IJLAI transactions on Science and Engineering

Quarterly Research Journal

ISSN: Applied

https://ijlaitse.com/index.php/site

Published by:

Ali Institute of Research & Skills Development (AIRSD)

Office No 1, Moiz Clinic Building, Khan Village Road,

New Gulgasht, Multan Pakistan.

Email: admin@ijlaitse.com

Brief Review of Medical Image Segmentation Based on Deep Learning

Lin Teng[1*]

1. Software College, Shenyang Normal University, Shenyang 110034 China

Email:tenglinheu@163.com

Corresponding author: Lin Teng

Abstract

In recent years, deep learning models based on Convolutional Neural Networks (CNN), such as U-Net, ResNet and VGG, have made outstanding achievements in the field of medical image segmentation, providing strong support for the diagnosis of related diseases. However, most of the existing models assume that training data and test data meet the requirement of Independent and Identically Distributed (IID), while in practice, medical images from different imaging devices do not meet this requirement, that is, there is a problem of domain migration. As a result, the accuracy and stability of segmentation results are greatly reduced. Therefore, how to solve the problem of domain migration is of great significance to the practical application of the model. In this paper, the application of deep learning algorithm in medical image segmentation, feature extraction and classification is described. Secondly, the algorithm of deep learning processing multi-modal medical images is summarized. Finally, the problems existing in medical image diagnosis are pointed out, and the future development direction is prospected.

Keywords: Deep learning; Convolutional Neural Networks; Medical image segmentation; Domain migration

1. Introduction

Deep learning is a relatively new and important area of research in machine learning that is closer to artificial intelligence. Traditional machine learning requires professional doctors to segment images, while deep learning can directly process images and is robust to changes in images, thus facilitating the automation of diagnosis [1]. Many scholars have introduced and summarized relevant papers on ultrasonic diagnosis of thyroid diseases based on deep learning. For example, Chen et al. [2] introduced the methods of thyroid segmentation and thyroid nodule segmentation in medical ultrasound images, and analyzed the correlation between the methods in detail. Cao et al. [3] summarized relevant studies on the classification and prediction performance of differentiated thyroid carcinoma (DTC) based on various imaging techniques, discussed the application and limitations of DTC, and mainly emphasized its practicability for DTC patients. Sharifi et al. [4] evaluated the diagnostic process of deep learning on ultrasound images of thyroid nodules and proposed several existing problems that need to be addressed in future work. In view of the above research, this paper not only introduces the segmentation methods of medical images, but also further elaborates from the aspects of image feature extraction and classification differentiation. Based on the deep learning algorithm, this paper reviews the multi-modal medical image research in more detail.

In short, this paper reviews the research progress of key technologies of medical image diagnosis based on deep learning proposed in recent years, and systematically reviews the domestic and foreign research status of deep learning in various processes of medical image analysis. This paper mainly summarizes and summarizes the representative literature on the segmentation, feature extraction, classification and differentiation of thyroid ultrasound images, and expounds the literature on the processing of medical images by deep learning under multi-modal images. Finally, it summarizes the challenges faced by deep learning in the application of medical image analysis and prospects the future development direction.

2. Medical Image Segmentation Based on Deep Learning

The goal of segmentation is to outline and separate different objects in the image to obtain the shape and boundaries of the diseased area. At present, the process of medical image segmentation generally includes three stages: image preprocessing, region of interest localization and image segmentation. Due to the inherent shortcomings of medical images, it is difficult for the computer to accurately identify the lesion area in the original medical image, so it is usually necessary to conduct image preprocessing before segmentation of medical images. The general preprocessing method is to denoise and enhance the image after marking the rough position of the nodule in the thyroid ultrasound image [5]. Region of interest localization refers to the detection of the general location of the focal area, which helps to reduce the calculation amount, improve the algorithm speed, reduce the background interference, and improve the algorithm accuracy.

The segmentation method based on deep learning can identify the size, shape, edge and other information of thyroid parenchyma and nodules in ultrasound images and segment them accurately, so as to diagnose thyroid ultrasound images more accurately. Convolutional neural network (CNN) is the most commonly used network architecture in deep learning algorithms, which can represent hierarchical features of images, which makes it highly adaptable in the field of medical image segmentation. Kumar et al. [6] proposed a novel multi-output CNN algorithm with extended convolutional layers. This algorithm could automatically detect and segment thyroid nodules and cystic components with an average Dice coefficient of 0.76, its performance is comparable to contemporary seed algorithms, but it could not segment very small cystic components.

The semantic segmentation model of CNN based on U-Net decod-coding network structure and U-Net network variant is also widely used in thyroid ultrasound image segmentation. By combining low resolution and high resolution feature maps, U-Net network effectively integrates low-level and high-level image features, and its typical research results are summarized in Table 1. Chu et al. [7] proposed a marker-guided UNet (MGU-Net) model for ultrasonic image segmentation of thyroid nodules. The overlap rate between the segmulated nodule region and the artificially depicted nodule region by this model was close to 100%, and the segmentation accuracy was as high as 97.85%. With less training data, MGU-Net model can significantly improve the segmentation accuracy of thyroid nodule, and provide a reference for clinical diagnosis and treatment. Wang et al. [8] proposed a method for ultrasound image segmentation of thyroid nodules based on joint up-sampling with U-Net as the backbone. This method improved the ability of mining global context information and achieves accurate location of the target with an accuracy of 93.19%. The Dice similarity coefficient (DSC) was 0.8558, which was superior to other existing thyroid nodule segmentation network models. However, this model was more complex than the U-Net model, so the calculation time was longer. Qi et al. [9] proposed a residual substructures and attention gates U-Net (ReAgU-Net) model, which embedded improved residual units into the skip connections between encoder paths. The attention mechanism was introduced to multiply the weight feature maps obtained from shallow layer and deep layer, and the accuracy reached 87.3%. This model solved the problem of spatial information loss due to increased network depth by adding a back-propagation gradient, but the model performed poorly when the contrast between the nodule and the background was low.

Table 1. Results of U-Net network in medical image segmentation

Model	Accuracy/%	DSC
MGU-Net	97.85	0.9576
U-Net model of joint upsampling	93.19	0.8558
ReAgU-Net	87.30	0.8690
DMU-Net	-	0.8107

3. Image Segmentation Based on Maximum Mean Difference

Maximum mean difference (MMD) [10] is a widely used metric for domain adaptation. First, the source domain and target domain data are mapped to the regenerated kernel Hilbert space, and then the sample means of the source domain and target domain are compared to reduce their distribution differences. Its definition is shown in formula (1). Where $\varphi()$ means mapping the data into a re-nucleated Hilbert space.

$$M_{MMD}(X^S, X^T) = || \frac{1}{n^S} \sum_{i=1}^{n^S} \varphi(x_i^S) - \frac{1}{n^T} \sum_{i=1}^{n^T} \varphi(x_i^T) || \quad (1)$$

Pan et al. [11] used MMD to measure the distance between two probability distributions and performed transfer learning by minimizing MMD. Borgwardt et al. [12] referred to the excellent performance of MMD in transfer learning, applied it to domain adaptation, and proposed domain adaptive neural network DANN, aiming to solve the problem that the model generalization performance was reduced due to the different distribution of source domain and target domain. In DANN, MMD is used to quantify the difference between source domain and target domain, and MMD regularization is turned into an effective deep domain adaptive tool. However, due to the relatively small number of layers in DANN network, its ability to extract transferable features is limited. Cao et al. [13] proposed the deep adaptive network architecture (DAN). Considering that the cross-domain mobility of high-level features would decrease significantly with the increase of inter-domain differences, DAN adopted the method of explicitly reducing the inter-domain distance to promote the cross-domain transfer of features. In addition, DAN extended the MMD metric to multi-core MMD (MK-MMD), further improving the domain adaptability of the model. Based on MK-MDD, Arbel et al. [14] proposed a new deep transfer learning method JAN, which aligns the probability distribution of input features and labels in the source domain and target domain in multiple cross-domain specific layers by combining the maximum mean difference (JMMD) criterion. Different from traditional transfer learning, JAN aligns both the edge probability distribution of two domains and the conditional probability distribution of two domains, which can better improve the transfer learning effect. Inspired by this, Wang et al. [15] proposed the deep transmission network DTN, which is mainly composed of a shared feature extraction layer and a discriminant layer. DTN uses MMD to measure the difference of data distribution between source domain and target domain, aligns the edge distribution of source domain and target domain data in the shared feature extraction layer, and aligns the conditional distribution of source domain and target domain data in the discriminant layer.

4. Feature Alignment Method

The feature space is a mathematical space used to represent and describe the features of the data. The hidden layer of deep convolutional neural network can map the input data to the feature space and perform feature extraction. It is found that low-dimensional features contain more position and detail information and have higher resolution, but lack some semantic information due to fewer convolutional layers. In contrast, high-dimensional features are processed by more convolutional layers and have stronger semantic information, so alignment of high-dimensional features helps to extract domain-invariant features and enhance model generalization performance.

The method of feature alignment in feature space originates from the field of image classification. In this field, Ganin et al. [16] proposed a domain adversarial neural network DANN based on representation learning, which was mainly composed of three components: feature extractor network, classifier network and discriminator network. Among them, the feature extractor network is responsible for extracting effective feature representations from input data, the classifier network uses feature representations for classification tasks, and the discriminator network is used to distinguish the source of input data. In order to learn shared feature representations between source and target domains, the optimization goal of discriminator networks is to minimize confusion loss, while the optimization goal of feature extractor networks is to maximize confusion loss. Considering the contradiction between the two optimization objectives, DANN introduces a gradient inversion layer to realize the adversarial training of the two networks, alternately training the feature extractor network and the discriminator network until the discriminator network cannot judge whether the input features come from the source domain or the target domain, so as to achieve the purpose of learning the invariant features of the domain. DANN can learn the feature representation with strong generalization performance, so as to improve the classification performance of the model in different fields.

5. Conclusions

Compared with natural images, medical images have more complex characteristics and application scenarios, and also face different social needs. Although the existing unsupervised domain adaptation technology has made some progress, it still faces many problems that need to be studied and solved. For example, the acquisition of medical image datasets is a complex and challenging task. First of all, these images need to be collected by expensive professional equipment (such as MRI, CT, etc.), due to equipment and technical limitations, the acquisition of medical image data becomes difficult, resulting in the scarcity of medical data. Secondly, the annotation of medical image data sets is a tedious and time-consuming work. The labeling process can take hours or even days for experienced doctors to complete. Due to the complexity of annotation work, annotation data is also scarce. In addition, medical image data sets often contain multiple modes, such as MRI, CT, etc., and problems such as data registration and feature extraction need to be considered when processing different mode data sets, which brings additional challenges to the development and application of models.

Acknowledgements

No

Conflict of Interest

The authors declare no conflict of interest.

References

[1] Chen J, You H, Li K. A review of thyroid gland segmentation and thyroid nodule segmentation methods for medical ultrasound images[J]. Computer methods and programs in biomedicine, 2020, 185: 105329.

[2] Cao Y, Zhong X, Diao W, et al. Radiomics in differentiated thyroid cancer and nodules: explorations, application, and limitations[J]. Cancers, 2021, 13(10): 2436.

[3] Sharifi Y, Bakhshali M A, Dehghani T, et al. Deep learning on ultrasound images of thyroid nodules[J]. Biocybernetics and Biomedical Engineering, 2021, 41(2): 636-655.

[4] Sagheer S V M, George S N. A review on medical image denoising algorithms[J]. Biomedical signal processing and control, 2020, 61: 102036.

[5] Kumar V, Webb J, Gregory A, et al. Automated segmentation of thyroid nodule, gland, and cystic components from ultrasound images using deep learning[J]. Ieee Access, 2020, 8: 63482-63496.

[6] Chu C, Zheng J, Zhou Y. Ultrasonic thyroid nodule detection method based on U-Net network[J]. Computer Methods and Programs in Biomedicine, 2021, 199: 105906.

[7] Wang R, Zhou H, Fu P, et al. A Multiscale Attentional Unet Model for Automatic Segmentation in Medical Ultrasound Images[J]. Ultrasonic Imaging, 2023, 45(4): 159-174.

[8] Qi F, Qiu M, Wei G. Review on ultrasonographic diagnosis of thyroid diseases based on deep learning[J]. Sheng wu yi xue Gong Cheng xue za zhi= Journal of Biomedical Engineering= Shengwu Yixue Gongchengxue Zazhi, 2023, 40(5): 1027-1032.

[9] Wang W, Li H, Ding Z, et al. Rethinking maximum mean discrepancy for visual domain adaptation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021.

[10] Pan Y, Hong R, Chen J, et al. Performance degradation assessment of wind turbine gearbox based on maximum mean discrepancy and multi-sensor transfer learning[J]. Structural Health Monitoring, 2021, 20(1): 118-138.

[11] Borgwardt K M, Gretton A, Rasch M J, et al. Integrating structured biological data by kernel maximum mean discrepancy[J]. Bioinformatics, 2006, 22(14): e49-e57.

[12] Nguyen T T, Gupta S, Venkatesh S. Distributional reinforcement learning with maximum mean discrepancy[J]. Association for the Advancement of Artificial Intelligence (AAAI), 2020.

[13] Arbel M, Korba A, Salim A, et al. Maximum mean discrepancy gradient flow[J]. Advances in Neural Information Processing Systems, 2019, 32.

[14] Wang W, Li H, Ding Z, et al. Rethink maximum mean discrepancy for domain adaptation[J]. ar**v preprint ar**v:2007.00689, 2020.

[15] Ganin Y, Ustinova E, Ajakan H, et al. Domain-adversarial training of neural networks[J]. The journal of machine learning research, 2016, 17(1): 2096-2030.