Class-Agnostic Self-Supervised Learning for Image Angle Classification

Hyeonseok Kim, Yeejin Lee*

Department of Electrical and Information Engineering,

Seoul National University of Science and Technology,

Seoul, 01811, Korea (khslab@seoultech.ac.kr, yeejinlee@seoultech.ac.kr*)

Introduction

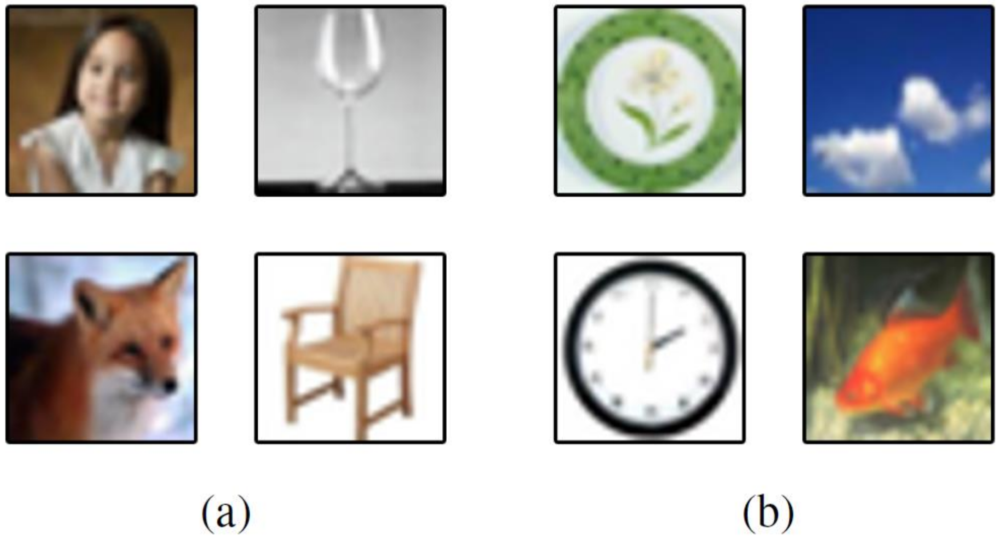
To train the prediction of image angles⋯

1. ~~Build new datasets~~ → ~~Require substantial resources and time~~

2. Use existing benchmark datasets ✓

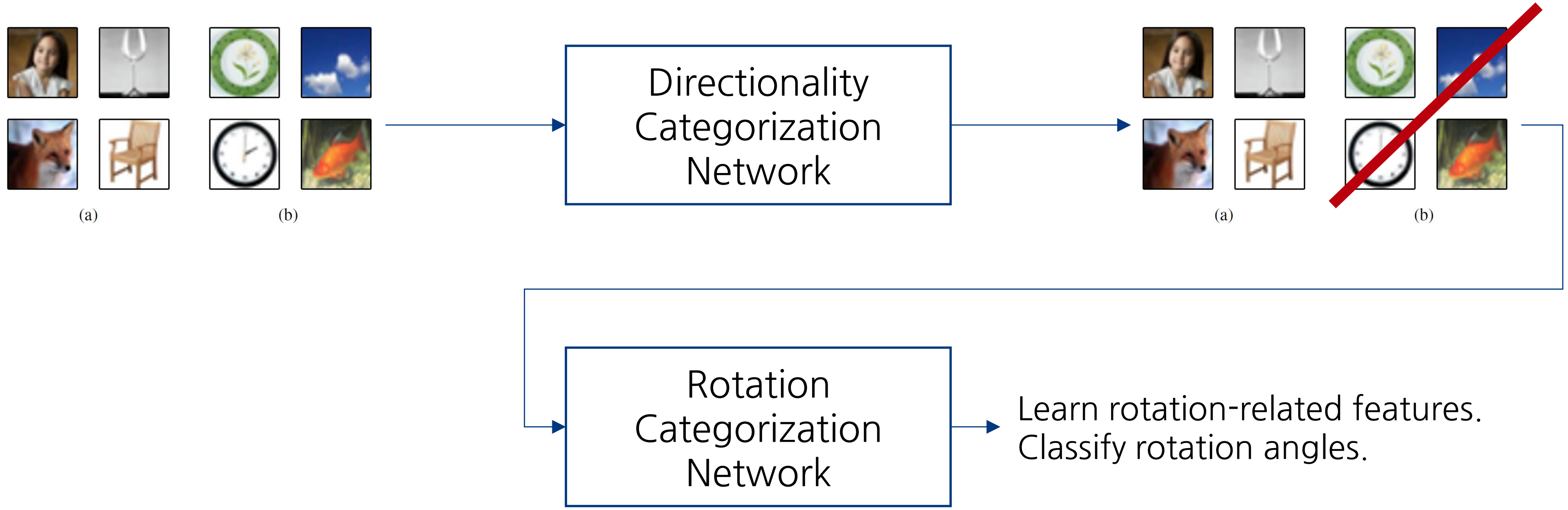
Introduction

When using existing datasets, the directionality issue arises.



Introduction

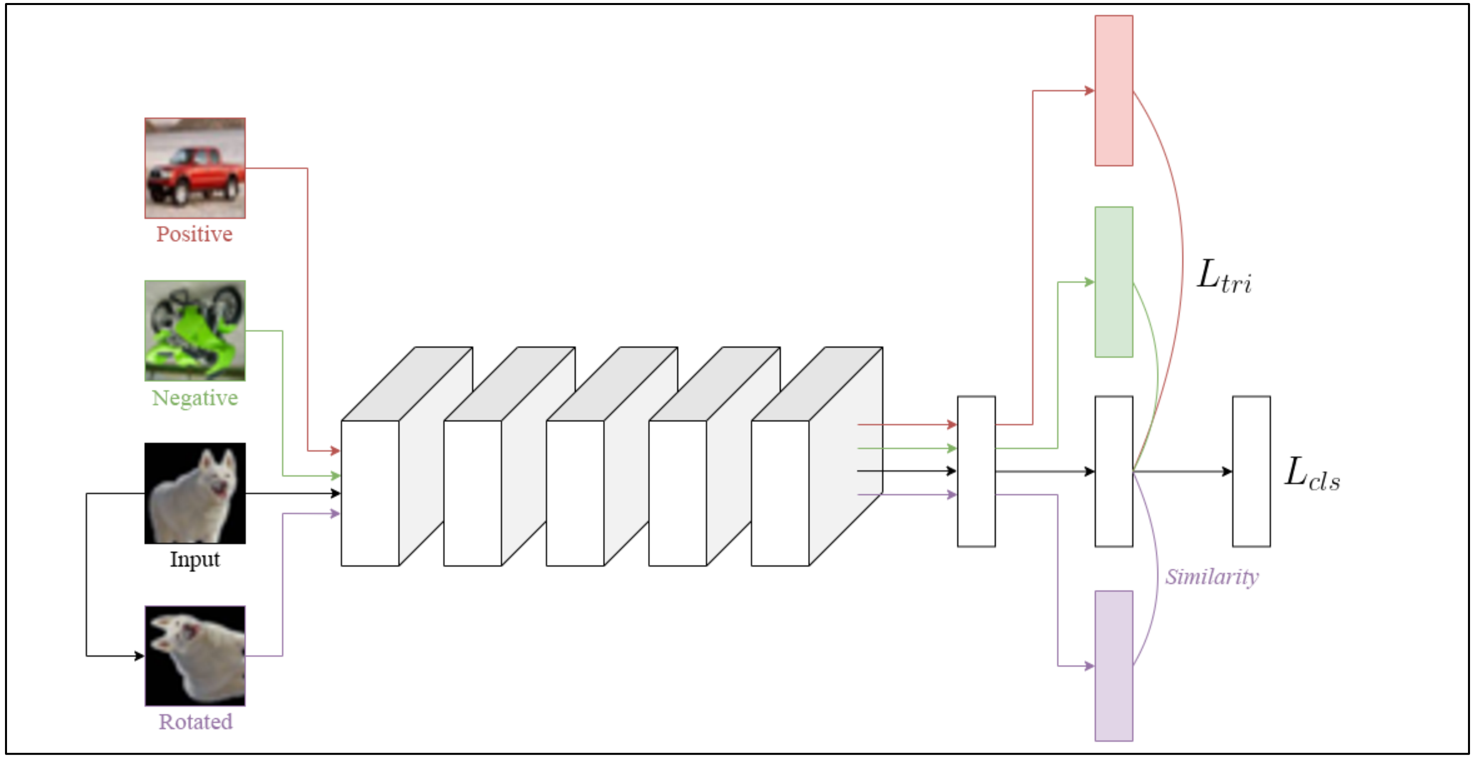
We propose a class-agnostic self-supervised angle prediction learning framework.



Learn rotation-related features.
Classify rotation angles.

Proposed method

The structure of the proposed network.



Proposed method

2.1 Directionality Categorization

We construct a dataset for angle prediction.

Existing dataset

$$D_{tr} = \left\{X_{tr}^{j}, Y_{tr}^{j}\right\}_{j=1}^{N_{tr}}$$

New dataset

$$D_{r} = \left\{X_{r}^{j}, k^{j}\right\}_{j=1}^{N_{r}}$$

k: [0∘, 90∘, 180∘, 270∘]

Proposed method

To learn discriminative features based on rotation angles⋯

Cross-entropy loss

$$\mathcal{L}_{cls}^{D} = -\frac{1}{N_r}\sum_{j=1}^{N_r}\sum_{i=1}^{4} k_i^j \log\left(M_D\left(X_r^j\right)\right), \qquad (1)$$

k: [0∘, 90∘, 180∘, 270∘]

N_r: the number of training samples
k_i^j: the rotated degree
M_D: the directionality categorization network
X_r^j: the training sample

Proposed method

To clarify the difference between features⋯

Triplet loss

$$\mathcal{L}_{tri}^{D} = \sum_{j=1}^{N_r} max\left\{ \left\| f_{x_r^j}^{D} - f_p^{D} \right\|^2 - \left\| f_{x_r^j}^{D} - f_n^{D} \right\|^2 + \alpha, 0 \right\}, \quad (2)$$

N_r: the number of training samples
$f_{x_r^j}^{D}$: the extracted feature vector of the input x_r^j
f_p^D: the feature vector of a positive input
f_n^D: the feature vector of a negative input
α: the margin

Proposed method

Total loss function.

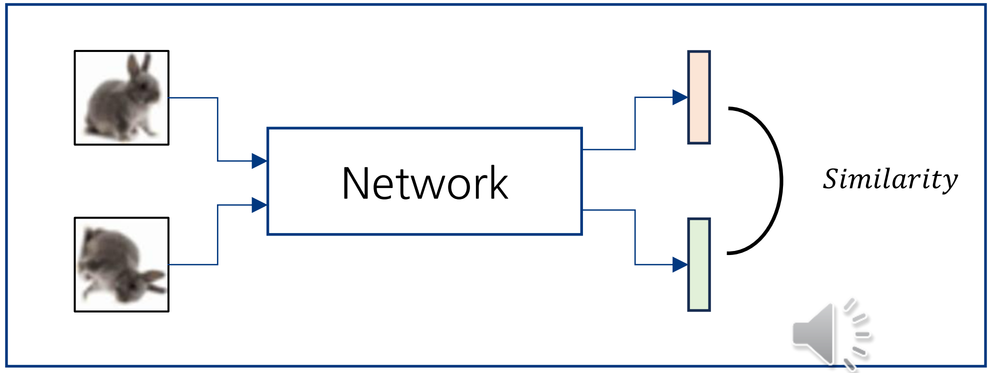
$$\mathcal{L}_{total}^{D} = \mathcal{L}_{cls}^{D} + \lambda_1 \mathcal{L}_{tri}^{D}, \qquad (3)$$

λ_1: the weighting coefficient

To distinguish the directionality of samples…

Cosine similarity

$$Similarity == \frac{f_{x_r^j}^{D} \cdot f_r}{\left\| f_{x_r^j}^{D} \right\| \cdot \left\| f_r \right\|}, \qquad (4)$$



Similarity > T : Non-directionality
Similarity <= T : Directionality

Proposed method

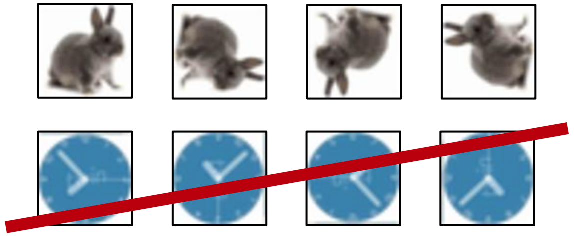
2.2 Rotation Categorization

We obtain a filtered dataset from the directionality categorization network.

New dataset

$$D_r = \{X_r^j, k^j\}_{j=1}^{N_r}$$

k: [0∘, 90∘, 180∘, 270∘]



Filtered dataset

$$D_f = \left\{X_f^j, k^j\right\}_{j=1}^{N_f}$$

Proposed method

Learning is similar to the directionality categorization network..

$$\mathcal{L}_{total}^{A} = \mathcal{L}_{cls}^{A} + \lambda_2 \mathcal{L}_{tri}^{A}, \tag{5}$$

$$\mathcal{L}_{cls}^{A} = -\frac{1}{N_f} \sum_{j=1}^{N_f} \sum_{i=1}^{4} k_i^j \log\left(M_A\left(X_f^j\right)\right), \tag{6}$$

$$\mathcal{L}_{tri}^{A} = \sum_{j=1}^{N_f} max\left\{\left\|f_{x_f^j}^A - f_p^A\right\|^2 - \left\|f_{x_f^j}^A - f_n^A\right\|^2 + \beta, 0\right\}. \tag{7}$$

Results

Comparison of classification accuracy (%)

Object Types	STL-10		CIFAR-100	
	Acurracy	ΔAcc	Acurracy	ΔAcc
All	72.4	-	58.2	-
Directional	***74.0***	***+1.6***	***63.6***	***+5.4***