International Journal on Magnetic Particle Imaging
Vol 9, No 1, Suppl 1, Article ID 2303001, 4 Pages

Infinite Science Publishing

Proceedings Article

Multi-purpose deep learning framework for MPI based on contrastive learning

Franziska Schrank [a,*] · Volkmar Schulz [a,b,c,d,*]

[a] Department of Physics of Molecular Imaging Systems, Institute for Experimental Molecular Imaging, RWTH Aachen University, Aachen, Germany
[b] Hyperion Hybrid Imaging Systems GmbH, Pauwelsstrasse 19, 52074 Aachen, Germany
[c] III. Physikalisches Institut B, Otto-Blumenthal-Strasse, 52074 Aachen, Germany
[d] Fraunhofer Institute for Digital Medicine MEVIS, Forckenbeckstrasse 55, Aachen, Germany
[*] Corresponding author, email: franziska.schrank@pmi.rwth-aachen.de, volkmar.schulz@pmi.rwth-aachen.de

Abstract

Deep learning can be used in many tasks for Magnetic Particle Imaging (MPI), prominently to reduce the calibration time of system matrix (SM)-based reconstruction by recovering undersampled SMs or directly reconstructing measurements without an SM. The success of supervised machine learning methods depends on the used training data, which should have high quality, match the distribution of the desired test cases and be cleanly labeled. For MPI, such data rarely exists. To find robust features in limited complex input data, the unsupervised method contrastive learning can be used. In this work, we show its applicability to MPI voltage signals to improve tasks like SM recovery and direct reconstruction of real MPI data in 2D. SM recovery is performed by predicting voltage signals of samples placed in the MPI field of view, which could also provide an alternative to classic simulation frameworks that cannot match real MPI measurements.

I. Introduction

The applicability and flexibility of Magnetic Particle Imaging (MPI) both in preclinical research as well as the translation to human scale is limited in big parts by the lack of realistic particle models. For image reconstruction, particle, environmental, and scanner characteristics can be implicitly stored in the system matrix (SM), which is acquired in a time-consuming calibration scan. Recently, deep learning has been used to reduce this calibration time by recovering the frequency components of an undersampled SM [1]. Another approach is the direct reconstruction from the voltage or frequency signal of an object measurement to its concentration distribution [2].

The success of both approaches highly depends on the available training data, where the size of the data set is equally important as the versatility of the training samples and the overlap with the test data distribution. As such datasets are still rare in MPI, we propose to adapt the approach of contrastive learning of music representations (CLMR) [3] as an unsupervised pretraining task on MPI data. With this approach, the MPI input data is encoded in latent representations that can then be used for multiple purposes - here used for SM recovery and direct reconstruction.

I.I. Contrastive learning

The key concept behind contrastive learning is to find features in complex natural input data that make any subsequent prediction task easier and more robust. In CLMR, which is based on a simple framework for contrastive learning of visual representations (SimCLR) [4], these features are learned by maximizing the agreement between

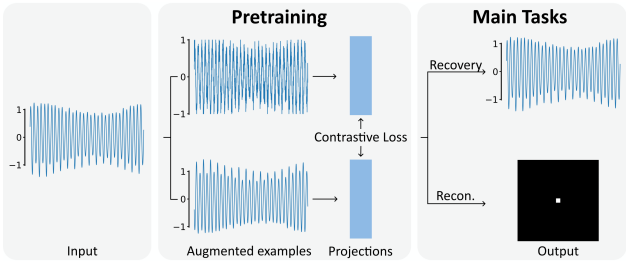

Figure 1: Framework including contrastive learning pretraining and either recovery or reconstruction as main task.

features of augmented views of the same input sample. Thus, the core components of the contrastive learning method are as follows: (1) A set of data augmentations that produce two correlated, augmented examples of one MPI input sample, the 'positive pair'. (2) An encoder and a projector network mapping the augmented samples to the latent space in which the contrastive loss is formulated, aiming to identify the positive pair from the negative examples in a batch. The used loss function is the normalized temperature-scaled cross-entropy loss, also called NT-Xent loss [4].

II. Materials and methods

In Figure 1, the complete framework including the main prediction tasks can be seen. For the training of the proposed framework, we use an SM from the OpenMPIData initiative [5], acquired with 2D excitation using perimag® particles on a grid of $19\times19\times19$ voxels, covering a volume of $38\,\text{mm}\times38\,\text{mm}\times19\,\text{mm}$. Only the main slice (slice 10) is used to reduce the whole framework to the complexity of the 2D plane.

All time domain training data is preprocessed with a bandpass filter from 80 kHz - 600 kHz to remove unwanted signal contributions. The data is then normalized by half of the total maximum amplitude of the dataset as this proved to yield good network performance. Both to enhance the variability of signals for the pretraining task as well as to emulate concentration distributions for the reconstruction task, single SM measurements are combined in weighted linear combinations. For the recovery task, the input consists of single SM columns and the network has to predict each column's 'neighbor', i.e. the measured voltage signal of the neighboring voxel in y-direction in the scanner. For each task (pretraining, recovery, reconstruction), the used input samples are split 80:20 for training and validation. For both recovery and reconstruction task, a mean squared error (MSE) loss is used during training, calculated based on the networks' output and the ground-truth for the respective task. For the additional contrastive NT-Xent loss, the temperature parameter is set to 0.5.

The data augmentation operations needed for the contrastive learning task are: Polarity inversion, i.e. the amplitude is multiplied by -1, additive white Gaussian noise with a signal-to-noise ratio between $[-25,-15]$ dB to the original signal, and gain reduction between $[-20,-1]$ dB. All augmentations are applied with a probability of 0.5.

Further training specifications are: batch size of 32; Adam optimizer [6] with a learning rate of 1e-4 and $\beta_1 = 0.9$ and $\beta_2 = 0.999$; He initialization [7] for all convolutional layers in the pretraining task; all trainings run for 1000 epochs but only the best epochs are evaluated. The architecture of the used models is (conv: convolutional layer, f: filters, fs: filter sizes, s: strides, p: paddings, lin: linear layer):

- Encoder netw.: 5×[conv1 - batch norm - ReLU] (f: 256, fs: [8, 5, 4, 4, 4], s: [4, 3, 2, 2, 2])
- Projection netw.: lin - ReLU - lin (out: 64)
- Recovery head: lin - ReLU - lin - conv1 - ReLU - conv1 (f: 8, fs: [3, 1], s: 1)
- Reconstruction head: lin (out: 19^2) - ELU(0.1) - reshape(19×19) - 2×[conv2 (f: 8, fs: 3, s: 1) - ELU(0.1)]

Testing of the proposed method is performed on another SM from the OpenMPIData initiative, using 2D excitation and perimag® particles but on a grid of $37\times37\times37$ voxels, covering a volume of $37\,\text{mm}\times37\,\text{mm}\times18.5\,\text{mm}$. Only the main slice (slice 18) is considered. For the recovery task, every fourth voxel measurement in y-direction (and every second in x-direction) is given as input to the network under the assumption that the network has learned to predict every second voxel measurement (in y-direction) based on the broader grid of the training SM. To evaluate the reconstruction performance, every second voxel measurement in x- and y-direction of the test SM slice is used as input to the reconstruction network and the output is compared to the ground-truth concentration distribution, which is one illuminated voxel. The benefit of contrastive learning is evaluated by training in two different settings: Training without the contrastive loss (wo-CL) and pretraining the encoder and projection network using the contrastive loss and then training the main tasks (CL).

III. Results and discussion

In Figure 2, the results for the recovery task can be seen for the SM slice at time step 80. The time domain is chosen for visualization as the networks' input and output are time domain signals but note that the network sees all time steps of one position at once and not all positions of one time step. Using the input, recovered and ground-truth SM, the OpenMPIData resolution phantom is reconstructed in frequency domain using the classic Kaczmarz algorithm with Tikhonov regularization ($\lambda = 0.01$;

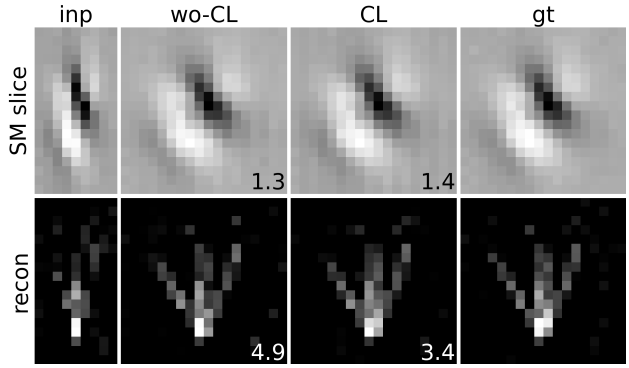

Figure 2: SM slice at time step 80 of the input SM (inp), recovered without (wo-CL), with (CL) contrastive loss and the ground-truth slice (gt). Below are the corresponding reconstruction results of the resolution phantom. NRMSE values are given in % in the images.

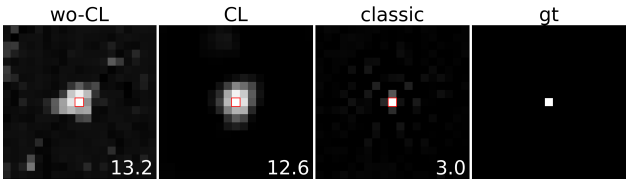

Figure 3: Reconstruction results of SM voxel (red square, ground truth (gt)) without (wo-CL) and with (CL) contrastive learning and using classic reconstruction. NRMSE values are given in % in the images.

3 iter.). The reconstruction results are also depicted in Figure 2.
The normalized root MSE (NRMSE) between recovered and ground-truth SM measurement over all time steps is: 2.2 % ± 0.4 % (wo-CL) and 1.9 % ± 0.4 % (CL).

For the reconstruction task, the NRMSE over all tested SM voxels is: 16.0 % ± 5.7 % (wo-CL) and 16.4 % ± 4.5 % (CL). An example for one reconstructed voxel can be seen in Figure 3 for wo-CL, CL and using classic reconstruction ($\lambda = 0$; 1 iter.) with the training SM.

For both tasks, the contrastive learning pretraining yields both quantitative and qualitative improvements. Visually, this is not as prominent in the recovery task (Figure 2) as in the reconstruction task (Figure 3) but the former task is also considerably easier. Nonetheless, the reconstruction task shows the possibility to generate MPI time traces with deep learning. As training and test data used here were acquired using the same particles and on a similar voxel grid, the generalizability of this approach has to be investigated further.

In terms of the reconstruction task, the results look promising, with a clear benefit from the contrastive learning pretraining, visible in Figure 3. It is also the first time that a learned direct reconstruction of real MPI data is performed in 2D, if only for single illuminated voxels and with a high NRMSE over all test samples.

This shows the general applicability of contrastive learning to MPI time traces but a thorough investigation, especially of the influence of the used augmentations, will be subject to future work. Other ablation studies are also necessary to investigate the influence of contrastive learning in contrast to using more data or e.g. combining both main tasks in a simultaneous training setting.

IV. Conclusion

In this work, we show SM recovery based on the prediction of 1D voltage signals and the direct reconstruction of MPI measurements to 2D concentration distributions. Both tasks use the same encoding network that can be pretrained using a contrastive learning approach, which showed a benefit for both main tasks. The pretraining can be performed unsupervised on any available MPI measurements. In the future, the presented framework could enable the prediction of high resolution SMs from very few calibration measurements. This would be an alternative to classic simulations that usually cannot sufficiently resemble real measurements.

Direct reconstruction without the need for an SM would be the fastest and most versatile solution for MPI but is difficult due to the multitude of parameters the measurement signal depends on. Here, the application of contrastive learning as pretraining shows promising results. Another possibility could be, rather than generating the tracer concentration distribution, only predicting parameters like particle binding status or temperature, from the measured voltage signal.

Author's statement

Research funding: We acknowledge the financial support by the German research foundation (DFG, grant number SCHU 2973/5-1).
Conflict of interest: Authors state no conflict of interest.

References

[1] F. Schrank, D. Pantke, and V. Schulz. Deep learning MPI super-resolution by implicit representation of the system matrix. *International Journal on Magnetic Particle Imaging*, 8(1, Suppl 1), 2022, doi:10.18416/IJMPI.2022.2203020.

[2] A. von Gladiss, R. Memmesheimer, N. Theisen, A. C. Bakenecker, T. M. Buzug, and D. Paulus, Reconstruction of 1D images with a neural network for Magnetic Particle Imaging, in *Bildverarbeitung für die Medizin 2022*, 247–252, Wiesbaden: Springer, 2022. doi:10.1007/978-3-658-36932-3_52.

[3] J. Spijkervet and J. A. Burgoyne. Contrastive Learning of Musical Representations. *International Society for Music Information Retrieval Conference*, 2021, doi:10.48550/arxiv.2103.09410.

[4] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton. A Simple Framework for Contrastive Learning of Visual Representations. *International Conference on Machine Learning*, 2020, doi:10.48550/arxiv.2002.05709.

[5] T. Knopp, P. Szwargulski, F. Griese, and M. Gräser. OpenMPIData: An initiative for freely accessible magnetic particle imaging data. *Data in Brief*, 28, 2020, doi:10.1016/j.dib.2019.104971.

[6] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations*, 2015, doi:10.48550/ARXIV.1412.6980.

[7] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *IEEE International Conference on Computer Vision*, pp. 1026–1034, 2015, doi:10.1109/ICCV.2015.123.

