Semi-Supervised Multi-Output Segmentation of Type-B Aortic Dissection

Denis Mikhailapov[0009-0009-1105-4358] and
Vladimir Berikov[0000-0002-5207-9764]

Sobolev Institute of Mathematics of the SB RAS, Novosibirsk 630090, Russia

Abstract. Convolutional neural networks (CNN) for multi-class segmentation of medical images are widely used today. Especially models with multiple outputs that can separately predict segmentation classes (regions) without relying on a probabilistic formulation of the segmentation of regions. These models allow for more precise segmentation by tailoring the network's components to each class (region). They have a common encoder part of the architecture but branch out at the output layers, leading to improved accuracy.
These methods are used to diagnose type B aortic dissection (TBAD), which requires accurate segmentation of aortic structures based on the ImageTBDA dataset, which contains 100 3D computed tomography angiography (CTA) images. These images identify three classes: true lumen (TL), false lumen (FL), and false lumen thrombus (FLT) of the aorta, which is critical for diagnosis and treatment decisions. In the dataset, 68 examples have a false lumen, while the remaining 32 do not, creating additional complexity for pathology detection.
However, implementing these CNN methods requires a large amount of high-quality labeled data. Obtaining accurate labels for the regions of interest can be an expensive and time-consuming process, particularly for 3D data. Semi-supervised learning methods allow models to be trained by using both labeled and unlabeled data, which is a promising approach for overcoming the challenge of obtaining accurate labels. However, these learning methods are not well understood for models with multiple outputs.
This paper presents a semi-supervised learning method for models with multiple outputs. The method is based on the additional rotations and flipping, and does not assume the probabilistic nature of the model's responses. This makes it a universal approach, which is especially important for multiple outputs architectures.

Keywords: 3D convolutional neural networks, multi-output segmentation, semi-supervised learning, type-B aortic dissection.

1 Introduction

Type B aortic dissection (TBAD) is a serious cardiovascular condition where damage to the aorta's inner layer creates two blood channels: the true lumen (TL) and false lumen (FL). Without timely treatment, it can be life-threatening.

Each year, approximately 3–4 per 100,000 people develop TBAD, and about 20% die without proper care [1]. Early diagnosis improves survival. Treatment options include medical therapy and thoracic endovascular intervention. With early accurate diagnosis, the 30-day mortality rate drops below 10% [2].

Computed tomography angiography (CTA) is commonly used for diagnosis. Segmentation of TL and FL is critical but manual slice-by-slice segmentation is labor-intensive and challenging due to complex structures or poor contrast, highlighting the need for automated methods.

Some rule-based strategies have shown promising results for TBAD segmentation [3, 4, 5]. However, they still require human intervention and lack detailed 3D information.

Deep learning methods have shown potential for accurate medical image segmentation using large annotated datasets [6, 7, 8]. Epifanov et al. [9] proposed a one-stage multitask deep learning approach for simultaneous vessel segmentation and centerline extraction using a hybrid convolutional-graph architecture with direct polyline prediction. The approach outperformed state-of-the-art methods on 142 CT angiography images and demonstrated robustness to input transformations compared to VMTK. Chen et al. [10] proposed a multi-stage framework for segmenting true lumen, false lumen, and branches in type B aortic dissection. They developed an aorta straightening method to simplify curvature before segmentation. Kesavuori et al. [11] compared 2D and 3D CNNs for aorta segmentation, finding that 2D CNN performed better on outer surface accuracy while overall accuracy was similar.

Cao et al. [12] explored three segmentation strategies: single one-task, single multi-task, and serial multi-task. The serial multi-task model, combining two subnets, achieved the best results.

However, these approaches require large labeled datasets. Collecting unlabeled or weakly-labeled data is easier. Semi-supervised learning (SSL) methods leverage unlabeled data via augmentations and additional models.

We propose a new SSL method for serial multi-task model training. It extracts useful information from data, reduces the impact of inaccurate masks, and lowers development costs. Our approach includes special preprocessing for ImageTBAD data and uses random rotations, flipping, and pseudo-label generation via an Exponential Moving Average (EMA) model.

Notably, existing TBAD segmentation works use private datasets, not multi-output models. We use an open dataset with 100 labeled examples. Therefore, we focus on developing and studying a semi-supervised segmentation method rather than directly comparing results.

The rest of this manuscript is organized as follows: In the next section, we observe the related work, specifically pseudo-labeling and consistency regulation methods. Section 3 describes the materials and methods. In particular, 3.1 and 3.2 describe the dataset and the preprocessing method, 3.3 and 3.5 describe the model and the evaluation metric, 3.4 describe our semi-supervised learning method, and 4 describe experiments. Results and conclusion in section 4 and 5, respectively.

2 Related work

Semi-supervised learning (SSL) is important for medical image segmentation where labeled data is scarce. SSL methods are divided into five categories: pseudo-labeling, consistency regularization, contrastive learning, GAN-based, and hybrid approaches [18]. We focus on pseudo-labeling and consistency regularization, as our method builds on them.

2.1 Pseudo-Labeling

Pseudo-labeling is a technique where an initial model, trained on labeled data, is employed to assign labels to unlabeled data. Common approaches include self-training [28, 29, 30, 31], where a model iteratively refines predictions, and co-training [32, 33], where multiple models train on complementary data views and share high-confidence predictions. A key challenge is pseudo-label quality; poor initial predictions introduce noise.

The Exponential Moving Average (EMA) technique [39] improves stability and reliability by aggregating model states over traning epochs, reducing noise in pseudo-labels.

2.2 Consistency Regularization

Consistency regularization [36, 37, 38] asserts that model predictions should remain stable under input perturbations. Oliver et al. [34] classify consistency regularization into three distinct categories: (1) data consistency — stable predictions under input transformations; (2) model consistency — consistent outputs across different models (e.g., Mean Teacher [41]); (3) task consistency — uniform outputs across tasks, improving robustness.

2.3 Other approaches

Contrastive learning [42, 43, 44] maximizes similarity between related instances. GAN-based methods [45, 46] use a discriminator to distinguish real from generated data. Hybrid methods combine consistency regularization with uncertainty estimation. These diverse strategies highlight the importance of SSL for medical image analysis.

3 Material and Methods

This section describes the materials and methods for a supervised 3D segmentation system trained on labeled and unlabeled data. Spatial transformations (rotation and flipping) were applied for data augmentation, and pseudo-labels were generated using an EMA model. Preprocessing included resizing to $128 \times 128 \times 128$ voxels and intensity transformations. The DICE coefficient was used for evaluation. Five experiments with varying labeled data proportions assessed the impact of unlabeled data with and without augmentation to improve segmentation accuracy for type B aortic dissection.

3.1 Dataset

The ImageTBAD dataset [14] contains 100 preoperative CT angiography (CTA) scans of type B aortic dissection (TBAD), acquired between January 2013 and April 2015 in the axial plane. The scan range extends from the neck to the brachiocephalic vessels. Segmentation labeling was performed by experienced cardiologists and radiologists, producing three substructures: true lumen (TL), false lumen (FL), and false lumen thrombus (FLT). Among the images, 68 contain FLT and 32 do not. Previous studies have reported that FLT is difficult to segment, with the best DICE score approximately 0.52.

The primary goal of this article is to investigate semi-supervised learning. To simplify the segmentation task and better explore the proposed method, a new class, conditionally denoted as "ALL", was formed as the union of the three substructures. The final dataset includes three segmentation classes: ALL, TL, and FL.

3.2 Data preprocessing

The source data was reduced in size to 192 pixels in the XY plane. Then, 32 pixels were removed from the image border in the same plane. As a result, the final image size was 128 by 128 pixels along the X and Y axes, and the Z axis was also reduced to 128 pixels. The final size of each 3D image became equal to (128, 128, 128).

There was also an intensity restriction on the Hounsfield scale, the application of standardization, exponential functions and final MinMax normalization. This helps to highlight important structures in the image and facilitate the segmentation process. An example of image processing code is shown in Listing 1.1. An example of the resulting image can be seen in Fig 1.

Listing 1.1: Image preprocessing code

```python
import numpy as np
from scipy import ndimage

def process_image(
    data_img: np.ndarray,
    vmin=1100,
    vmax=1600,
    exp_coef=1.3
):
    data_img[data_img < vmin] = 0
    data_img[data_img > vmax] = 0
    mask = ndimage.grey_erosion(data_img, size=(2, 2, 1))
    data_img[mask == 0] = 0

    data_img -= 2 * np.mean(data_img)
    data_img /= np.std(data_img + 1e-6)
    data_img = np.exp(exp_coef * data_img)
```

```
    data_img = data_img - data_img.min()
    data_img /= (data_img.max() - data_img.min())

    return data_img
```

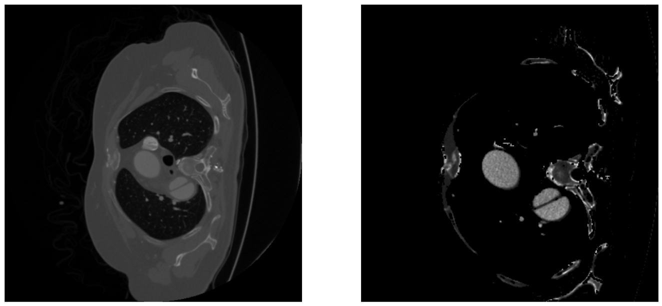

Fig. 1: An example of data preprocessing, before and after.

3.3 Model

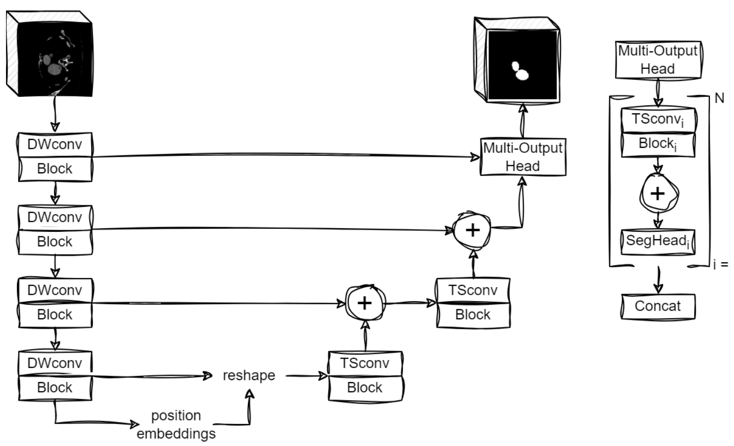

Fig. 2: Multi-Output Slim UNETR Architecture.

In this work, we used Slim UNETR model [19] with multi-output head (Fig. 2). The architecture consists of three main components: the down-scale part, the bottleneck and the up-scale part. The down-scale part includes several convolutional blocks that are used in the Slim UNETR architecture, then bottleneck applies positional encoding to downscale feature maps, and finally, the up-scale part decodes the segmentation mask from the feature maps. A key feature of the multi-output architecture is the use of separate output blocks for each class in a segmentation mask. This can improve segmentation accuracy because it differentiates the specialized segmentation mask of the decoding class; however, the resulting segmentation masks are not adjusted relative to each other.

3.4 Proposed Method

In this section, we introduce our proposed semi-supervised learning (SSL) method, which combines supervised and unsupervised approaches. Supervised learning uses labeled 3D samples to train the model for accurate 3D segmentation, optimized with Generalized Dice Loss (GDL) and Focal Loss (FL). In unsupervised learning, pseudo-labels are generated using an Exponential Moving Average (EMA) model and data augmentations, where the EMA model relies on the main model's weights. The learning method is detailed below.

Mathematical definitions This section provides the Mathematical definitions used throughout the remainder of this paper.

PROVE SET Let $x \in \mathcal{X} \subset \{-1024, \ldots, +3071\}^{H \times W \times D}$ denote a CT scan of size $H \times W \times D$, and let $y \in \mathcal{Y} \subset \{0, 1\}^{H \times W \times D \times C}$ denote the corresponding ground-truth annotation provided by specialists, of size $H \times W \times D \times C$. Furthermore, let y_p denote a pseudo-label of the same dimensions obtained via the semi-supervised learning (SSL) algorithm. Here, C represents the additional dimension corresponding to the segmented classes, and $y_{i,j,k,c}$ denotes a specific value in the mask at indices i, j, k, c.

The labeled dataset is defined as $D_l = \{(x_l, y_l)\}$, and the unlabeled dataset is defined as $D_u = \{(x_u)\}$.

Let f_θ denote the segmentation model with trainable parameters θ, and let $\hat{y} = f_\theta(x)$ be the model prediction for input x. To generate pseudo-labels on unlabeled data, an exponential moving average (EMA) update of the model parameters is employed, producing predictions $f_{\theta^{\mathrm{EMA}}}(x_u) = y_p$.

The loss function is denoted by $\mathcal{L}(\theta, D)$ and $L(y, \hat{y})$. The former notation emphasizes the parameters being optimized and the data dependency, while the latter is used as a direct measure of the discrepancy between segmentation masks.

Supervised learning The weights of the neural network are optimized by minimizing a loss function using the backpropagation algorithm. However, the data in the dataset may be unbalanced, which means that some classes are underrepresented. For example, The FLT class is only present in 68 out of the 100 examples in our dataset, and the TL and FL classes also have different numbers of examples. It is therefore important to use loss functions that are able to account for this class imbalance.

In our work, we use two loss functions: GeneralizedDiceLoss (GDL) [21] and FocalLoss (FL) [20]. The study [22] shows that using a mixed loss function can improve the quality of the final model.

To directly enhance the Dice metric, we use the Generalized Dice loss function:

$$L_{GDL}(y, \hat{y}) = 1 - 2\frac{\sum_c w_c \sum_{i,j,k} \hat{y}_{i,j,k,c} \cdot y_{i,j,k,c} + \varepsilon}{\sum_c w_c (\sum_{i,j,k,c} \hat{y}^2_{i,j,k,c} + \sum y^2_{i,j,k,c}) + \varepsilon}.$$

where i, j, k are the voxel indices, c is the class index, $w_c = 1/(\sum_{i,j,k} y_{i,j,k,c})^2$.

To work with unbalanced classes, we use Focal loss function:

$$L_{FL}(y,\hat{y}) = -\sum_{i,j,k,c} \alpha_c [y_{i,j,k,c}(1-\hat{y}_{i,j,k,c})^{\gamma} \cdot$$
$$\cdot log(\hat{y}_{i,j,k,c}) + (1-y_{i,j,k,c})\hat{y}^{\gamma}_{i,j,k,c} \cdot log(1-\hat{y}_{i,j,k,c})]$$

The hyperparameter γ addresses the issue of class imbalance in the dataset. A higher value of γ reduces the impact of easy-to-classify examples on the training process, while a lower value increases it. In addition, we use a class weight multiplier α_c to improve accuracy for small classes.

As the final supervised loss function, a combination of L_{FL} and L_{GDL} was used:

$$L(y,\hat{y}) = w_{GDL} L_{DICE}(y,\hat{y}) + w_{FL} L_{FL}(y,\hat{y}).$$

Hyperparameters w_{GDL} and w_{FL} are used to balance the two loss terms because L_{GDL} and L_{FL} differ in scale and convergence dynamics.

Values for all hyperparameters, including w_{GDL}, w_{FL}, γ, and α_c, can be found in the section **3.7 Implementation Details**.

Unsupervised learning The process of unsupervised learning is similar to supervised learning, but it does not require ground truth samples.

EMA model: To obtain these ground truth (pseudo-labels), the EMA model is used: First, create a copy of the main model, including its architecture and parameters. Then, during each training epoch, the EMA model predicts the labels for the unlabeled samples, producing the pseudo-labels and then updates its own parameters using the main model's parameters according to the following formula:

$$\theta^{EMA}_t = \mu\theta^{EMA}_{t-1} + (1-\mu)\theta_t$$

where θ^{EMA}, θ parameters of models, $\mu \in [0, 1]$ represents a smoothing factor that usually approximately equals 0.95.

This formula can be rewritten as follows:

$$\theta^{EMA}_t = \theta_t + \mu(\theta^{EMA}_{t-1} - \theta_t)$$

here we can note that the term $\theta^{EMA}_{t-1} - \theta_t$ is the delta of the weights of the models. Hence the parameter μ adjusts the delay rate of the model.

The EMA approach provides several advantages for predicting pseudo-labels in semi-supervised learning. It stabilizes the training process, reducing noise and enhancing model generalization. Additionally, the EMA model provides consistent and reliable targets for pseudo-labeling by acting as a form of models aggregation. Its smoothing effect also helps reduce overfitting, which is particularly beneficial when working with limited labeled data.

Description of the algorithm for obtaining pseudo-labels: The general algorithm for generating pseudo-labels can be described as follows:

1. First, applying random rotation and flip augmentations to the unlabeled image. Then, the prediction of the model is calculated ($\hat{y}_u$).

2. After that, the pseudo-label is calculated using the EMA model. And then, the same augmentations are applied to this pseudo-label without randomness. Finally, it is binarized, so that each pixel has a value of 0 or 1 (y_p).
3. At the end, the loss function is calculated as in the supervised part.

The Figure 3 shows a diagram of the method.

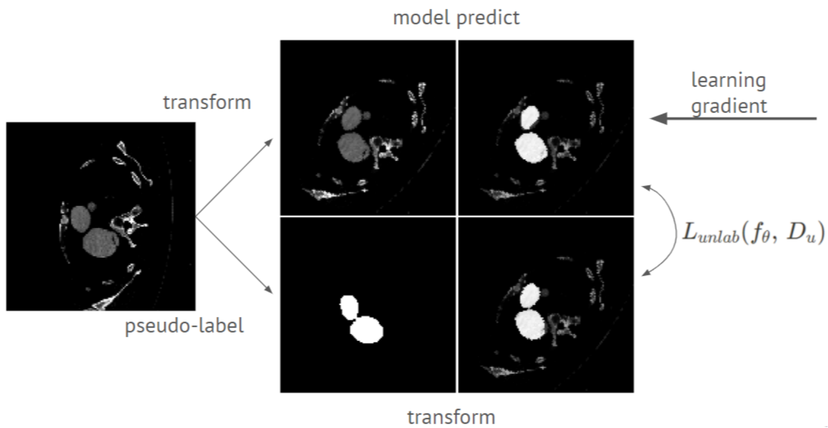

Fig. 3: The scheme of the semi-supervised method.

Note that unlabeled data is used in the learning process not from the beginning of learning, but after a certain epoch. This is necessary for the model to learn the initial understanding of the data and the segmentation tasks.

The summary of our semi-supervised learning method is presented in Algorithm 1. Note that service function $rand\_seed(a, b)$ returns a random value in the range $[a\, b]$, no_grad_θ is a function that disables gradient computation during training, and $\mathcal{T}_s(x)$ is a spatial random transformation (rotation and flip) parameterized by s.

3.5 Metric of Model quality

The quality of predictions made by the constructed models was evaluated using the DICE coefficient:

$$DICE_c\,(y, \hat{y}) = \frac{2 \sum_{i,j,k} \hat{y}_{i,j,k,c} \cdot y_{i,j,k,c} + \varepsilon}{\sum_{i,j,k} \hat{y}^2_{i,j,k,c} + \sum_{i,j,k} y^2_{i,j,k,c} + \varepsilon}$$

where $y = \{y_{i,j,k,c}\}$ corresponds to the segmentation mask delineated by an expert; $\hat{y} = \{\hat{y}_{i,j,k,c}\}$ is the predicted mask generated by the model; i, j, k are voxel indices, c is class index and ε is a small positive value.

3.6 Experiments performed

In this study, the following experiments were performed:

Algorithm 1 Semi-Supervised Learning

Input:
$D_l = \{(x_b^L, y_b^L)\}_{b=1}^N$: labeled dataset
$D_u = \{x_b^U\}_{b=1}^M$: unlabeled dataset
T: total epochs
$\hat{T} < T$: epoch to start unsupervised loss
$\mu \in [0,1]$: EMA decay
Output: model parameters θ_T

    θ_0: initial parameters
    $\theta_0^{\text{EMA}} := \theta_0$
    for $t = 1, \ldots, T$ **do**
        $s' := \text{randint}(0, 100)$ ▷ random seed for augmentation
        $\mathcal{L}_l(\theta_{t-1}) = \sum_{(x_l, y_l) \in D_l} \ell\big(f(\theta_{t-1}, \mathcal{T}_{s'}(x_l)),\, y_l\big)$
        $\theta_t := \theta_{t-1} + \eta \nabla_\theta \mathcal{L}_l(\theta_{t-1})$
        if $t > \hat{T}$ **then**
            $\theta_t^{EMA} := no\_grad_{\theta EMA}(\theta_t + \mu(\theta_{t-1}^{EMA} - \theta_{t-1}))$ ▷ stop gradients here
            $s'' := \text{randint}(0, 100)$
            $\mathcal{L}_u(\theta_t, \theta_t^{\text{EMA}}) = \sum_{x_u \in D_u} \ell\Big(f(\theta_t, \mathcal{T}_{s''}(x_u)),\, f(\theta_t^{\text{EMA}}, \mathcal{T}_{s''}(x_u))\Big)$
            $\theta_t := \theta_t + \eta \nabla_\theta \mathcal{L}_u(\theta_t, \theta_t^{\text{EMA}})$
    return θ_T

1. Training model on a full labeled dataset.
2. Training model on 50% of the labeled data.
3. Application of Algorithm 1 on 50% of the labeled data.
4. Applying Algorithm 1 on 50% of the labeled data with data augmentation.

3.7 Implementation details

The number of training epochs is 800, the size of the batch is 6, $\gamma = 2., w_{GDL} = 1., w_{FL} = 0.8$, a class weight multipliers $\alpha = [0.8, 0.9, 1.5]$, for each class, respectively. Data splitting into training and validation is 80/20%, AdamW is used as an optimizer. The training code is originally taken from the official implementation of Slim-UNETR on github[1]. The training was conducted on an Nvidia Tesla P100 16gb GPU.

4 Results

Table 1 presents DICE metrics for three classes. We validated our model using Shuffle&Split cross-validation with four random splits. The "±" symbol indicates the standard deviation across trials.

Supervised learning on 50% of data shows lower accuracy compared to full-data training, confirming that data scarcity degrades model performance. Metrics behave similarly until approximately the 100th epoch, after which divergence begins [40]. This suggests the end of the neural network optimization drift phase and the start of the diffusion phase.

Adding unlabeled data improves DICE metrics, especially for the "false lumen" (FL) class. Notably, metrics temporarily drop when unlabeled data is

[1] https://github.com/aigzhusmart/Slim-UNETR

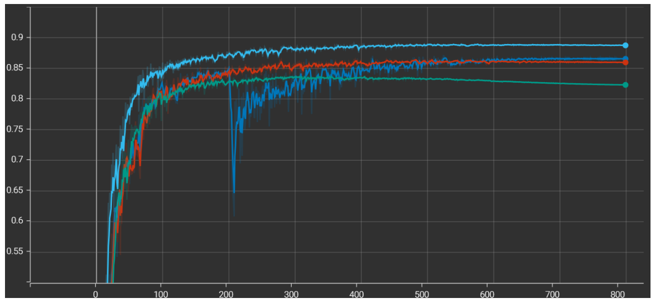

Fig. 4: The DICE (ALL) metric in the learning process (same random seed): red – experiment 1, green – experiment 2, blue – experiment 3, light blue – experiment 4

introduced — likely because the model must adapt to new examples, requiring more training time. This is visible in Figure 4 at the 200th epoch. However, by the end of training, the metric rises to the level achieved with the full dataset.

Table 1: Class-wise DICE score for different experiment settings

№	ALL	TL	FL
1	$86.51_{\pm 0.28}$	$76.32_{\pm 0.54}$	$60.84_{\pm 0.47}$
2	$83.04_{\pm 0.3}$	$72.2_{\pm 0.48}$	$58.04_{\pm 0.36}$
3	$85.82_{\pm 0.51}$	$77.36_{\pm 0.8}$	$63.9_{\pm 0.9}$
4	$87.38_{\pm 0.21}$	$78.42_{\pm 0.73}$	$63.37_{\pm 1.2}$

In this work, we also compared the model's accuracy for different start epochs of the semi-supervised method (see Figure 5). It turned out that, on average, the best option is to start the learning process unsupervised with epoch 200. At the start epoch of 100, the model does not have time to study the labeled data, however, at the start epoch of 300, the model no longer has time to study the unlabeled data.

Moreover, the high standard deviation of metrics makes it difficult to provide a clear answer.

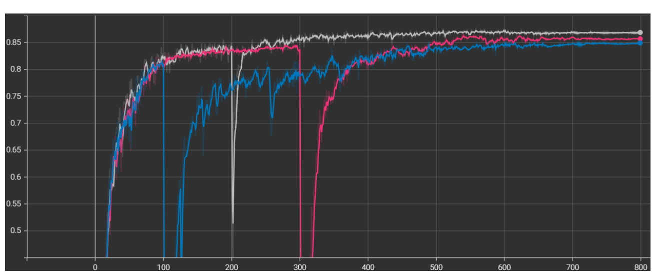

Fig. 5: The DICE (ALL) metric for different start epoch for unlabeled training: blue – start epoch 100, gray – start epoch 200, pink – start epoch 300

Table 2: Compare DICE score for different unlabeled learning start epoch (100, 200 and 300)

Start epoch	ALL	TL	FL
100	$85.02_{\pm 0.8}$	$76.35_{\pm 0.4}$	$60.92_{\pm 0.8}$
200	$85.82_{\pm 0.3}$	$77.36_{\pm 0.8}$	$63.9_{\pm 0.8}$
300	$86.15_{\pm 0.3}$	$77.06_{\pm 0.6}$	$61.18_{\pm 0.6}$

5 Conclusion

n this study, we introduced a novel semi-supervised learning approach for the 3D multi-output segmentation model. Accurate segmentation of the true lumen, false lumen, and thrombus is essential for TBAD diagnosis and treatment planning, which motivates our work. By leveraging both labeled and unlabeled data, our method addresses the challenges associated with obtaining extensive high-quality labeled datasets, which can be both costly and time-consuming, particularly in 3D imaging contexts.

Our approach employs data augmentation techniques, such as random rotation and reflection, alongside an Exponential Moving Average (EMA) model for generating pseudo-labels. Importantly, the method is probability-free – it does not assume a probabilistic interpretation of the model's outputs. This makes it universally applicable to multi-output architectures, where standard probabilistic assumptions may not hold. The combination of these techniques mitigates the impact of inaccurate segmentation masks and enhances the robustness of model training.

The evaluation of our method on the ImageTBAD dataset, comprising 100 CTA images, illustrates its effectiveness in improving segmentation accuracy. Overall, our findings highlight the potential of employing semi-supervised learning strategies in medical image segmentation, demonstrating that such methods can significantly reduce the dependency on annotated data while maintaining reliable and precise segmentation outputs. In the future, it is worth testing the method on more popular datasets, as well as considering cases with a smaller percentage of labeled data.

Acknowledgements The work of Denis Mikhailapov was supported by the State Contract of Sobolev institute of mathematics, project No. FWNF-2024-0002[2], to develop and conduct experiments. The work of Vladimir Berikov was supported by the State Contract of Sobolev institute of mathematics, project No. FWNF-2026-0005 for experiments develop and manuscript preparation.

[2] https://ai-biolab.ru/

Bibliography

[1] Karthikesalingam, A., Holt, P.J.E., Hinchliffe, R.J., Thompson, M.M., Loftus, I.M. (2010). The diagnosis and management of aortic dissection. Vascular and Endovascular Surgery, 44(3), 165–169.

[2] Hagan, P.G., Nienaber, C.A., Isselbacher, E.M., Bruckman, D., Karavite, D.J., Russman, P.L., Evangelista, A., Fattori, R., Suzuki, T., Oh, J.K., et al. (2000). The International Registry of Acute Aortic Dissection (IRAD): new insights into an old disease. JAMA, 283(7), 897–903.

[3] Fetnaci, N., Łubniewski, P., Miguel, B., Lohou, C. (2013). 3D segmentation of the true and false lumens on CT aortic dissection images. In Three-Dimensional Image Processing (3DIP) and Applications 2013 (Vol. 8650, pp. 176–190). SPIE.

[4] Morais, P., Vilaça, J.L., Queir'os, S., Bourier, F., Deisenhofer, I., Tavares, J.M.R.S., D'hooge, J. (2017). A competitive strategy for atrial and aortic tract segmentation Based on deformable models. Medical Image Analysis, 42, 102–116.

[5] Vassilevski, Y. V., Simakov, S. S., Gamilov, T. M., Salamatova, V. Y. E., Dobroserdova, T. K., Kopytov, G. V. E., ... & Shipilov, A. A. (2022). Personalization of mathematical models in cardiology: obstacles and perspectives. Computer research and modeling, 14(4), 911-930.

[6] Yang, H., Shan, C., Kolen, A.F., de With, P.H.N. (2022). Weakly-supervised learning for catheter segmentation in 3D frustum ultrasound. Computerized Medical Imaging and Graphics, 96, 102037.

[7] Cherikbayeva, L., Berikov, V., Melis, Z., Yeleussinov, A., Baigozhanova, D., Tasbolatuly, N., ... & Mikhailapov, D. (2025). Ensembling Transformer-Based Models for 3D Ischemic Stroke Segmentation in Non-Contrast CT. Applied Sciences, 15(17), 9725.

[8] Dobshik, A.V., Verbitskiy, S.K., Pestunov, I.A., Sherman, K.M., Tulupov, A.A., Berikov, V.B., et al. (2023). Acute ischemic stroke lesion segmentation in non-contrast CT images using 3D convolutional neural networks. Computer Optics, 47(5), 770–777.

[9] Epifanov, R., Fedotova, Y., Dyachuk, S., Gostev, A., Karpenko, A., & Mullyadzhanov, R. (2025). The Robust Vessel Segmentation and Centerline Extraction: One-Stage Deep Learning Approach. Journal of Imaging, 11(7), 209.

[10] Chen, D., Zhang, X., Mei, Y., Liao, F., Xu, H., Li, Z., Xiao, Q., Guo, W., Zhang, H., Yan, T., et al. (2021). Multi-stage learning for segmentation of aortic dissections using a prior aortic anatomy simplification. Medical Image Analysis, 69, 101931.

[11] Kes"avuori, R., Kaseva, T., Salli, E., Raivio, P., Savolainen, S., Kangasniemi, M. (2023). Deep learning-aided extraction of outer aortic surface from CT angiography scans of patients with Stanford type B aortic dissection. European Radiology Experimental, 7(1), 35.

[12] Cao, L., Shi, R., Ge, Y., Xing, L., Zuo, P., Jia, Y., Liu, J., He, Y., Wang, X., Luan, S., et al. (2019). Fully automatic segmentation of type B aortic dissection from CTA images enabled by deep learning. European Journal of Radiology, 121, 108713.
[13] Peng, J., Pedersoli, M., Desrosiers, C. (2020). Mutual information deep regularization for semi-supervised segmentation. In Medical Imaging with Deep Learning (pp. 601–613). PMLR.
[14] Yao, Z., Xie, W., Zhang, J., Dong, Y., Qiu, H., Yuan, H., Jia, Q., Wang, T., Shi, Y., Zhuang, J., et al. (2021). Imagetbad: A 3d computed tomography angiography image dataset for automatic segmentation of type-b aortic dissection. Frontiers in Physiology, 12, 732711.
[15] Peng, J., Pedersoli, M., Desrosiers, C. (2021). Boosting semi-supervised image segmentation with global and local mutual information regularization. arXiv preprint arXiv:2103.04813.
[16] Basak, H., Bhattacharya, R., Hussain, R., Chatterjee, A. (2022). An exceedingly simple consistency regularization method for semi-supervised medical image segmentation. In 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI) (pp. 1–4). IEEE.
[17] Isaksson, L.J., Summers, P., Raimondi, S., Gandini, S., Bhalerao, A., Marvaso, G., Petralia, G., Pepa, M., Jereczek-Fossa, B.A. (2022). Mixup (sample pairing) can improve the performance of deep segmentation networks. Journal of Artificial Intelligence and Soft Computing Research, 12(1), 29–39.
[18] Han, K., Sheng, V.S., Song, Y., Liu, Y., Qiu, C., Ma, S., Liu, Z. (2024). Deep semi-supervised learning for medical image segmentation: A review. Expert Systems with Applications, 123052.
[19] Pang, Y., Liang, J., Huang, T., Chen, H., Li, Y., Li, D., Huang, L., Wang, Q. (2023). Slim UNETR: Scale hybrid transformers to efficient 3D medical image segmentation under limited computational resources. IEEE Transactions on Medical Imaging.
[20] Ross, T-Y.L.P.G., Doll'ar, G.K.H.P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2980–2988).
[21] Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (pp. 240–248). Springer.
[22] Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848.
[23] Ran, L., Li, Y., Liang, G., Zhang, Y. (2024). Pseudo Labeling Methods for Semi-Supervised Semantic Segmentation: A Review and Future Perspectives. IEEE Transactions on Circuits and Systems for Video Technology.
[24] Lee, D.-H., et al. (2013). Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In Workshop on challenges in representation learning, ICML (Vol. 3, No. 2, p. 896). Atlanta.

[25] Yao, H., Hu, X., Li, X. (2022). Enhancing pseudo label quality for semi-supervised domain-generalized medical image segmentation. In Proceedings of the AAAI conference on artificial intelligence (Vol. 36, No. 3, pp. 3099–3107).
[26] Wang, X., Yuan, Y., Guo, D., Huang, X., Cui, Y., Xia, M., Wang, Z., Bai, C., Chen, S. (2022). SSA-Net: Spatial self-attention network for COVID-19 pneumonia infection segmentation with semi-supervised few-shot learning. Medical Image Analysis, 79, 102459.
[27] Zeng, L.-L., Gao, K., Hu, D., Feng, Z., Hou, C., Rong, P., Wang, W. (2023). SS-TBN: A semi-supervised tri-branch network for COVID-19 screening and lesion segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(8), 10427–10442.
[28] Chaitanya, K., Erdil, E., Karani, N., Konukoglu, E. (2023). Local contrastive loss with pseudo-label based self-training for semi-supervised medical image segmentation. Medical Image Analysis, 87, 102792.
[29] Qiu, Z., Gan, H., Shi, M., Huang, Z., Yang, Z. (2023). Self-training with dual uncertainty for semi-supervised medical image segmentation. arXiv preprint arXiv:2304.04441.
[30] Wu, H., Li, X., Lin, Y., Cheng, K.-T. (2023). Compete to win: Enhancing pseudo labels for barely-supervised medical image segmentation. IEEE Transactions on Medical Imaging, 42(11), 3244–3255.
[31] Xu, M., Zhou, Y., Jin, C., de Groot, M., Alexander, D.C., Oxtoby, N.P., Hu, Y., Jacob, J. (2024). Expectation maximisation pseudo labels. Medical Image Analysis, 94, 103125.
[32] Xia, Y., Liu, F., Yang, D., Cai, J., Yu, L., Zhu, Z., Xu, D., Yuille, A., Roth, H. (2020). 3d semi-supervised learning with uncertainty-aware multi-view co-training. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 3646–3655).
[33] Peng, J., Estrada, G., Pedersoli, M., Desrosiers, C. (2020). Deep co-training for semi-supervised image segmentation. Pattern Recognition, 107, 107269.
[34] Oliver, A., Odena, A., Raffel, C.A., Cubuk, E.D., Goodfellow, I. (2018). Realistic evaluation of deep semi-supervised learning algorithms. Advances in Neural Information Processing Systems, 31.
[35] Horlava, N., Mironenko, A., Niehaus, S., Wagner, S., Roeder, I., Scherf, N. (2020). A comparative study of semi-and self-supervised semantic segmentation of biomedical microscopy data. arXiv preprint arXiv:2011.08076.
[36] Qiao, P., Li, H., Song, G., Han, H., Gao, Z., Tian, Y., Liang, Y., Li, X., Zhou, S.K., Chen, J. (2022). Semi-supervised CT lesion segmentation using uncertainty-based data pairing and SwapMix. IEEE Transactions on Medical Imaging, 42(5), 1546–1562.
[37] Wang, Y., Xiao, B., Bi, X., Li, W., Gao, X. (2023). Mcf: Mutual correction framework for semi-supervised medical image segmentation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 15651–15660).
[38] Gao, S., Zhang, Z., Ma, J., Li, Z., Zhang, S. (2023). Correlation-aware mutual learning for semi-supervised medical image segmentation. In Interna-

tional Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 98–108). Springer.

[39] Cai, Z., Ravichandran, A., Maji, S., Fowlkes, C., Tu, Z., Soatto, S. (2021). Exponential moving average normalization for self-supervised and semi-supervised learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 194–203).

[40] Shwartz-Ziv, R., Tishby, N. (2017). Opening the black box of deep neural networks via information. arXiv preprint arXiv:1703.00810.

[41] Tarvainen, A., Valpola, H. (2017). Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Advances in Neural Information Processing Systems, 30.

[42] Pandey, P., Pai, A., bhatt, N., Das, P., Makharia, G., AP, P., et al. (2021). Contrastive semi-supervised learning for 2D medical image segmentation. arXiv preprint arXiv:2106.06801.

[43] Hu, X., Zeng, D., Xu, X., Shi, Y. (2021). Semi-supervised contrastive learning for label-efficient medical image segmentation. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2021 (pp. 481–490). Springer.

[44] Wang, K., Zhan, B., Zu, C., Wu, X., Zhou, J., Zhou, L., Wang, Y. (2022). Semi-supervised medical image segmentation via a tripled-uncertainty guided mean teacher model with contrastive learning. Medical Image Analysis, 79, 102447.

[45] Tan, Y., Wu, W., Tan, L., Peng, H., Qin, J. (2022). Semi-supervised medical image segmentation based on generative adversarial network. Journal of New Media, 4(3), 155.

[46] Wang, X., Li, J., Liu, Q., Zhao, W., Li, Z., Wang, W. (2022). Generative adversarial training for supervised and semi-supervised learning. Frontiers in Neurorobotics, 16, 859610.