

A Study of Tissue Image Segmentation Based on “Deep Learning”

Dhyanendra Jain^{1*} Dr. P. K. Bharti² Dr. Prashant Singh³

¹ Research Scholar, Department of Computer Science and Engineering, Shri Venkateshwara University Gajraula Amroha, Uttar Pradesh

² Vice Chancellor, Shri Venkateshwara University, Gajraula Amroha, Uttar Pradesh

³ Department of Information Technology, Dr. Akhilesh Das Gupta Institute of Technology and Management, Delhi

Abstract – Tissue segmentation aims at partitioning an image into segments corresponding to different tissue classes. In healthy subjects, these classes are biologically defined as specific types of tissue, whole organs, or sub-regions of organs, deep learning is mainly used in data dimensionality reduction, handwritten number recognition, pattern recognition and other fields. Such as image recognition, image repair, image segmentation, object tracking, scene analysis, etc., showing very high effectiveness and the mainly study which discussed about Image Segmentation, Applications of Image Segmentation, Image Segmentation and Deep Learning, Most Popular Image Segmentation Datasets, Overview of Deep Learning , Convolutional Neural Networks, 2D and 3D Convolutional Neural Networks, Image Processing Using Deep Learning

Keyword – Image Segmentation, Deep Learning, CNN

-----X-----

INTRODUCTION

Image segmentation is an important and difficult part of image processing. It has become a hotspot in the field of image understanding. This is also a bottleneck that restricts the application of 3D reconstruction and other technologies. Image segmentation divides the entire image into several regions, which have some similar properties. Simply put, it is to separate the target from the background in an image. At present, image segmentation methods are developing in a faster and more accurate direction. By combining various new theories and new technologies, we are finding a general segmentation algorithm that can be applied to kind of images.

Nuclear magnetic resonance imaging gives a clear and high-resolution image of brain tissues. It is a common method for clinical examination of brain diseases. Human brain structure is very complicated. Important tissues include grey matter, white matter, and cerebrospinal fluid. These tissues play a key role in memory, cognition, awareness, and language. Cerebral atrophy/ expansion and leukodystrophy are serious brain dysfunction diseases that have a high incidence in infants and elderly people. However, crucial tissues such as cerebrospinal fluid, grey matter, and white matter are hard to differentiate due to blurry boundaries, especially in the cross-sectional

images that do not show the center of the brain, as shown in. As a result, it is hard for doctors to analyze them separately and find the location of the disease. With the popularization of image-aided medical diagnosis, computer-aided doctors can improve the efficiency of segmenting the grey matter and white matter of the brain MRI. In MR imaging, different signal intensities and weighted images can make the image display at different grey levels. Since the T1 brain magnetic image shows that the soft tissue is better, the experiment selects the brain magnetic resonance T1-W image as the experimental sample. Many approaches have been made to segment the brain image automatically. Segmentation algorithms based on regional, texture, and histogram thresholds are simple but lack accuracy. .resold is a simple but effective way to segment images. However, there are some limits regarding only using this method for segmentation. First, the grey scale of tissues may not be restricted in one range. .is means that, if we simply use threshold to locate the tissues, it may fail to separate all the parts. Secondly, the threshold usually does not consider the spatial properties of an image. For example, the skull is a round structure that covers the other tissues. .is can help us to determine the location of tissues and get more accurate segmentation images. As a result, threshold determination is often considered as an early stage sequential image process. Later, methods related to

fuzzy c-means (FCM) and machine learning are introduced. The atlas-based method is also widely used for brain image segmentation. It has a relatively complete system framework. However, explicit information such as intensity and spatial features is required in order to get accurate results. Spatial and intensity features could be avoided by using convolution neural networks (CNNs).

Image Segmentation

Image segmentation is a classic problem in computer vision research and has become a hotspot in the field of image understanding. The so-called image segmentation refers to the division of an image into several disjointed areas according to features such as grayscale, color, spatial texture, and geometric shapes. So that these features show consistency or similarity in the same area, but between different areas shows a clear difference. Image segmentation is divided into semantic segmentation, instance segmentation and panoramic segmentation according to the different course and fine granularity of segmentation. Segmentation of medical images is regarded as a semantic segmentation task. At present, there are more and more research branches of image segmentation, such as satellite image segmentation, medical image segmentation, autonomous driving etc. With the large increase in the proposed network structure, the image segmentation method is improved step by step to obtain more and more accurate segmentation results. However, for different segmentation examples, there is no universal segmentation algorithm that is suitable for all images.

Traditional image segmentation methods can no longer be compared with the segmentation methods based on deep learning in effect, but the ideas are still worth learning. Like the proposed threshold-based segmentation method, region based image segmentation method, and edge detection-based segmentation method. These methods use the knowledge of digital image processing and mathematics to segment the image. The calculation is simple and the segmentation speed is fast, but the accuracy of the segmentation cannot be guaranteed in terms of details. At present, methods based on deep learning have made remarkable achievements in the field of image segmentation. Their segmentation accuracy has surpassed traditional segmentation methods. The fully convolutional network was the first to successfully use deep learning for image semantic segmentation. This was the pioneering work of using convolutional neural networks for image segmentation. The authors proposed the concept of full convolutional networks. Then there are outstanding segmentation networks such as U-Net, Mask R-CNN, RefineNet, and DeconvNet, which have a strong advantage in processing fine edges.

Applications of Image Segmentation

Image segmentation plays a central role in a broad range of real-world computer vision applications, including road sign detection, biology, the evaluation of construction materials, or video surveillance. Also, autonomous vehicles and Advanced Driver Assistance Systems (ADAS) need to detect navigable surfaces or apply pedestrian detection.

Furthermore, image segmentation is widely applied in medical applications, such as tumor boundary extraction or measurement of tissue volumes. Here, an opportunity is to design standardized image databases that can be used to evaluate fast-spreading new diseases and pandemics (for example, for AI vision applications of coronavirus control).

Deep Learning-based Image Segmentation has been successfully applied to segment satellite images in the field of remote sensing, including techniques for urban planning or precision agriculture. Also, images collected by drones (UAVs) have been segmented using Deep Learning based techniques, offering the opportunity to address important environmental problems related to climate change.

Image Segmentation and Deep Learning

Multiple image segmentation algorithms have been developed. Earlier methods include thresholding, histogram-based bundling, region growing, k-means clustering, or watersheds. However, more advanced algorithms are based on active contours, graph cuts, conditional and Markov random fields, and sparsity-based methods.

Over the last years, Deep Learning models have introduced a new segment of image segmentation models with remarkable performance improvements. Deep Learning based image segmentation models often achieve the best accuracy rates on popular benchmarks, resulting in a paradigm shift in the field.

Most Popular Image Segmentation Datasets

Due to Deep Learning models' success in a wide range of vision applications, there has been a substantial amount of research aimed at developing image segmentation approaches using Deep Learning. At present, there are many general datasets related to image segmentation. The most popular image segmentation datasets are:

PASCAL VOC

The PASCAL Visual Object Classes (VOC) Challenge provides publicly available image datasets and annotations. The PASCAL VOC is one of the most popular datasets in computer vision, with annotated images available for 5 tasks—

classification, segmentation, detection, action recognition, and person layout. A high number of popular segmentation algorithms have been evaluated on this dataset.

For segmentation tasks, the PASCAL VOS supports 21 classes of object labels: vehicles, household, animals, airplane, bicycle, boat, bus, car, motorbike, train, bottle, chair, dining table, potted plant, sofa, TV/monitor, bird, cat, cow, dog, horse, sheep, and person.

Pixels are labeled as background if they do not belong to any of these classes. The training/validation data of the PASCAL VOC has 11'530 images containing 27'450 ROI annotated objects and 6'929 segmentations.

MS COCO

The Microsoft Common Objects in Context (MS COCO) is a large-scale object detection, segmentation, and captioning dataset. COCO includes images of complex everyday scenes containing common objects in their natural contexts. Therefore, COCO is based on a total of 2.5 million labeled segmented instances in 328k images, containing photos of 91 object types that would be recognized easily by a 4-year-old person. For more information about COCO, check out our article what is the COCO Dataset? What you need to know.

.Cityscapes

The large-scale database focuses on the semantic understanding of urban street scenes. It contains a diverse set of stereo video sequences recorded in street scenes from 50 cities, 5'000 fully annotated images and a set of 20'000 weakly annotated frames. Also, the collection time spans several months, which covers seasons of spring, summer, and fall. Cityscapes include semantic and dense pixel annotations of 30 classes, grouped into 8 categories (flat surfaces, humans, vehicles, constructions, objects, nature, sky, and void). The dataset is especially important for autonomous driving applications.

ADE20K

ADE20K offers a standard training and evaluation platform for scene parsing algorithms. The ADE20K dataset contains over 20'000 scenecentric images annotated with objects and object parts, and it provides 150 semantic categories. Unlike other datasets, ADE20K includes object segmentation mask and parts segmentation mask. There are 20'210 images in the training set, 2'000 images in the validation set, and 3'000 images in the testing set.

YouTube-Objects

The YouTube-Objects Dataset is composed of videos collected from YouTube by querying for the names of 10 object classes. In particular, it includes objects from the 10 PASCAL VOC class's airplane, bird, boat, car, cat, cow, dog, horse, motorbike, and train.

The original dataset was developed for object detection with weak annotations and did not contain pixel-wise annotations. Therefore, a fully annotated YouTube Video Object Segmentation dataset (YouTube-VOS) was released containing 4'453 YouTube video clips and 94 object categories.

KITTI

The KITTI dataset is one of the most popular datasets for mobile robotics and autonomous driving. It contains hours of videos of traffic scenarios captured by driving around the mid-sized city of Karlsruhe (on highways and in rural areas). Averagely, in every image, up to 15 cars and 30 pedestrians are visible.

The main tasks of this dataset are road detection, stereo reconstruction, optical flow, visual odometry, 3D object detection, and 3D tracking. The original dataset does not contain ground truth for semantic segmentation, but researchers have manually annotated parts of the dataset.

Overview of Deep Learning

Network Deep learning is a research trend in the rise of machine learning and artificial intelligence. It uses deep neural networks to simulate the learning process of the human brain and extract features from large-scale data in an unsupervised manner. A neural network is composed of many neurons. Each neuron can be regarded as a small information-processing unit. The neurons are connected to each other in a certain way to form the entire deep neural network. The emergence of neural networks makes end-to-end image processing possible. When the hidden layers of the network develop to multiple layers, it is called deep learning. In order to solve the difficult problem of deep network training, layer-by-layer initialization and batching are required, which makes deep learning the protagonist of the era and the research boom. In the field of computer vision, deep learning is mainly used in data dimensionality reduction, handwritten number recognition, pattern recognition and other fields. Such as image recognition, image repair, image segmentation, object tracking, scene analysis, etc., showing very high effectiveness.

Convolutional Neural Networks

The convolutional neural network is a classic model produced by the combination of deep learning and image-processing technology. As one of the most representative neural networks in the field of deep learning technology, it has made many breakthroughs in the field of image analysis and processing. In the standard image annotation set Image Net, which is commonly used in academia, many achievements have been made based on convolutional neural networks, including image feature extraction and classification, pattern recognition, etc. The convolutional neural network is a deep model with supervised learning. The basic idea is to share the weights of feature mapping in different positions of the previous layer network, and to reduce the number of parameters by using spatial relative relationships to improve training performance. From the proposal of the convolutional neural network to the current wide application, it has roughly experienced the stage of theoretical budding, experimental development, large-scale application and in-depth research. The proposal of receptive fields and neurocognitive machines in human visual information is an important theory in the embryonic stage of theory. In 1962, Hubel et al. showed through biological research that the transmission of visual information in the brain from the retina is accomplished through multilevel receptive field excitation. This is the first proposed the concept of receptive field. In 1980, Fukushima proposed a neurocognitive machine based on the concept of receptive fields. It is regarded as the first implementation network of convolutional neural networks. In 1998, LéCun et al. proposed LeNet5 using a gradient-based back propagation algorithm for supervised training of the network, which entered the experimental development stage. The academic circle's attention to convolutional neural networks also began with the proposal of the LeNet5 network and successfully applied to handwriting recognition. After the LeNet5 network, the convolutional neural network has been in the experimental development stage. It was not until the introduction of the Alex Net network in 2012 that the position of convolutional neural networks in deep learning applications was established. The Alex Net proposed by Krizhevsky et al. was the most successful at image classification of the training set of Image Net, making convolutional neural networks become the key research object in computer vision, and this research continues to deepen.

2D Convolutional Neural Networks

Convolutional Neural Networks consists of an input layer, an output layer, and several hidden layers. Each layer in the hidden layer performs a specific operation, such as convolution, pooling, and activation. The input layer is connected to the input image, and the number of neurons in this layer is the pixel of the input image. The middle convolutional layer performs feature extraction on the input data

through a convolution operation to obtain a feature map. The result of the convolution operation depends on the setting of the parameters in the convolution kernel. The pooling layer behind the convolutional layer filters and selects feature maps, simplifying the computational complexity of the entire network. Through the fully connected layer, all neurons in the previous layer are fully connected. The obtained output value is sent to the classifier, which gives the classification result. The general convolutional neural network is 2D CNN. Its input image is 2D and the convolution kernel is a 2D convolution kernel, such as ResNet, VGG (Visual Geometry Group), etc. Suppose the input image size is $H \times W$ with three channels, RGB. The convolution kernel of size (c, h, w) slides on the spatial dimension of the input image, where c, h, w denote the number of channels, the height and the width of the convolution kernel, respectively. The value of the image and the value of (h, w) are entered on each channel to perform a convolution operation to obtain a value.

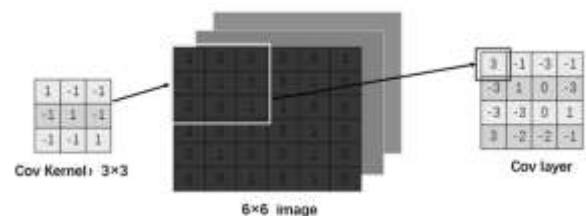


Figure 1: Two-dimensional convolutional neural network (2D CNN) convolution

3D Convolutional Neural Networks

Most images in medical images are usually 3D, such as CT and MRI. Although the CT image we usually see is a 2D image, it is just a slice of it. Therefore, if you want to segment some diseased tissues, you must use a 3D convolution kernel. For example, the convolution kernel used by the segmentation network 3D U-Net is 3D. It changed the 2D convolution kernel in the U-Net network to a 3D convolution kernel, which is suitable for 3D medical image segmentation. 3D Convolutional Neural Networks can extract a more powerful volume representation on the three axes of X, Y, and Z. The use of three-dimensional information in segmentation makes full use of the advantages of spatial information. The 3D convolution kernel has one more depth than the 2D convolution kernel, which means the number of 2D slices of medical images. Given a 3D image $C \times N \times H \times W$ where C, N, H and W represent the number of channels, the number of slice layers, the height and width of the convolution kernel. Like the 2D convolution operation, a value is obtained by sliding the window on the height, width, and number of layers on each channel.

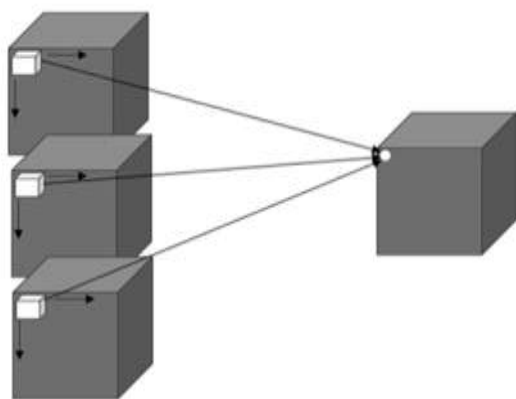


Figure 2: 3D CNN convolution.

Image Processing Using Deep Learning

The first outcome regards the difference between natural images and medical images in Deep Learning. Natural images are plenty of very different objects with very different structures. This allows the network to learn very complex and different filters especially in the deeper layers. Deep learning is one of the promising fields focusing on medical image analysis. Deep learning opens possibilities in areas like bio-imaging, neuron-imaging and DNA sequencing. Deep learning algorithms help in automatic medical imaging segmentations with focus on various features extracted from the medical image dataset. It provides a new way of identifying abnormalities and brings plausible outcomes with better diagnosis. As disclosed previously, Supervised Learning is an approach to learning that requires a known dataset. This set is provided of both inputs and correct outputs for the algorithm used. Starting from this set of examples, the program is guided to describe a model able to predict the correct output. At this point, the prediction model must be validated with another known dataset independent from the training set. Only when the validation phase is satisfactory, the algorithm can be considered reliable for use on unknown data.

Algorithm selection: The first step is to choose the supervised algorithm to use. Every method has different strengths and weak points. The choice depends on the particular problem and on the kind and amount of available data. Some of these algorithms are: Support Vector Machine (SVM), Decision Tree, Artificial Neural Network and Deep Learning.

Training: The training phase is probably the most important one, as the final performances depend on the predictive model built.

- A known dataset is selected; must be as more representative of the problem as possible. Using a dataset not general enough can lead to overfitting and to bad

performances. This set, the training set, must provide an output (label) for each listed input.

- The algorithm is trained with the selected dataset. The aim of this phase is trying to build a model able to fit the data provided, that is, predict the correct output for each input provided as best as possible.

Validation: The validation phase is important to test the performances achieved by the prediction model built in the previous phase. Another known dataset, called test set, is prepared. The dataset must provide, as the training set, reliable input and output for each example. An important property of this set is that it should be as independent as possible from the training one. The previously trained algorithm is here used to predict the input data of the test set. Only the input is used and the output is predicted by the algorithm and stored. The fundamental difference from the train step is that, in this one, the output label is not used to improve the prediction capabilities of the model, but only to evaluate its performances. The predicted outputs are validated using the known outputs. The performance is hence evaluated and analyzed. If they are satisfactory, it is possible to go to the final step, otherwise the algorithm or the training phase must be reviewed with different precautions or parameters. **Model Deployment** Once the algorithm is trained and validated; it is possible to use it as an automatic system to solve the original problem on new data.

CONCLUSION

To conclude, we presented a method to successfully segment brain tissues from MRI using the convolutional neural network. This is a breakthrough since artificial intelligence and machine learning have become more and more widely used in research. By introducing deep learning into the therapeutic field, the speed and accuracy can be improved. This is because machines can automatically analyze the data, which can be much faster and accurate than the manual and semiautomatic analysis. For future work, we can visualize the contours of the borders of different tissues in 3D so that it can be integrated with optical simulation software such as MCVM for low-level light therapy. Our work has a great potential in the medical field, and we hope that our technique can be a criterion of judgment for diagnosing. Medical images are different from natural images. There are differences between different medical images. This difference also affects the adaptability of the deep learning model during segmentation. The noise and artifacts of medical images are also a major problem in data preprocessing. The deep learning model has its own flaws. It mainly focuses on three aspects: network structure design, 3D data segmentation model design and loss function design. The design of the network structure is worth exploring. The

effect of modifying the network structure is significant and can be easily migrated to other tasks. 3D medical data can more accurately capture the geometric information of the target, which may be lost when the 3D data is sliced slice by slice. Therefore, a researchable direction is the design of 3D convolution models to process 3D medical image data. The design of loss function has always been a difficult point in deep learning research

REFERENCE

1. M Manoj krishna, M Neelima (2018) “Image classification using Deep learning” International Journal of Engineering & Technology, 7 (2.7) pp. 614-617 International Journal of Engineering & Technology Website: www.sciencepubco.com/index.php/IJET Research Paper
2. Patitapaban Rath (2017). ‘Contribution of image processing and machine learning for automated analysis of retinal vessels: A Review’, International Journal of Recent Innovation in Engineering and Research, vol. 02, no. 02, pp. 01-07.
3. Sandra Morales, Kjersti Engan, Valery Naranjo & Adrian Colomer (2017). ‘Retinal disease screening through local binary patterns’, IEEE Journal of Biomedical and Health Informatics, vol. 21, pp. 184-192.
4. Sallam Osman Fageeri, Shyma Mogtaba Mohammed Ahmed, Sahar Abdalla Almubarak & Abubakar Aminu Muazu (2017). ‘Eye refractive error classification using machine learning techniques’, Proceedings of International Conference on Communication, Control, Computing and Electronics Engineering, pp. 01-06.
5. Minal B Wankhade & Gurjar, AA (2016). ‘Detection of retinal blood vessels for disease diagnosis’, International Journal of Computer Science and Mobile Computing, vol. 06, pp. 2295-2297.
6. Muhammad Salman Haleem, Liangxiu Han, Jano van Hemert, Baihua Li & Alan Fleming (2015). ‘Retinal area detector from Scanning Laser Ophthalmoscope (SLO) images for diagnosing retinal diseases’, IEEE Journal of Biomedical and Health Informatics, Vol. 19, No. 4, pp. 1472-1482.
7. Hassanien, AE, Emary, E & Zawbaa, MH (2015). ‘Retinal blood vessels localization approach based on bee colony swarm optimization, fuzzy c-means and pattern search’, Journal of Visions and Communication Image Representation, vol. 31, no. 5, pp. 185-196.
8. Siva, Sundhara, Raja, D, Vasuki, S & Rajesh Kumar, D (2014). ‘Performance analysis of retinal image blood vessel segmentation’, Advanced Computing: An International Journal, vol. 5, no. 02/03, pp. 17-23.
9. Gehad Hassan, Nashwa El Bendary, Aboul Ella Hassanien, Ali Fahmy, Abullah M Shoeb & Vaclav Snasel (2015). ‘Retina blood vessel segmentation approach based on mathematical morphology’, Procedia Computer Science, vol. 65, pp. 612-622.
10. Sun, K, Chen, Z & Jiang, S (2012). ‘Local morphology fitting active contour for automatic vascular segmentation’, IEEE. Trans. Biomed. Eng., Vol. 59, No. 02, pp. 464-473.
11. Ahmed Mahfouz, E & Ahmed Fahmy, S (2010). ‘Fast localization of the optic disc using projection of image features’, IEEE Transactions on Image Processing, vol. 19, no. 12, pp. 3285-3289.
12. Delibasis, KK, Kechriniotis, AI, Tsonos, C, Assimakis, N & Gang, L (2010). ‘Automatic model-based tracing algorithm for vessel segmentation and diameter estimation’, Computer Methods and Programs in Biomedicine, vol. 100, no. 2, pp. 108-122

Corresponding Author

Dhyanendra Jain*

Research Scholar, Department of Computer Science and Engineering, Shri Venkateshwara University Gajraula Amroha, Uttar Pradesh