EdgeAL: An Edge Estimation Based Active Learning Approach for OCT Segmentation

Md Abdul Kadir[1][0000−0002−8420−2536], Hasan Md Tusfiqur Alam[1][0000−0003−1479−7690], and Daniel Sonntag[1,2][0000−0002−8857−8709]

[1] German Research Center for Artificial Intelligence (DFKI), Germany
[2] University of Oldenburg, Oldenburg, Germany
{abdul.kadir, hasan.alam, daniel.sonntag}@dfki.de

Abstract. Active learning algorithms have become increasingly popular for training models with limited data. However, selecting data for annotation remains a challenging problem due to the limited information available on unseen data. To address this issue, we propose EdgeAL, which utilizes the edge information of unseen images as *a priori* information for measuring uncertainty. The uncertainty is quantified by analyzing the divergence and entropy in model predictions across edges. This measure is then used to select superpixels for annotation. We demonstrate the effectiveness of EdgeAL on multi-class Optical Coherence Tomography (OCT) segmentation tasks, where we achieved a 99% dice score while reducing the annotation label cost to 12%, 2.3%, and 3%, respectively, on three publicly available datasets (Duke, AROI, and UMN). The source code is available at https://github.com/Mak-Ta-Reque/EdgeAL

Keywords: Active Learning · Deep Learning · Segmentation · OCT

1 Introduction

In recent years, Deep Learning (DL) based methods have achieved considerable success in the medical domain for tasks including disease diagnosis and clinical feature segmentation [20, 28]. However, their progress is often constrained as they require large labelled datasets. Labelling medical image data is a labour-intensive and time-consuming process that needs the careful attention of clinical experts. Active learning (AL) can benefit the iterative improvement of any intelligent diagnosis system by reducing the burden of extensive annotation effort [19, 25].

Ophthalmologists use the segmentation of ocular Optical Coherence Tomography (OCT) images to diagnose, and treatment of eye diseases such as Diabetic Retinopathy (DR) and Diabetic Macular Edema (DME) [6]. Here, we propose a novel Edge estimation-based Active Learning EdgeAL framework for OCT image segmentation that leverages prediction uncertainty across the boundaries of the semantic regions of input images. The Edge information is one of the image's most salient features, and it can boost segmentation accuracy when integrated into neural model training [13]. We formulate a novel acquisition function that leverages the variance of the predicted score across the gradient surface of

the input to measure uncertainty. Empirical results show that EdgeAL achieves *state-of-the-art* performance with minimal annotation samples, using a seed set as small as 2% of unlabeled data.

2   Related Work

Active learning is a cost-effective strategy that selects the most informative samples for annotation to improve model performance based on uncertainty [11], data distribution [22], expected model change [4], and other criteria [1]. A simpler way to measure uncertainty can be realized using posterior probabilities of predictions, such as selecting instances with the least confidence [9, 11], or computing class entropy [14].

Some uncertainty-based approaches have been directly used with deep neural networks [24]. Gal et al. [7] propose dropout-base Monte Carlo (MC) sampling to obtain uncertainty estimation. It uses multiple forward passes with dropout at different layers to generate uncertainty during inference. Ensemble-based methods also have been widely used where the variance between the prediction outcomes from a collection of models serve as the uncertainty [18, 23, 27].

Many AL methods have been adopted for segmentation tasks [8, 15, 18]. Gorriz et al. [8] propose an AL framework Melanoma segmentation by extending Cost-Effective Active Learning (CEAL) [26] algorithm where complimentary samples of both high and low confidence are selected for annotation. Mackowiak et al. [15] use a region-based selection approach and estimate model uncertainty using MC dropout to reduce human-annotation cost. Nath et al. [18] propose an ensemble-based method where multiple AL frameworks are jointly optimized, and a *query-by-committee* approach is adopted for sample selection. These methods do not consider any prior information to estimate uncertainty. Authors in [24] propose an AL framework for multi-view datasets [17] segmentation task where model uncertainty is estimated based on *Kullback-Leibler* (KL) divergence of posterior probability distributions for a disjoint subset of prior features such as depth, and camera position.

However, viewpoint information is not always available in medical imaging. We leverage edge information as a prior for AL sampling based on previous studies where edge information has improved the performance of segmentation tasks [13]. To our knowledge, there has yet to be any exploration of using image edges as an *a priori* in active learning.

There has not been sufficient work other than [12] related to Active Learning for OCT segmentation. Their approach requires foundation models [10] to be pre-trained on large-scale datasets in similar domains, which could be infeasible to collect due to data privacy. On the other hand, our method requires a few samples ($\sim$2%) for initial training, overcoming the limitation of the need for a large dataset.
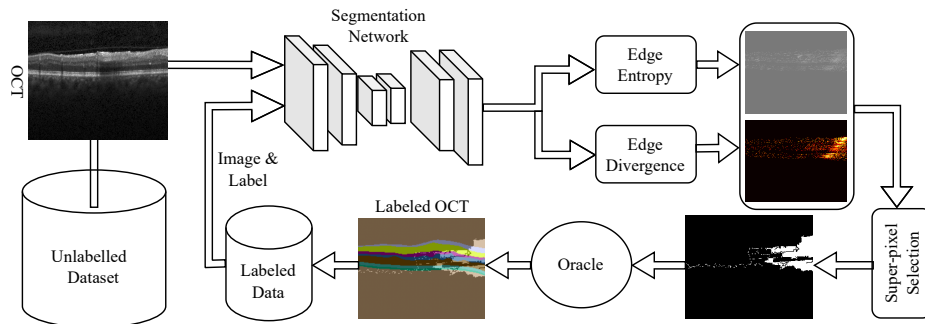
Fig. 1: The figure above illustrates the workflow of our AL framework. It first computes an OCT image's *edge entropy* and *edge divergence* maps. Later, it calculates the overlaps between superpixels based on the divergence and entropy map to recommend an annotation region.

3 Methodology

Figure 1 shows that our active learning technique consists of four major stages. First, we train the network on a subset of labeled images, usually a tiny percentage of the total collection (e.g., 2%). Following that, we compute uncertainty values for input instances and input areas. Based on this knowledge, we select superpixels to label and obtain annotations from a simulated oracle.

3.1 Segmentation network

We trained our OCT semantic segmentation model using a randomly selected small portion of the labeled data D_s, seed set, keeping the rest for oracle imitation. We choose Y-net-gen-ffc (YN^*) without pre-retrained weight initialization as our primary architecture due to its superior performance [6].

3.2 Uncertainty in prediction

EdgeAL seeks to improve the model's performance by querying uncertain areas on unlabeled data D_u after training it on a seed set D_s. To accomplish this, we have created a novel edge-based uncertainty measurement method. We compute the *edge entropy* score and *edge divergence* score - to assess the prediction ambiguity associated with the edges. Figure 2 depicts examples of input OCT, measured *edge entropy*, and *edge kl-divergence* corresponding to the input.

Entropy Score on Edges Analyzing the edges of raw OCT inputs yields critical information on features and texture in images. They may look noisy, but they summarize all the alterations in a picture. The Sobel operator can be used to identify edges in the input image [13]. Let us define the normalized absolute value of edges of an image I_i by S_i. $|\nabla I_i|$ is the absolute gradient.

$$S_i = \frac{|\nabla I_i| - min(|\nabla I_i|)}{max(|\nabla I_i|) - min(|\nabla I_i|)}$$

To determine the probability that each pixel in an image belongs to a particular class c, we use the output of our network, denoted as $P_i^{(m,n)}(c)$. We adopt Monte Carlo (MC) dropout simulation for uncertainty sampling and average predictions over $|D|$ occurrence from [7]. Consequently, an MC probability distribution depicts the chance of a pixel at location (m, n) in picture I_i belonging to a class c, and C is the set of segmentation classes. We run MC dropouts $|D|$ times during the neural network assessment mode and measure $P_i^{(m,n)}(c)$ using Equation 1.

$$P_i^{(m,n)}(c) = \frac{1}{|D|} \sum_{d=1}^{D} P_{i,d}^{(m,n)}(c) \tag{1}$$

Following Zhao et al. [30], we apply contextual calibration on $P_i^{(m,n)}(c)$ by S_i to prioritize significant input surface variations. Now, S_i is linked with a probability distribution, with $\phi_i^{(m,n)}(c)$ having information about the edges of input. This formulation makes our implementation unique from other active learning methods in image segmentation.

$$\phi_i^{m,n}(c) = \frac{e^{P_i^{(m,n)}(c) \bullet S_i(m,n)}}{\sum_{k \in C} e^{P_i^{(m,n)}(k) \bullet S_i^{(m,n)}}} \tag{2}$$

We name $\phi_i^{m,n}(c)$ as contextual probability and define our *edge entropy* by following entropy formula of [14].

$$EE_i^{m,n} = -\sum_{c \in C} \phi_i^{m,n}(c) \log(\phi_i^{m,n}(c)) \tag{3}$$

Divergence Score on Edges In areas with strong edges/gradients, *edge entropy* reflects the degree of inconsistency in the network's prediction for each input pixel. However, the degree of this uncertainty must also be measured. *KL-divergence* is used to measure the difference in inconsistency between $P_i^{(m,n)}$ and $\phi_i^{(m,n)}$ for a pixel (m, n) in an input image based on the idea of self-knowledge distillation I_i [29]. The *edge divergence* $ED_i^{m,n}$ score can be formalized using equation 1 and 2.

$$ED_i^{(m,n)} = D_{KL}\big(P_i^{(m,n)} || \phi_i^{(m,n)}\big)$$

Where $D_{KL}\big(P_i^{(m,n)} || \phi_i^{(m,n)}\big)$ measures the difference between model prediction probability and contextual probability for pixels belonging to edges of the input (Figure 2c).

3.3 Superpixel Selection

Clinical images have sparse representation, which can be beneficial for active learning annotation [15]. We use a traditional segmentation technique, SEEDS [2], to leverage the local structure from images for finding superpixels. Annotating superpixels and regions for active learning may be more beneficial to the user than annotating the entire picture [15].

We compute mean *edge entropy* EE_i^r and mean *edge divergence* ED_i^d for a particular area r within a superpixel.

$$EE_i^r = \frac{1}{|r|} \sum_{(m,n)\in r} EE_i^{(m,n)} \tag{4}$$

$$ED_i^r = \frac{1}{|r|} \sum_{(m,n)\in r} ED_i^{(m,n)} \tag{5}$$

(a) OCT   (b) Entropy map   (c) Divergence map   (d) Annotation area
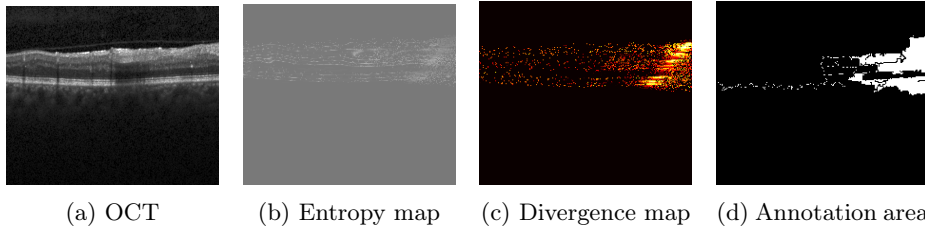
Fig. 2: The figures depict an example of (a) OCT slice with corresponding (b) *edge entropy* map, (c) *edge divergence* map, (d) query regions for annotation by our EdgeAL. The figures reveal that there is less visibility of retinal layer separation lines on the right side of the OCT slice, which could explain the model's high uncertainty in that region.

Where $|r|$ is the amount of pixels in the superpixel region. We use regional entropy to find the optimal superpixel for our selection strategy and pick the one with the most significant value based on the literature [24].

$$(i,r) = \arg\max_{(j,s)} \quad EE_j^s \tag{6}$$

Following [24], we find the subset of superpixels in the dataset with a 50% overlap (r,i). Let us call it set R. We choose the superpixels with the largest *edge divergence* to determine the ultimate query (sample) for annotation.

$$(p,q) = \arg\max_{(j,s)\in R}\big\{ED_j^s \quad | \quad (j,s)\cap(i,r); (i,r)\in D_u)\big\} \tag{7}$$

After each selection, we remove the superpixels from R. The selection process runs until we have K amount of superpixels being selected from R.

After getting the selected superpixel maps, we receive the matching ground truth information for the selected superpixel regions from the oracle. The model is then freshly trained on the updated labeled dataset for the next active learning iteration.

4   Experiments and Results

This section will provide a detailed overview of the datasets and architectures employed in our experiments. Subsequently, we will present the extensive experimental results and compare them with other state-of-the-art methods to showcase the effectiveness of our approach. We compare our AL method with nine well-known strategies: softmax margin (**MAR**) [9], softmax confidence (**CONF**) [26], softmax entropy (**ENT**) [14], MC dropout entropy (**MCDR**) [7], Core-set selection (**CORESET**) [23], (**CEAL**) [8], and regional MC dropout entropy (**RMCDR**) [15], maximum representations (**MAXRPR**) [27], and random selection (**Random**).

4.1   Datasets and Networks

To test EdgeAL, we ran experiments on Duke [3], AROI [16], and UMN [21] datasets in which experts annotated ground truth segmentations. Duke contains 100 B-scans from 10 patients, AROI contains 1136 B-scans from 24, and UMN contains 725 OCT B-scans from 29 patients. There are nine, eight, and two segmentation classes in Duke, AROI, and UMN, respectively. These classes cover fluid and retinal layers. Based on convention and dataset guidelines [6, 16], we use a 60:20:20 training: testing: validation ratio for the experiment without mixing one patient's data in any of the splits. Further, we resized all the images and ground truths to 224×224 using Bilinear approximation. Moreover, we run a 5-fold cross-validation (CV) on the Duke dataset without mixing individual patient data in each fold's training, testing, and validation set. Table 1 summarizes the 5-fold CV results.

Table 1: The Table summarizes 5-fold cross-validation results (mean dice) for active learning methods and EdgeAL on the Duke dataset. EdgeAL outperforms other methods, achieving 99% performance with just 12% annotated data.

GT(%)	RMCDR	CEAL	CORESET	**EdgeAL**	MAR	MAXRPR
2%	**0.40 ±0.05**	**0.40 ±0.05**	0.38 ±0.04	**0.40 ±0.05**	**0.40 ±0.09**	**0.41 ±0.04**
12%	0.44 ±0.04	0.54 ±0.04	0.44 ±0.05	**0.82 ±0.03**	0.44 ±0.03	0.54 ±0.09
22%	0.63 ±0.05	0.54 ±0.04	0.62 ±0.04	**0.83 ±0.03**	0.58 ±0.04	0.67 ±0.07
33%	0.58 ±0.07	0.55 ±0.06	0.57 ±0.04	**0.81 ±0.04**	0.67 ±0.03	0.61 ±0.03
43%	0.70 ±0.03	0.79 ±0.03	0.69 ±0.03	**0.83 ±0.02**	0.70 ±0.04	0.80 ±0.04
100%	0.82 ±0.03	0.82 ±0.03	0.82 ±0.03	0.82 ±0.02	0.83 ±0.02	0.83 ±0.02

Table 2: The table summarizes the test performance (mean dice) of various active learning algorithms on different deep learning architectures, including pre-trained weights, trained on only 12% actively selected data from the Duke dataset. The results (mean ±sd) are averaged after running two times in two random seeds. Superscript 'r' represents ResNet, 'm' represents MobileNet version 3 backbones, and '†' indicates that the networks are initialized with pre-trained weights from ImageNet [5].

| Arch.	p100	**EdgeAL**	CEAL	CORESET	RMCDR	MAXRPR
YN^*[6]	**0.83 ±0.02**	**0.83 ±0.01**	0.52 ±0.01	0.45 ±0.02	0.44 ±0.01	0.56 ±0.01
YN [6]	0.82 ±0.02	**0.81 ±0.02**	0.48 ±0.01	0.47 ±0.02	0.45 ±0.01	0.53 ±0.01
UN[10]	0.79 ±0.02	**0.80 ±0.01**	0.39 ±0.01	0.48 ±0.02	0.63 ±0.01	0.51 ±0.01
DP-V3^r	0.74 ±0.04	**0.74 ±0.02**	0.62 ±0.01	0.49 ±0.01	0.57 ±0.01	0.61 ±0.01
DP-V3^m	0.61 ±0.01	**0.61 ±0.01**	0.28 ±0.02	0.25 ±0.01	0.59 ±0.02	0.51 ±0.01
DP-V3^{r,†}	0.78 ±0.01	**0.79 ±0.01**	0.29 ±0.01	0.68 ±0.01	0.68 ±0.01	0.73 ±0.01
DP-V3^{m,†}	0.78 ±0.01	**0.79 ±0.01**	0.18 ±0.01	0.57 ±0.01	**0.79 ±0.02**	0.75 ±0.02
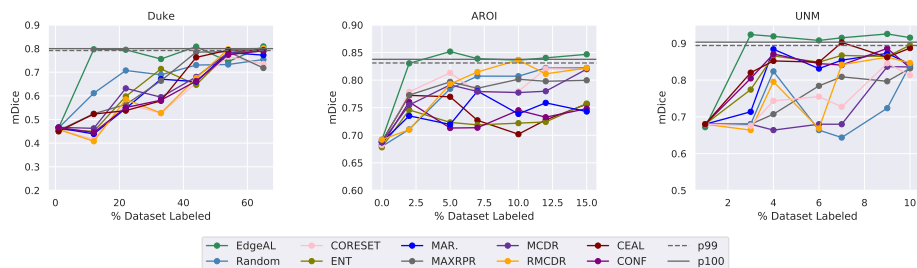
Fig. 3: EdgeAL's and other AL methods' performances (mean dice score) compared to baselines for Duke, AROI, and UNM datasets. Solid and dashed lines represent model performance and 99% of it with 100% labeled data.

We run experiments using Y-net(YN) [6], U-net (UN) [10], and DeepLab-V3 (DP-V3) [24] with ResNet and MobileNet backbones [10]. We present the results in Table 2. No pre-trained weights were employed in the execution of our studies other than the ablation study presented in Table 2. We apply mixed loss of dice and cross-entropy and Adam as an optimizer, with learning rates of 0.005 and weight decay of 0.0004, trained for 100 epochs with a maximum batch size of 10 across all AL iterations. We follow the hyperparameter settings and evaluation metric (dice score) of [6], which is the baseline of our experiment.

4.2   Comparisons

Figure 3 compares the performance of EdgeAL with other contemporary active learning algorithms across three datasets. Results show EdgeAL outperforms other methods on all 3 datasets. Our method can achieve 99% of maximum model performance consistently with about 12% (∼8 samples), 2.3% (∼16 samples), and 3% (∼14 samples) labeled data on Duke, AROI, and UNM datasets.

Other AL methods, CEAL, RMCDR, CORESET, and MAR, do not perform consistently in all three datasets. We used the same segmentation network YN* and hyperparameters (described in Section 3) for a fair comparison.

Our 5-fold CV result in Table 1 also concludes similarly. We see that after training on a 2% seed set, all methods have similar CV performance; however, after the first active selection at 12% training data, EdgeAL reaches close to the performance of full data training while outperforming all other active learning approaches.

Furthermore, to scrutinize if EdgeAL is independent of network architecture and weight initialization, we run experiments on four network architectures with default weight initialization of PyTorch (LeCun initialization)[1] and image-net weight initialization. Table 2 presents the test performance after training on 12% of actively selected data. These results also conclude that EdgeAL's performance is independent of the architecture and weight choices, while other active learning methods (RMCDR, MAXRPR) only perform well in pre-trained models.
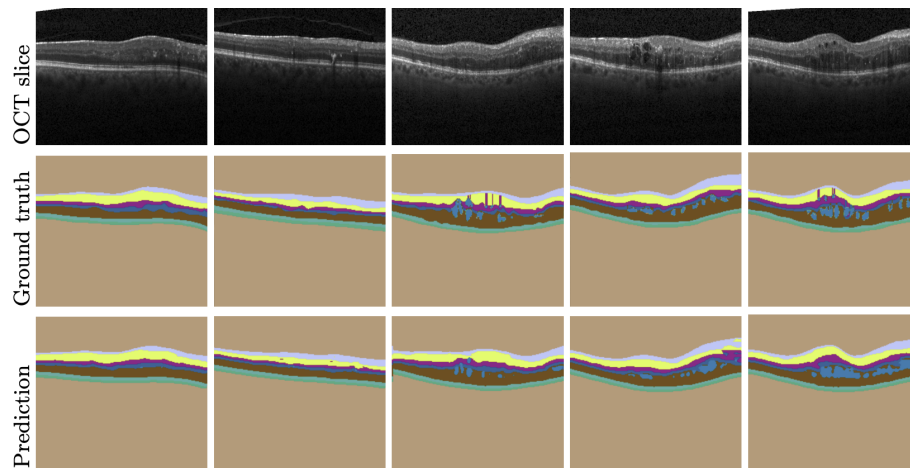

Fig. 4: Figures show sample OCT (Duke) test images with human-annotated ground truth segmentation maps and our prediction results, trained on just 12% of the samples.

5   Conclusion

EdgeAL is a novel active learning technique for OCT image segmentation, which can accomplish results similar to full training with a small amount of data by utilizing edge information to identify regions of uncertainty. Our method can

---

[1] https://pytorch.org

reduce the labeling effort by requiring only a portion of an image to annotate and is particularly advantageous in the medical field, where labeled data can be scarce. EdgeAL's success in OCT segmentation suggests that a significant amount of data is not always required to learn data distribution in medical imaging. Edges are a fundamental image characteristic, allowing EdgeAL to be adapted for other domains without significant modifications, which leads us to future works.

References

1. Bai, F., Xing, X., Shen, Y., Ma, H., et al.: Discrepancy-based active learning for weakly supervised bleeding segmentation in wireless capsule endoscopy images. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII, pp. 24–34, Springer (2022)
2. Van den Bergh, M., Boix, X., Roig, G., de Capitani, B., et al.: Seeds: Superpixels extracted via energy-driven sampling. ECCV (7) **7578**, 13–26 (2012)
3. Chiu, S.J., Allingham, M.J., Mettu, P.S., Cousins, S.W., et al.: Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema. Biomedical optics express **6**(4), 1172–1194 (2015)
4. Dai, C., Wang, S., Mo, Y., Zhou, K., et al.: Suggestive annotation of brain tumour images with gradient-guided sampling. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part IV 23, pp. 156–165, Springer (2020)
5. Deng, J., Dong, W., Socher, R., Li, L.J., et al.: Imagenet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition pp. 248–255 (2009)
6. Farshad, A., Yeganeh, Y., Gehlbach, P., Navab, N.: Y-net: A spatiospectral dual-encoder network for medical image segmentation. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part II, pp. 582–592, Springer (2022)
7. Gal, Y., Islam, R., Ghahramani, Z.: Deep bayesian active learning with image data. In: International conference on machine learning, pp. 1183–1192, PMLR (2017)
8. Gorriz, M., Carlier, A., Faure, E., Giro-i Nieto, X.: Cost-effective active learning for melanoma segmentation. arXiv preprint arXiv:1711.09168 (2017)
9. 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)
10. Khan, A., Sohail, A., Zahoora, U., Qureshi, A.S.: A survey of the recent architectures of deep convolutional neural networks. Artificial intelligence review **53**, 5455–5516 (2020)

11. Lee, B., Paeng, K.: A robust and effective approach towards accurate metastasis detection and pn-stage classification in breast cancer. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part II 11, pp. 841–850, Springer (2018)

12. Li, X., Niu, S., Gao, X., Liu, T., et al.: Unsupervised domain adaptation with self-selected active learning for cross-domain oct image segmentation. In: Neural Information Processing: 28th International Conference, ICONIP 2021, Sanur, Bali, Indonesia, December 8–12, 2021, Proceedings, Part II 28, pp. 585–596, Springer (2021)

13. Lu, F., Tang, C., Liu, T., Zhang, Z., et al.: Multi-attention segmentation networks combined with the sobel operator for medical images. Sensors **23**(5), 2546 (2023)

14. Luo, W., Schwing, A., Urtasun, R.: Latent structured active learning. Advances in Neural Information Processing Systems **26** (2013)

15. Mackowiak, R., Lenz, P., Ghori, O., Diego, F., et al.: Cereals-cost-effective region-based active learning for semantic segmentation. arXiv preprint arXiv:1810.09726 (2018)

16. Melinščak, M., Radmilovič, M., Vatavuk, Z., Lončarić, S.: Aroi: Annotated retinal oct images database. In: 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO), pp. 371–376 (2021)

17. Muslea, I., Minton, S., Knoblock, C.A.: Active learning with multiple views. Journal of Artificial Intelligence Research **27**, 203–233 (2006)

18. Nath, V., Yang, D., Landman, B.A., Xu, D., et al.: Diminishing uncertainty within the training pool: Active learning for medical image segmentation. IEEE Transactions on Medical Imaging **40**(10), 2534–2547 (2020)

19. Nath, V., Yang, D., Roth, H.R., Xu, D.: Warm start active learning with proxy labels and selection via semi-supervised fine-tuning. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VIII, pp. 297–308, Springer (2022)

20. Nguyen, D.M.H., Ezema, A., Nunnari, F., Sonntag, D.: A visually explainable learning system for skin lesion detection using multiscale input with attention u-net. In: KI 2020: Advances in Artificial Intelligence: 43rd German Conference on AI, Bamberg, Germany, September 21–25, 2020, Proceedings 43, pp. 313–319, Springer (2020)

21. Rashno, A., Nazari, B., Koozekanani, D.D., Drayna, P.M., et al.: Fully-automated segmentation of fluid regions in exudative age-related macular degeneration subjects: Kernel graph cut in neutrosophic domain. PloS one **12**(10), e0186949 (2017)

22. Samrath, S., Sayna, E., Trevor, D.: Variational adversarial active learning. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE (2019)

23. Sener, O., Savarese, S.: Active learning for convolutional neural networks: A coreset approach. arXiv preprint arXiv:1708.00489 (2017)

24. Siddiqui, Y., Valentin, J., Nießner, M.: Viewal: Active learning with viewpoint entropy for semantic segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 9433–9443 (2020)

25. Tusfiqur, H.M., Nguyen, D.M.H., Truong, M.T.N., Nguyen, T.A., et al.: Drg-net: Interactive joint learning of multi-lesion segmentation and classification for diabetic retinopathy grading (2022), https://doi.org/10.48550/ARXIV.2212.14615

26. Wang, K., Zhang, D., Li, Y., Zhang, R., et al.: Cost-effective active learning for deep image classification. IEEE Transactions on Circuits and Systems for Video Technology **27**(12), 2591–2600 (2016)

27. Yang, L., Zhang, Y., Chen, J., Zhang, S., et al.: 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)

28. Yuan, W., Lu, D., Wei, D., Ning, M., et al.: Multiscale unsupervised retinal edema area segmentation in oct images. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part II, pp. 667–676, Springer (2022)

29. Yun, S., Park, J., Lee, K., Shin, J.: Regularizing class-wise predictions via self-knowledge distillation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (June 2020)

30. Zhao, Z., Wallace, E., Feng, S., Klein, D., et al.: Calibrate before use: Improving few-shot performance of language models. In: Meila, M., Zhang, T. (eds.) Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research, vol. 139, pp. 12697–12706, PMLR (18–24 Jul 2021)