Interventional Contrastive Learning with Meta Semantic Regularizer

Wenwen Qiang*, Jiangmeng Li*, Changwen Zheng, Bing Su, Hui Xiong

ISCAS

中国科学院大学
University of Chinese Academy of Sciences

广州海洋实验室

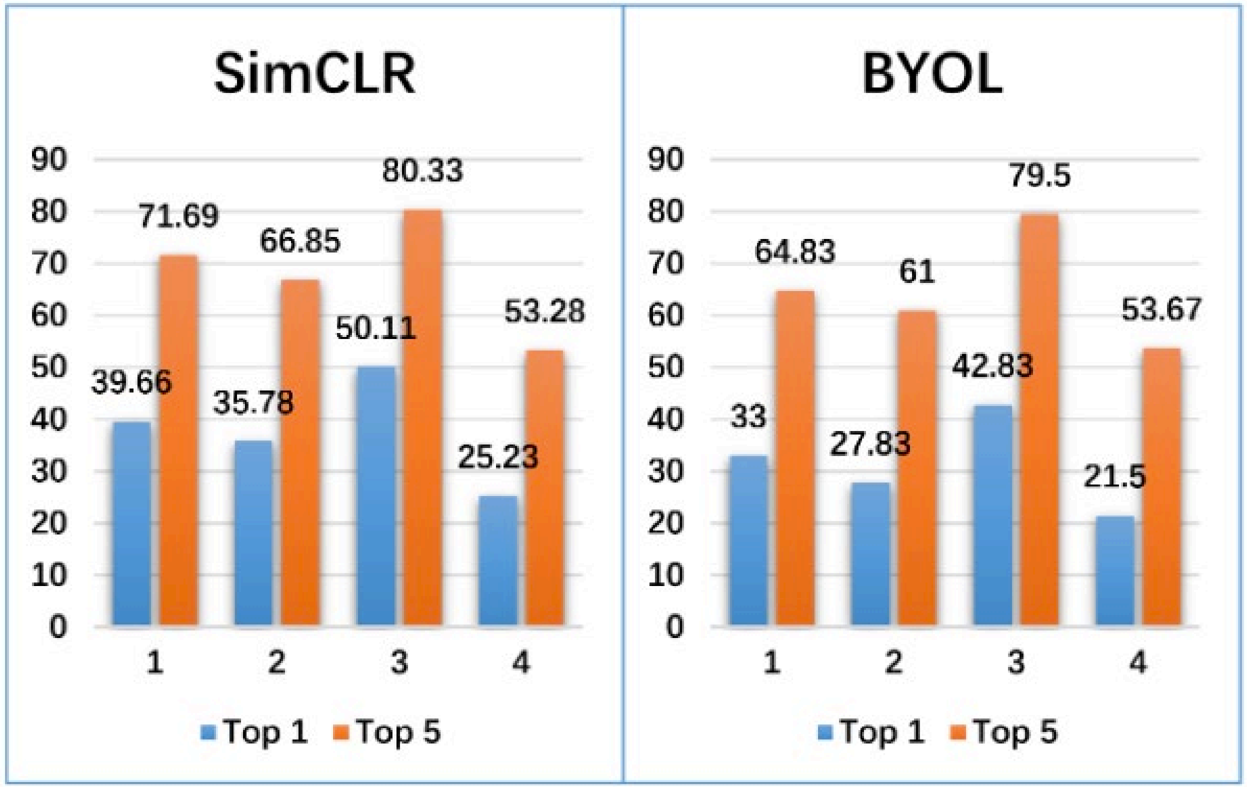
RENMIN UNIVERSITY OF CHINA 1937 中国人民大学

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

ICML | 2022

Motivation

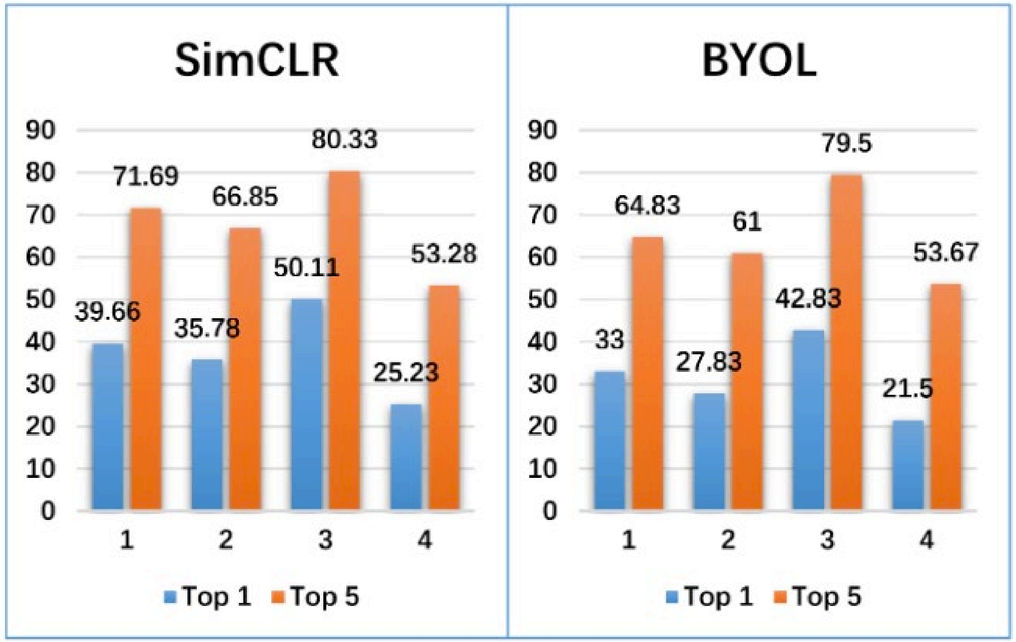
- Often-Overlooked Characteristic of Current Contrastive Learning Methods



- 1: training and testing on full images
- 2: training on full images and testing on foreground images
- 3: training and testing on foreground images
- 4: training on foreground images and testing on full images.

Motivation

- Often-Overlooked Characteristic of Current Contrastive Learning Methods



Observation: background-related information degrades the performance of the CL models.

Explanation: the feature extractor trained on full images so that it extracts background-dependent semantic features. But contrastive learning strives to be adaptable to a variety of downstream tasks. Only foreground-related semantic information can ensure the robustness of the learned features to various tasks.

Intuition

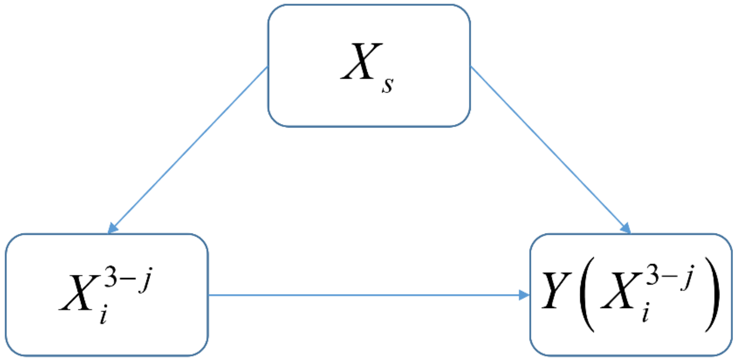
1. To capture the causal links between semantic information, positive sample, and anchor, we establish a Structural Causal Model (SCM).

2. We propose a new method by implementing backdoor adjustments to the planned SCM.

Problem Formulation

➢ Structural Causal Model

- The nodes in SCM represent the abstract data variables and the directed edges represent the (functional) causality



- X_s : semantic information
- X_i^{3-j} : positive sample
- $Y(X_i^{3-j})$: anchor (or label)

Problem Formulation

➢ **Causal Intervention via Backdoor Adjustment**

- he backdoor adjustment assumes that we can observe and stratify the confounder

$$P(Y(X_i^{3-j})|do(X_i^{3-j})) = \sum_{i=1}^{n} P(Y(X_i^{3-j})|X_i^{3-j}, Z_s^i)P(Z_s^i)$$

- Z_s^i : a stratification of semantic feature
- $P(Y(X_i^{3-j})|do(X_i^{3-j}))$: the true causality between $Y(X_i^{3-j})$ and X_i^{3-j}.

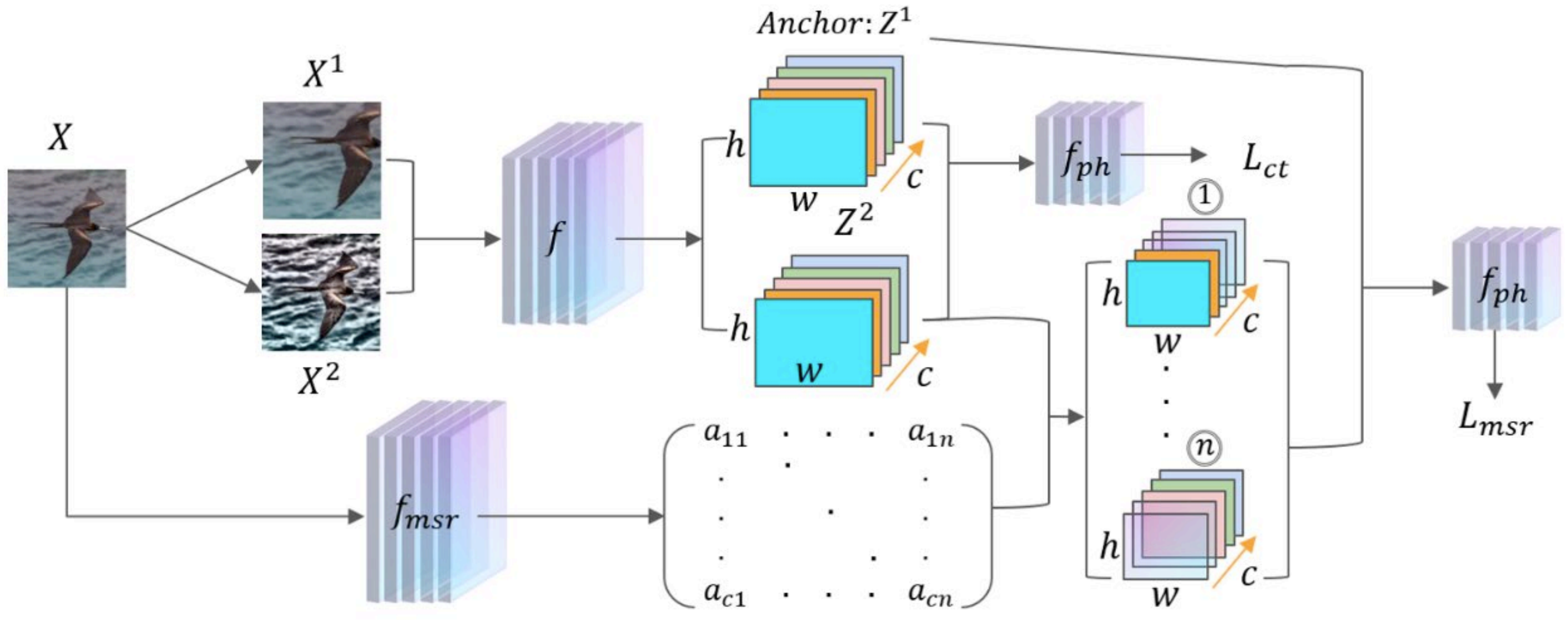
Meta Semantic Regularizer

- we present the implementation of the backdoor adjustment during the training phase

$$P(Y(X_i^{3-j})|do(X_i^{3-j})) = \sum_{t=1}^{n} \frac{\exp\left(\frac{\mathrm{sim}\left(Z_i^j, a_t \odot Z_i^{3-j}\right)}{\tau}\right) \times \frac{1}{n}}{\exp\left(\frac{\mathrm{sim}\left(Z_i^j, a_t \odot Z_i^{3-j}\right)}{\tau}\right) + \sum_{\substack{k=1,\\ k\neq i}}^{N} \sum_{\substack{l=1,\\ l\neq j}}^{2} \exp\left(\frac{\mathrm{sim}\left(Z_i^j, Z_k^l\right)}{\tau}\right)}$$

Meta Semantic Regularizer

- The meta semantic regularizer is trained alongside the feature extractor, with two stages per epoch



- In the first stage, f and f_{msr} are learned using the two augmented training set X_{tr}^{aug}, and the semantically relevant weight matrix A_s. In the second stage, f_{msr} is updated by computing its gradients with respect to the contrastive loss.

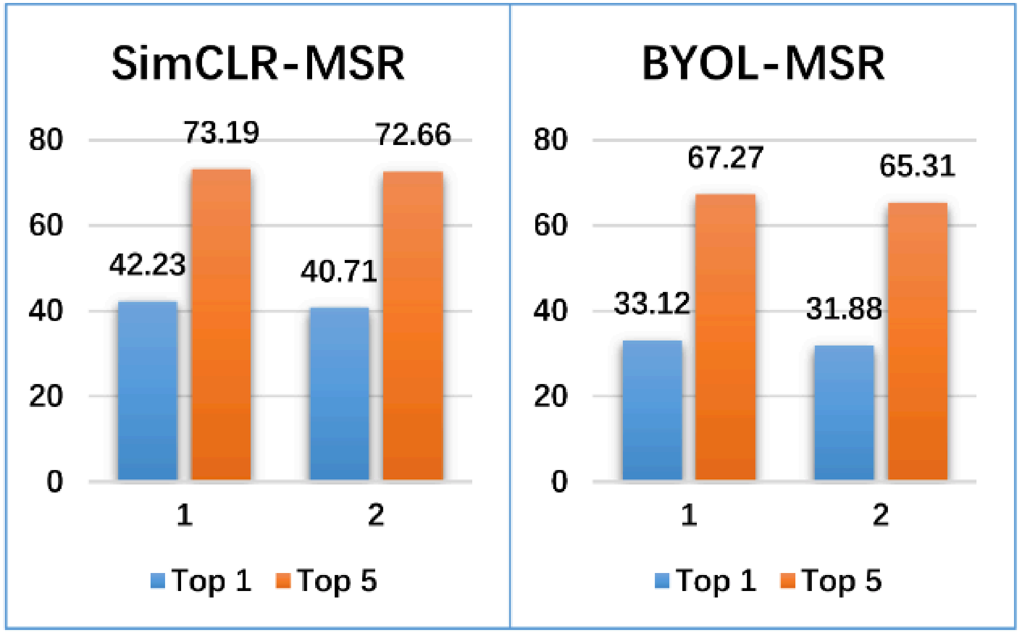
Evaluation

- Comparison with self-supervised learning methods

Methods	CIFAR-10		CIFAR-100		STL-10		Tiny ImageNet	
	linear	5-nn	linear	5-nn	linear	5-nn	linear	5-nn
SimCLR (Chen et al., 2020a)	91.80	88.42	66.83	56.56	90.51	85.68	48.84	32.86
BYOL (Grill et al., 2020)	91.73	89.45	66.60	56.82	91.99	88.64	51.00	36.24
W-MSE (Ermolov et al., 2021)	91.99	89.87	67.64	56.45	91.75	88.59	49.22	35.44
ReSSL (Zheng et al., 2021)	90.20	88.26	63.79	53.72	88.25	86.33	46.60	32.39
LMCL (Chen et al., 2021a)	91.91	88.52	67.01	56.86	90.87	85.91	49.24	32.88
SSL-HSIC (Li et al., 2021)	91.95	89.99	67.23	57.01	92.09	88.91	51.37	36.03
RELIC (Mitrovic et al., 2021)	91.96	89.35	67.24	56.88	91.15	86.21	49.17	32.97
ICL-MSR(SimCLR + MSR)	92.34	89.47	67.59	57.64	92.03	86.94	50.12	32.88
ICL-MSR(BYOL + MSR)	92.26	**90.12**	66.97	**57.97**	**93.22**	**89.36**	52.54	**37.54**
ICL-MSR(LMCL + MSR)	**92.45**	89.38	**67.99**	57.71	91.56	87.73	**52.61**	32.35
ICL-MSR(ReSSL + MSR)	91.77	89.06	65.12	55.07	89.91	88.06	47.17	33.03

Evaluation

- The experimental results for two kinds of ICL-MSR models



- 1: training and testing on full images
- 2: training on full images, and testing on foreground images