

Few Shot Domain Adaptation for *in situ* Macromolecule Structural Classification in Cryo-electron Tomograms

– Supplementary Document

Liangyong Yu¹, Ran Li², Xiangrui Zeng¹, Hongyi Wang³, Jie Jin⁴, Ge Yang⁴, Rui Jiang²,
and Min Xu^{*1}

¹Computational Biology Department, Carnegie Mellon University, Pittsburgh, 15213, USA

²Department of Automation, Tsinghua University, Beijing, 100084, China

³Department of Electronic Engineering, Tsinghua University, Beijing, 100084, China

⁴Institute of Automation, Chinese Academy of Science, Beijing, 100190, China

S1 Related work

Current analysis of cryo-ET includes template matching [2]. First we create the templates for every class; and for every subtomogram, we calculate the matching score between the template and itself. The method is straight-forward and easy to realize, but the computation complexity is unbearable, especially on the data which has countless classes and more dimensions compared to traditional images. What’s more, because of the intense disruption of noise, the error rate of this method is very high.

Another method utilizes unsupervised subtomogram classification (e.g. [8]). Now there are a set of subtomograms which correspond to k classes. First we initialize k class centers, each of which represents the average of subtomograms in each class, and therefore all of subtomograms can be classified by computing the distance to k class centers. Second, after labeling all of subtomograms, we redirect k class centers by calculating average of labeled subtomograms in each class. By computing the two above steps iteratively, we can approximately obtain the label of every subtomogram. This method doesn’t require the label of any subtomogram, reducing the workload of labeling our data. Nevertheless, even if some tactics eliminating noises has been elaborated in this paper, the noises are still remaining a severe problems in our subtomogram classification, which extremely affect the performance in this method adversely.

There is another straightforward resolution that we can implement transfer learning into subtomogram classification. We generate simulated dataset in the computer as source domain and set real dataset as target domain. With the development of Neural Network, many Deep Learning models use this tactic to solve this problem. [4] proposed an unsupervised classification method with transfer learning. First we train a CNN model by simulated dataset. In the second stage, we remove the last layer from the original model and then extract the feature vector of the real dataset. In the end, we apply k-means clustering to the feature vector of the real dataset.

Recently [3] applies Unsupervised Domain Adaptation to subtomogram classification, in order to resolve the situation, in which source domain(train dataset) and target domain(test dataset) have different image intensity distribution. Even though it reaches a desirable performance in target domain, there still remains limitations because no label in target domain is utilized; what’s more, adversarial training is used in this paper. In Section 2.2, we have discussed in detail that adversarial training is hard to be convergent when using Cryo-ET. Comparing to this paper, we success to utilize the label information in target domain and further improve its performance.

*Corresponding author email: mxu1@cs.cmu.edu

There are two mainstream ways for current Unsupervised Domain Adaptation methods to decrease image intensity distribution. First, the training dataset is used to optimize our model and later the parameter of this model is fine-tuned by test dataset. Even if the label information of test dataset isn't available, some of its global features, such as mean and covariance, can still be calculated. This kind of information is crucial for us to fine-tune the parameters of our model. For instance, [6] and [7] use this way as domain adaptation. The second way is transforming the data in target domain, making its distribution more similar to the data in source domain. Compared to the first way, parameters of the model would not be fine-tuned. For instance, [1] use this way as domain adaptation. In this paper, whitening and re-coloring, which utilizes the covariance of data in source domain and data in target domain, are applied to data in source domain. The source data being transformed are used to train the classification model. Because transformed source data has the similar distribution with target data, the model can reach a desirable result on target domain.

Compared to Unsupervised Domain Adaptation, Few-shot Domain Adaptation utilizes the whole data in source domain and very few labeled data in target domain. The core idea is very similar to Few Shot Learning[5]: we require our model to learn the features in very few images. However, the two fields still have very significant difference. Few shot learning needs to learn the image features whose labels aren't presented in training dataset, while few shot domain adaptation needs to learn the image features whose domain are different from training dataset.

S2 Time complexity

We test the time complexity of FSFT and other Deep Learning methods, which is presented in Table S1.

Table S1: This table lists cost time of five Deep Learning method. FSFT costs less time than SWD and Fine-tune while costs more time than FADA and CORAL.

Model	Time Cost(s)
CORAL	323.66
SWD	1797.37
Fine-tune	1002.08
FADA	554.80
FSFT	921.36

From the table, CORAL, as the simplest method, costs the least time. Compared to FADA, FSFT add Deep CORAL as one of crucial stage, and its model is more complex. These changes introduce more computation, in order to have a better performance in subtomogram classification.

S3 Result Analysis

We conduct some experiments to analyze to verify the effectiveness of FSFT. We generate 23 classes for the simulated datasets S_1 and S_2 which are mentioned in Section 3.1.1. In this section, we want to verify the effectiveness of each stage and each contribution we proposed. All the experiments in this section are conducted on these datasets.

Firstly, in the task of subtomogram classification, we split the whole training procedure into 3 stages which are tightly linked. We verify that each stage plays an important role in improving the classification precision. Table S2 presents the improvement of each stage in FSFT. Unsupervised Domain Adaptation is used in Stage 2 and Supervised Domain Adaptation is used in Stage 3. The combination of them enable encoders $f^0 \circ f^t$ to adapt the target domain.

Secondly, the contributions mentioned in Section 1 are effective in improving the performance of FSFT. In order to verify their effectiveness, we remove each contribution in FSFT and test its performance in target domain. In table S3, FSFT method we proposed realizes the best performance, while others can't reach the optimal accuracy compared to FSFT.

Table S2: We calculate the classification accuracy in each stage. This table shows that in Stage 2.1, Deep CORAL method improves the accuracy from 37.2% to 70.9%; In Stage 2.2, we only update the parameter of discriminator, so the prediction accuracy is the same as Stage 2.1. In Stage 3, Fine-tune the encoder f_ϕ improves the accuracy from 70.9% to 95.3%

Stage	Accuracy
Stage 1	37.2%
Stage 2.1	70.9%
Stage 2.2	70.9%
Stage 3	95.3%

Table S3: Accuracy of ablation study. Row 1 corresponds to FSFT we proposed; Row 2 corresponds to FSFT without Stage 2.1; Row 3 corresponds to FSFT which use GAN to train discriminator and encoder; Row 4 corresponds to FSFT which only use 1D discriminator.

Model	Accuracy
FSFT	95.3%
FSFT without CORAL	79.1%
FSFT with GAN	94.2%
FSFT without 3D Discriminator	95.1%

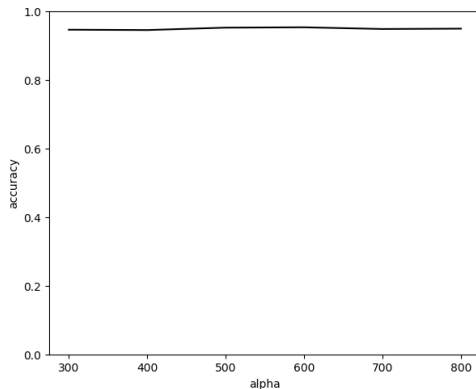


Figure S1: This picture shows how α affect the prediction accuracy.

S4 Hyper-parameter Adjustment

In this section, we discuss how to choose the hyper-parameter, for example, α in equation 6. In figure S1, the best result is 0.954%. Accuracy, as α changes, the accuracy nearly stays constant. The value of α will have little effect on the performance of our model.

References

- [1] Md Jahangir Alam, Gautam Bhattacharya, and Patrick Kenny. Speaker verification in mismatched conditions with frustratingly easy domain adaptation. In *Odyssey*, pages 176–180, 2018.
- [2] Martin Beck, Johan A Malmström, Vinzenz Lange, Alexander Schmidt, Eric W Deutsch, and Ruedi Aebersold. Visual proteomics of the human pathogen leptospira interrogans. *Nature methods*, 6(11):817, 2009.
- [3] Ruogu Lin, Xiangrui Zeng, Kris Kitani, and Min Xu. Adversarial domain adaptation for cross data source macromolecule in situ structural classification in cellular electron cryo-tomograms. *Bioinformatics*, 35(14):i260–i268, 2019.
- [4] Emmanuel Moebel. *New strategies for the identification and enumeration of macromolecules in 3D images of cryo electron tomography*. Theses, Université de Rennes 1, February 2019.
- [5] Jake Snell, Kevin Swersky, and Richard Zemel. Prototypical networks for few-shot learning. In *Advances in neural information processing systems*, pages 4077–4087, 2017.
- [6] Baochen Sun and Kate Saenko. Deep coral: Correlation alignment for deep domain adaptation. In *European Conference on Computer Vision*, pages 443–450. Springer, 2016.
- [7] Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7167–7176, 2017.
- [8] Min Xu, Martin Beck, and Frank Alber. High-throughput subtomogram alignment and classification by fourier space constrained fast volumetric matching. *Journal of structural biology*, 178(2):152–164, 2012.