An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis

Levy Chaves¹, Alceu Bissoto¹, Eduardo Valle², Sandra Avila¹

¹Institute of Computing  ²School of Electrical and Computing Engineering

Recod.ai, University of Campinas (UNICAMP), Brazil

Seventh ISIC Skin Image Analysis Analysis Workshop @ ECCV2022
Yann LeCun: AI Doesn’t Need Our Supervision

Meta’s AI chief says self-supervised learning can build the metaverse and maybe even human-level AI
Self-Supervised Learning

Bastanlar, Yalin, and Semih Orhan. "Self-Supervised Contrastive Representation Learning in Computer Vision." (2022).
Pretext-task examples

Gidaris et al., 2018, Predicting Image Rotations
Pretext-task examples

Zhang, Richard, Phillip Isola, and Alexei A. Efros. "Colorful image colorization." ECCV. 2016.
How Transferable are Self-supervised Features in Medical Image Classification Tasks?

TUAN.TRUONG@BAYER.COM
SADEGH.MOHAMMADI@BAYER.COM
MATTHIAS.LENG@BAYER.COM

Big Self-Supervised Models Advance Medical Image Classification

Shekoofeh Azizi, Basil Mustafa, Fiona Ryan*, Zachary Beaver, Jan Freyberg, Jonathan Deaton, Aaron Loh, Alan Karthikesalingam, Simon Kornblith, Ting Chen, Vivek Natarajan, Mohammad Norouzi
Google Research and Health†

Abstract

Self-supervised pretraining followed by supervised fine-tuning has seen success in image recognition, especially when labeled examples are scarce, but has received limited attention in medical image analysis. This paper studies the effectiveness of self-supervised learning as a pre-training strategy for medical image classification. We con-

ON THE IMPACT OF SELF-SUPERVISED LEARNING IN SKIN CANCER DIAGNOSIS

Maria Rita Verdelho and Catarina Barata
Institute for Systems and Robotics, Instituto Superior Técnico, Lisboa, Portugal

ABSTRACT

Deep neural networks (DNNs) are the standard approach for image classification. However, they require a large amount of data and corresponding annotations. Collecting

A Systematic Benchmarking Analysis of Transfer Learning for Medical Image Analysis

Mohammad Reza Hosseinzadeh Taher1, Fatemeh Haghighi1, Ruibin Feng2, Michael B. Gotway3, and Jianming Liang1

1 Arizona State University, Tempe, AZ 85281, USA
{mossei2,fhaghigh,jianming.liang}@asu.edu
2 Stanford University, Stanford, California 94305, USA
ruibin@stanford.edu
3 Mayo Clinic, Scottsdale, AZ 85259, USA
Gotway.Michael@mayo.edu

Abstract. Transfer learning from supervised ImageNet models has been frequently used in medical image analysis. Yet, no large-scale evaluation has been conducted to benchmark the efficacy of newly-developed

Self-supervised learning (SSL) has emerged as a strategy

(1) Self-supervised learning on unlabeled natural images

(2) Supervised learning on labeled natural images
## What were they missing?

| Work $_{year}$ | Out-of-distribution Evaluation | Low-data Evaluation |
|----------------|--------------------------------|---------------------|
| Azizi et al. 2021 | ✗                              | ✓                   |
| Hosseinzadeh et al. 2021 | ✗                              | ✗                   |
| Truong et al. 2021 | ✗                              | ✓                   |
| Verdelho et al. 2022 | ✗                              | ✗                   |
| Ours 2022           | ✓                              | ✓                   |

Chaves et al., "An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis ″, ISIC Workshop @ ECCV 2022
Experimental Design & Preliminary results
Standard Evaluation Protocol

- Model Selection
- Fine-tuning
- In-Distribution Data

Supervised (baseline)
Self-supervised (5 candidates)
Evaluated self-supervised learning methods

Fine-tuning results on ISIC 2019 (Melanoma vs. benign)

| Method    | AUC (%) |
|-----------|---------|
| SimCLR [1] | 95.6    |
| SwAV [2]  | 95.3    |
| BYOL [3]  | 94.6    |
| MoCo-V2 [4]| 94.4    |
| InfoMin [5]| 93.9    |

[1] Chen, Ting, et al. "A simple framework for contrastive learning of visual representations.". ICML 2020.
[2] Caron, Mathilde, et al. "Unsupervised learning of visual features by contrasting cluster assignments.". NeurIPS 2020
[3] Grill, J. B, et al. Bootstrap your own latent-a new approach to self-supervised learning. NeurIPS. 2020
[4] Chen, Xinlei, et al. "Improved baselines with momentum contrastive learning." arXiv preprint arXiv:2003.04297. 2020.
[5] Tian, Yonglong, et al. "What makes for good views for contrastive learning?.". NeurIPS 2020.
Evaluated self-supervised learning methods

Fine-tuning results on ISIC 2019 (Melanoma vs. benign)

| Method       | AUC (%) |
|--------------|---------|
| Sup. Baseline| 94.8    |
| SimCLR [1]   | 95.6    |
| SwAV [2]     | 95.3    |
| BYOL [3]     | 94.6    |
| MoCo-V2 [4]  | 94.4    |
| InfoMin [5]  | 93.9    |

Strong Baseline!

[1] Chen, Ting, et al. "A simple framework for contrastive learning of visual representations.". ICML 2020.
[2] Caron, Mathilde, et al. "Unsupervised learning of visual features by contrasting cluster assignments.". NeurIPS 2020.
[3] Grill, J. B, et al. Bootstrap your own latent-a new approach to self-supervised learning. NeurIPS 2020.
[4] Chen, Xinlei, et al. "Improved baselines with momentum contrastive learning." arXiv preprint arXiv:2003.04297. 2020.
[5] Tian, Yonglong, et al. "What makes for good views for contrastive learning?.". NeurIPS 2020.
Our Evaluation Protocol

Model Selection → In-domain pre-training? → Yes → Contrastive Learning → Fine-tuning → Testing

No → In-Distribution Data → Training Dataset

Contrastive Learning → Fine-tuning → Testing

Model Selection → Training Dataset → Fine-tuning → Testing

Training Dataset → Training Dataset → Fine-tuning → Testing

Contrastive Learning → Fine-tuning → Testing

In-Distribution Data → Training Dataset → Fine-tuning → Testing

Out-of Distribution Data → Training Dataset → Fine-tuning → Testing
Our pipelines

Self-supervised pre-training on unlabeled natural images

Contrastive learning pre-training on labeled or unlabeled skin-lesion images

Supervised training on labeled natural images

Contrastive learning pre-training on labeled or unlabeled skin-lesion images

Supervised fine-tuning on labeled skin images

SSL \rightarrow SCL \rightarrow FT

SSL \rightarrow UCL \rightarrow FT

SSL \rightarrow FT

SUP \rightarrow FT

Chaves et al., "An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis ", ISIC Workshop @ ECCV 2022
Contrastive Learning

\[(t, t') \in T \quad \text{Set of transformations}\]

Representation Space
Contrastive Learning

Unsupervised Contrastive Learning (UCL) -> Image augmentations to create positive views

Supervised Contrastive Learning (SCL) -> Label class to create positive views

(t, t') ∈ T — Set of transformations

Representation Space
Full-data evaluation

| Training Data | 100 % |
|---------------|-------|

| Full-data evaluation | |
Low-data evaluation
Low-data evaluation
Out-of-Distribution Evaluation

| Train          | Test                  |
|----------------|-----------------------|
| ISIC 2019      | ISIC 2020             |
| ISIC 2019      | PAD-UFES-20           |
| Derm7pt-dermato| Derm7pt-dermato       |
| Derm7pt-clinical| Derm7pt-clinical     |
| Additional benign Diagnosis | Additional benign Diagnosis |
Results
Full data and out-of-distribution performance

100% of training data — 14,805 samples

Chaves et al., “An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis“, ISIC Workshop @ ECCV 2022
Full data and out-of-distribution performance

100% of training data — 14,805 samples

Chaves et al., “An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis”, ISIC Workshop @ ECCV 2022
Low-data and out-of-distribution performance

Chaves et al., “An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis”, ISIC Workshop @ ECCV 2022
Low-data and out-of-distribution performance

`1% of training data — 148 samples`

Chaves et al., “An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis “, ISIC Workshop @ ECCV 2022
Qualitative Analysis

100% of training data

True Positive

10% of training data

Supervised | SimCLR | SwAV | BYOL

Supervised | SimCLR | SwAV | BYOL

1% of training data

Supervised | SimCLR | SwAV | BYOL

Chaves et al., "An Evaluation of Self-Supervised Pre-Training for Skin-Lesion Analysis", ISIC Workshop @ ECCV 2022
Conclusion

• The advantage of self-supervised pipelines was particularly positive in the low-data scenarios
Conclusion

• The advantage of self-supervised pipelines was particularly prominent in the low-data scenarios

• Models pre-trained in a self-supervised manner felt easier to optimize
Conclusion

• The advantage of self-supervised pipelines was particularly prominent in the low-data scenarios.

• Models pre-trained in a self-supervised manner felt easier to optimize.

• Understanding what circumstances make self-supervised competitive from a theoretical perspective is a promising research area.
Limitations

- Explored just one training dataset and model architecture
Limitations

• Explored just one training dataset and model architecture

• Extensive exploration is necessary to evaluate if self-supervised is reinforcing data biases
Code and data available on Github!
https://github.com/VirtualSpaceman/ssl-skin-lesions

Thank you!

Levy Chaves levy.chaves@ic.unicamp.br
Alceu Bissoto alceubissoto@ic.unicamp.br
Eduardo Valle dovalle@dca.fee.unicamp.br
Sandra Avila sandra@ic.unicamp.br