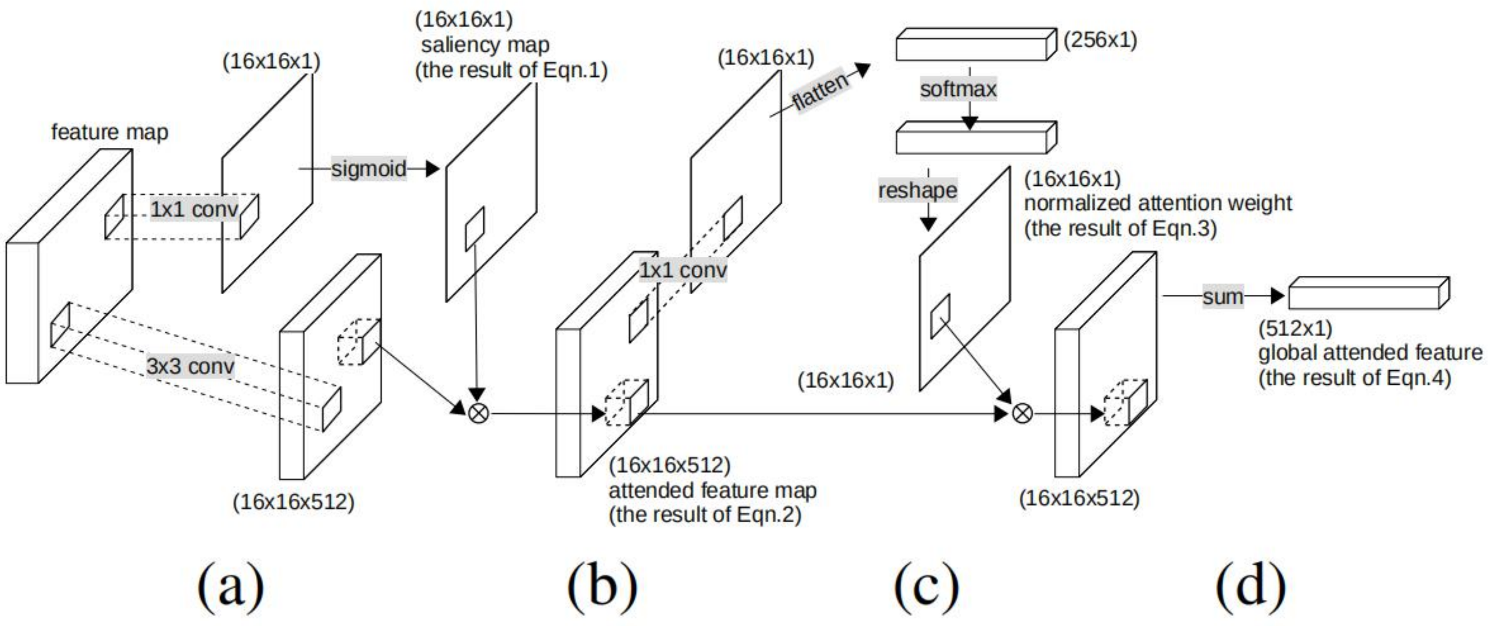


Figure 3. The details of the attention module.

基于多源弱监督的显著性检测 CVPR2019

CNET

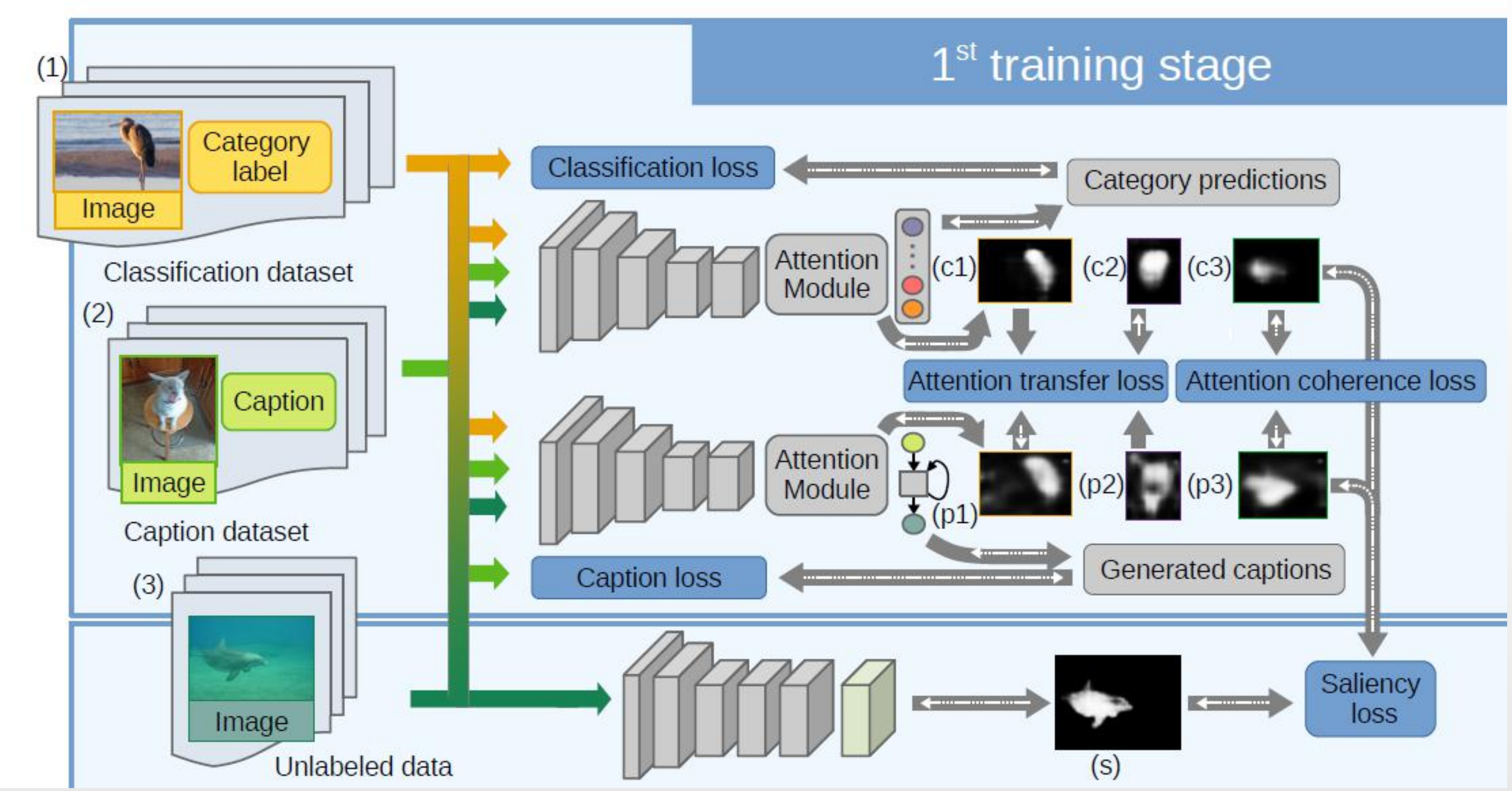
Feature extractors+ Attention module+全连接层

PNET

Feature extractors+ Attention module+LSTM

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term

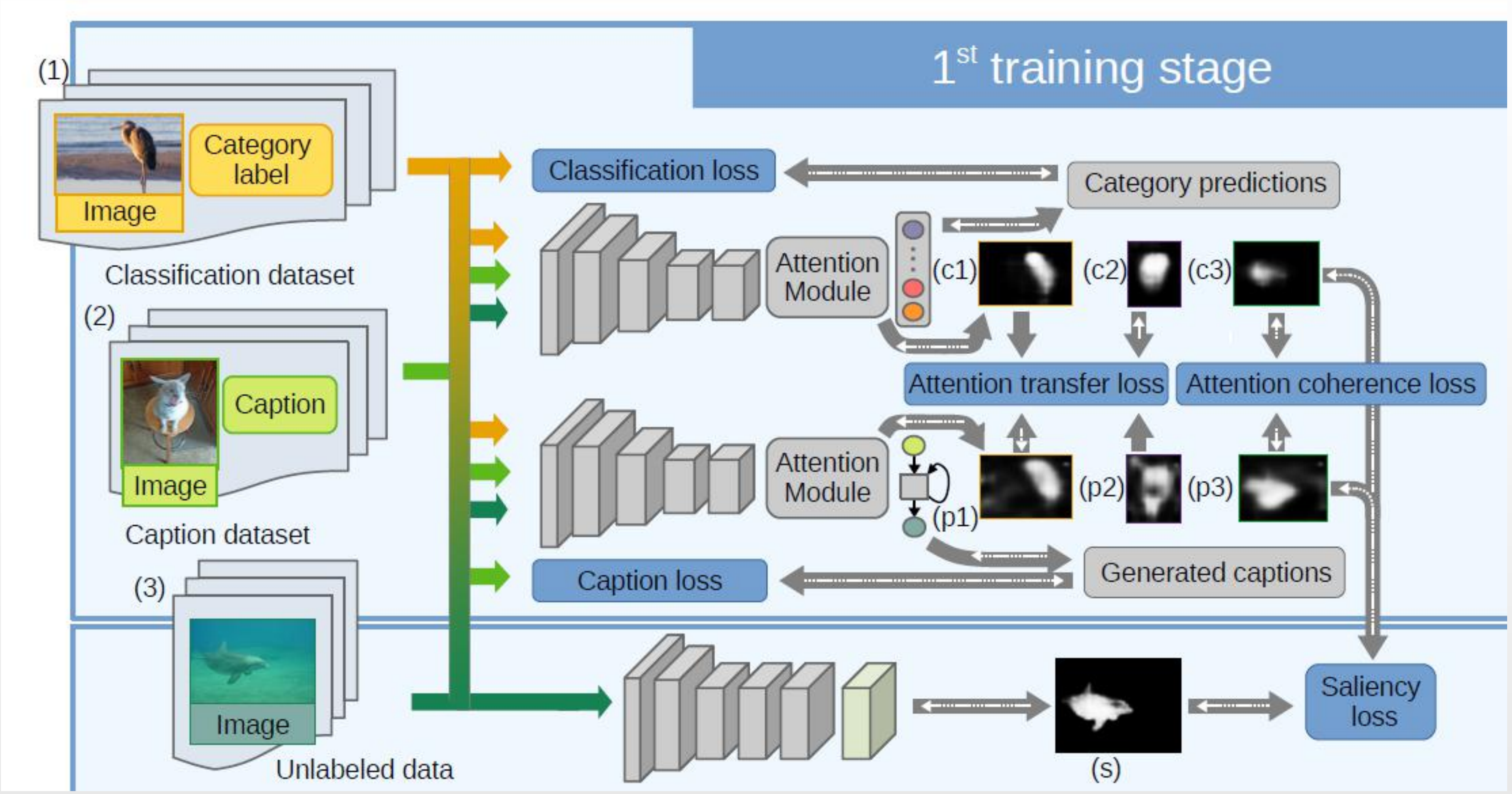
memory. Neural computation, 9(8):1735–1780, 1997.



基于多源弱监督的显著性检测 CVPR2019

SNet

由一个特征抽取器、四个膨胀率分别为6、12、18、24的卷积层和一个反褶积层组成。四个展开卷积层以特征图为输入，预测出四个显著性图。然后将四个显著性图叠加在一起，通过反褶积层对输入图像进行上采样。



基于多源弱监督的显著性检测 CVPR2019

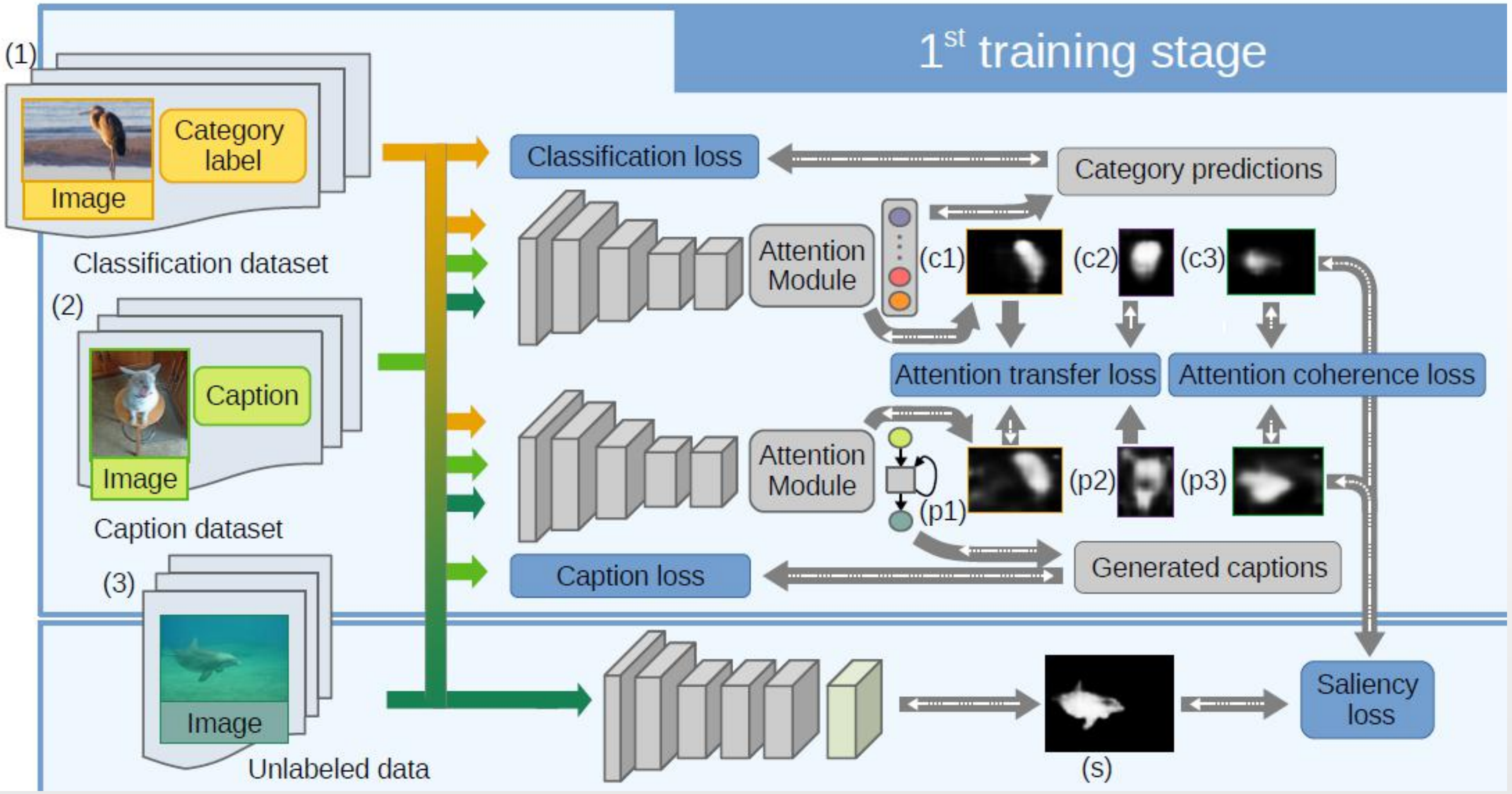
$$s_i = \sigma(\boldsymbol{w}_s^\intercal \boldsymbol{v}_i + b_s),$$

$$L_c = -\frac{1}{N_c} \sum_{(X,\boldsymbol{y})\in\mathcal{D}_c} \left[\sum_{j=1}^{C} \log p(y_j|X) + \beta \sum_{s\in S_c} \log(1-s) \right]$$

$$L_p = -\frac{1}{N_p} \sum_{(X,y_{1:T})\in\mathcal{D}_p} \left[\sum_{t=1}^{T} \log p(y_t|y_{1:t-1}, X) + \beta \sum_{s\in S_p} \log(1-s) \right],$$

$$L_{at} = -\frac{1}{N_c} \sum_{(X,\boldsymbol{y})\in\mathcal{D}_c} \left[\sum_{i\in I_c^+} \log sp_i + \sum_{i\in I_c^-} \log(1-sp_i) \right] - \frac{1}{N_p} \sum_{(X,y_{1:T})\in\mathcal{D}_p} \left[\sum_{i\in I_p^+} \log sc_i + \sum_{i\in I_p^-} \log(1-sc_i) \right]$$

$$L_{ac} = -\frac{1}{N_u} \sum_{X\in\mathcal{D}_u} \left[\sum_{i\in I_u^+} \log sc_i + \log sp_i + \sum_{i\in I_u^-} \log(1-sc_i) + \log(1-sp_i) \right].$$



$$L = L_c + L_p + \lambda L_{at} + \lambda L_{ac},$$

$$L_b(S,Y) = -\sum_i [\delta y_i + (1-\delta)a_i] \log s_i + [\delta(1-y_i) + (1-\delta)(1-a_i)] \log(1-s_i),$$

基于多源弱监督的显著性检测
CVPR2019

数据集

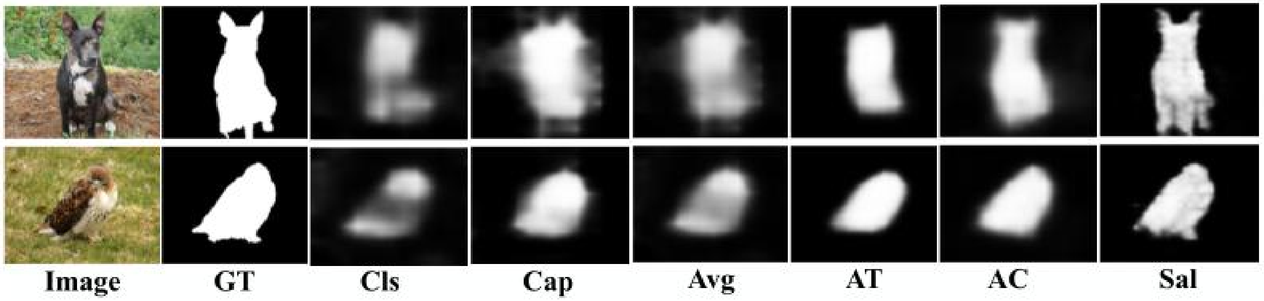
ECSSD

PASCAL-S

SOD

MSRA5K

DUT-OMRON



GT：基本事实。
Cls:CNet训练的结果。
Cap：用字幕定位丢失Lp训练PNet的结果。

Avg：Cls和Cap的平均结果。

AT：使用注意转移损失Lat对两个网络进行联合训练。
AC：联合训练两个网络，并使用未标记的数据进行正则化。未标记数据上的损失是注意力一致性损失Lac。

Sal：用CNet和PNet生成的伪标签训练SNet。

Cls	Cap	AT	AC	Sal	max F_β
√					0.720
	√				0.730
√	√				0.762
√	√	√			0.786
√	√	√	√		0.820
√	√	√	√	√	0.878

基于多源弱监督的显著性检测
CVPR2019

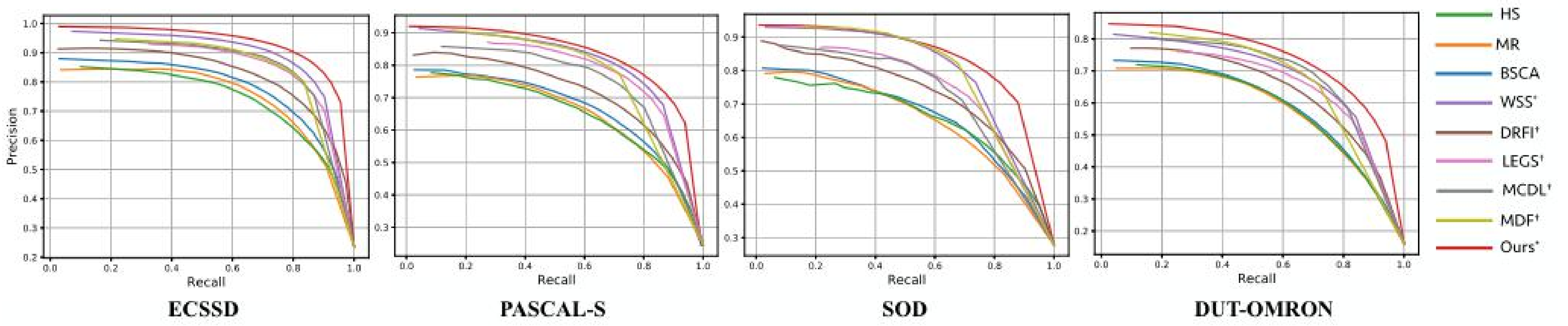


Table 2. Comparison with weakly supervised (marked with * and unsupervised methods in terms of maximum F-measure (the larger the better) and MAE (the smaller the better). The best scores are in bold.

	ECSSD		PASCAL-S		SOD		MSRA5K		DUT-OMRON	
Methods	max F_β	MAE	max F_β	MAE	max F_β	MAE	max F_β	MAE	max F_β	MAE
BSCA	0.758	0.182	0.663	0.223	0.656	0.252	0.829	0.132	0.613	0.196
MB+	0.736	0.193	0.673	0.228	0.658	0.255	0.822	0.133	0.621	0.193
MST	0.724	0.155	0.657	0.194	0.647	0.223	0.809	0.098	0.588	0.161
MR	0.742	0.186	0.650	0.232	0.644	0.261	0.821	0.128	0.608	0.194
HS	0.726	0.227	0.644	0.264	0.647	0.283	0.815	0.162	0.613	0.233
WSS*	0.856	0.104	0.778	0.141	0.780	0.170	0.877	0.076	0.687	0.118
Ours*	**0.878**	**0.096**	**0.790**	**0.134**	**0.799**	**0.167**	**0.890**	**0.071**	**0.718**	**0.114**