Boundary Contrastive Learning for Label-Efficient Medical Image Segmentation

Satoshi Kamiya[1]
180442042@ccalumni.meijo-u.ac.jp
Kota Yamashita[2]
200442179@ccalumni.meijo-u.ac.jp
Kazuhiro Hotta[2]
kazuhotta@meijo-u.ac.jp

[1] Mitsubishi Electric Advanced Techhnology R&D Center
8-1-1 Tsukaguchi-honmachi
Amagasaki, Hyogo 661-8661, Japan

[2] Meijo University
1-501 Shiogamaguchi, Tempaku-ku,
Nagoya 468-8502, Japan

Abstract

Recent advances in deep learning have significantly improved the accuracy of medical image segmentation, yet the need for extensive annotations remains a challenge due to the high costs associated with practical implementation. In this study, we introduce Boundary Contrastive Learning for Label-Efficient (BCLL), a novel label-efficient learning method. The primary innovation of BCLL is the extension of contrastive loss concept used in conventional label-efficient learning methods like CLLE. We propose Boundary Contrastive Learning (BCL), which applies average pooling filters to annotated images to capture positional information about the boundary and internal regions of classes. By using features extracted from these regions, BCL computes various combinations of contrastive losses in a single image. Not only does this method bring features of the same class closer and push those of different classes apart, but it also designs contrastive losses to draw boundary region features closer to those of internal regions. This approach significantly enhances the accuracy of segmenting challenging boundary parts using only a small set of labeled data. Additionally, we have incorporated a new similarity function based on the Generalized Gaussian Distribution (GGD), named GGD-vMF, for the similarity calculations. This new similarity loss function enables enhanced learning with only minimal supervised data. Our experiments on Automatic Cardiac Diagnosis Challenge (ACDC), Synapse multi-organ segmentation (SMO), and Covid19 datasets demonstrated that BCLL achieves superior accuracy compared to the baseline and other label-efficient medical image segmentation methods. Specifically, BCLL showed an improvement in mIoU of 7.17% with 5% labels on ACDC, 7.74% with 5% labels on SMO, and 0.44% with 10% labels on Covid19 in comparison with baseline U-Net.

1 Introduction

Medical image segmentation using deep learning has achieved high precision in recent years and is increasingly applied in the real world, such as in diagnostics using CT and MRI scans. However, training these models still requires a substantial amount of images and annotations. The necessity for extensive training data in medical imaging represents a significant

drawback, as data collection is often limited by privacy concerns. Moreover, the human and financial costs of creating annotations hinder further real-world application of medical image segmentation. Recently, label-efficient learning methods [7], which train models from a limited number of annotations, have gained attention. CLLE [11] utilizes contrastive learning to train models by attracting features that belong to the same class and repelling features from different classes, using only a limited number of labeled images. The discrimination of feature classes is facilitated using the information from corresponding annotations. Nevertheless, CLLE faces challenges. Firstly, feature sampling for contrastive learning can be biased due to class imbalance, reducing segmentation accuracy, especially for small objects within images. Secondly, although contrastive learning generally requires extensive data, CLLE's reliance on limited labeled data can decrease the discriminative accuracy around complex object boundaries.

In this paper, we propose a novel learning method, Boundary Contrastive Learning for Label-Efficient (BCLL), which effectively learns from a limited number of labeled images. This method addresses two key challenges. First, to tackle the issue of poor segmentation accuracy for small objects due to class imbalance, BCLL creates a feature head for each class and performs equitable sampling from features specialized to each class. These features are then processed using the subsequent Boundary Contrastive Learning (BCL). Second, to improve the segmentation accuracy around object boundaries, we propose BCL, a new approach to contrastive learning. BCL applies an average pooling filter to annotated images to capture positional information of class boundary and interior regions. By using features extracted from these regions, BCL computes various combinations of contrastive losses within a single image. The design of these contrastive losses not only attracts features of the same class and repels those from different classes, as in CLLE, but also aims to bring features from boundary regions closer to those from interior regions. This is beneficial as the accuracy tends to be higher in interior regions than in boundary areas. This approach significantly enhances the segmentation of boundary areas by using only a limited number of labeled images. Furthermore, by incorporating GGD-vMF, a new similarity function based on the generalized Gaussian distribution, into the similarity calculations of BCL, we can facilitate more effective learning than with the normal cosine similarity or t-vMF [4] used in conventional contrastive learning.

In experiments, we demonstrate that our proposed method, BCLL, achieves superior accuracy compared to baseline U-Net and other label-efficient medical image segmentation methods on three datasets: Automatic Cardiac Diagnosis Challenge (ACDC) [1], Synapse multi-organ segmentation (SMO) [5], and Covid19 [10] datasets. Specifically, BCLL outperformed U-Net by 7.17% on the ACDC dataset with only 5% labeled data, 7.74% on the Synapse multi-organ segmentation with only 5% labeled data, and 0.44% on the Covid19 dataset with only 10% labeled data.

Our primary contributions are as follows:

1. We tackle the issue of poor segmentation accuracy for small objects due to class imbalance, BCLL creates a feature head for each class and performs equitable sampling from features specialized to each class.

2. We propose a new contrastive learning method, BCL, which calculates diverse combinations of cosine similarity losses within a single image. BCL aims to bring features from boundary regions closer to those from interior regions. This approach significantly enhances the segmentation performance of boundary areas by using only a limited number of labeled images.

3. The new similarity function, GGD-vMF, improves learning effectiveness relative to conventional cosine similarity and t-vMF metrics.

4. We demonstrate that BCLL achieves the highest accuracy compared to baseline U-Net and other label-efficient methods on the ACDC, SMO, and Covid19 datasets.

This paper is organized as follows. Section 2 discusses related works on label-efficient methods and cosine similarity. Section 3.1 details our BCLL method. Section 4 describes experimental setup and results on three datasets. Finally, Section concludes our paper.

2 Related Works

2.1 Label-efficient learning

In recent years, the field of label-efficient learning, which focuses on training models with limited labeled data, has gained prominence. This area is closely related to few-shot learning, but differs primarily in the classification methods used. While few-shot learning compares features of template and target data to classify similar items into the same group, label-efficient learning employs standard classification techniques and typically utilizes more data than few-shot learning. Label-efficient methods encompass various strategies, including techniques for only minimal labeled data training [11], specialized methods for medical imaging [9, 12], and adaptations for point clouds [2, 8]. For instance, CLLE, foundational to this study, uses a small number of labeled images for contrastive learning. Specifically, it employs contrastive loss to draw features within the same class closer and separate features from different classes. Class feature discrimination is enabled through the use of information from corresponding annotations. Similarly with CLLE, many label-efficient approaches leverage contrastive learning frameworks, and this paper follows precedent in proposing an advanced label-efficient method using a contrastive learning framework.

2.2 Cosine similarity

In machine learning, cosine similarity is commonly used to measure the similarity between vectors, but it often suffers from slow learning rates. Kobayashi [4] proposed an enhanced similarity measure based on the von Mises-Fisher distribution, termed t-vMF similarity. The vMF similarity function ϕ is defined as follows.

$$\phi(x;\kappa) = 2\frac{\psi(x;\kappa) - min\psi(x;\kappa)}{max\psi(x;\kappa) - min\psi(x;\kappa)} - 1 \tag{1}$$

where x represents the cosine similarity, κ is a hyperparameter, and $\psi(x;\kappa)$ is a profile function based on the probability density function. As shown in Figure 3, by narrowing the 'peak' and widening the 'base' of the similarity function, t-vMF compute higher similarity for closer features and lower similarity for more distant features. In this paper, we define a new profile function based on the generalized Gaussian distribution instead of the student's t-function, proposing a similarity function with an even narrower peak and a wider base compared to t-vMF.

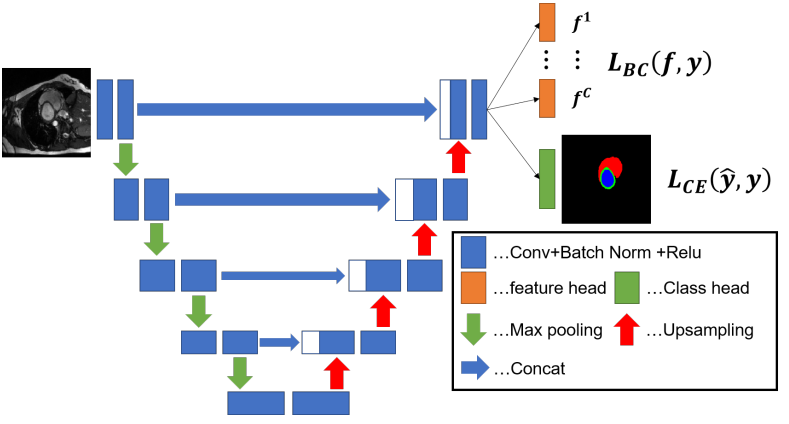

Figure 1: Overview of BCL. For segmentation, the U-Net model is employed. The term L_{CE} represents the cross-entropy loss, while L_{BCL} denotes the proposed BCL's loss. f is the output from the feature head, y is the label, and $\hat{y}$ indicates the output from the class head.

3 Proposed Method

We propose a new learning method, BCLL, which effectively learns from a small number of labeled images. BCLL addresses the first issue mentioned in Section 1, namely the poor segmentation accuracy for small objects due to class imbalance, by creating a feature extractor for each class and equally sampling class-specific features for subsequent BCL. For the second issue highlighted in Section 1—the low segmentation accuracy around object boundaries—BCL is introduced. This new contrastive learning approach extends the contrastive loss concept used in conventional label-efficient methods like CLLE, which brings together features within the same class and separates those between different classes. BCL utilizes average pooling filters on annotated images to gather positional data regarding class boundaries and internal areas. By leveraging features extracted from these regions, BCL calculates diverse contrastive losses within the same image. This technique not only pulls features of identical classes nearer while distancing those from distinct classes but also configures the contrastive losses to align features of boundary regions more closely with those of internal areas. Such an approach greatly improves the segmentation accuracy of challenging boundary parts with merely a limited collection of labeled data. Moreover, by incorporating a new profile function based on the generalized Gaussian distribution (GGD), called GGD-vMF, BCL advances beyond conventional contrastive learning metrics like cosine similarity and t-vMF, facilitating more effective learning.

3.1 Feature Head

Figure 1 illustrates the architecture of our proposed BCLL. The segmentation model utilizes U-Net [6], with class heads performing segmentation and feature heads outputting features for contrastive learning. The number of feature heads corresponds to the number of classes C. By equally sampling the class-specific features $f = \{f^c\}_{c=1}^{C}$ from each feature head and applying BCL, we address the issue of poor segmentation accuracy for small objects due to

class imbalance. The loss function for training the final model is defined as follows.

$$L(\hat{y},y,f) = L_{CE}(\hat{y},y) + L_{BCL}(f,y) \tag{2}$$

where L_{CE} represents cross-entropy loss, and L_{BCL} is the proposed contrastive loss. $y \in 0,1^{C\times H\times W}$ denotes labels, and $\hat{y} \in [0,1]^{C\times H\times W}$ represents the outputs from the class heads.

3.2 Boundary Contrastive Learning

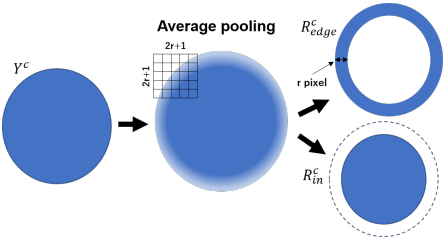

Figure 2: Method for obtaining the boundary of object.

Contrastive Loss. The process for obtaining boundary areas is shown in Figure 2. Average pooling with filter size $2r+1$ is applied to the annotation Y^c for class c. Areas where values change are defined as the boundary R^c_{edge}, and those without changes are defined as R^c_{in}. The r is a hyperparameter. By using the positional information from these labels, BCL identifies whether the features are associated with object boundaries or internal regions.

$$S^c_{edge}(h,w) = \begin{cases} \frac{1}{C|R^c_{edge}|} & \{h,w\} \in R^c_{edge}, \\ 0 & \{h,w\} \notin R^c_{edge}. \end{cases} \tag{3}$$

$$S^c_{in}(h,w) = \begin{cases} \frac{1}{C|R^c_{in}|} & \{h,w\} \in R^c_{in}, \\ 0 & \{h,w\} \notin R^c_{in}. \end{cases} \tag{4}$$

where h,w represents pixel positions, and S^c_{edge} and S^c_{in} are the probability densities for the boundary and interior regions respectively.

In our study, we define a novel contrastive loss function, $L_{BCL}(f,y)$, for improving segmentation accuracy, particularly around object boundaries, by leveraging a limited amount of labeled data. The function is structured as follows.

$$L_{BCL}(f,y) = \frac{1}{C}\sum_{c=1}^{C} L_{edge\text{-}in}(f^c,y^c) + L_{in\text{-}in}(f^c,y^c) \tag{5}$$

$$L_{edge\text{-}in}(f^c,y^c) = \frac{1}{N^2}\sum_{\substack{\{h,w\} \\ \sim S^c_{edge}}}^{N}\sum_{\substack{\{i,j\} \\ \sim S^c_{in}}}^{N} (\phi_{GGD}(f^c_{h,w}\cdot f^c_{i,j};\kappa) - \mathbb{1}[y^c_{h,w},y^c_{i,j}])^2 \tag{6}$$

$$L_{in\text{-}in}(f^c,y^c) = \frac{1}{N^2}\sum_{\substack{\{h,w\} \\ \sim S^c_{in}}}^{N}\sum_{\substack{\{i,j\} \\ \sim S^c_{in}}}^{N} (\phi_{GGD}(f^c_{h,w}\cdot f^c_{i,j};\kappa) - \mathbb{1}[y^c_{h,w},y^c_{i,j}])^2 \tag{7}$$

where N is the number of samples, y represents labels, and ϕ_{GGD} is the similarity function based on the generalized Gaussian distribution. The feature vectors f^c_{hw} and $f^c_{i,j}$ correspond to the features at pixel positions $\{h,w\}$ and $\{i,j\}$ within class c, respectively. The indicator function $\mathbb{1}$ is defined as

$$\mathbb{1}[y^c_{h,w}, y^c_{i,j}] = \begin{cases} 1 & \text{if } y^c_{h,w} = 1 \text{ and } y^c_{i,j} = 1, \\ -1 & \text{otherwise.} \end{cases} \tag{8}$$

Function $L_{edge\text{-}in}$ aims to bring the features from boundary regions closer to those from internal region in same class, enhancing feature consistency across these zones. Conversely, $L_{in\text{-}in}$ acts to reduce the variance among features within the internal regions, preventing the pull of accurately classified interior features towards the boundary features. This contrastive loss mechanism calculates multiple similarity losses from an image, efficiently improving segmentation accuracy by aligning boundary features more closely with internal features.

GGD-vMF. In BCL, we propose GGD-vMF to enhance the effectiveness of similarity calculations, which are commonly used in conventional contrastive learning. Unlike conventional methods such as t-vMF, our GGD-vMF significantly improves learning efficiency by refining the similarity function's profile. As illustrated in Figure 3, the GGD-vMF similarity function achieves a narrower 'peak' and a broader 'base,' resulting in higher similarity values for proximal features and lower similarity for those that are more distant. Furthermore, by narrowing the 'peak' of the similarity function, the 'slope' of the mountain becomes steeper, which facilitates rapid learning progression due to higher gradients achieved when data points have high similarity. The wider 'base' helps prevent gradient vanishing as gradients approach zero. While t-vMF exhibits similar effects, as demonstrated in the experiments presented in Section 4.2, our proposed GGD-vMF achieves superior performance.

The similarity function ϕ_{GGD} is defined as

$$\psi_{GGD}(x;\kappa) = \exp(-(2-2x)^{\frac{1}{2\kappa}}) \tag{9}$$

$$\phi_{GGD}(x;\kappa) = 2\frac{\psi_{GGD}(x;\kappa) - \exp(-2^{\frac{1}{\kappa}})}{1-\exp(-2^{\frac{1}{\kappa}})} - 1 \tag{10}$$

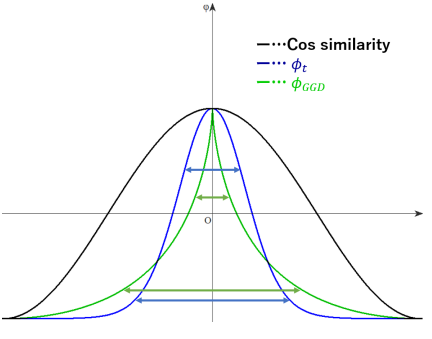

Figure 3: Overview of GGD-vMF. The vertical axis represents the output value of similarity function, while the horizontal axis represents the angle between vectors. The black, blue, and green lines indicate cos similarity, t-vMF, and GGD-vMF, respectively.

4 Experiments

4.1 Setting

Datasets. In our experiments, we utilized three datasets: Automatic Cardiac Diagnosis Challenge (ACDC) [1], Synapse multi-organ segmentation (SMO) [5], and Covid19 [10]. **ACDC** dataset comprises cardiac MRI images categorized into four classes: background, RV cavity, myocardium, and LV cavity. For training, validation, and testing, we use 690, 95, and 190 images respectively, treating these 3D MRI images as 2D data. This dataset is prone to segmentation errors at object boundaries due to the presence of thin objects. **SMO** dataset contains CT images of the lower abdomen, comprising 13 classes. It is a large dataset with 2474 training, 613 validation, and 692 testing images, notable for its significant class imbalance. **Covid19** dataset includes CT images of the lungs, divided into four classes: background, lungs other, ground glass, and consolidations. We use 70, 10, and 20 images for training, validation, and testing, respectively. This dataset includes small objects, such as the consolidations class.

Training and evaluation details. We employed U-Net [6] as the baseline. Adam and Cosine Annealing are used for training 3000 epochs. The hyperparameters were set as follows: pooling size $r = 3$, sampling number $N = 512$, and $\kappa = 0.5$ in equation 10. We measured the validation accuracy for κ at intervals of 0.5 within the range of 0.5 to 5.0, and selected the value that yielded the best validation accuracy. To assess the efficiency of learning with limited labeled data, we trained using only 5%, 10%, 30%, and 50% of the training datasets, prioritizing images containing a wide range of classes. The evaluation metric was mean Intersection over Union (mIoU), with results averaged over five trials using five-hold cross validation.

4.2 Results

Table 1: Experimental results on three datasets. We show mIoU results from five-hold cross validation. The baseline is U-Net with cross-entropy loss. Due to extremely limited training data available for Covid19, evaluations with only 5% of the data are not conducted.

dataset	method	5%	10%	30%	50%	100%
ACDC	baseline	64.97%	77.08%	81.17%	84.00%	86.96%
	CPC	66.45%	75.04%	80.97%	84.11%	
	CLLE	68.64%	77.22%	81.49%	84.59%	
	ours	**72.14%**	**77.30%**	**82.10%**	**85.89%**	
SMO	baseline	39.33%	41.43%	53.74%	61.82%	66.19%
	CPC	40.14%	40.97%	56.36%	63.06%	
	CLLE	39.59%	40.93%	56.63%	62.25%	
	ours	**47.07%**	**52.28%**	**60.87%**	**68.09%**	
Covid19	baseline	-	39.39%	43.89%	46.16%	45.65%
	CPC	-	38.52%	43.14%	46.27%	
	CLLE	-	34.65%	41.15%	45.27%	
	ours	-	**39.83%**	**47.07%**	**47.12%**	

Table 1 shows mIoU on three datasets, while Figure 4 illustrates their segmentation results. In ACDC which contains many thin objects and is prone to boundary identification

errors, our method outperformed the baseline, and existing label-efficient methods CPC [3] and CLLE by 7.17%, 5.69%, and 3.50% respectively with 5% of the dataset. This demonstrates the effectiveness of our approach, particularly in accurately segmenting near object boundaries, as clearly visible in Figure 4 for ACDC. Our method also shows superior performance with 50% of the ACDC dataset.

For SMO dataset, which exhibits significant class imbalance, our method surpasses the baseline, CPC, and CLLE by 7.74%, 6.93%, and 7.48% with 5% of the dataset, effectively addressing class imbalance. With 50% of SMO, our method continues to perform exceptionally well, outperforming others by 6.27%, 5.03%, and 5.84% respectively. The segmentation results for SMO in Figure 4 also indicate the superior segmentation capability of our method.

Next, we discuss the results of Covid19 dataset. Due to the extremely limited number of images, models were trained using either 10% or 50% of the training data. For training dataset in Covid19 comprising 70 images, only 10% training means just 7 images. In this subset containing small objects, when our method is compared to the baseline and existing label-efficient methods such as CPC and CLLE, our proposed method showed improvements of 0.44%, 1.31%, and 5.18%, respectively. The existing label-efficient methods were outperformed by U-Net, indicating their ineffectiveness with severely limited data as in this trial. For such a small dataset, few-shot learning methods would be deemed optimal. However, our proposed method utilizing BCL to compute multiple contrastive loss from a single image, surpasses the performance of U-Net even with minimal training data. In the 50% Covid19 dataset, while CPC exceeds U-Net's accuracy, CLLE remains below it. Our method consistently outperforms the baseline and the methods CPC and CLLE, proving to be effective even in datasets containing small objects like the Covid19 dataset. The superior segmentation capabilities of our method are clearly evident in Figure 4 for Covid19. In summary, across all trials on three datasets, our method consistently achieved the highest accuracy in comparison with the baseline U-Net, CPC, and CLLE, effectively addressing boundary regions, class imbalances, and small objects.

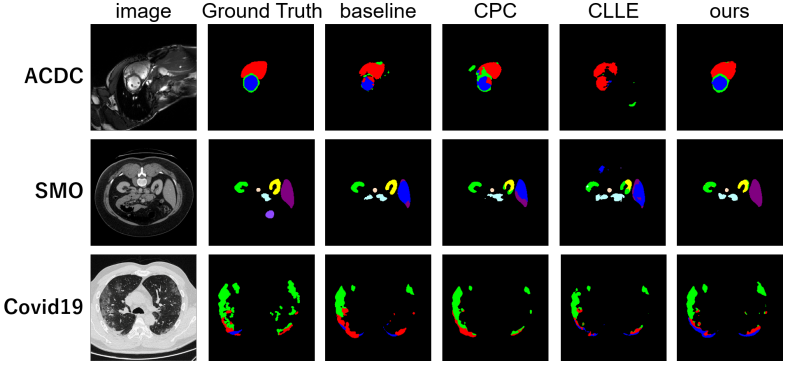

Figure 4: Segmentation results. The results of ACDC and SMO are shown for models trained with 5% of data, while results for Covid19 are shown for models trained with 10% of data.

Table 2 presents the results of an ablation study on the features and similarity used in contrastive learning, showing the average mIoU across five cross-validation trials utilizing 5% of the ACDC dataset. The "features" used in contrastive learning include "class imbalance": features in a state of class imbalance, "class balance": features achieved by equal sampling across classes to eliminate class imbalance, and "+ edge": features that, in addition

to resolving class imbalance, also differentiate between the boundary and internal regions of objects (our proposed method). "Similarity" refers to the similarity function used for computing contrastive learning. The hyperparameter κ was calibrated in 0.5 increments ranging from 0.5 to 5.0 and the model with the highest validation accuracy was selected for testing. The top row represents the baseline, U-Net, which uses only L_{CE}.

When 'feature' was changed to our proposed method (4th row from the top), there was a 5.81% improvement in accuracy over the baseline. Furthermore, changing 'similarity' from the conventional cosine similarity to our proposed GGD-vMF resulted in an additional 1.36% increase in accuracy. These results confirm the effectiveness of our proposal to resolve class imbalance and bring the features of boundary regions closer to those of internal regions. Additionally, our new similarity function, GGD-vMF, characterized by a narrow peak and a broad base, proves to be highly effective.

Table 2: Results of ablation study on features and similarity used in contrastive learning.

w/L_{BCL}	feature	similarity	mIoU
-	-	-	64.97%
✓	class imbalance	Cos	67.33%
✓	class balance	Cos	70.06%
✓	+ edge	Cos	70.78%
✓	+ edge	t-$vMF(\kappa = 0.5)$	71.17%
✓	+ edge	GGD-$vMF(\kappa = 0.5)$	**72.14%**

5 CONCLUSIONS

In this paper, we introduced a new label-efficient learning method called BCLL. To addresses the issue of poor segmentation accuracy in small objects due to class imbalance, BCLL creates a feature head for each class and performs equitable sampling from features specialized to each class. Furthermore, new contrastive learning approach in BCLL, named BCL, resolve the issue of poor segmentation accuracy in object boundary areas using only a limited number of labeled data. By applying average pooling filters to annotated images, BCL obtains positional information about the boundary and internal regions of classes. By using features extracted from these regions, BCL computes various combinations of contrastive losses in a single image. BCL not only brings features of the same class closer and distances features of different classes but also designs contrastive losses to align features in boundary regions closer to features in internal regions, thereby enhancing the segmentation accuracy of challenging boundary parts with few labeled data. Additionally, we introduced a new similarity function, GGD-vMF, which facilitates learning with limited supervised data. Our experiments on three datasets demonstrated that BCLL surpasses the baseline U-Net and other label-efficient medical image segmentation methods, achieving the highest accuracy. Specifically, BCLL improved accuracy by 7.17% on ACDC with 5% labels, 7.74% on SMO with 5% labels, and 0.44% on Covid19 with 10% labels compared to U-Net.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number 24K15020.

References

[1] Olivier Bernard, Alain Lalande, Clement Zotti, Frederick Cervenansky, Xin Yang, Pheng-Ann Heng, Irem Cetin, Karim Lekadir, Oscar Camara, Miguel Angel Gonzalez Ballester, et al. Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE transactions on medical imaging*, 37(11):2514–2525, 2018.

[2] Matheus Gadelha, Aruni RoyChowdhury, Gopal Sharma, Evangelos Kalogerakis, Liangliang Cao, Erik Learned-Miller, Rui Wang, and Subhransu Maji. Label-efficient learning on point clouds using approximate convex decompositions. In *European Conference on Computer Vision*, pages 473–491. Springer, 2020.

[3] Junichiro Iwasawa, Yuichiro Hirano, and Yohei Sugawara. Label-efficient multi-task segmentation using contrastive learning. In *International MICCAI Brainlesion Workshop*, pages 101–110. Springer, 2020.

[4] Takumi Kobayashi. T-vmf similarity for regularizing intra-class feature distribution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6616–6625, 2021.

[5] Bennett Landman, Zhoubing Xu, J Igelsias, Martin Styner, T Langerak, and Arno Klein. Miccai multi-atlas labeling beyond the cranial vault–workshop and challenge. In *Proc. MICCAI Multi-Atlas Labeling Beyond Cranial Vault—Workshop Challenge*, volume 5, page 12, 2015.

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[7] Wei Shen, Zelin Peng, Xuehui Wang, Huayu Wang, Jiazhong Cen, Dongsheng Jiang, Lingxi Xie, Xiaokang Yang, and Qi Tian. A survey on label-efficient deep image segmentation: Bridging the gap between weak supervision and dense prediction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

[8] Xian Shi, Xun Xu, Ke Chen, Lile Cai, Chuan Sheng Foo, and Kui Jia. Label-efficient point cloud semantic segmentation: An active learning approach. *arXiv preprint arXiv:2101.06931*, 2021.

[9] Xueying Shi, Yueming Jin, Qi Dou, and Pheng-Ann Heng. Semi-supervised learning with progressive unlabeled data excavation for label-efficient surgical workflow recognition. *Medical Image Analysis*, 73:102158, 2021.

[10] Xingyi Yang, Xuehai He, Jinyu Zhao, Yichen Zhang, Shanghang Zhang, and Pengtao Xie. Covid-ct-dataset: a ct scan dataset about covid-19. *arXiv preprint arXiv:2003.13865*, 2020.

[11] Xiangyun Zhao, Raviteja Vemulapalli, Philip Andrew Mansfield, Boqing Gong, Bradley Green, Lior Shapira, and Ying Wu. Contrastive learning for label efficient semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10623–10633, 2021.

[12] Lei Zhu, Kaiyuan Yang, Meihui Zhang, Ling Ling Chan, Teck Khim Ng, and Beng Chin Ooi. Semi-supervised unpaired multi-modal learning for label-efficient medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 394–404. Springer, 2021.