Semi-supervised Domain Generalization for Medical Image Analysis

IEEE ISBI 2022
International Symposium on Biomedical Imaging
28-31 March 2022, ITC Royal Bengal, Kolkata, India

Ruipeng Zhang[1,2], Qinwei Xu[1,2], Chaoqin Huang[1,2], Ya Zhang[1,2], Yanfeng Wang[1,2]

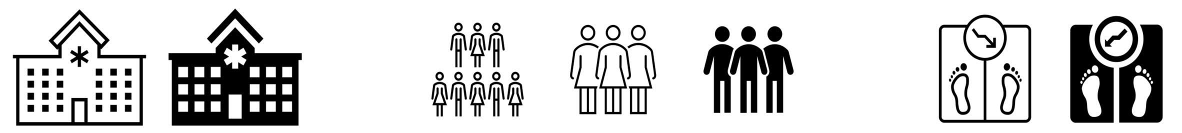
[1] Cooperative Medianet Innovation Center, Shanghai Jiao Tong University [2] Shanghai AI Laboratory

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

上海人工智能实验室 Shanghai Artificial Intelligence Laboratory

Introduction:

Domain shift in medical applications:

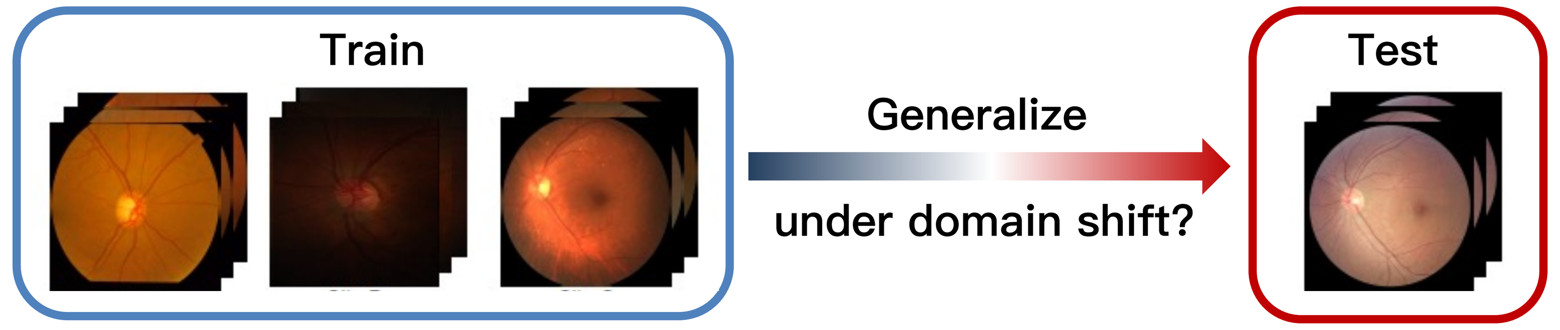


Cause Limited generalization ability for machine learning methods

Performance degradation under domain shift –> unreliable for medical

Domain Generalization (DG)

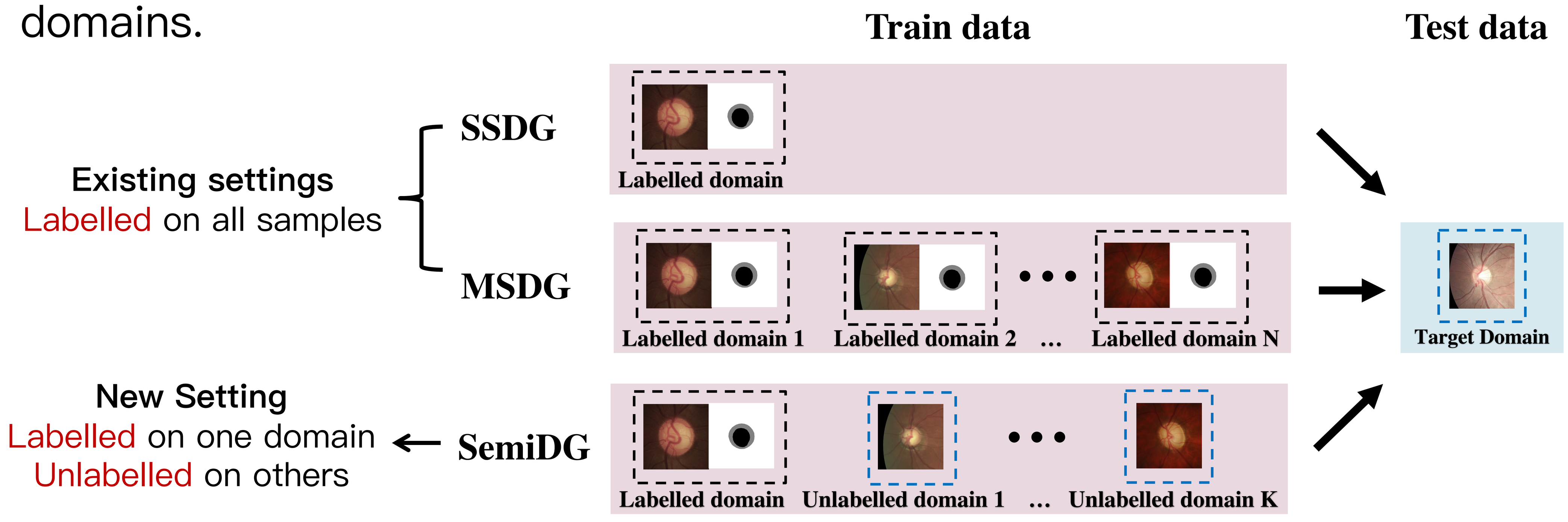
How to learn generalizable models under domain shift?



Semi-supervised Domain Generalization (SemiDG):

Single-Source DG / Multi-Source DG: all source domain(s) are labelled.

SemiDG: One labelled source domain and several unlabelled source domains.



SemiDG only needs annotations from one domain
and aims to achieve comparable performance as MSDG.

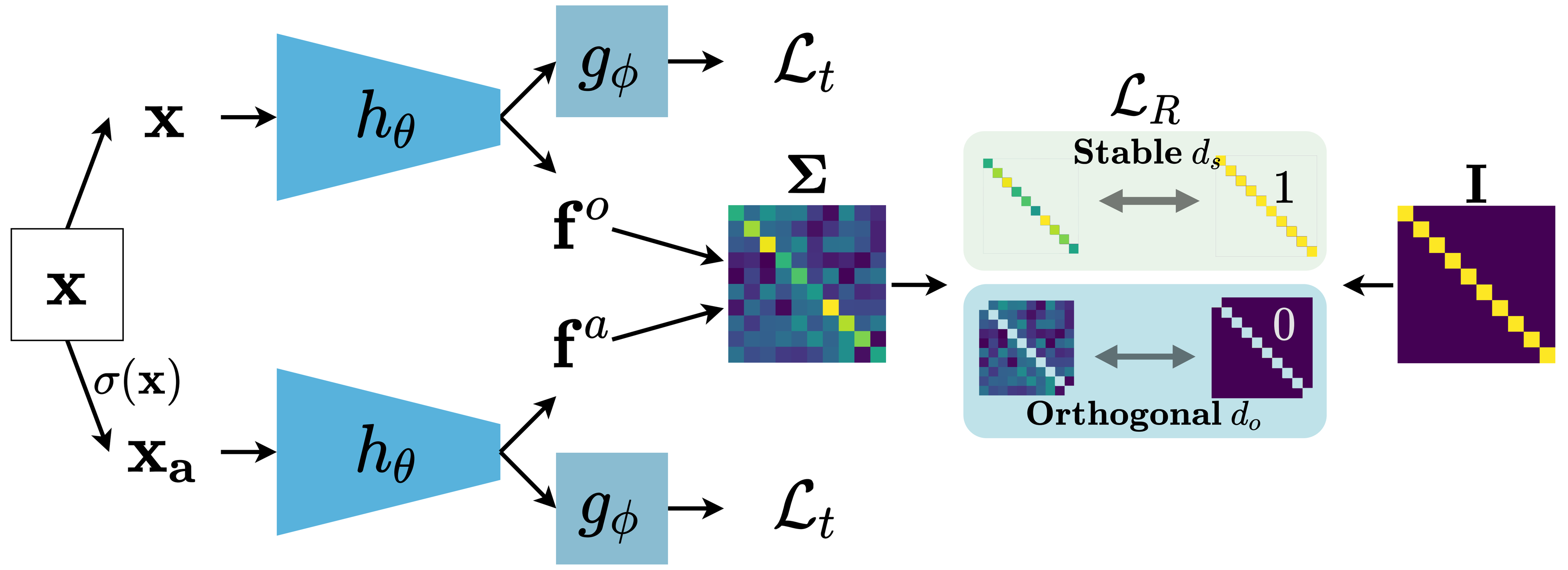
Method:

Stable and Orthogonal Feature Regularization (SOFR)

Two characteristics to achieve generalizable on features

- **Stability**: be robust across different domains. $\rightarrow$ $d_s = \sum_{i=1}^{C}(1-\cos(\mathbf{f}_i^o, \mathbf{f}_i^a))$
- **Orthogonality**: be of minimal redundancy to avoid overfitting $\rightarrow$ $d_o = \sum_{i \neq j} |\cos(\mathbf{f}_i^o, \mathbf{f}_j^o)|.$

SOFR: $$\mathcal{L}_R(\mathbf{f}^o, \mathbf{f}^a) = \overbrace{\sum_{i=j}(1-\cos(\mathbf{f}_i^o, \mathbf{f}_i^a))}^{\text{stablility}} + \overbrace{\sum_{i \neq j}|\cos(\mathbf{f}_i^o, \mathbf{f}_j^a)|}^{\text{orthogonality}},$$



Stable ✓ and Orthogonal ✗–>dummy solutions on unlabelled samples!

Stable ✗ and Orthogonal ✓–>hard to obtain domain-invariant predictions!

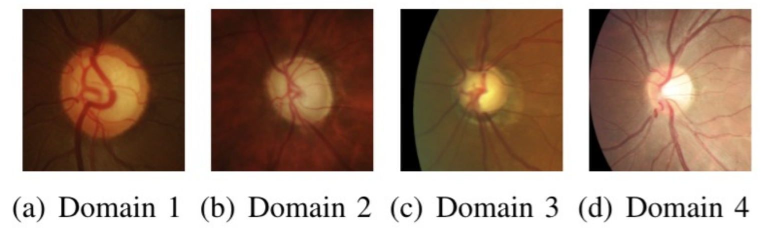
New Benchmarks for SemiDG:

Source domains for train:

- One labelled + Two unlabelled

Target domain for test:

- The left one domain



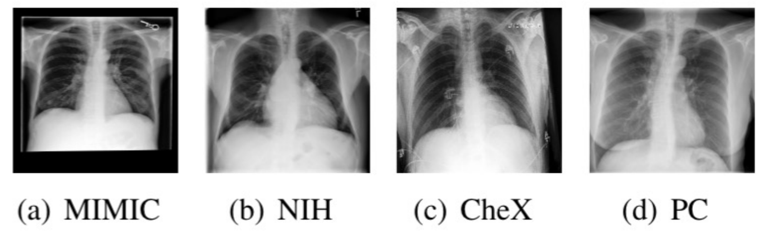
Fundus Images: Optic Cup/Disc Segmentation Task

Evaluation for Fundus benchmark:

- Averaged on 12 different permutation of domains

Evaluation for Chest X-Ray benchmark:

- Train: MIMIC (labelled) NIH+CheX (unlabelled)
- Test: PC



Chest X-Ray Images: Diagnosis Prediction Task

Experimental Results:

Table 1: Comparison with baselines on two benchmarks.

Method	Fundus (DSC)			CXR (mAUC)
	Disc	Cup	avg	
SSDG	0.8434	0.6662	0.7548	0.8166
MT	0.8441	0.6771	0.7606	0.8149
CORAL	0.8466	0.6781	0.7624	0.8248
RSC	0.8470	0.6793	0.7631	0.7897
Self-training	0.8499	0.6863	0.7681	0.7955
DANN	0.8543	0.6887	0.7715	0.8284
TCSM	0.8596	0.6853	0.7725	-
EntMin	0.8633	0.6908	0.7771	0.7993
SOFR + Jitter	0.8634	0.7055	0.7844	0.8334
SOFR + AM	**0.8822**	**0.7277**	**0.8050**	**0.8443**
MSDG	*0.9125*	*0.7822*	*0.8473*	*0.8317*

Achieve high performance on different domain randomizations!

Perform better than MSDG with less annotations!
(On CXR benchmark)

Table 2: The analysis of different components in SOFR on Fundus (Using AM as domain randomization).

SFR	OFR	Self-OFR	Disc DSC	Cup DSC	avg
-	-	-	0.8617	0.7039	0.7828
✓	-	-	0.8663	0.7064	0.7864
-	✓	-	0.8649	0.7186	0.7917
-	-	✓	0.8710	0.7071	0.7891
✓	-	✓	0.8724	0.7244	0.7984
✓	✓	-	**0.8822**	**0.7277**	**0.8050**

With only both stable and orthogonal can achieve high generalizable performance!

Table 3: The impact of samples from unsupervised domains in SOFR on Fundus (Using AM as domain randomization).

#Samples@U1	0	400	0	200	400
#Samples@U2	0	0	400	200	400
avg DSC	0.7185	0.7455	0.7657	0.7711	**0.7730**

Unlabelled samples from more domains are more effective!

Conclusions:

We propose a **new realistic setting** for domain generalization named SemiDG by introducing unlabelled domain into the training process.

We design a **regularization-based** SemiDG method which constrains the feature to be both stable and orthogonal to improve the generalization ability under domain shift.

We propose **two new benchmarks** for SemiDG and the experimental results show the effectiveness of our proposed method.

Contact Us:

Zhang Ruipeng, PhD student, SJTU

Email: zhangrp@sjtu.edu.cn

Homepage: https://mediabrain.sjtu.edu.cn/