euspen's 23rd International Conference & Exhibition, Copenhagen, DK, June 2023
www.euspen.eu

Ultrasound image tracking using deep learning mask R-CNN in radiotherapy

Yu-Chia Wang[1], Sen-Ting Cheng[1], Ai-Ho Liao[2] and Ho-Chiao Chuang[1]

[1]*Department of Mechanical Engineering, National Taipei University of Technology, Taipei 106344, Taiwan*
[2]*Graduate Institute of Biomedical Engineering, National Taiwan Univerity of Science and Tchnology, Taipei 106355, Taiwan*

hchuang@mail.ntut.edu.tw

Abstract
In this study, 1800 ultrasonic images of human diaphragm were captured for training Mask R-CNN. Then, an ultrasonic image tracking algorithm (UITA) was developed to calculate the mean pixel coordinates of the diaphragm detected by Mask RCNN. The coordinates are further input to the Respiratory Motion Compensation System (RMCS) for compensation of breathing introduced motion. The tracking similarity verification experiment of M-UITA was performed. In the experiment, the correlation between the input signal and the signal tracked by M-UITA was compared, and the computed average discrete Fréchet distance was less than 4mm. A respiratory displacement compensation experiment was performed. The proposed method was compared to the UITA by calculating the compensation rates of three different respiratory signals. The experimental results showed that the compensation rate of the proposed method was at most 6.22% higher than that of UITA. This study proposed a novel method, M-UITA, which not only has a high tracking precision and a great tracking stability in the process of tracking diaphragm movement, but also has the advantage of no additional manual adjustment for parameters during operations

Deep learning, mask R-CNN, respiratory motion, object detection, ultrasound image, tumor motion

1. Introduction

Radiation therapy is a common method for cancer treatment and is a local treatment where high-energy photons only hit the cells in the treatment area. However, the treatment is accompanied by side effects of radiation to the healthy tissue surrounding the tumor. Common side effects include Skin inflammation, feeling fatigue, depraved appetite or dysphagia; some severe side effects include spondylitis and myocarditis. During the treatment, the respiratory motion will cause organs to move [1, 2, 3], which results in excess displacement of tumor, especially for lung tumors. Mask R-CNN is one of the SOTA object detection frameworks which was proposed by Kaiming He et al. in 2017 [4]. This framework can perform object classification, localization and pixel-level instance segmentation. Therefore, the main purpose of this study is to build a new ultrasound image tracking technology based on Mask R-CNN, expecting to have a good stability while tracking the detected motion of diaphragm; moreover, we apply it to RMCS to evaluate the effeteness of compensation.

2. Materials and Methods

In order to explore the feasibility of applying M-UITA to RMCS, we took CT images for verification at Taipei Medical University Hospital. Experiments were performed to verifiy the tracking displacement accuracy of M-UITA and respiration motion compensation. The compensation rates tracked with UITA and M-UITA were compared to see the advantage of the proposed technology.

2.1. Experimental apparatus

The experimental apparatus used in this study include Respiratory Motion Simulation System (RMSS) [5], Respiratory Motion Compensation System (RMCS), ultrasound equipment, diaphragm phantom. RMSS and diaphragm phantom were used for the simulation of human respiratory motion. RMSS would carry the diaphragm phantom and move according to the received pre-recorded respiratory signals. This design could better simulate the actual movement when a human breathes. A rubber belt was placed on the wall for the simulation of the diaphragm, and ultrasonic wave would bounce off the belt to form an ultrasound image (Fig. 1). Since the rubber belt was difficult to be seen under CT, a metal wire was closely attached to the rubber belt for a clear contour. The CT images were analyzed with Tracker. We manually located feature point of the diaphragm phantom in every frame of the series CT images and acquired the phantom's movement trajectories. Fig. 2 shows the overall experiment setup.

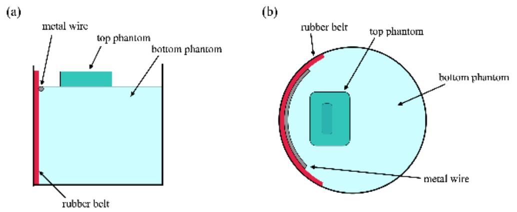

Figure 1. (a) Side view of the diaphragm phantom. (b) Top view of the diaphragm phantom.

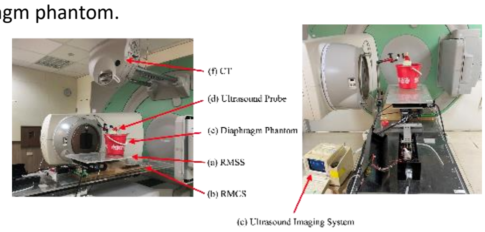

Figure 2. The overall experiment setup.

2.2. Dataset and model training

The dataset of 1,800 human diaphragm images was obtained with ultrasonic equipment, and the size of each image is

800×600. Deformed shapes of the diaphragm under severe breathing were intentionally recorded in order to increase the variety of the dataset for model train. The 1800 images were divided into training set and validation set at the ratio of 80% and 20%. In this work, VIA (VGG Image Annotator) was used for data labeling. Because the number of samples in this study was not large, a pre-trained model based on the COCO dataset [6-7] was used. During training the model, learning rate was set to 0.001, batch size was set to 5, and epoch was set to 100. Fig. 3 shows the screenshots of working M-UITA. The red dots in the figures are the visualization of the mean coordinates calculated by M-UITA.

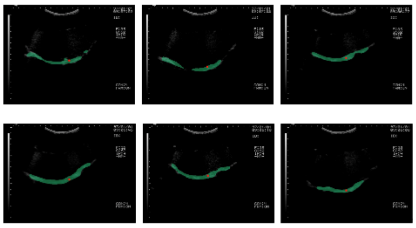

Figure 3. Screenshots of working M-UITA.

3. Results

When tracking specific objects in noisy scenes such as ultrasound images, deep learning-based tracking technology can bring greater robustness. Fig. 4 shows the racking trajectories of each input signals. It can be seen from Fig. 4 (b) that after 70 seconds in Pattern C, the tracking trajectory of UITA is extremely unstable, which may be because of the feature point of UITA was lost while tracking. The reason could be the contact surface between the ultrasonic probe and the upper diaphragm phantom existed a little gap during the compensation process, which makes the contour of the diaphragm disappear. However, tracking with M-UITA did not have this issue and showed a stable tracking process. In Pattern D, the compensation rate using M-UITA is quite close to that of UITA. Due to the lower respiratory frequency, the uncertain factors such as speckles and contact surface gap in the ultrasound image are less, which may not highlight M-UITA's advantages. The residual signal at the turning point has larger concaves (closer to zero), which means M-UITA can track the displacement of the diaphragm more accurately at the turning points.

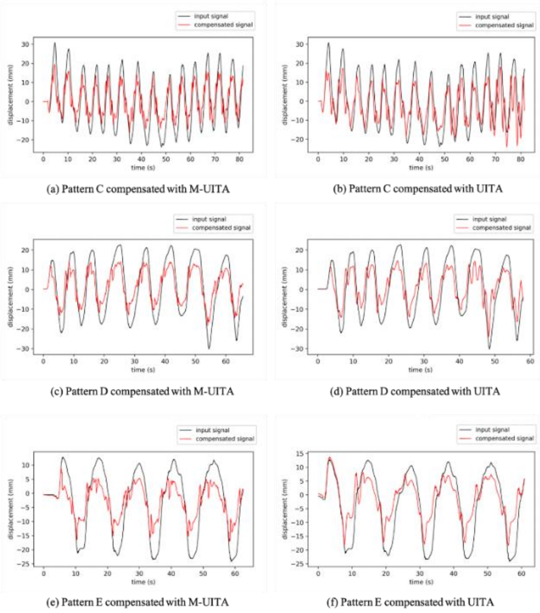

Figure 4. The tracking trajectories of: (a) Pattern C compensated with M-UITA; (b) Pattern C compensated with UITA; (c) Pattern D compensated with M-UITA; (d) Pattern D compensated with UITA; (e) Pattern E compensated with M-UITA; (f) Pattern E compensated with UITA.

4. Discussion

Since UITA is a traditional computer vision based on rules, image processing parameters (such as binarize, erode) need to be manually set before tracking. If the conditions of the scene do not correspond to the setting, it will not be able to track effectively. However, it is difficult to assure that the quality of ultrasonic purity of ultrasound pictures and the shape of diaphragm are clean and completed when detecting different patients. During the treatment, the RMCS will continuously moving, which makes the angle of the ultrasonic probe and the situation of contact with body surface cannot be perfectly remained as the initial position. The above factors increase the uncertainty while tracking the diaphragm with UITA, and also make the tracking process more difficult. M-UITA does not have the above problems. M-UITA does not need to manually adjust the parameters to fit the condition, and the tolerance for uncertainty in the image is also much higher. UITA tracks a single feature point of the diaphragm while M-UITA tracks the mean pixels of the detected diaphragm's area. As long as the outline of the diaphragm still exists, even if the shape is incomplete or distorted, M-UITA can still detect the diaphragm by virtue of the good generalization of the model. Tracking the mean pixel coordinates of the area is more representative of the overall displacement than tracking a single specific point.

5. Conclusions

In this work, the ultrasonic image tracking technology, M-UITA, was developed based on Mask R-CNN and was applied to the respiratory motion compensation system, RMCS. A total of 1800 ultrasonic images of real diaphragm were obtained by ultrasonic equipment as a data set for the model training. M-UITA's function is to detect and calculate the pixel average coordinates of the diaphragm segmented by Mask RCNN, and transmit the data to RMCS. In order to verify the feasibility of M-UITA, we performed the compensation experiment. The compensation rate was calculated to evaluate the effeteness of the proposed method and compared the results of M-UITA and UITA. We found that the compensation rate of the proposed method is at most 6.22% higher than that of UITA. In general, M-UITA not only has a high tracking precision and a great tracking stability in the process of tracking diaphragm movement, but also has the advantage of no additional manual adjustment for parameters during operation.

References

[1] Goitein and Michael 2004 Organ and tumor motion: an overview *Seminars in Radiation Oncol.* **14**. WB Saunders.

[2] Siebenthal V and Martin 2007 4D MR imaging of respiratory organ motion and its variability *Phys. Med.& Bio.* **52.6** 1547.

[3] Pham D 2014 A review of kidney motion under free, deep and forced-shallow breathing conditions: implications for stereotactic ablative body radiotherapy treatment." *Tech. cancer research & treatment* **13.4** 315-323.

[4] Kaiming H 2017 Mask R-CNN *IEEE International Conference on Computer Vision (ICCV)* 2980-2988.

[5] Chuang HC 2014 A respiratory compensating system: design and performance evaluation *J appl. Clinic. med. Phys.* **15** 4710.

[6] Lin TY 2014 Microsoft COCO: Common Objects in Context ECCV.

[7] Kaiming H 2016 Deep Residual Learning for Image Recognition *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 770-778.