SBC-AL: Structure and Boundary Consistency-based Active Learning for Medical Image Segmentation

Taimin Zhou[1], Jin Yang[2], Lingguo Cui[1], Nan Zhang[3], Senchun Chai[1]

[1] The School of Automation, Beijing Institute of Technology, Beijing, China
zhoutm3797@gmail.com
[2] Department of Radiology, Washington University in St. Louis, St. Louis, MO, USA
[3] The Department of Radiology, Beijing An Zhen Hospital: Capital Medical University Affiliated Anzhen Hospital, Beijing, China

Abstract. Deep learning-based (DL) models have shown superior representation capabilities in medical image segmentation tasks. However, these representation powers require DL models to be trained by extensive annotated data, but the high annotation costs hinder this, thus limiting their performance. Active learning (AL) is a feasible solution for efficiently training models to demonstrate representation powers under low annotation budgets. It is achieved by querying unlabeled data for new annotations to continuously train models. Thus, the performance of AL methods largely depends on the query strategy. However, designing an efficient query strategy remains challenging due to limited information from unlabeled data for querying. Another challenge is that few methods exploit information in segmentation results for querying. To address them, first, we propose a Structure-aware Feature Prediction (SFP) and Attentional Segmentation Refinement (ASR) module to enable models to generate segmentation results with sufficient information for querying. The incorporation of these modules enhances the models to capture information related to the anatomical structures and boundaries. Additionally, we propose an uncertainty-based querying strategy to leverage information in segmentation results. Specifically, uncertainty is evaluated by assessing the consistency of anatomical structure and boundary information within segmentation results by calculating Structure Consistency Score (SCS) and Boundary Consistency Score (BCS). Subsequently, data is queried for annotations based on uncertainty. The incorporation of SFP and ASR-enhanced segmentation models and this uncertainty-based querying strategy into a standard AL strategy leads to a novel method, termed Structure and Boundary Consistency-based Active Learning (SBC-AL). Experimental evaluations conducted on the ACDC dataset and KiTS19 dataset demonstrate the superior performance of SBC-AL on efficient model training under low annotation budgets over other AL methods. Our code is available at https://github.com/Tmin16/SBC-AL.

Keywords: Active Learning · Medical Image Segmentation · Query Metrics · Consistency Scores · Uncertainty.

1 Introduction

The superior representation capabilities of deep learning-based (DL) segmentation methods require them to be trained by extensive annotated data. However, manual annotation is labor-intensive and time-consuming, resulting in a bottleneck to limit their segmentation performance. Active learning (AL) is a feasible solution to enhance the performance of DL-based methods while minimizing the burden of extensive annotation efforts [9,13,18]. Recognizing the great potential of AL, it has been increasingly adopted for medical image segmentation tasks [3]. AL enables models to query the most informative samples from unlabeled data and subsequently annotate these data for model retraining. Thus, the effectiveness of AL significantly depends on the query strategy employed. Various strategies have been proposed to enable models to annotate the most informative samples and maximize the benefits of these annotations. For instance, a query strategy is designed to select samples with high uncertainty evaluated by cosine similarity [19]. Similarly, another uncertainty-based query strategy is employed to select samples by evaluating uncertainty in attention maps [12]. In EdgeAL, the query strategy is designed to evaluate uncertainty by measuring the divergence and entropy in subject edges [8]. In PID-AL, uncertainty is evaluated by measuring divergence during querying [17]. However, few AL-based methods exploit information in segmentation results for querying. Additionally, the effectiveness of querying is often constrained by insufficient information from unlabeled data.

To address these limitations, first, we propose a Structure-aware Feature Prediction (SFP) and Attentional Segmentation Refinement (ASR) module to enable the segmentation models to generate results with sufficient information for querying. Segmentation results from models can provide some information for querying, but the incorporation of SFP and ASR enhances the models to capture more information related to the anatomical structures and boundaries. Additionally, we propose an uncertainty-based querying strategy to efficiently leverage information in segmentation results. Specifically, due to the enhancement of SFP and ASR modules, uncertainty can be evaluated by assessing the consistency of anatomical structure and boundary information within segmentation results. This consistency is assessed by calculating Structure Consistency Score (SCS) and Boundary Consistency Score (BCS). Following this strategy, samples with low consistency in segmentation results will be queried for annotations. Models retrained by these samples can learn sufficient information about anatomical structures and organ boundaries. Finally, we propose a novel AL method for medical image segmentation by integrating the SFP and ASR-enhanced segmentation models, and this uncertainty-based querying strategy, termed Structure and Boundary Consistency-based Active Learning (SBC-AL). These designs effectively enable SBC-AL to exploit sufficient information about anatomical structures and organ boundaries for querying. Experimental results on ACDC and KiTS19 datasets demonstrate the superior performance of SBC-AL compared to other state-of-the-art AL methods.

Our contribution has threefolds: (i) We propose **Structure-aware Feature Prediction** (SFP) and **Attentional Segmentation Refinement** (ASR) modules. They enable models to generate segmentations with sufficient information about anatomical structures and boundaries, thus facilitating the evaluation of consistency and querying. (ii) We propose an uncertainty-based query strategy. It evaluates the uncertainty by measuring the consistency scores for anatomical structure and boundary information, including **Structure Consistency Score** (SCS) and **Boundary Consistency Score** (BCS). (iii) We propose a **Structure and Boundary Consistency-based Active Learning** (SBC-AL) method for efficient medical image segmentation under low annotation budgets. It exploits information on anatomical structures and organ boundaries from segmentation results for querying. Extensive experimental results on ACDC and KiTS19 datasets demonstrate that our SBC-AL method outperforms other AL methods.

2 Methodology

The SBC-AL method consists of four major stages (Figure 1). First, we train the segmentation network with a few labeled data. Second, this trained network generates segmentation results for unlabeled data. Specifically, coarse segmentation results are generated by the backbone, while Structure-aware Feature Prediction (SFP) and Attentional Segmentation Refinement (ASR) modules are employed to refine these coarse results. Subsequently, sufficient information is extracted for querying via distance transformation from segmentation results, and uncertainty is evaluated based on this information by calculating the Structural Consistency Score (SCS) and the Boundary Consistency Score (BCS). Finally, based on evaluation results, we select data and obtain annotations from an Oracle.

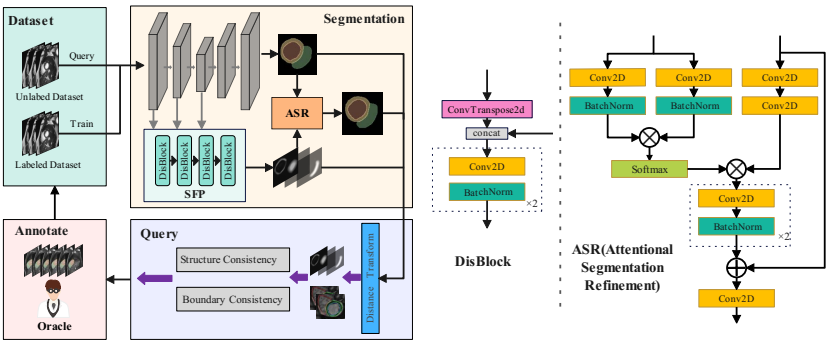

Fig. 1. The overall architecture of the SBC-AL as well as the detailed architecture of the DisBlock and the ASR module.

2.1 Segmentation Network

The segmentation network consists of a U-Net backbone [14], and two additional modules, Structure-aware Feature Prediction (SFP) and Attentional Segmentation Refinement (ASR) modules. Firstly, the segmentation network is trained on a few labeled data. Subsequently, this trained network is utilized to generate segmentation results for unlabeled data during querying.

Structure-aware Feature Prediction (SFP) To enable the segmentation network to capture multi-scale structure-related features, we utilize a Structure-aware Feature Prediction (SFP) module. The SFP module is built by cascading four Distance blocks (DisBlocks) (Figure 1). The DisBlock is utilized to capture structure-related features from semantic contextual information in the encoding path. By employing the Disblock in both high levels and low levels, the SFP module extracts multi-scale structure-aware features progressively and ultimately outputs structure-related information. The structure-related information captured by the SFP module is utilized twofold: First, they are used to refine coarse results from the backbone within the ASR module. Second, they are provided as additional information to improve uncertainty evaluation during querying.

Attentional Segmentation Refinement (ASR) To utilize structure-related features extracted from the SFP module to refine coarse results from the backbone, we propose an Attentional Segmentation Refinement (ASR) module (Figure 1). Specifically, structure-related features are converted to attention maps by two parallel convolution and batch normalization layers, and a Softmax function. Subsequently, these attention maps are utilized to adaptively refine spatial features from the backbone and thus highlight important anatomical structures and boundary regions.

2.2 Uncertainty Evaluation and Consistency Score

In the querying stage, SBC-AL aims to query unlabeled data by evaluating the uncertainty in their segmentation results predicted by the network. Uncertainty is evaluated by calculating two consistency scores, termed **Structure Consistency Score** (SCS) and **Boundary Consistency Score** (BCS). These two scores are introduced to evaluate consistency in significant anatomical structures and boundary information, respectively.

Distance Transformation First, we convert segmentation results to structure-aware distance heatmaps for consistency calculation via a distance transformation. Consider results from the segmentation network $X = \{x_1, x_2, ..., x_{H\times W}\} \in \mathbb{R}^{H\times W}$, a distance transformation is performed on their corresponding ground

truth $Y = \{y_1, y_2, ..., y_{H\times W}\} \in \mathbb{R}^{H\times W}$. For any label $i \in \{0, 1, ..., c\}$, a one-hot heatmap L_i is generated from the set of pixels p labeled as i in ground truth Y.

$$L_i = \begin{cases} 1, p = i \\ 0, p \neq i \end{cases}$$

Subsequently, the set of boundary pixels is constructed by selecting pixels from boundary regions of the one-hot heatmap L_i for any label i. Considering $m \in M$ are pixels in the boundary of L_i and $n \in N$ are arbitrary pixels in L_i, the structure-aware distance heatmap D_i is calculated.

$$D_i = \begin{cases} \min||m - n||_2, & L_i = 1 \\ 0, & L_i \neq 1 \end{cases}$$

Consistency Scores Consistency score (CS) is calculated by combining SCS and BCS, where w_s and w_b are balancing hyper-parameters to determine the relative significance of the SCS and the BCS.

$$CS = w_s SCS + w_b BCS.$$

Structure Consistency Score (SCS) The Structure Consistency Score (SCS) is calculated to evaluate the uncertainty in the anatomical structures of segmentation results. Specifically, a structure-aware heatmap is generated from the output of the SFP module via distance transformation, and subsequently, the SCS is calculated based on this map and results from the segmentation network. SCS consists of two metrics, cosine similarity-based consistency score $Q_{cos}(x)$ and dice coefficient-based consistency score $Q_{dice}(x)$. When the structure map is described as a multi-dimensional distribution of the prediction result, the cosine similarity-based consistency score $Q_{cos}(x)$ is calculated to assess the consistency of the distribution between the segmentation results $d_s(y|x)$ and output of the SFP module $d_o(x)$ for any unlabeled data as follows:

$$Q_{cos}(x) = \text{Cos}(d_s(y|x), d_o(x)) = -\frac{\sum \left(d_s(y|x) \times d_o(x)\right)}{||d_s(y|x)||_2^2 \times ||d_o(x)||_2^2}.$$

Dice coefficient-based consistency score $Q_{dice}(x)$ is derived from the dice similarity coefficient (DSC). The dice similarity coefficient is usually used to evaluate the segmentation accuracy by calculating the overlap percentage between the prediction and ground truth. Thus, dice coefficient-based consistency score $Q_{dice}(x)$ is used to assess pixel-level consistency by evaluating the spatial similarity between the segmentation results and the output of the SFP module.

$$Q_{dice}(x) = \text{Dice}(d_s(y|x), d_o(x)) = 1 - \frac{2 \times \sum \left(d_s(y|x) \times d_o(x)\right)}{\sum d_s(y|x) + \sum d_o(x)}.$$

Now, Structure Consistency Score (SCS) can be calculated where w_c and w_d are used to balance the significance of the two metrics.

$$SCS = w_c Q_{cos}(x) + w_d Q_{dice}.$$

Boundary Consistency Score (BCS) Boundary regions provide important contextual information about the geometric shapes of organs. However, the boundaries of organs are hard to segment, so boundary regions often have high uncertainty. Thus, evaluating the uncertainty in the boundary regions is an effective way to evaluate uncertainty. To achieve this, we propose a Boundary Consistency Score (BCS) based on the Hausdorff distance (HD). HD is used in evaluating the difference in boundaries between segmentation results [6]. Boundary Consistency Score is calculated from coarse segmentation results $d_c(y_c|x_c)$ and segmentation results refined by ASR $d_s(y_s|x_s)$. Subsequently, HD-based BCS is used to measure inconsistency and evaluate the uncertainty in the boundary regions between these two results.

$$\mathrm{BCS} = \max(\mathrm{h}(y_s, y_c), \mathrm{h}(y_c, y_s)).$$

where $\mathrm{h}(y_s, y_c) = \max\limits_{x_s \in y_s} \min\limits_{x_c \in y_c} ||x_s - x_c||$ and $\mathrm{h}(y_c, y_s) = \max\limits_{x_c \in y_c} \min\limits_{x_s \in y_s} ||x_c - x_s||$ are calculated by the Euclidean distance.

2.3 Selective Annotation

After evaluating uncertainty for unlabeled data, selective annotation is implemented. At each selection, the K unlabeled samples with the highest uncertainty evaluated by consistency scores are selected to be accurately annotated by Oracle. These samples with their accurate labels are combined with other labeled data to re-train the segmentation network to improve its segmentation performance. Then we will repeat the iteration (segmentation, uncertainty evaluation, and selective annotations) on other unlabeled data. This active learning iteration runs until all data are annotated or the segmentation accuracy by the segmentation network on unlabeled data reaches requirements.

2.4 Loss function

We use a multi-categorical cross-entropy loss L_{SEG} to evaluate segmentation results where y_{ic} and p_{ic} are the real label of a pixel point $i \in \{1, 2, ..., N\}$ for a category $c \in \{1, 2, ..., K\}$ and its predicted probability as follows:

$$L_{SEG} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{K} y_{ic} \log p_{ic}.$$

The output of the SFP module is a distance-transformed map of the target region. Thus, its optimization is a regression task and the mean-square error is used as the loss function L_{SFP}. y_{ij} and p_{ij} are the true distance value of a pixel point $i \in \{1, 2, ..., N\}$ on the channel $j \in \{1, 2, ..., M\}$ and its predicted value.

$$L_{SFP} = \frac{1}{MN} \sum_{i=1}^{N} \sum_{j=1}^{M} (p_{ij} - y_{ij})^2.$$

The output of the ASR module is the final refined segmentation result, so we use the loss function L_{ASR} which is a combination of multi-categorical cross-entropy loss L_{SEG} and the region-based Dice loss L_{DICE}.

$$L_{ASR} = L_{SEG} + L_{DICE}.$$

Thus, the whole model is trained end-to-end by the final joint loss function L, where α, β, and γ are hyper-parameters to balance the importance of each loss function.

$$L = \alpha L_{SEG} + \beta L_{SFP} + \gamma L_{ASR}.$$

3 Experiments

Datasets. We used two publicly available datasets to evaluate SBC-AL. The first one is the Automated Cardiac Diagnosis Challenge dataset (ACDC) [2]. It consists of 150 MR images with three labels, right ventricle (RV), myocardium (MYO), and left ventricle (LV). The second one is the 2019 Kidney Tumor Segmentation dataset (KiTS19) [5]. It consists of 210 CT images with labels of kidney and mass regions. To avoid the effects of unbalanced data, only kidney labels were used. In pre-processing, 2D patches with the dimension 256×256 were extracted from 3D volumes.

Experimental details. Our models were implemented by Pytorch[4]. We optimized the network using the SGD with a momentum of 0.9 and weight decay of 0.0001. The joint optimization of the regression and classification functions is unstable. To address this issue, we set initial learning rates to 0.01 and 0.0001 for the backbone and the SFP module, respectively. Then learning rates gradually reduced by a factor of 0.9 every 500 iterations. In ACDC, the batch size was set to 16, and the training epoch was 60, while in KiTS19 these were 8 and 100, respectively. In our best practice, hyperparameters for consistency scores were set as $w_s = 0.8$, $w_b = 0.2$, $w_c = 0.6$ and $w_d = 0.4$. We compared the segmentation performance of the SBC-AL method with six well-known AL strategies, including Random, Maximal Entropy (MaxEntropy) [16], Least Confidence (LC) [1,10], Softmax Margin (Margin) [7], Mean Standard Deviation (Mean STD) [4], and Variational Adversarial Active Learning (VA-AL) [15]. The initial labeled set was generated randomly and was the same for all AL methods in each experiment. Subsequently, we repeated experiments five times with different initial sets. The dice similarity coefficient was used to evaluate segmentation results, and average values and their standard deviation among these experiments were shown in experimental results.

Main results. Table 1 and Table 2 compared the performance of SBL-AL with other recent state-of-the-art AL methods in ACDC and KiTS19 datasets,

[4] https://pytorch.org/

Table 1. Comparison of segmentation performance between SBL-AL and other AL methods on the ACDC dataset. (**Bold** represents the best result).

Methods	6.67%	10.00%	13.33%	16.67%	20.00%	23.33%	26.67%	30.00%	33.33%	100%
Random	66.11 ± 0.04	72.70 ± 0.03	80.50 ± 0.02	84.70 ± 0.01	84.89 ± 0.03	86.74 ± 0.01	86.33 ± 0.01	87.43 ± 0.01	87.92 ± 0.02	91.37 ± 0.01
MaxEntropy	66.11 ± 0.04	75.58 ± 0.02	81.34 ± 0.02	85.57 ± 0.01	87.07 ± 0.01	87.41 ± 0.01	88.00 ± 0.01	88.65 ± 0.02	88.78 ± 0.01	91.37 ± 0.01
LC	66.11 ± 0.04	74.65 ± 0.01	83.32 ± 0.01	84.38 ± 0.01	86.16 ± 0.02	87.47 ± 0.01	88.52 ± 0.01	89.03 ± 0.01	89.08 ± 0.02	91.37 ± 0.01
Margin	66.11 ± 0.04	75.05 ± 0.01	82.71 ± 0.01	86.01 ± 0.03	86.74 ± 0.01	87.63 ± 0.01	88.10 ± 0.01	88.75 ± 0.02	89.13 ± 0.02	91.37 ± 0.01
Mean STD	66.11 ± 0.02	78.51 ± 0.01	81.92 ± 0.01	83.02 ± 0.02	85.46 ± 0.01	87.52 ± 0.01	88.22 ± 0.02	88.16 ± 0.02	88.62 ± 0.01	91.37 ± 0.01
VA-AL	**66.40** ± 0.05	78.60 ± 0.01	82.64 ± 0.01	86.10 ± 0.03	87.00 ± 0.01	87.89 ± 0.02	88.20 ± 0.01	88.65 ± 0.01	89.05 ± 0.02	91.37 ± 0.01
SBC-AL	66.20 ± 0.05	**81.43** ± 0.01	**84.55** ± 0.01	**87.55** ± 0.02	**88.16** ± 0.01	**89.14** ± 0.01	**89.27** ± 0.01	**89.99** ± 0.01	**91.05** ± 0.01	91.37 ± 0.01

Table 2. Comparison of segmentation performance between SBL-AL and other AL methods on the KiTS19 dataset. (**Bold** represents the best result).

Methods	6.67%	10.00%	13.33%	16.67%	20.00%	23.33%	26.67%	30.00%	100%
Random	54.97 ± 0.04	67.30 ± 0.01	84.89 ± 0.01	85.70 ± 0.04	88.00 ± 0.02	83.41 ± 0.02	89.67 ± 0.03	92.56 ± 0.03	96.94 ± 0.01
MaxEntropy	54.97 ± 0.04	81.07 ± 0.02	86.48 ± 0.01	86.45 ± 0.01	89.65 ± 0.03	93.67 ± 0.02	92.71 ± 0.01	94.88 ± 0.02	96.94 ± 0.01
LC	54.97 ± 0.04	79.31 ± 0.05	87.09 ± 0.03	92.57 ± 0.01	91.85 ± 0.02	93.84 ± 0.02	95.07 ± 0.01	95.72 ± 0.01	96.94 ± 0.01
Margin	54.97 ± 0.04	81.84 ± 0.04	89.40 ± 0.01	92.55 ± 0.02	93.62 ± 0.01	94.45 ± 0.03	93.69 ± 0.01	95.19 ± 0.02	96.94 ± 0.01
SBC-AL	54.97 ± 0.03	**86.51** ± 0.01	**89.61** ± 0.01	**93.59** ± 0.01	**94.52** ± 0.02	**95.19** ± 0.01	**95.81** ± 0.02	**96.73** ± 0.01	96.94 ± 0.01

respectively, and Figure 2 demonstrates the qualitative comparison results on the ACDC dataset. Figure 3 shows the qualitative results from SFP and ASR modules and samples with high CS and low CS. The results of using 100% data for training were the upper limit for the model performance. Using different ratios of annotated data for training, SBC-AL achieved the highest segmentation accuracy than other AL methods at all stages except at 6.6% data. The results of using 6.6% data showed the performance of different AL methods trained by the same initial set, and VAAL utilized a U-Net encoder for initial self-encoding training. SBC-AL achieved over 95% performance with only 16.67% annotated data in both two datasets.

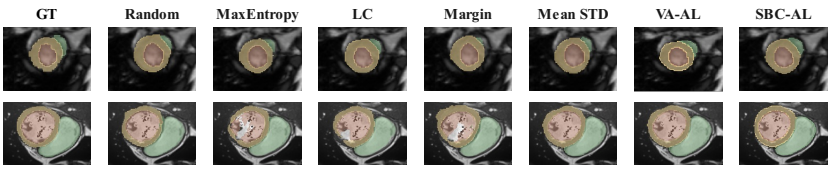

Fig. 2. The qualitative comparison of SBC-AL and other methods using 33.3% of the annotated data from the ACDC dataset.

Ablation study. To evaluate the architectural effectiveness of ASR, different from ASR where the results from the SFP were converted to an attention map, ASR (A) and ASR (C) were designed where the coarse results from the backbone were combined with the results from the SFP module via addition and concatenation, respectively. Models incorporated with ASR showed superior performance than models incorporated with ASR (A), ASR (C), or the model without ASR (Table 3). To evaluate the superior performance of SBC-AL

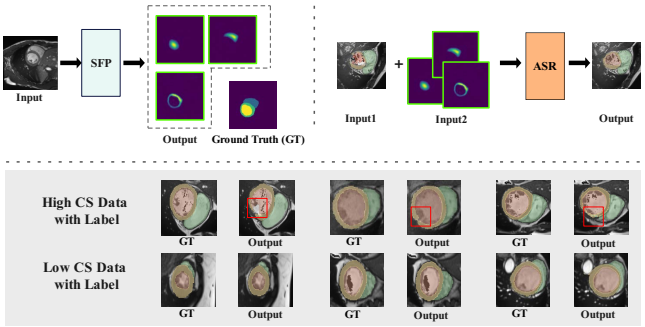

Fig. 3. The qualitative representations of output from SFP and ASR modules, and qualitative results with high CS and low CS.

Table 3. The results of the ablation study on the ACDC dataset.

Designs	RV	MYO	LV	Avg.
No ASR	88.36	85.64	94.36	89.45
ASR (A)	90.59	87.14	95.63	90.79
ASR (C)	90.90	87.21	95.10	90.09
ASR	**91.28**	**87.60**	**95.23**	**91.37**

Table 4. The results of the ablation study on the KiTS19 dataset.

Methods	UNet	UNet++	Att-UNet
Random	67.30	82.23	80.15
MaxEntropy	81.07	89.29	87.44
LC	79.31	91.86	89.36
Margin	81.84	88.68	86.84
SBC-AL	**86.51**	**92.86**	**91.85**

on different segmentation networks than other AL methods, we implemented an ablation study in three segmentation networks, including UNet [14], UNet++ [20], and Attention U-Net (Att-UNet) [11]. Table 4 demonstrates the results of using 10% labeled data from the KiTS19 dataset. When different networks were used as backbones, SBC-AL always achieved the best segmentation accuracy than other AL methods.

4 Conclusion

We propose a novel Structure and Boundary Consistency Active Learning (SBC-AL) method for efficient medical image segmentation. It enhances the segmentation performance of DL models with limited annotated data available. SBC-AL demonstrates superior performance than other popular AL methods, and we believe it can achieve promising segmentation performance on various tasks.

Acknowledgments. No acknowledgments or competing interests.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Agrawal, A., Tripathi, S., Vardhan, M.: Active learning approach using a modified least confidence sampling strategy for named entity recognition. Progress in Artificial Intelligence **10**, 113–128 (2021)
2. Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.A., Cetin, I., Lekadir, K., Camara, O., Ballester, M.A.G., 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)
3. Budd, S., Robinson, E.C., Kainz, B.: A survey on active learning and human-in-the-loop deep learning for medical image analysis. Medical Image Analysis **71**, 102062 (2021)
4. Gal, Y., Islam, R., Ghahramani, Z.: Deep bayesian active learning with image data. In: International conference on machine learning. pp. 1183–1192. PMLR (2017)
5. Heller, N., Sathianathen, N., Kalapara, A., Walczak, E., Moore, K., Kaluzniak, H., Rosenberg, J., Blake, P., Rengel, Z., Oestreich, M., et al.: The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. arXiv preprint arXiv:1904.00445 (2019)
6. Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J.: Comparing images using the hausdorff distance. IEEE Transactions on pattern analysis and machine intelligence **15**(9), 850–863 (1993)
7. Joshi, A.J., Porikli, F., Papanikolopoulos, N.: Multi-class active learning for image classification. In: 2009 ieee conference on computer vision and pattern recognition. pp. 2372–2379. IEEE (2009)
8. Kadir, M.A., Alam, H.M.T., Sonntag, D.: Edgeal: An edge estimation based active learning approach for oct segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 79–89. Springer (2023)
9. Konyushkova, K., Sznitman, R., Fua, P.: Learning active learning from data. Advances in neural information processing systems **30** (2017)
10. Lewis, D.D., Catlett, J.: Heterogeneous uncertainty sampling for supervised learning. In: Machine learning proceedings 1994, pp. 148–156. Elsevier (1994)
11. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al.: Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999 (2018)
12. Qu, C., Zhang, T., Qiao, H., Tang, Y., Yuille, A.L., Zhou, Z., et al.: Abdomenatlas-8k: Annotating 8,000 ct volumes for multi-organ segmentation in three weeks. Advances in Neural Information Processing Systems **36** (2024)
13. Ren, P., Xiao, Y., Chang, X., Huang, P.Y., Li, Z., Gupta, B.B., Chen, X., Wang, X.: A survey of deep active learning. ACM computing surveys (CSUR) **54**(9), 1–40 (2021)
14. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. pp. 234–241. Springer (2015)
15. Sinha, S., Ebrahimi, S., Darrell, T.: Variational adversarial active learning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 5972–5981 (2019)
16. Stoyanov, D., Taylor, Z., Carneiro, G., Syeda-Mahmood, T., Martel, A., Maier-Hein, L., Tavares, J.M.R., Bradley, A., Papa, J.P., Belagiannis, V., et al.: Deep learning in medical image analysis and multimodal learning for clinical decision

support: 4th international workshop, dlmia 2018, and 8th international workshop, ml-cds 2018, held in conjunction with miccai 2018, granada, spain, september 20, 2018, proceedings, vol. 11045. Springer (2018)
17. Tang, Y., Hu, Y., Li, J., Lin, H., Xu, X., Huang, K., Lin, H.: Pld-al: Pseudo-label divergence-based active learning in carotid intima-media segmentation for ultrasound images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 57–67. Springer (2023)
18. Wang, K., Zhang, D., Li, Y., Zhang, R., Lin, L.: Cost-effective active learning for deep image classification. IEEE Transactions on Circuits and Systems for Video Technology **27**(12), 2591–2600 (2016)
19. Yang, L., Zhang, Y., Chen, J., Zhang, S., Chen, D.Z.: Suggestive annotation: A deep active learning framework for biomedical image segmentation. In: Medical Image Computing and Computer Assisted Intervention- MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III 20. pp. 399–407. Springer (2017)
20. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. pp. 3–11. Springer (2018)