



Abstract: Self-supervised Pre-training for Dealing with Small Datasets in Deep Learning for Medical Imaging

Evaluation of Contrastive and Masked Autoencoder Methods

Daniel Wolf^{1,2}, Tristan Payer¹, Catharina S. Lisson², Christoph G. Lisson²,
Meinrad Beer², Michael Götz^{2,3}, Timo Ropinski^{1,3}

¹Visual Computing Research Group, Institute of Media Informatics, Ulm University, Germany

²Experimental Radiology Research Group, Department for Diagnostic and Interventional Radiology, Ulm University Medical Center, Germany

³These two authors contributed equally to this work

daniel.wolf@uni-ulm.de

Deep learning in medical imaging has the potential to minimize the risk of diagnostic errors, reduce radiologist workload, and accelerate diagnosis. Training such deep learning models requires large and accurate datasets, with annotations for all training samples. However, in the medical imaging domain, annotated datasets for specific tasks are often small due to the high complexity of annotations, limited access, or the rarity of diseases. To address this challenge, deep learning models can be pre-trained on large image datasets without annotations using methods from the field of self-supervised learning. After pre-training, small annotated datasets are sufficient to fine-tune the models for a specific task. The most popular self-supervised pre-training approaches in medical imaging are based on contrastive learning. However, recent studies in natural image processing indicate a strong potential for masked autoencoder approaches. Our work [1] compares state-of-the-art contrastive learning methods with the recently introduced masked autoencoder approach "SparK" for convolutional neural networks (CNNs) on medical images. Therefore, we pre-train on a large unannotated CT image dataset and fine-tune on several CT classification tasks. Due to the challenge of obtaining sufficient annotated training data in medical imaging, it is of particular interest to evaluate how the self-supervised pre-training methods perform when fine-tuning on small datasets. By experimenting with gradually reducing the training dataset size for fine-tuning, we find that the reduction has different effects depending on the type of pre-training chosen. The SparK pre-training method is more robust to the training dataset size than the contrastive methods. Based on our results, we propose the SparK pre-training for medical imaging tasks with only small annotated datasets.

References

1. Wolf D, Payer T, Lisson CS, Lisson CG, Beer M, Götz M et al. Self-supervised pre-training with contrastive and masked autoencoder methods for dealing with small datasets in deep learning for medical imaging. Nat Sci Rep. 2023;13(1):20260.