Adversarial learning with data selection for cross-domain histopathological breast Cancer segmentation

Zhenrong Lin[1] · Jidong Li[1] · Qipeng Yao[1] · Haocheng Shen[2] · Lihang Wan[1]

Abstract

Histopathology plays an important role in the clinical diagnosis of breast diseases. Early diagnosis and adjuvant therapy are of great help to patients. With the development of deep learning, fully convolutional networks (FCNs) have achieved remarkable results in the field of segmentation. However, this approach suffers from obtaining sufficient labelled data and underperforms when it comes to a new domain. Recently, adversarial learning becomes prevalent in domain adaptation, which is able to transfer learned knowledge between domains and greatly reduces the workload of labeling. In this paper, we propose a new domain adaptation method, which consists of three steps: adversarial learning, data selection and pseudo-label model refinement. Our method combines the advantages of adversarial learning and pseudo-labelling for domain adaptation. We also introduce a new data selection method to select target domain data with their pseudo-label for model refinement, considering prediction confidence and representativeness, which further strengths the model capability in target domain. We evaluate our method on private HE- and IHC-stained datasets. In order to strength the robustness, the color augmentation is utilized in this paper, the cross-domain prediction performance has been improved from 0.213 *Dice* to 0.703 *Dice*. The experimental results show that with only using unlabeled data, the proposed method can achieve 0.846 *Dice* on target domain, which outperforms the state-of-the-art method by 1.8%.

Highlights

- The dataset is standard collected and labeled by the professional cooperative doctors and hospital, which means the dataset has a high research and reference value.
- The influence of color disturbance on the training effect of segmentation network is analyzed, and gave a through study to this problem.

✉ Qipeng Yao
  qpyao@ncu.edu.cn

1 School of Information Engineering, Nanchang University, Nanchang 330031, People's Republic of China

2 AI Lab, Tencent, Shenzhen 518000, People's Republic of China

- Domain-adaptive method was utilized to relieve the scarcity of data caused cross-domain histopathological segmentation problem, and achieved remarkable performance on cross-domain breast cancer segmentation.
- In order to ensure the data has a highly reliability, an entropy sorting-based data sorting method is proposed.
- A representative selection method is proposed to select the high representativeness and reference data, which can make our proposed network achieve enough discriminative and robustness.

Keywords Breast cancer segmentation · Domain adaptation · Pseudo label

1 Introduction

With the rapid development and widespread popularity of artificial intelligence, deep learning is widely regarded as one of the most representative and promising methods. It attracts numerous attentions as impressive results have been achieved by this approach in many fields in recent years [46, 52].

Histopathology plays an important role in the clinical diagnosis of breast diseases. Early diagnosis and adjuvant therapy are of great help to patients. Hematoxylin-eosin (HE) and Immunohistochemical (IHC) staining techniques are widely utilized for breast cancer histopathological diagnosis. The former is used for observing pathological changes in tissues and the latter is used for evaluating severity as well as choosing therapy methods [28]. From the perspective of images, HE-stained images are normally in purple and red colors while IHC-stained images are in brown or blue colors. The examples of IHC- and HE-stained images are shown in Fig. 1.

Reliable and automatic segmentation of breast cancer regions on HE- and IHC-stained images would be of considerable value for histopathological analysis. However, the task is challenging due to large diversity of nuclei size, appearance and staining procedure. Although fully convolutional networks (FCNs) have achieved breakthrough performance on several biomedical image segmentation tasks [13], they are data-driven
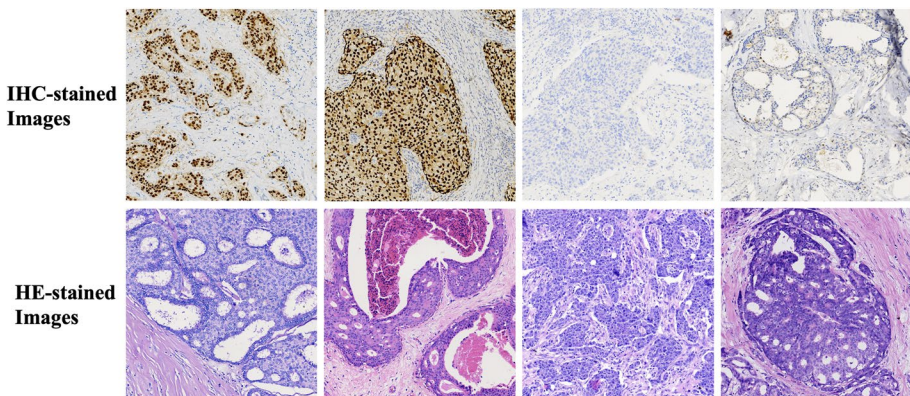

Fig. 1 The samples of IHC- and HE-stained images

methods and require a large amount of labelled images for training. Manually annotating the images is very cost in terms of time and labor, especially in medical imaging, which requires specific domain knowledge. In addition, when FCNs comes into a new domain, the segmentation performance usually drops dramatically because of the existence of the domain gaps. In recent years, domain adaptation has been widely used to alleviate this problem: 1) The strategy of adversarial learning [16] is to make two networks compete with each other. One of them is generator network, which constantly captures the probability distribution of real pictures in the training library and transforms the input random noise into new samples. The other is the discriminator network, which can observe the real and fake data at the same time and judge whether the data is real or not. Through repeated confrontation, the ability of the generator and discriminator will continue to increase until a balance is reached. Finally, the generator can generate high-quality images.2) Pseudo label technology [23] is a process that uses the model trained in labeled data to predict the unlabeled data, filters the samples according to the prediction results, and inputs them into the model again for training.

In this paper, we propose a novel domain adaptation framework for breast cancer segmentation on histopathological images. The aim is to train the model with the labelled data in source domain and the unlabelled data in target domain, and the trained model adapts and performs well in target domain. The proposed method consists of three steps: 1) apply adversarial learning for segmenting target domain data; 2) select target domain data with the highest prediction confidence and the most representativeness; 3) add the selected data with the corresponding pseudo-labels into the training set and refine the model with both source and target domain data.

In our work, we consider the HE-stained images as the source domain (with labels available) and the IHC-stained images as the target domain (no labels available). This can be attributed to two reasons: firstly, as stated before, HE-stained images are normally used for checking pathological changes in tissues, thus it is easier to recognize and annotate the cancer regions on it than on IHC-stained images; secondly, HE-stained images, as a conventical staining protocol, are easier to access in daily diagnosis routine. Thus, it has more practical value to develop a domain adaptation method transferring from HE domain to IHC domain. We build both HE- and IHC-stained datasets with the cancer regions delineated by the pathologists and evaluated our proposed method on it. The experimental results show that by only using unlabeled data, our method can achieve 0.846 *Dice* score on the target domain images, which has 1.8% improvement over the state-of-the-art method.

Entropy sorting and representativeness sorting based data selection method is proposed. Meanwhile, the domain adaptation method is utilized to segmentate cross-domain histopathology samples. The influence of color augmentation for the segmentation performance is analyzed. The main contributions of this paper are summarized as follows:

1. We propose a novel domain adaptation framework for histopathological breast cancer segmentation, which combines the advantages of adversarial learning and pseudo-labelling;
2. We introduce a new data selection method, considering both prediction confidence and representativeness, and add the selected target domain data with their corresponding pseudo-label into the training set and then refine the model, which further improves the segmentation performance in the target domain;

3. We evaluate our proposed method on private HE- and IHC-stained datasets. The experimental results demonstrate the effectiveness of each component in the framework and our method also outperforms the state-of-the-art domain adaptation method.

The rest of this paper is organized as follows: The related works are introduced in Section 2. In Section 3, we describe our proposed method in details. The experiment results are shown in Section 4. Finally, we make a conclusion and discuss the future work in Section 5.

2 Related works

In the clinical examinations, some appearance can reflect the patients' health conditions, such as face, gait and iris [1, 2].

Diagnosis of breast cancer requires strong expertise, and pathologists need to spend a lot of time in diagnosing the disease, and pathologists use the results of optical microscopic evaluation of tissue sections for manual morphological evaluation and tumor grading. A fast, accurate and robust diagnostic algorithm for breast tumor pathology slides is urgently needed. Computer technology is widely used in medical diagnostics because of its high accuracy, low cost, and robustness. Traditional algorithms have been applied to image segmentation of pathological tissues. Pathology is an image-based discipline, which is generally studied by light-field microscopy. For example, Qu et al. [30] proposed a pixel-based SVM classifier method that performs well in segmenting HE-stained histopathological images of IDC. Boykov [5] proposed that graph cuts to find globally optimal segmentation of N-dimensional images. The multiphase level set framework proposed by Vese [47] has a good effect in segmenting various structures in different types of pathological histological images. Taher et al. [40, 41] used Bayesian classification, computer-aided diagnosis (CAD) system was proposed for early detection of lung cancer. Zheng [54] proposed the method of Gabor Cancer Detection for breast cancer detection. Leng et al. [24]. proposed a lightweight network framework. Tahmoush [42] proposed an image similarity-based method for breast cancer diagnosis. The segmentation method of ROI in breast cancer pathology images has been very extensively studied. Kong et al. [20] proposed an expectation-maximizing ROI segmentation method based on color texture features. Ruiz et al. [33] proposed an efficient GPU implementation method for segmentation of neuroblastoma. Kong et al. [21] proposed an expectation-maximizing ROI segmentation method based on the color texture features. Foran [15] implemented a fast and accurate image segmentation algorithm. The delineation algorithm and a learning-based tumor region segmentation approach which utilizes multiple scale texton histograms are introduced. Lahoura proposed an extreme learning machine based breast cancer diagnostic framework based on cloud computing [22]. Nguyen et al. [29] proposed a method for automatic segmentation and classification of glandular tissue based on morphological features. However, these methods rely on the prior knowledge of engineers and have poor generalization.

The diagnosis of breast cancer requires precise localization of the pathological area, lesion area and potential lesion area. Deep learning has been widely used in recent years for various image tasks, like tracking [50], detection [19], classification [49] and diagnosis [43]. Deep learning has achieved impressive results in the segmentation field and is widely used in the localization of various diseases. Beeravolu et al. [4] preprocessed breast cancer images and create data sets for Deep CNN. Malebary proposed an automatic breast

mass classification system based on deep learning and ensemble learning [27]. Bayramo-glu et al. [3] proposed both single-task CNN to predict malignancy and multitask CNN to predict malignancy and image magnification simultaneously. Recent developments in deep learning methods have shown that in many cases, deep learning functions are better than what is learned from large datasets for specific tasks. For some specific cases, it can even reach human-level performance. However, these existing methods rely heavily on manually labeled data for learning, which is undoubtedly a tedious and time-consuming task.

However, segmentation networks have a high demand for data and the segmentation task is very difficult to label, so a self-learning segmentation approach is of great research importance. Saber applies transfer learning technology to automatic detection and classi-fication of breast cancer [34]. Shen et al. [35] combined deep active learning and self-paced learning paradigm to propose a new framework for large-scale detection learning. The approach improves the speed of model training with fewer annotation samples, active learning greatly reduces the physician's annotation work, and self-paced learning mitigates the ambiguity of the data. However, the approach still requires partial labeling and multiple rounds of training. In terms of domain adaptation, fine-tuning an existing model using a small sample of the target domain is a widely used method, and recent studies [9, 18, 26, 38, 44, 45, 48] have used labeled source domain data and unlabeled target domain data to shift the recognition region of the model from a supervised learning source domain to an unsupervised target domain. Vu et al. [48] proposed the concept of entropy minimization to reduce inter-domain differences, but this method relys on the initial entropy information selection, which would lead to some limitations. In medical imaging, the use of adversarial generative learning is increasing. Courty et al. [9] found domain adaptation of the com-putational distance between data distributions. Tzeng [45] et al. used the maximum mean difference MMD method to minimize the difference in data distribution between source and target domains, Sun et al. [38] minimized the difference in data distribution between source and target domains by correlation distance. Because of the emergence of generative adversarial networks [16], the domain adaptation task can be learned implicitly by select-ing the adversarial loss to minimize the domain offset [26], Tsai [44] et al. proposed multi-level production adversarial learning with game road scene data as the source domain and real road scene data as the target domain to achieve domain adaptation of different road scenes, and achieved good results on real road scene classification. Kamnitsas [18] used generative adversarial networks to segment MR data from patients with traumatic brain injury to achieve good effect. Dou [12] et al. optimized the domain adaptive and domain annotator modules through adversarial loss to adapt MRI image trained models to unpaired CT data for cardiac Structural Segmentation. Unfortunately, despite the importance of this direction, it has not evolved rapidly in breast tumor sections due to the lack of data sets. At the same time, the accuracy is not satisfactory. Hence, two data selection methods are pro-posed, and the two methods are combined with pseudo labeling techniques.

3 Materials and methods

3.1 Overview of the framework

Cancer region segmentation plays an important role in the diagnosis of breast cancer which assist in the assessment of the severity of the breast cancer according to the proportion of the cancer area. Although deep neural networks have achieved promising results in many

segmentation tasks, such as U-net [32], FCN [25], Deeplab_v2 [8], Pspnet [53], Linknet [7], when a model trained by the deep convolution neural network enters a new domain, the segmentation performance drops dramatically due to the domain gaps. Annotating the images in the new domain manually from scratch is very time and labor consuming, especially in medical imaging domain, which requires experts' knowledge. In this paper, we propose a novel domain adaptation framework for segmenting the cancer regions on IHC-stained images with only annotations on HE-stained images available. The framework is divided into three stages: 1. Adversarial learning for domain adaptation; 2. Target domain data selection based on entropy and representativeness ranking; 3. Model refinement with the selected data and pseudo labels.

3.2 Adversarial learning

The adversarial learning framework consists of one segmentor and two discriminators. The segmentor is a normal network for segmentation task, such as Linknet, Deeplab, etc. The purpose of the segmentor is to generate segmentation outputs for both source and target domain while the discriminators plays an adversarial role to distinguish whether these outputs are from the source domain or the target domain. During training, the segmentor is also trained to fool the discriminators thus it can gradually produce similar outputs for source and target domain, ignoring the domain-specific information. In our work, we place the two discriminators in the last two layers of the segmentation network to consider multi-scale image information, which is usually helpful for the segmentation task. The network structure is shown in Fig. 2. The last two layers of our method are used to calculate the losses.

Specifically, we use Deeplab_v2 structure as the segmentation network with ResNet101 [17] (pre-trained on ImageNet dataset [10]) as the backbone, the classification layer of ResNet101 was discard, the filter size of the last two convolution layer is 1×1 the size of the output feature maps is 1/8 of the input image size. In the conv4 and conv5 layers, the sizes of the filters are set to 2×2 and 4×4, respectively to enlarge the receptive field. After the last layer, we used Atrous Spatial Pyramid Pooling (ASPP) module to encode multi-scale information in the feature maps, followed by an upsampling layer with softmax output, which upsamples the output dimensions to the input dimensions. Cross-entropy loss was calculated based on the source ground truth to obtain segmentation losses L_{seg}.
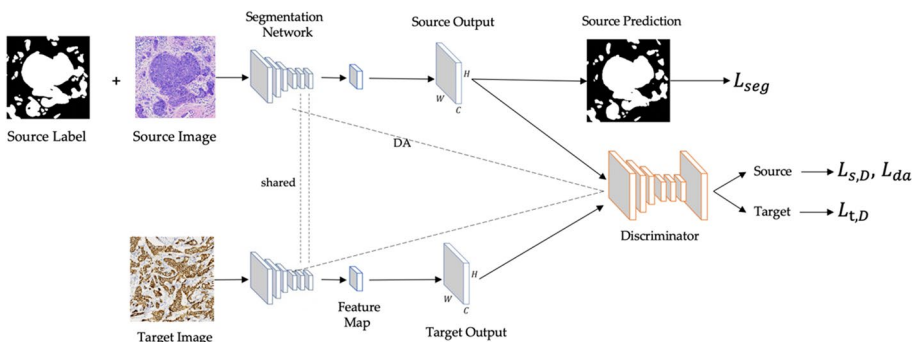

Fig. 2 Network structure of adversarial learning for domain adaptation

$$L_{seg} = -G_s \log(P_s) \tag{1}$$

where G_s is the ground truth annotations of the source domain data S. P_s is the output of the segmentation network in the source domain data S.

For discriminators, we use full convolutional layers to retain spatial information. The network is composed of 5 convolutional layers, with 4×4 kernel size and stride $= 2$. The number of channels is 64, 128, 256, 512 and 1, respectively. There is a leaky ReLU at the end of each of the first four layers, and the parameter is 0.2. In the last convolutional layer, we do not use the upsampling layer and the discriminator results are calculated directly.

As stated before, discriminators D is used to determine whether the input image is from the source domain or the target domain. The images from the source domain is marked as label 1, and the discriminator prediction result of the source domain is used to calculate the cross-entropy loss as the discriminator loss $L_{s,D}$.

$$L_{s,D} = -\log\left(D(P_s)\right)^{(h,w,1)} \tag{2}$$

Where h, w is the length and width of the input image. Target domain data T passes through the same segmentation network to predict the softmax output P_t. We feed the predicted results of the target domain P_t through discriminator D, and the cross-entropy between P_t and marked label 0 is calculated as the adaptive loss L_{da}.

$$L_{da} = -\log\left(D(P_t)^{(h,w,0)}\right) \tag{3}$$

The adversarial loss $L_{t,D}$, is obtained by calculating cross-entropy between the target discriminator result and marked label 1.

$$L_{t,D} = -\log\left(D(P_t)\right)^{(h,w,1)} \tag{4}$$

The generator loss $L_{(S,T)}$ includes two modules. We optimize the segmentation network by segmentation loss L_{seg} and adversarial loss $L_{t,D}$. λ_{da} is the loss equilibrium weight.

$$L_{(s,T)} = L_{seg} + \lambda_{da} L_{t,D} \tag{5}$$

For optimizing the discriminators, we put the segmentation outputs into full convolution discriminators to distinguish the source and target domain based on the cross-entropy loss L_D, which can be written as:

$$L_D = L_{s,D} + L_{da} \tag{6}$$

The segmentor and the discriminators are trained simultaneously. Based on these two levels of adversarial learning, we constantly optimize the segmentation network and discriminators to improve the segmentation ability of the network to the target domain images.

3.3 Data selection

After adversarial learning, the model can produce reasonable segmentation results for the target domain. In the next step, we would like to select some data in the target domain automatically, and refine the model with the selected data and its pesudo-label to further strength the model capability for the target domain. Specifically, among all the unannotated

images I_u , our task is to select a subset of M images, $I_m \subseteq I_u$. We hope that I_m has high prediction confidence and is representative over I_u as well. Firstly, we rank all the unlabelled images in the target domain by calculating the image entropy and select k images, $k = 60$ in this paper, with lowest entropy to form I_k, $I_k \subseteq I_u$, and then we refine I_k by using the representative selection to get subset I_m Finally, the model prediction results are taken as pseudo labels, and the selected data and pseudo labels are added to the training set for training.

3.3.1 Entropy rank

We feed all the target domain data through the segmentation network and calculate the entropy map E_t. We use a simple way to rank data: directly calculate the mean value S_T of its entropy map, such as:

$$S_T = 1/HW \sum_{h,w} E_t^{(h,w)} \tag{7}$$

Normally, the images with low S_T potentially have higher confidence with less noise, resulting in a better prediction effect. Thus, we select the data with 40% lowest S_T to form I_k with highest confidence. The predicted results can be added to the segmentation network as pseudo labels to obtain a more favorable learning effect. The cost of computational of the proposed entropy rank approach is only determined by h and w, so the time complexity of this method is $O(h \times w)$. Considering the sizes of the feature maps are smaller than 512, so the time complexity is always satisfactory and can be considered as $O(n)$.

3.3.2 Representative selection

The data selected by the entropy rank is believed to have a high confident level. However, these data may be similar. The main reasons can be explained as follows: 1. The selected patches could be from a single patient; 2. The appearance of patches from different patients could also be similar. On the other hand, the selected images are expected to contain different characteristics as many as possible to maximize the effectiveness of these data. Thus, we propose to a representativeness criterion to select the most representative images from I_k.

Similar as [14, 51]. In order to calculate the representativeness of I_m for I_u , we use the Max-cover [14] method. Firstly, I_m was defined as a representation of an image $O_j \in I_u$ as: $f(I_m, O_x) = \max_{O_i \in I_m}(sim(O_i, O_j))$, where $sim(O_i, O_j)$ is the similarity estimate between O_i, O_x. In our opinion, O_j is represented by the most similar image in I_m and measured by similarity $sim(O_i, O_j)$. We define the representativeness of I_m for I_u as: $F(I_m, I_u) = \sum_{O_j \in I_u} f(I_u, O_j)$, I_m is a good representation of all the images in I_u. In order to calculate the subset I_m that maximizes $F(I_m, I_u)$, I_m belongs to I_u, we should ensure that: 1. The selected images should be similar to many unannotated images in I_u; 2. Include different conditions (e.g. adding two slices of the same patient does not significantly increase $F(I_m, I_u)$). We use an approximate greedy [25] method iteratively put in $F(I_m \bigcup O_i, I_u)$ and overwrite S until the number of images in I_m reaches the value M we preset.

To calculate the similarity between O_i and O_x, we extract the last layer of the output of the block as the advanced features and calculate the channel-wise mean, represented by finally using cosine similarity as the two image similarity measure, namely $sim(O_i, O_x)$. The computational cost is a convex operation between O_i and O_x, benefit from the strong

parallel computational ability, the time cost of representative selection approach is only O(1).

$$sim(O_i, O_x) = y_i \bullet y_x^T / (\parallel y_i \parallel \bullet \parallel y_x \parallel) \tag{8}$$

3.4 Model refinement

After data selection, the selected images are supposed to have high confidence and representativeness among all the unannotated target domain images, which can make the model better for domain adaptation learning.

The selected images and its pseudo- labels were added to the original source training set together, and a separate segmentation network was used for training. The network structure is shown in Fig. 3.

4 Experiments and discussion

The proposed framework was implemented in Pytorch. The segmentation learning rate λ_{seg} used in the experiment was 0.0025, and the discriminator learning rate λ_D is set to 0.0001, the specific gravity of the last two layers is 0.1 when calculating the loss, and the batch size of the source domain and the target domain is 1, and the input size of the image is 512×512.

In order to improve the generalization ability of the model, we use albumentations library [6] for data augmentation, which includes horizontal and vertical flip, random cropping, random rotation and random color distortion. The color distortion consists of saturation, brightness and contrast adjustments. We also made an attempt to explore the effectiveness of color augmentation (e.g., adjust the hue of images) for the cross-domain segmentation, which will be discussed later.

The key of this paper is how to learn enough information from the imbalanced data and dimension reduction [31, 37, 39], the proposed method is evaluated on an in-house histopathology breast cancer dataset. The dataset contains 400 patch images extracted from hot spot areas in whole slide tissue images (WSIs) from over 100 patients, including 200 IHC and 200 HE patches. The IHC-stained protocols include estrogen receptor (ER) and
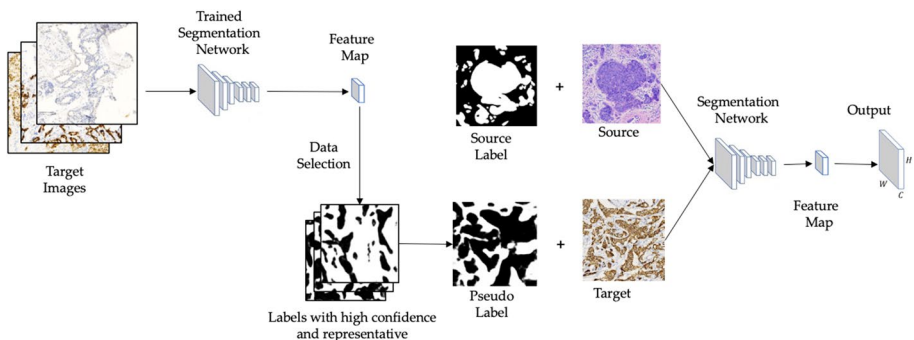

Fig. 3 The framework of adding pseudo labels in training the segmentation network

progesterone receptor (PR). For both HE- and IHC-stained images, we use 75% of the data (150 images) as the training set and the rest (50 images) as the validation set. There is no overlap of patients between training and validation set. All the images were resized to have 0.848 um/pixel, which is equivalent to 10 × objective magnification of a normal microscope. The images used in this paper are all manually annotated by junior pathologists and reviewed and modified by senior pathologists, and the all operations are following the standard clinical operations.

For quantitative evaluation, we calculate *Dice* score between predicted segmentation results and the ground truth

$$Dice = \frac{2 * |P \cap G|}{|P| + |G|} \tag{9}$$

where $\cap$ denotes the intersection operation, P and G represent the predicted region and the ground truth, respectively. *Dice* is a commonly measurement for evaluation segmentation accuracy [11, 36]. Its value is in the range of [0, 1] where high *Dice* values stand for good segmentations while low values may indicate segmentation failures.

4.1 Naive segmentation without adaptation

In order to investigate the influence of domain gap for image segmentation, we conduct two experiments: 1) training on HE images only (with label masks); 2) training on IHC images only (with label masks), and evaluate on both HE and IHC validation sets. We experimented three different networks, namely Deeplab_v2, Pspnet and Linknet. Without color augmentation, the result is shown in Fig. 4.

The model trained on HE-stained images is very poor in predicting IHC-stained images, and the model trained on IHC-stained images can predict HE-stained images better. The main reasons can be concluded as:1. The HE-stained images is entirely close to red or purple color (see Fig. 1b), making the model overfit to HE images easily, while the color range of IHC-stained images cover more than HE as it has more kinds of colors (e.g., brown and blue), which strengths the model generalization capability to other domain implicitly. 2. The spatial feature distribution of HE-stained images is simpler than IHC-stained images,
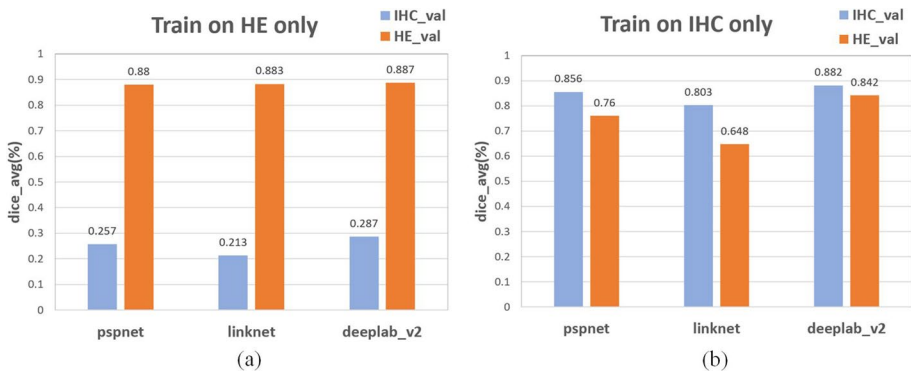

Fig. 4 The segmentation results based on different domains. **a** - **b** represent the networks are trained based on HE and IHC-stained images, respectively

and the model learns fewer features from HE-stained images than IHC-stained images. 3. HE-stained images are naturally used for observing organ tissues, thus it is easier to segment cancer regions on HE-stained images than on IHC-stained images through visual appearance.

4.2 The effectiveness of color augmentation

IHC-stained images are mainly brown and blue, and HE-stained images are mainly purple. If the random range is greater than 40%, the selection range of HE-stained images can be considered as the subset of the selection range of IHC-stained images after color augmentation. Hence, the segmentation performance is generally satisfactory. Unfortunately, the selection range of IHC-stained images cannot contain the selection range of HE-stained images after color augmentation, so the transfer segmentation performance is terrible. Hence, in order to improve the robustness of our method, we must ensure that the selection range of IHC-stained/HE-stained after color augmentation cannot contain any information of HE-stained/IHC-staifpzned images.

As explained above, the main reason for the poor segmentation performance in a different domain is the color discrepancy between HE- and IHC-stained images. Thus, we conducted another experiment to evaluate the effectiveness of color augmentation, using the same training set, but with color perturbations of different degrees, and test the model on the same validation set. Specifically, we applied color augmentation by adjusting the hue value of the image. The hue value was randomly chosen from the range of [0, H] where larger H means heavier color augmentation. Figure 5 shows the exemplars of HE- and
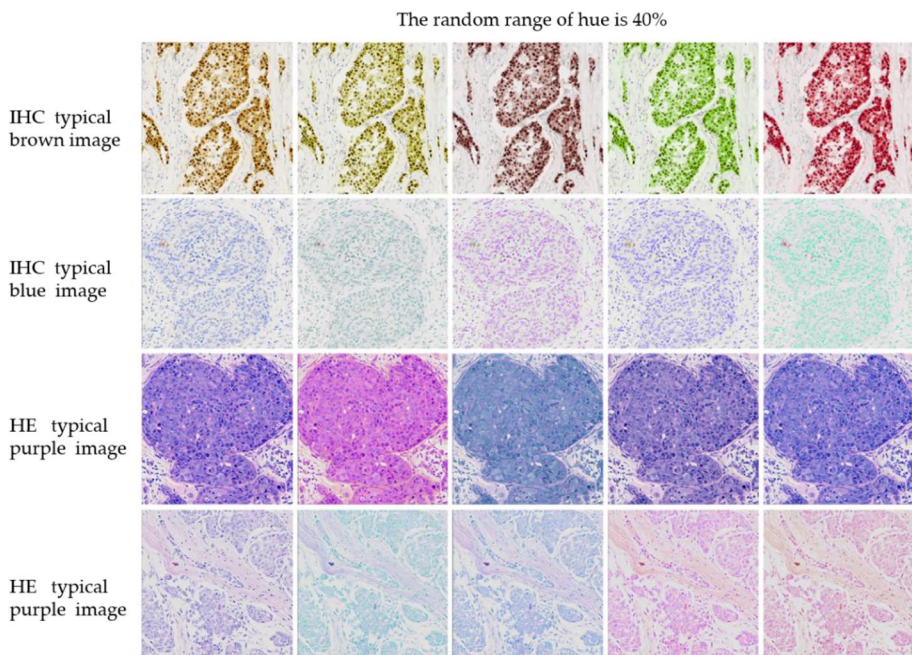

Fig. 5 The examples of changing the color of the images in data augmentation pipeline

IHC-stained images with color augmentation (hue = 0.2). Linknet was used as the segmentation network. Figure 6 shows the segmentation performance under different degrees (setting hue from 0 to 0.5) of color augmentation.

The experimental results show that color augmentation has positive effects on cross-domain segmentation. Increasing hue range improves orange bar scenario (training on HE and validate on IHC) gradually. For example, increasing hue range from 0.1 to 0.5 make the average *Dice* improve from 0.261 to 0.703, but less obvious for the blue bar scenario (training on IHC and validate on HE).

IHC-stained images are mainly brown and blue, and HE-stained images are mainly purple. When adjusting hue of blue images, they can be potentially transformed near purple color, which helps the model to predict HE-stained images. However, when adjusting hue of purple HE-stained images, they can be potentially transformed near blue color only, but not near brown color. Therefore, it is difficult for the model to predict the brown IHC-stained images (see Fig. 1), so the segmentation performance of IHC-stained images by using the model trained on HE-stained image is not ideal (e.g., only 0.703) even when the heaviest color augmentation is applied.

4.3 The effectiveness of adversarial learning

Firstly, we use the labeled IHC-stained images as the training set, and train a Deeplab_v2 network model. In this way, the *Dice* of the model trained by fully supervised learning can reach 89.2% on the IHC-stained validation set, which is the upper bound of domain adaptation from HE to IHC.

Then, we apply the adversarial learning described in Section 3.2 for the domain adaptation task from HE-stained images to IHC-stained images (see Table 1). In other words, the HE-stained image is used as the source domain with labels, and the IHC-stained image is used as the target domain without labels. The *Dice* value of our model could achieve 88.1% on the validation set of the source domain. The *Dice* value of adversarial learning framework achieve 82.8% on the validation set of the target domain. It can be considered that the improvement space of the domain adaptation task from HE-stained images to IHC-stained images is 6.4%.

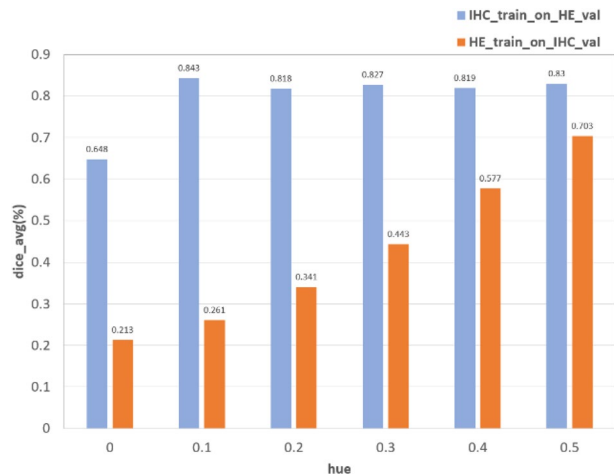**Fig. 6** Segmentation performance under different degrees of color augmentation

Table 1 The results of adversarial learning (HE as the source domain, IHC as the target domain)

| Method	Source	Target	Hue(0–0.5)	Dice on (HE)	Dice on (IHC)
SegNaiveAug (Deeplab_v2)	IHC	/	0	/	0.892
AdversarialLearning	HE	IHC	0	0.881	0.828
AdversarialLearning	HE	IHC	0.2	0.880	0.831

Similarly, we use the labeled HE-stained images as the training set, and use Deeplab_ v2 network model. In this way, the *Dice* value of the model trained by fully supervised learning can reach 88.7% on the IHC-stained validation set, which is the upper bound of domain adaptation from IHC to HE.

We use the same network structure for the domain adaptation task from IHC-stained images to HE-stained images (see Table 2). The IHC-stained image is used as the source domain, and HE-stained image is used as the target domain. The *Dice* value of this model could achieve 88.2% on the validation set of the target domain. The *Dice* value of this model achieve 86.4% on the validation set of the source domain, and it can be considered that the improvement space of the domain adaptation task from IHC-stained images to HE-stained images is 0.5%.

We also conduct an experiment to investigate the effects of color augmentation for adversarial learning. The performances are similar in the output space under training with color augmentation and without. Figure 7 shows that applying color augmentation has little effect on the segmentation performance, but could speed up the model convergence during training phase.

Table 2 The results of adversarial learning (IHC as the source domain, HE as the target domain)

Method	Source	Target	Hue(0–0.5)	Dice on (HE)	Dice on (IHC)
SegNaiveAug (Deeplab_v2)	HE	/	0	0.887	/
AdversarialLearning	IHC	HE	0	0.882	0.864
AdversarialLearning	IHC	HE	0.2	0.884	0.861
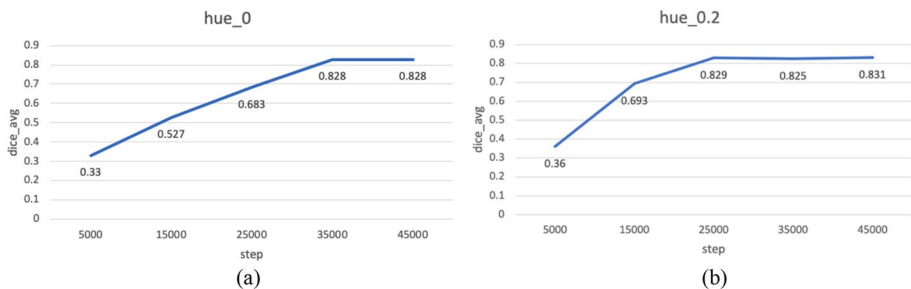
Fig. 7 The segmentation performance curves of adversarial learning **a** without color augmentation and **b** with color augmentation

4.4 The effectiveness of data selection and model refinement

After the target domain data is ranked by entropy and representativeness selection criteria, the selected images and their pseudo labels are added to the training set. In our experiment, we set the number of selected images to 30. Figure 8 shows the validation *Dice* curve along with training steps of model refinement.

From Fig. 8, it can be observed that after refining the model, the average *Dice* was increased from 82.8% to 84.6%, resulting in an 1.8% improvement. The visual segmentation results after model refinement are shown in the 4th column of Fig. 10. After adding selected target data and pseudo labels, the training set contains the target domain data for fully supervised learning. Although the pseudo labels are not the real ground truth, by our entropy selection criterion, they are among the high prediction confidence samples and supposed to be close to the real ground truth. In addition, by our representativeness criterion, the selected data are supposed to contain as many characteristics as possible, while not limited to a few of patients' specific features. Therefore, the refined model will have a better generalization capability to the target domain than only applying adversarial learning. Note that no real target domain label is used during training, our method is indeed an unsupervised domain adaption framework.

At present, in the field of annotation, it is much more difficult to label IHC-stained images than HE- stained images. Our method could alleviate this problem thus it has a practical value for histopathology research.

4.5 Comparison with the state-of-the-art methods

4.5.1 Direct minimize entropy

Vu et al. [48] propose to directly minimize entropy loss to maximize the predictive certainty in the target domain. This entropy loss is added to the segmentation network to constrain the model. Thus, the model is supposed to produce high-confidence predictions for the source domain as well as the target domain.

We re-implement this approach in our framework and from our experimental results, this method does not seem to be suitable for domain adaptation tasks from HE-stained images to IHC -stained images (see MinEnt in Table 3). We extract the feature layer of the IHC image from the model, and generate a probability map through softmax. Figure 9

Fig. 8 The segmentation performance curves of model refinement with selected data and pseudo labels
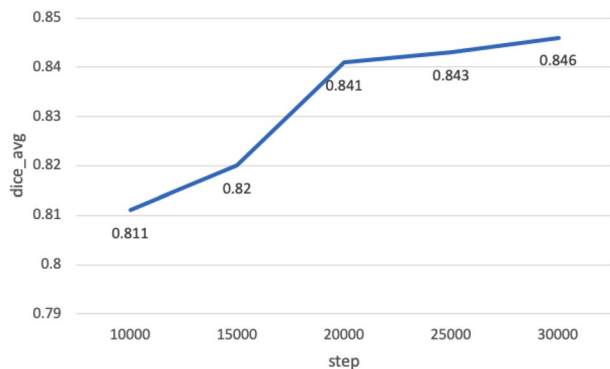
Table 3 Results of all experiments. We first use three segmentation networks without color augmentation to compare the segmentation results(SegNaiveAug)

Method	Source	Target	Hue(0–0.5)	Dice on (HE)	Dice on (IHC)
SegNaiveAug (Deeplab_v2)	HE	/	0	0.887	0.287
SegNaiveAug (Pspnet)	HE	/	0	0.88	0.257
SegNaiveAug (Linknet)	HE	/	0	0.883	0.213
SegColorAug1	HE	/	0.1	0.843	0.261
SegColorAug2	HE	/	0.2	0.818	0.341
SegColorAug3	HE	/	0.3	0.827	0.443
SegColorAug4	HE	/	0.4	0.819	0.577
SegColorAug5	HE	/	0.5	0.830	0.703
MinEnt [48]	HE	IHC	0	0.886	0.295
MinAdv [48]	HE	IHC	0	0.872	0.332
AdversarialLearning [44]	HE	IHC	0	0.881	0.828
Ours	HE	IHC	0	0.879	0.846

Then we compare the results of different degrees of color augmentation based on Linknet(SegColorAug). Finally, we show the experimental data using MinEnt, MinAdv, AdversarialLearning and our proposed method
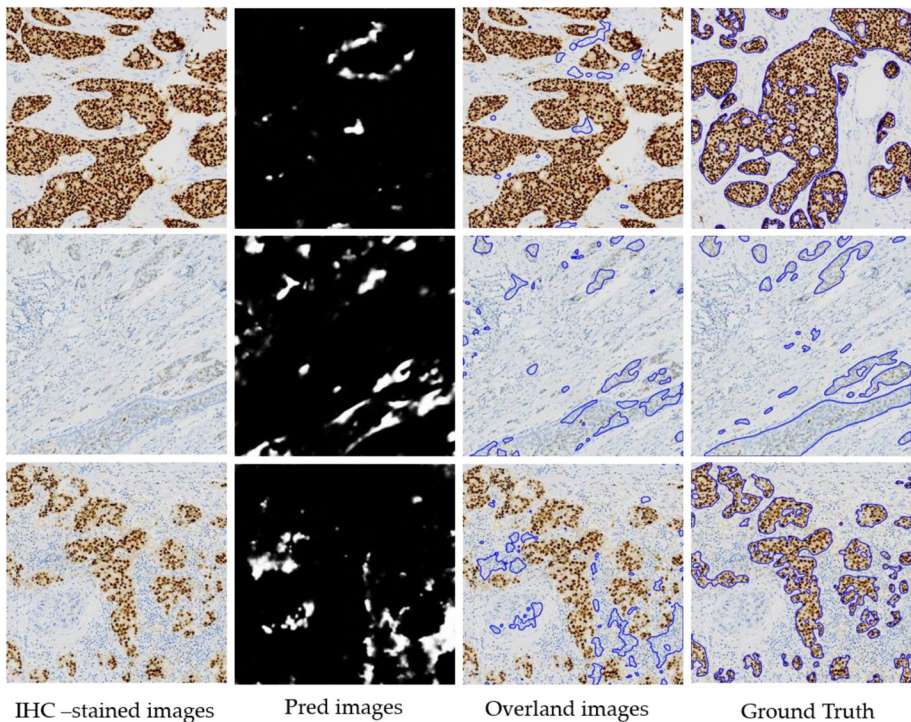

| IHC –stained images | Pred images | Overland images | Ground Truth |

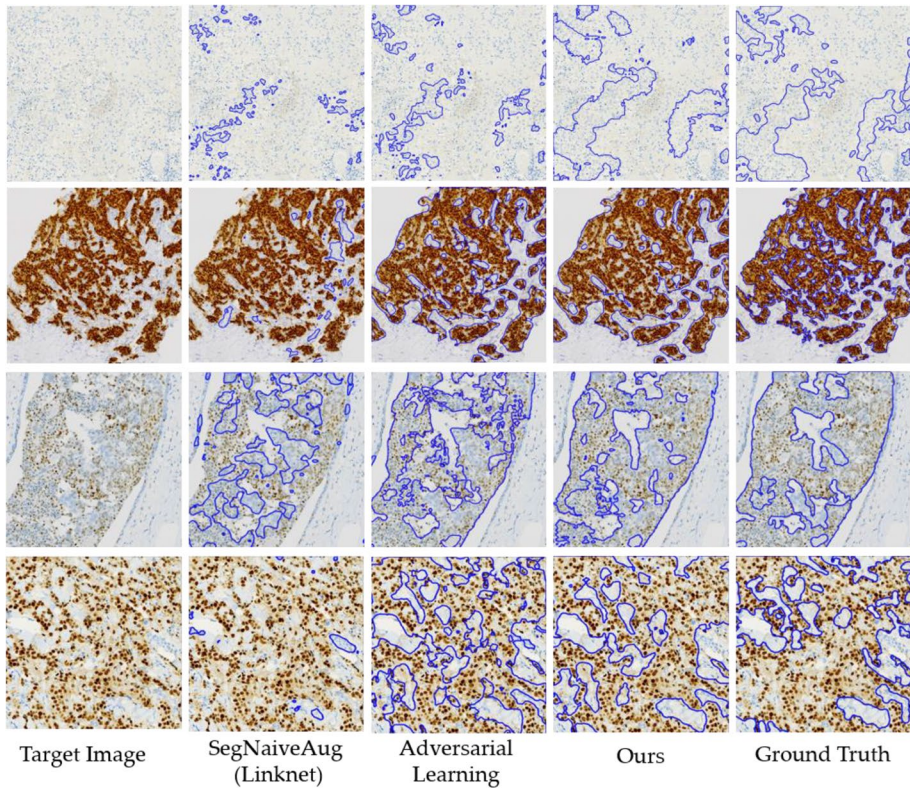**Fig. 9** Bad cases of direct minimize entropy method

Fig. 10 Example results of adapted segmentation for HE-stained images to IHC -stained images. For each target image, we show results SegNaiveAug(Linknet) with Adversarial Learning and our method in the output space

shows some poor predictions on IHC-stained images by directly minimizing entropy. From the probability map (2nd column of Fig. 9), it can be observed that if the initial prediction on IHC-stained images is far away from the ground truth, the entropy loss would have a negative influence for segmentation: it will potentially strength the confidence of wrong prediction, which is difficult to be corrected by the model during training, resulting in the poor performance in target domain.

4.5.2 Minimize entropy and adversarial learning

The direct minimization of entropy loss ignores the structural dependence between local semantics. Adversarial learning reduces the difference between the source domain and the target domain through discriminator loss and segmentation loss This method defines the training constraint as the sum of minimizing entropy loss and adversarial loss. In our experiment, the segmentation performance (see MinAdv in Table 3) are similar to direct entropy minimization (MinEnt) which also attributes to the negative impact of entropy loss as described above.

We list all the experimental results as well as the state-of-the-art methods in Table 3, which include naive segmentation with Linknet, Pspnet, Deeplab_v2, different degrees

Table 4 Comparison results of running time between different methods

	Linknet	Pspnet	Ours
Image number	200	200	200
Time	13.36	21.86	17.24
FPS	14.97	9.15	11.60

of color augmentation, direct minimize entropy (MinEnt), minimize entropy with adversarial learning (MinAdv) and our proposed method.

Table 3 shows evaluation results of the proposed algorithm against the state-of-the-art methods [9, 26, 38, 44, 45, 48] that use domain adaptation. The models trained by naïve segmentation networks without color enhancement have poor prediction results for target domain. The different degree of color disturbance based on Linknet is helpful to the training of segmentation network, due to the increasement in coverage between domains. Entropy minimization depends on the initial entropy information, which leads to the poor effective in the output space.

Figure 10 visually shows the segmentation results of naive segmentation, adversarial learning only and our proposed method. It can be seen that the model trained by our method is better than other methods on IHC image prediction. Lastly, the experimental results show that our model works better in the output space. The selected data are representative samples with high prediction confidence, close to the real ground truth value, it has better generalization ability. Our method is the best in IHC validation set, which is 1.8% higher than the state-of-the-art. Meanwhile, our method is able to give a competitive performance on the HE validation set as well.

In order to assess the proposed method, the computational time is also tested in this paper, and shown in Table 4. Our method achieved 11.6 frame per second (FPS), which is satisfactory. Although our method is not the fastest method, but compared with Linknet, our method can achieve a greater accuracy. Meanwhile, compared with the traditional manual clinical operation, our method is satisfactory.

5 Conclusions

In this paper, we present a novel domain adaptation framework for cross-domain histopathological breast cancer segmentation. Through a well-designed segmentation scheme based on domain adaptation, the proposed method can segment the cancer region in the target domain accurately. The accuracy is improved by 1.8% compared with the latest method, and the computational time is also satisfactory, benefit from the proposed entropy selection criterion, the selected pseudo labels are among the high prediction confidence samples. In addition, by our representativeness criterion, the selected data are supposed to contain as many characteristics as possible, while not limited to a few of patients' specific features.

In summary, our method alleviates the burden of manual annotation by only using unlabelled data in target domain, which has large practical value for medical application. It is worth emphasizing that our method is general and not constrained to HE and IHC images. It can be adapted to other similar tasks through simple fine tuning. In future work, we will attempt to evaluate our methods on a larger dataset and on other image domains as well.

References

1. Achanta SDM, Karthikeyan T (2019) A wireless IOT system towards gait detection technique using FSR sensor and wearable IOT devices. Int J Intell Unmanned Syst 8(1):43–54
2. Achanta SDM, Karthikeyan T, Vinothkanna R (2019) A novel hidden Markov model-based adaptive dynamic time warping (HMDTW) gait analysis for identifying physically challenged persons. Soft Comput 23(18):8359–8366
3. Bayramoglu N, Kannala J, Heikkilä J (2016) Deep learning for magnification independent breast cancer histopathology image classification. In: Proceeding of the International Conference on Pattern Recognition, Cancun, Mexico, pp 2440–2445
4. Beeravolu AR, Azam S, Jonkman M, Shanmugam B, Kannoorpatti K, Anwar A (2021) Preprocessing of breast Cancer images to create datasets for deep-CNN. IEEE Access 9:33438–33463
5. Boykov YY, Jolly MP (2001) Interactive graph cuts for optimal boundary and region segmentation of objects in ND images. In: Proceeding of International Conference on Pattern Recognition, Vancouver, British Columbia, Canada, pp 105–112
6. Buslaev A, Parinov A, Khvedchenya E, Iglovikov VI, Kalinin A. A. (2018) Albumentations: fast and flexible image augmentations. arXiv, arXiv:1809.06839.
7. Chaurasia A, Culurciello E (2017) LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation. IEEE Visual Communications and Image Processing, St. Petersburg, FL, USA, pp 1–4
8. Chen LC, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2017) Deeplab: semantic image segmentation with deep convolutional Cets, Atrous convolution, and fully connected Crfs. IEEE Trans Pattern Anal Mach Intell 40(4):834–848
9. Courty N, Flamary R, Tuia D, Rakotomamonjy A (2017) Optimal transport for domain adaptation. IEEE Trans Pattern Anal Mach Intell 39(9):1853–1865
10. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. (2009) Imagenet: A large-scale hierarchical image database. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Sparkle East, USA, pp 248–255
11. Dice LR (1945) Measures of the amount of ecologic association between species. ECY Ecol 26(3):297–302
12. Dou Q, Ouyang C, Chen C, Chen H, Heng P-A (n.d.) Unsupervised cross-modality domain adaptation of convnets for biomedical image segmentations with adversarial loss. arXiv2018, arXiv:1804.10916
13. Drozdzal M, Vorontsov E, Chartrand G, Kadoury S, Pal C (2016) The importance of skip connections in biomedical image segmentation. In: Deep Learning and Data Labeling for Medical Applications, Springer: Cham, pp. 179–187
14. Feige U (1998) A threshold of ln n for approximating set cover. J ACM 45(4):634–652
15. Foran DJ, Yang L, Tuzel O, Chen W, Hu J, Kurc TM, Ferreira R, Saltz JH (2009) A Cagrid-Enabled Learning based Image Segmentation Method for Histopathology Specimens. In: Proceeding of the International Symposium on Biomedical Imaging, Boston, USA, pp 1306–1309
16. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y (2014) Generative adversarial nets. Proceedings of the Advances in Neural Information Processing Systems, Montreal, QC, Canada, pp. 2672–2680
17. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, pp. 770–778
18. Kamnitsas K, Baumgartner C, Ledig C, Newcombe V, Simpson J, Kane A, Menon D, Nori A, Criminisi A, Rueckert D, Glocker B (2017) Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. Proceedings of the International Conference on Information Processing in Medical Imaging, Boone, NC, USA, pp. 597–609
19. Khaki S, Pham H, Han Y, Kuhl A, Kent W, Wang L (2020) Convolutional neural networks for image-based corn kernel detection and counting. Sensors 20:2721
20. Kong J, Shimada H, Boyer K, Saltz J, Gurcan M (2007) Image analysis for automated assessment of grade of neuroblastic differentiation. In: Proceeding of the International Symposium on Biomedical Imaging, Washington, USA, pp. 61–64

21. Kong H, Gurcan M, Belkacem-Boussaid K (2011) Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting. IEEE Trans Med Imaging 30(9):1661–1677

22. Lahoura V, Singh H, Aggarwal A, Sharma B, Mohammed MA, Damaševičius R, Cengiz K (2021) Cloud computing-based framework for breast Cancer diagnosis using extreme learning machine. Diagnostics 11(2):241

23. Lee DH (2013) Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. Workshop on Challenges in Representation Learning, ICML 3(2)

24. Leng L, Yang Z, Kim C, Zhang Y (2020) A light-weight practical framework for feces detection and trait recognition. Sensors 20:2644

25. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, Massachusetts, pp. 3431–3440

26. Long M, Zhu H, Wang J, Jordan MI (2016) Unsupervised domain adaptation with residual transfer networks. Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 136–144

27. Malebary SJ, Hashmi A (2021) Automated breast mass classification system using deep learning and ensemble learning in digital mammogram. IEEE Access 9:55312–55328

28. Nagao T, Sato E, Inoue R, Oshiro H, Takahashi RH, Nagai T, Yoshida M, Suzuki F, Obikane H, Yamashina M, Matsubayashi J (2012) Immunohistochemical analysis of salivary gland tumors: application for surgical pathology practice. Acta Histochem Cytochem 45(5):269–282

29. Nguyen K, Jain AK, Allen RL (2010) Automated Gland Segmentation and Classification for Gleason Grading of Prostate Tissue Images. In Proceeding of International Conference on Pattern Recognition, Istanbul, Turkey, pp. 1497–1500

30. Qu A, Chen J, Wang L, Yuan J, Yang F, Xiang Q, Maskey N, Yang G, Liu J, Li Y (2015) Segmentation of hematoxylin-eosin stained breast Cancer histopathological images based on pixel-wise SVM classifier. Sci China Inf Sci 58:1–13

31. Roccetti M, Delnevo G, Casini L, Mirri S (2021) An alternative approach to dimension reduction for Pareto distributed data: a case study. J Big Data 8(1):1–23

32. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. Comput Vis Pattern Recognit arXiv:1505.04597

33. Ruiz A, Kong J, Ujaldon M, Boyer K, Saltz J, Gur-Can M (2008) Pathological image segmentation for neuroblastoma using the GPU. In: Proceeding of the International Symposium on Biomedical Imaging, Paris, France, pp. 296–299

34. Saber A, Sakr M, Abo-Seida OM, Keshk A, Chen H (2021) A novel deep-learning model for automatic detection and classification of breast Cancer using the transfer-learning technique. IEEE Access 9:71194–71209

35. Shen R, Yan K, Tian K, Jiang C, Zhou K (2019) Breast mass detection from the digitized x-ray mammograms based on the combination of deep active learning and self-paced learning. Futur Gener Comput Syst 101:668–679

36. Shen H, Tian K, Dong P, Zhang J, Yan K, Che S, Yao J, Luo P, Han X (2020) Deep Active Learning for Breast Cancer Segmentation on Immunohistochemistry Images. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp 509–518

37. Somasundaram A, Reddy US (2016) Data imbalance: Effects and Solutions for Classification of Large and Highly Imbalanced Data. In International Conference on Research in Engineering, Computers and Technology (ICRECT), Tiruchirappalli, India, pp 1–16

38. Sun B, Saenko K (2016) Deep coral: correlation alignment for deep domain adaptation. Proceedings of the European Conference on Computer Vision, Amsterdam Netherlands, pp. 443–450

39. Sun Y, Wong AK, Kamel MS (2009) Classification of imbalanced data: a review. Int J Pattern Recognit Artif Intell 23(04):687–719

40. Taher F, Werghi N, Al-Ahmad H, Donner C (2013) Extraction and segmentation of sputum cells for lung Cancer early diagnosis. Algorithms 6:512–531

41. Taher F, Werghi N, Al-Ahmad H (2015) Computer aided diagnosis system for early lung Cancer detection. Algorithms 8:1088–1110

42. Tahmoush D (2009) Image similarity to improve the classification of breast Cancer images. Algorithms 2:1503–1525

43. Theriot CM, Joshua RF (2019) Human fecal Metabolomic profiling could inform Clostridioides difficile infection diagnosis and treatment. J Clin Invest 129:3539–3541

44. Tsai YH, Hung WC, Schulter S, Sohn K, Yang MH, Chandraker M (2018) Learning to adapt structured output space for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, Utah, USA, pp 7472–7481
45. Tzeng E, Hoffman J, Zhang N, Saenko K, Darrell T (2014) Deep domain confusion: Maximizing for domain invariance. arXiv:1412.3474
46. Van Opbroek A, Achterberg HC, Vernooij MW, De Bruijne M (2019) Transfer learning for image segmentation by combining image weighting and kernel learning. IEEE Trans Med Imaging:213–224
47. Vese LA, Chan TF (2002) A multiphase level set framework for image segmentation using the Mumford and Shah model. Int J Comput Vis 50:271–293
48. Vu TH, Jain H, Bucher M, Cord M, Pérez P (2019) Advent: Adversarial Entropy Minimization for Domain Adaptation in Semantic Segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, California, USA, pp. 2517–2526
49. Vununu C, Lee S-H, Kwon K-R (2020) A strictly unsupervised deep learning method for HEp-2 cell image classification. Sensors 20:2717
50. Wilkowski A, Stefańczyk M, Kasprzak W (2020) Training data extraction and object detection in surveillance scenario. Sensors 20:2689
51. Yang L, Zhang Y, Chen J, Zhang S, Chen DZ (2017) Suggestive annotation: a deep active learning framework for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Quebec, Canada, pp 399–407
52. Yang Z, Leng L, Kim B-G (2019) StoolNet for color classification of stool medical images. Electronics 8:1464
53. Zhao H, Shi J, Qi X, Wang X, Jia J (2017) Pyramid scene parsing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, Hawaii, USA, pp 2881–2890
54. Zheng Y (2010) Breast Cancer detection with Gabor features from digital mammograms. Algorithms 3:44–62