 https://doi.org/10.1051/bioconf/202412910037

Comparative Analysis of Self-Supervised Learning Techniques for Electron Microscopy Images

Dr.-ing. Bashir Kazimi[1], Prof. Dr. Stefan Sandfeld[1,2]

[1]Forschungszentrum Jülich, Institute for Advanced Simulation – Materials Data Science and Informatics (IAS-9), Aachen, Germany, [2]RWTH Aachen University, Chair of Materials Data Science and Informatics, Aachen, Germany

Background incl. aims

Deep learning has revolutionized a wide array of tasks across different domains, including electron microscopy (EM) image analysis, by leveraging large labeled datasets for training. However, the scarcity of such labeled datasets in EM necessitates the exploration of alternative methods. Self-supervised learning (SSL) emerges as a promising approach to leverage unlabeled data, featuring techniques such as, e.g., masked image modeling (MIM) — which predicts missing parts of the input data, as well as contrastive learning — which learns by distinguishing between similar and dissimilar pairs of data. This study aims to investigate the impact of these SSL techniques on EM images, providing a case study on the effectiveness of leveraging unlabeled data in a domain where labeled datasets are limited and expensive to create.

Methods

Utilizing the “NFFA dataset”, which comprises 21,169 Scanning EM images across 10 categories, we established a baseline by training models from scratch with random weight initialization. We then pre-trained models using two SSL approaches: Masked Autoencoders (MAE) for MIM and Momentum Contrast V3 (MoCoV3) for contrastive learning, followed by fine-tuning on the NFFA dataset. Another pixel-based MIM technique, Multi-level Feature Fusion (MFF), was also tested. The performance of each SSL technique was evaluated based on accuracy improvements and convergence speeds relative to the baseline. Our analysis highlights the distinctions between MIM and contrastive learning approaches in handling EM images.

Results

The baseline model yielded an accuracy of 77.42%. Upon employing SSL techniques, significant improvements were observed: finetuning with MAE weights achieved an accuracy of 92.84%, MFF led to 93.86%, and MoCoV3 led to an accuracy of 92.56%. MFF, in particular, demonstrated a superior ability to enhance feature learning from unlabeled data, indicating its impact in the task of EM image classification. Furthermore, all SSL-pretrained models showcased accelerated convergence rates compared to the baseline.

Conclusion

This study confirms the viability and potential of SSL techniques in EM images. MIM, exemplified by MFF, outperformed contrastive learning in this domain, suggesting that methods focusing on reconstructing or predicting unseen parts of the image are particularly beneficial for EM tasks. The results advocate for a targeted selection of SSL strategies based on specific dataset characteristics and task requirements, highlighting a path forward for efficient model training in EM image analysis and beyond. The influence of SSL pretraining was studied in this research and experiments were conducted on SEM image classification. Further research in this direction includes investigating the influence of SSL pretraining on dense, pixel-wise classification (i.e., semantic segmentation) tasks in EM.

Keywords:

Self-Supervised-Learning, Electron-Microscopy, Image-Classification, Masked-Image-Modeling, Contrastive-Learning

Reference:

1. He, K., Chen, X., Xie, S., Li, Y., Dollár, P. and Girshick, R., 2022. Masked autoencoders are scalable vision learners. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 16000-16009).
2. Liu, Y., Zhang, S., Chen, J., Yu, Z., Chen, K. and Lin, D., 2023. Improving pixel-based mim by reducing wasted modeling capability. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 5361-5372).
3. Chen, X., Xie, S. and He, K., 2021. An empirical study of training self-supervised vision transformers. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 9640-9649).
4. Aversa, R., Modarres, M. H., Cozzini, S., & Ciancio, R., 2018. NFFA-EUROPE - 100% SEM Dataset
5. Conrad, R. and Narayan, K., 2021. CEM500K, a large-scale heterogeneous unlabeled cellular electron microscopy image dataset for deep learning. Elife, 10, p.e65894.