Revisiting Contrastive Learning through the Lens of Neighborhood Component Analysis: an Integrated Framework

Ching-Yun Ko[1] , Jeet Mohapatra[1],
Sijia Liu[2,3], Pin-Yu Chen[2],
Luca Daniel[1], Lily Weng[2,4]

[1]MIT, [2]MIT IBM Watson AI Lab/IBM Research, [3]MSU, [4]UCSD

★**arXiv:** https://arxiv.org/abs/2112.04468

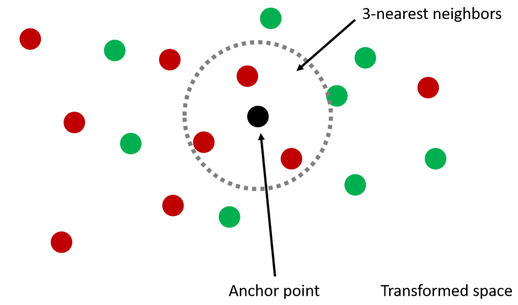
Long version

Background

- A Simple Framework for Contrastive Learning of Visual Representations —— SimCLR[1]

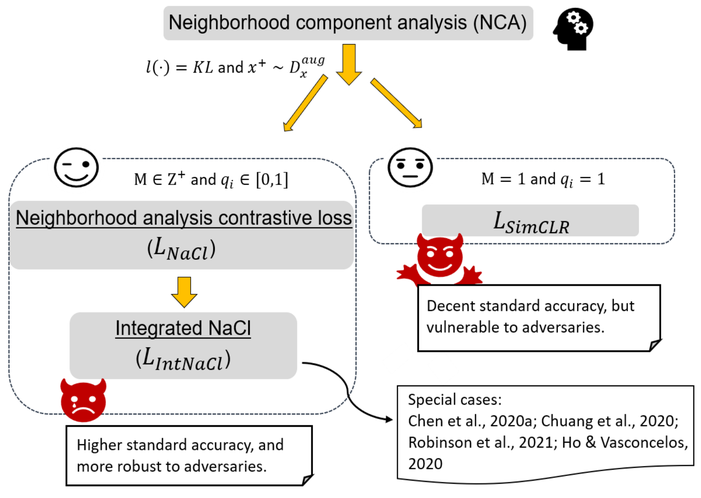
$$\min_f \mathbb{E}_{x\sim\mathcal{D},x^+\sim\mathcal{D}_x^{\text{aug}},x_i^-\sim\mathcal{D}_{\setminus x}^{\text{aug}}}\left[-\log\left(\frac{e^{f(x)^T f(x^+)}}{e^{f(x)^T f(x^+)}+Ng_0(x,\{x_i^-\}^N)}\right)\right]$$

- Neighborhood components analysis (NCA) [2] is a supervised learning method that learns a transformation A such that the average leave-one-out (LOO) classification performance is maximized in the transformed space.
- Maximizing the LOO performance is equivalent to minimizing the L_1 distance or KL-divergence between the predicted class distribution and the true class distribution.



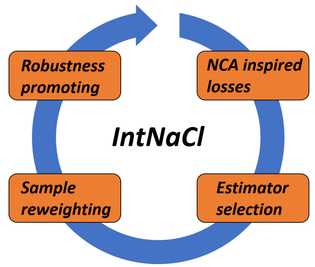
Motivations

- The reduction from the NCA formulation to SimCLR requires:
  1. Estimating the expectation over the D_x^{aug} by only 1 sample
     → L_{VAR}
  2. The expected relative density of positives in the underlying data distribution is $1/N$
     → L_{BIAS}
  3. The probability induced by representation network are all equal to 1
     → L_{MIXUP}



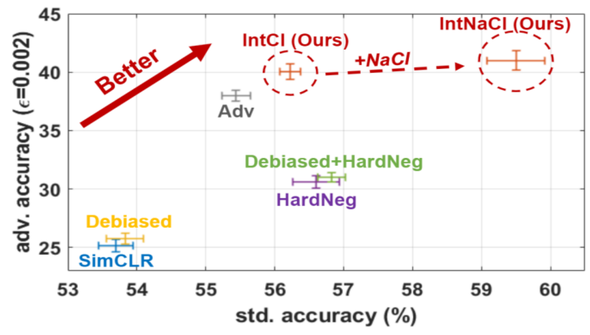
Integrated Neighborhood analysis Contrastive loss (IntNaCl)

$$\mathcal{L}_{\text{IntNaCl}}(\mathcal{L}_{\text{NaCl}}(g^1,M,\lambda),\alpha,\mathcal{L}_{\text{Robust}}(g^2,w)) := \mathcal{L}_{\text{NaCl}}(g^1,M,\lambda)+\alpha\mathcal{L}_{\text{Robust}}(g^2,w)$$

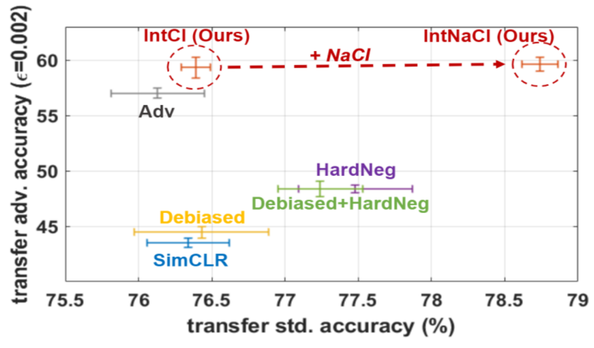


$\mathcal{L}_{\text{VAR}}(g^1,M)$	$\mathbb{E}_{x\sim\mathcal{D},x_j^+\sim\mathcal{D}_x^{\text{aug}},x_{ij}^-\sim\mathcal{D}_{\setminus x}^{\text{aug}}}[-\frac{1}{M}\sum_{j=1}^{M}\log\frac{e^{f(x)^T f(x_j^+)}}{e^{f(x)^T f(x_j^+)}+Ng^1(x,\{x_{ij}^-\}_i^N)}]$
$\mathcal{L}_{\text{BIAS}}(g^1,M)$	$\mathbb{E}_{x\sim\mathcal{D},x_j^+\sim\mathcal{D}_x^{\text{aug}},x_i^-\sim\mathcal{D}_{\setminus x}^{\text{aug}}}[-\log\frac{\sum_{j=1}^{M}e^{f(x)^T f(x_j^+)}}{\sum_{j=1}^{M}e^{f(x)^T f(x_j^+)}+Ng^1(x,\{x_i^-\}^N)}]$
$\mathcal{L}_{\text{MIXUP}}(g^1,M,\lambda)$	$\mathbb{E}_{x\sim\mathcal{D},x^+\sim\mathcal{D}_x^{\text{aug}},x_{i_1}^-,x_{i_2j}^-,x_j^-\sim\mathcal{D}_{\setminus x}^{\text{aug}}}[-\log\frac{e^{f(x)^T f(x^+)}}{e^{f(x)^T f(x^+)}+Ng^1(x,\{x_{i_1}^-\}^N)}$ $-\frac{\lambda}{M-1}\sum_{j=1}^{M-1}\log\frac{e^{f(x)^T f(\lambda x^+ +(1-\lambda)x_j^-)}}{e^{f(x)^T f(\lambda x^+ +(1-\lambda)x_j^-)}+Ng^1(x,\{x_{i_2j}^-\}_{i_2}^N)}$ $-\frac{1-\lambda}{M-1}\sum_{j=1}^{M-1}\log\left(1-\frac{e^{f(x)^T f(\lambda x^+ +(1-\lambda)x_j^-)}}{e^{f(x)^T f(\lambda x^+ +(1-\lambda)x_j^-)}+Ng^1(x,\{x_{i_2j}^-\}_{i_2}^N)}\right)]$
$g_0(x,\{x_i^-\}_i^N)$	$\frac{1}{N}\sum_{i=1}^{N}e^{f(x)^T f(x_i^-)}$
$g_1(x,\{u_i\}^n,\{v_j\}^m)$	$\max\{\frac{1}{1-\tau^+}(\frac{1}{n}\sum_{i=1}^{n}e^{f(x)^T f(u_i)}-\tau^+\frac{1}{m}\sum_{j=1}^{m}e^{f(x)^T f(v_j)}),e^{-1/t}\}$
$g_2(x,\{u_i\}^n,\{v_j\}^m)$	$\max\{\frac{1}{1-\tau^+}(\frac{\sum_{i=1}^{n}e^{(\beta+1)f(x)^T f(u_i)}}{\sum_{i=1}^{n}e^{\beta f(x)^T f(u_i)}}-\tau^+\frac{1}{m}\sum_{j=1}^{m}e^{f(x)^T f(v_j)}),e^{-1/t}\}$
$\hat{w}(x)$	$-\log\frac{e^{f(x)^T f(x^+)}}{e^{f(x)^T f(x^+)}+Ng(x,\cdot)}$

Experimental Results



(a) CIFAR100



(b) CIFAR10

Better standard, transfer, and adversarial performance. We train the representation network on CIFAR100 and test on both CIFAR10 and CIFAR100.

References

[1] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in ICML'20.
[2] J. Goldberger, G. E. Hinton, S. Roweis, and R. R. Salakhutdinov, "Neighbourhood componentsanalysis," in NeurIPS'04.
[3] C.-Y. Chuang, J. Robinson, L. Yen-Chen, A. Torralba, and S. Jegelka, "Debiased contrastive125learning,"in NeurIPS'20.
[4] C.-H. Ho and N. Vasconcelos, "Contrastive learning with adversarial examples,"arXiv preprint arXiv:2010.12050, 2020.
[5] J. D. Robinson, C.-Y. Chuang, S. Sra, and S. Jegelka, "Contrastive learning with hard negative samples," in ICLR'21.

Ching-Yun Ko