The Instance Segmentation of Medical Ultrasound Images

Chien-Hsing Huang

raylan@isu.edu.tw
Department of Information Engineering, I-Shou University, Taiwan, R.O.C.

Abstract

Ultrasound diagnosis is a medical imaging diagnosis technology based on ultrasound, which is used to visualize soft tissues such as muscles and internal organs, allowing medical staff to understand the patient's condition more intuitively. Although ultrasound hardware equipment has become popular, medical ultrasound images require professional interpretation and even multiple additional operations to obtain useful medical information. They cannot produce a complete human body like computed tomography(CT) or magnetic resonance imaging(MRI). Bioinformatics. This article proposes the use of artificial intelligence technology for image instance segmentation to simplify medical ultrasound image interpretation, reduce medical manpower consumption, and improve medical efficiency.

Keywords: Medical Ultrasound Imaging, Objection Detection, Image Segmentation

1. INTRODUTION

Ultrasound exams use ultra-high-frequency sound waves to pass through the body. Due to the influence of the transmission medium on the speed of sound waves, the material elasticity and density of the medium are collected and calculated to correspond to human tissue. This can be distinguished by different color tones in medical ultrasound images. Doctors use tissues in the body to determine whether organs and tissues are normal or abnormal.[1]

With the advancement of science and technology, ultrasound equipment is becoming more and more popular. At the same time, due to its low cost and no radiation, it is less harmful to the human body than X-rays, computed tomography or nuclear magnetic resonance. It's a low-cost, fairly safe and non-invasive method commonly used to check what's going on inside a patient's body.

However, because ultrasound cannot penetrate bone and air, it is difficult to detect lesions within bones as well as lesions within the gastrointestinal tract (stomach, large intestine, small intestine) and respiratory tract (lungs, trachea, etc.) that contain air. In addition, differences in physical characteristics such as fat content or muscle density can weaken the intensity of ultrasound, making it more difficult to see deeper organ tissue. Specialized doctors are needed to effectively interpret medical ultrasound images.[2]

Ultrasound images are used to examine many parts of the body, depending on the area being examined:

- Abdomen: Mainly examine the liver, gallbladder, part of the pancreas, spleen, kidneys and other organs in the upper abdomen. It can screen for fatty liver, cirrhosis, liver tumors, liver cysts, gallstones, gall polyps, bile duct stones, kidney stones, renal cysts or tumors, splenomegaly, etc.
- Breasts: Examine the breasts for fibrocysts, tumors, or other abnormalities.

- Female pelvic cavity: Pathological changes in the uterus, ovaries and other organs can be examined, such as uterine fibroids, endometrial thickening, endometrial cancer, ovarian cysts, ovarian cancer, etc.
- Carotid arteries: Check the surface and interior of the carotid artery walls on both sides for atherosclerosis. This can assess the condition of the major blood vessels that flow into the brain and give an idea of the extent of vascular disease. Since arteriosclerosis is the main cause of stroke (blockage of blood vessels in the brain), carotid artery ultrasound examination is an important health examination for preventing stroke.
- Heart: Cardiac ultrasound is an important test tool for examining the heart. This test evaluates the structure and function of the heart, learns its size and contraction, and determines the activity of the heart valves.
- Male Prostate: The full name is Transanal Prostate Ultrasound. It puts an ultrasound probe into the rectum and measures the size and shape of the prostate from the rectum, as well as whether there is prostate cancer, etc., related to the anus. Both digital exams and blood PSA tests are powerful tools for checking your prostate.

Applying image segmentation technology to ultrasound medical images can help doctors interpret image content, discover abnormal areas of organs or tissues in the body, and simplify the diagnosis and treatment process. Even when ultrasound equipment becomes popular in home medical equipment such as thermometers and sphygmomanometers, non-medical professional users can simply perform self-examination and detect physical abnormalities in time for early medical treatment.

However, Similar image distribution in ultrasound images of different body parts will correspond to different organs and tissues, requiring interpretation by professional physicians. Simply applying object detection or image segmentation techniques to general images can lead to errors in principle.

2. Medical Ultrasound Imaging

On average, the human ear can hear sounds with frequencies between 16Hz and 20KHz. Sound waves that are too high in frequency to be heard by the human ear are defined as ultrasonic waves. Ultrasound waves can travel through vibrations in air, liquids or solids. Generally divided into destructive ultrasound and non-destructive ultrasound.

When ultrasonic waves are vertically incident from medium 1 (density $\rho 1$, sound speed c1) to medium 2 (density $\rho 2$, sound speed c2), the reflectivity R and transmittance T of the sound pressure are determined by the density ρ and sound speed c of the medium. This product is called the natural acoustic impedance and is defined as follows:

$$R = \frac{\rho_2 \cdot c_2 - \rho_1 \cdot c_1}{\rho_1 \cdot c_1 + \rho_2 \cdot c_2} \tag{1}$$

$$T = \frac{2 \cdot \rho_2 \cdot c_2}{\rho_1 \cdot c_1 + \rho_2 \cdot c_2} \tag{2}$$

In medicine, destructive ultrasound can be used in ultrasonic hyperthermia, which uses focused ultrasound to heat cancer cells to treat cancer. Non-destructive ultrasound can be applied to ultrasonic body scanning: the use of ultrasound pulse reflection signal imaging to detect the condition of human organs and tissues has become a commonly used imaging diagnostic method in medicine (as mentioned in the previous section).

Ultrasound waves pass through a medium or tissue that can conduct sound waves. When they encounter tissue interfaces of different densities, some of the energy is reflected back. The reflected acoustic energy is then received by the probe's piezoelectric crystal and converted into electrical energy. The signal is then converted into a grayscale image and presented on the screen. The frequency of clinical ultrasound is generally between 2MHz and 30MHz. If the influence of bubbles or bones is not considered, the transmission speed between human soft tissues is approximately c=1540 m/s. The table1 lists the speed of sound waves in different media.

Table 1: The Speed of Sound Waves in Different Media

Materials	Speed of Sound (m/s)
Air	330
Water	1480
Fat	1450
Blood	1560
Muscle	1600
Skull Bone	4080
Soft Tissue	1540
Lung	600
Liver	1555
Kidney	1565
Myocardial	1550

3. Image Segmentation

Image segmentation is usually used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of labeling each pixel in an image so that pixels with the same label share some common visual characteristics. The result of image segmentation is a collection of sub-regions on the image (all these sub-regions cover the entire image), or a collection of contours extracted from the image (e.g. edge detection). It can be divided into the following categories according to the subdivision method: [3][4]

- Image segmentation model based on fully convolutional network: including convolutional layer, downsampling layer and fully connected layer. In order to produce segmentation results that are consistent with the input image size, early deep learning-based image segmentation models discarded fully connected layers and introduced a series of upsampling operations. Therefore, the current model aims to solve how to better restore the information loss lost by convolutional downsampling.
- Image segmentation model based on context module: It gradually evolved from restoring pixel information to how to use context information more effectively. On this basis, a series of network structures for extracting multi-scale features were designed.
- Transformer-based image segmentation model: any layer can achieve a global receptive field and establish global dependencies; feature extraction can be achieved without extensive downsampling and more image information can be retained.

4. Medical Ultrasound Images Instance Segmentation

Image segmentation uses image region characteristics to determine whether pixels should be merged into the same object or separated into different objects. Artificial intelligence image segmentation uses convolutional neural networks to determine macroscopically which pixels should be merged into objects. These methods are based on human vision rules. During development, ultrasound image interpretation differed from general image segmentation. Although in theory, since waves entering different materials will have different acoustic impedances, these materials can be distinguished based on their impedance. But in fact, the average acoustic impedance of human organs is very close, but it also varies from person to person. Unless the acoustic impedance is particularly extreme, for example, the largest bone or the smallest air can be clearly distinguished, it is necessary to first set the possible organs according to the shooting position, and then distinguish the position of the organ through color distribution and outline the outline of the organ.

In the convolutional neural network, the image is convolved to obtain the regional correlation of adjacent pixels, and the wide-area image feature distribution is obtained through the pooling operation. Wide-area image features are suitable for image classification, and regional image features are suitable for identifying pixel clusters. Therefore, the purpose of ultrasonic medical image segmentation is to first distinguish the photographed human body parts, then train the image segmentation of the organs and tissues of the parts according to different human body parts, and finally perform image segmentation according to the corresponding organ tissue parts of different human body parts. After the image is segmented, the gray-scale intensity of the segmented area is used to distinguish what kind of tissue or organ the area should be. [5][6][7][8]

This article uses multiple learning modules for the following training:

- The input image of size N is divided into sub-images of size m, and a fully connected neural network is used to train the foreground and background of the labeled blocks.
- After inputting an image of size N, use convolution and pooling downsampling to obtain feature maps of size N/2, N/4, and N/8 respectively, and use a fully connected model to classify the feature map of size N/4. Train to identify which organs of the human body should be included in the image.
- Input an image of size N, and use convolution and pooling downsampling to obtain feature maps of size N/2, N/4, and N/8 respectively. Calculate the relative area ratio of the marked area in the feature map of size N/8. According to the direction of the image center, the distance to the image center and the marked block image histogram, these features are sent to the corresponding image classification fully connected neural network to train marked block classification.

5. Experimental Results

Take cardiac ultrasound images as an example. Cardiac ultrasound images need to segment the left atrium, left ventricle, right atrium, and right ventricle to check for hydropericardium. As shown in Figure 1, the transmission speeds of ultrasound in the three media of myocardium, blood and water are 1550m/s, 1560m/s and 1480m/s respectively. Therefore, despite the different oxygen content of the blood in the atria and ventricles, the grayscale images still have very similar tones. Therefore, the image is first distinguished into cardiac ultrasound based on global features, and then six regions are segmented based on the brightness of the image. Finally, these six regions were defined as ventricles (blue, indigo, green, and yellow regions) and pericardial effusion (red region) based on their brightness characteristics and physiological locations of the tissues.

IJRP.ORG
International Journal of Research Publications
ISSN: 2708-3578 (Online)

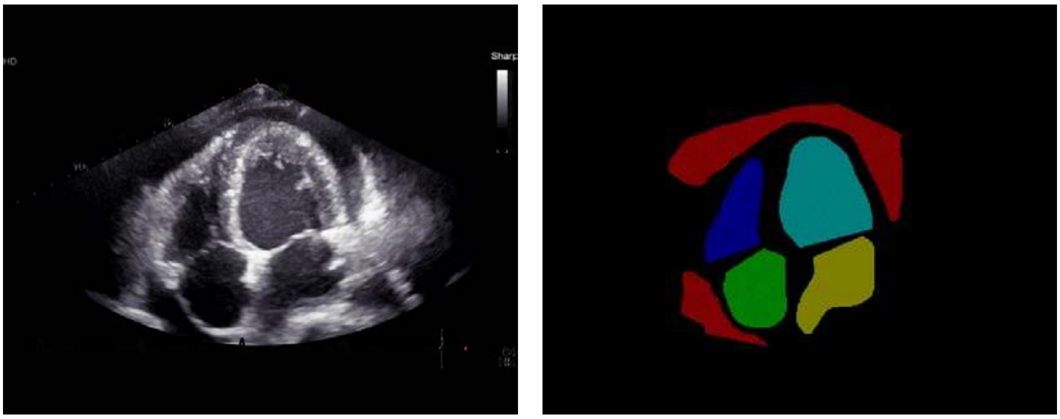

Figure 1: The left image is a cardiac ultrasound image, and the right image is the segmentation results: cardiac chambers (blue, indigo, green and yellow areas) and pericardial effusion (red area).

Ultrasound imaging of the liver is used to detect liver abnormalities. Abnormal shadowed areas on images may indicate tumors, abscesses, or blisters. As shown in Figure 2, abnormal shadows were found in the uniform liver area, and based on the image brightness histogram, it was determined to be hepatocellular carcinoma (HCC).

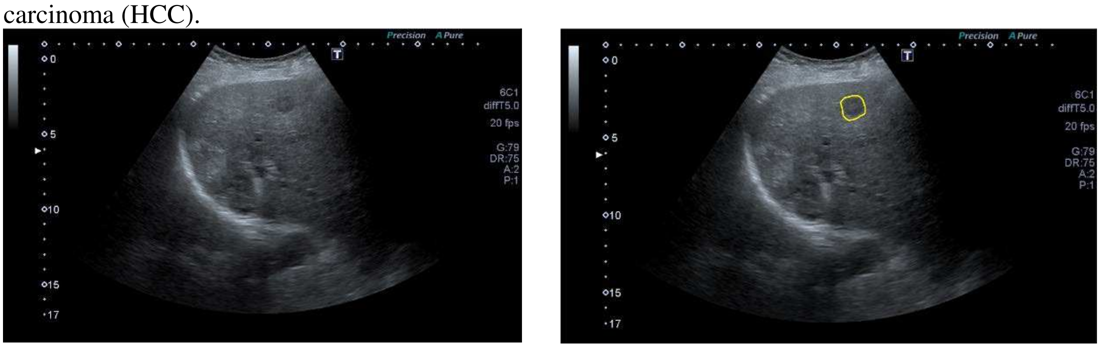

Figure 2: The left image is an ultrasound image of the liver, and the right image is the segmentation result of hepatocellular carcinoma (HCC).

6. Conclusions

By using hierarchical classification to refine and segment ultrasound medical images, the rules of neural network training can be clearly analyzed: which part of the human body is captured in the image, the tissue in that human body part that needs to be distinguished, or something in between. of separation. wait. When errors occur in segmented images, the training of error rules can be strengthened to avoid principle errors such as liver detection rules apply in cardiac ultrasound images.

REFERENCES:

[1] Powers, J., & Kremkau, F. (2011). Medical ultrasound systems. Interface focus, 1(4), 477-489.
[2] Rumack, C. M., & Levine, D. (Eds.). (2023). Diagnostic ultrasound. Elsevier Health Sciences.
[3] Pal, N. R., & Pal, S. K. (1993). A review on image segmentation techniques. Pattern recognition, 26(9),

1277-1294.
[4] Kaur, D., & Kaur, Y. (2014). Various image segmentation techniques: a review. International Journal of Computer Science and Mobile Computing, 3(5), 809-814.
[5] Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems, 33(12), 6999-7019.
[6] Anwar, S. M., Majid, M., Qayyum, A., Awais, M., Alnowami, M., & Khan, M. K. (2018). Medical image analysis using convolutional neural networks: a review. Journal of medical systems, 42, 1-13.
[7] Jha, D., Riegler, M. A., Johansen, D., Halvorsen, P., & Johansen, H. D. (2020, July). Doubleu-net: A deep convolutional neural network for medical image segmentation. In 2020 IEEE 33rd International symposium on computer-based medical systems (CBMS) (pp. 558-564). IEEE.
[8] Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., & Chen, M. (2014, December). Medical image classification with convolutional neural network. In 2014 13th international conference on control automation robotics & vision (ICARCV) (pp. 844-848). IEEE.