MAGAN: Mask Attention Generative Adversarial Network for Liver Tumor CT Image Synthesis

Yang Liu
Shenyang Institute of Automation
Lu Meng (menglu1982@gmail.com)
Northeastern State University - Broken Arrow Campus https://orcid.org/0000-0003-2442-8354
Jianping Zhong
Northeastern University

Research article

Keywords: generative adversarial network, attention map, medical image synthesis, liver CT image, liver tumor

Posted Date: July 16th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-41685/v1

Abstract

Background: For deep learning, the size of the dataset greatly affects the final training effect. However, in the field of computer-aided diagnosis, medical image datasets are often limited and even scarce.

Methods: We aim to synthesize medical images and enlarge the size of the medical image dataset. In the present study, we synthesized the liver CT images with a tumor based on the mask attention generative adversarial network (MAGAN). We masked the pixels of the liver tumor in the image as the attention map. And both the original image and attention map were loaded into the generator network to obtain the synthesized images. Then the original images, the attention map, and the synthesized images were all loaded into the discriminator network to determine if the synthesized images were real or fake. Finally, we can use the generator network to synthesize liver CT images with a tumor.

Results: The experiments showed that our method outperformed the other state-of-the-art methods, and can achieve a mean peak signal-to-noise ratio (PSNR) as 64.72dB.

Conclusions: All these results indicated that our method can synthesize liver CT images with tumor, and build large medical image dataset, which may facilitate the progress of medical image analysis and computer-aided diagnosis.

Background

Medical image analysis and processing is the core of computer-aided diagnosis, which has been greatly prompted by deep learning. And the training of deep learning can be extensively influenced by the size of the dataset, that is, the more datasets can be obtained, the better the performance the trained deep learning model can achieve. However, in the field of computer-aided diagnosis, the medical image is very limited and even scarce, due to the privacy of patients, the expense of medical image acquisition, and so on. Therefore, synthesized medical images can be seen as the only feasible way to solve this problem, and generative adversarial networks (GAN) provides us a powerful tool to realize it.

GAN was firstly proposed by Goodfellow and colleagues in 2014 and was widely used in various fields, such as image processing, natural language processing, and even medical image synthesis[1]. For skin lesion images, Baur and colleagues synthesized the images of skin lesions with GAN[2], which enlarged the skin image dataset and improved the performance of lesion segmentation. For liver CT images, GAN was mainly used for expanding the dataset of the liver lesion[3] or image deoising[4], but the focus of GAN was only on the liver lesion, not on the whole liver CT images. For brain images[5], there are many image modules, such as CT images, magnetic resonance (MR) image, positron emission tomography (PET), and different modules have different image acquisition methods and different influence on human brains. Dong Nie and colleagues used GAN to synthesize 7T images from 3T MR images[6], because 7T MR images were very rare due to the expensive image acquisition costs and the side effects of high magnetic field strength. Moreover, some studies proposed to train a GAN to generate CT images from MR images to

avoid the radiation from the CT image acquisition[7, 8]. For retinal images, the image resolutions were generally smaller than 100 × 100, and the image contents were only limited to single color background and vessels. Based on the characteristics, some studies[9] used GAN to synthesize the whole retinal image to enlarge the retinal image dataset, but the method cannot be generalized to other medical image modules with bigger image resolution and more organs, such as liver CT image or brain MR image.

Above all, all these medical image synthesis methods can be categorized into three types: (1) transformation of different modules, such as from CT images to MR images; (2) transformation between different parameter of image acquisition, such as from 3T MR images to 7T MR images; (3) image synthesis of the small resolution, such as skin and retinal images. As far as we know, there are no image synthesis methods for high-resolution medical images with lesions or tumors, such as 512 × 512 liver CT images with tumors. Therefore, in the present study, we proposed a novel image synthesis model for liver CT images with tumors based on mask attention generative adversarial network (MAGAN). Using this model, we can build a liver CT image dataset consisting of thousands of synthesized 512 × 512 slices, furthermore, it also can facilitate the progress of computer-aided diagnosis and the training of deep learning models.

The main contributions of our work are as follows, (1) we combined GAN with attention mechanism and proposed a novel MAGAN model; (2) we proposed an effective method of enlarging the existing medical image dataset.

Methods

In the present study, we synthesized liver CT images with tumors based on the mask attention generative adversarial network model[10], whose framework was shown in Fig.1. Firstly, all the pixels of liver tumors in the original image were labeled by the white color and used as the attention map. According to the attention mechanism, liver tumors were the highlighted relevant features of the CT images, and the attention map was also the key part of the success of the proposed algorithm. In the procedure of image synthesis, the liver tumor was the saliency map in the whole liver CT image, which meant that all the pixels of the liver tumors were masked by the attention map. The original image and the attention map were paired together and called “pairing A”. Then, the original image and the attention map were loaded into the generator network to obtain a synthesized image, and the attention map and the synthesized image were paired together and called “pairing B”. Next, pairing A and pairing B were both loaded into the discriminator network to determine if the synthesized image was real or fake. The generator network and the discriminator network were trained with adversarial learning so that both of them can become more and more powerful. After training, the generator network can fill the pixels of attention map with similar gray values, texture, and shape of liver tumors, to synthesize liver CT images with tumors. More details of our model can be obtained from 2.1∼2.3 sections.

2.1 Attention model

All liver CT images used in our method were from a public liver CT dataset, which was Liver Tumor Segmentation (LiTS) [11, 12]. And all the tumors in the liver CT images were manually labeled by radiologists. We aimed to synthesize liver CT images with tumors, and the synthesized materials were

from two aspects, liver CT images from healthy controls and liver tumor CT images from patients. Moreover, the liver tumor was the most salient region for clinicians and was also the most difficult part of the whole synthesis procedure. Therefore, according to the tumor labels from the LiTS dataset, the image values of all the corresponding pixels in the tumors were changed to 4096, which meant “white color”, and represented as an attention map in our model. Based on attention mechanism, the original image and the attention map were transformed into feature maps A and B by using 1×1 convolution, respectively, and then all these feature maps were concatenated by using matrix multiplication, shown in Fig.2.

$$S_{i,j} = A_i^T B_j \tag{1}$$

Then, we performed softmax on the concatenated feature maps $S_{i,j}$ to calculate the distribution of attention $D_{i,j}$ on the ith position of the jth synthetic region.

$$D_{i,j} = \frac{\exp(s_{i,j})}{\sum_{i=1}^{N} \exp(s_{i,j})} \tag{2}$$

Therefore, the liver tumor mask images were used as attention maps to efficiently find the liver tumors’ internal and external characteristics of the images.

2.2 Generator network

The structure of our generator network was shown in Fig.3, which consisted of two contracting paths and an expansive path, showing the U-shape architecture[13]. The input of these two contracting paths were original image and attention map, respectively, and both of them consisted of nine blocks, and each block was composed of the ReLu layer, convolutional layer, and batch normalization(BN) layer.

In the contracting path, the image resolution was reduced but the feature information was increased. To overcome the drawback of a regular convolution operator, whose receptive field was small, we used a dilated convolution operator[14] in the first four layers of the contracting path, so that we can capture image features from a larger scale. And we used regular convolution operator in the other five layers of the contracting path because the sizes of the images were already smaller than 32×32, which can not support dilated convolution operator. The feature maps from both of the two contracting paths were firstly loaded as input to the attention model, whose framework was shown in Fig.2, and then the distribution of attention value was transferred via residual connections. In the expansive path, the spatial information and the feature information were combined through a sequence of up-convolutions layer, BN layer, ReLu layer and residual connections with high-resolution features from the attention model. Residual connections played important roles in MAGAN, which were used to bypass the nonlinear

transformation, accelerate the training speed and upgrade the performance of our model in the training of the deep CNN.

512×512 original image and attention map were loaded as inputs into the generator network, and the image resolution was reduced by half while passing each block in the contracting path. After nine blocks in the contracting path, the input images became 1×1 with 1024 feature maps. Then, these feature maps were up-convolved in the expansive path, and the size of the image increased one time while passing each block in the expansive path. After nine blocks in the expansive path, the image was restored as a 512×512 resolution image. In the generator network, the whitened regions in the liver CT images can be transformed into tumor regions. The loss function of our generator network was shown as formula (3)

$$L_{adv}(G) = \mathrm{E}_{v,r\sim p_{data}(v,r)}[\| r - G(v) \|_1] \tag{3}$$

r denoted the real image, v denoted the concatenated image, $G(v)$ denoted the synthesized image calculated by the generator network.

2.3 Discriminator network

The structure of our discriminator network was shown in Fig.4, which consisted of six blocks, and each block was composed of a convolutional layer, ReLu layer, BN layer or sigmoid layer.

The inputs of the discriminator network were two pairings, which were pairing A (original image, attention map) and pairing B (synthesized image, attention map). Inspired by PatchGAN[10], all the 512×512 resolution images were divided into 900 patches, whose size was 142×142. After going through six blocks of discriminator network, the size of output probabilities maps were 30×30, which indicated each pixel in the output probabilities maps corresponded to one patch of the input images. The mean value of all the pixels in the probabilities maps can be recognized as the result of the discriminator network.

The loss function of our discriminator network was shown as formula (4).

$$L_{adv}(D) = \mathrm{E}_{v,r\sim p_{data}(v,r)}[\log D(v,r)_{real}] + \mathrm{E}_{v\sim p_{data}(v)}[\log(1 - D(v, G(v,r))_{fake})] \tag{4}$$

r denoted the real image, v denoted the attention map, $G(v,r)$ denoted the synthesized image calculated by the generator network, $D(v, r)$ denoted the discrimination probability calculated by the discriminator network.

The total loss function of our GAN was shown as formula(5).

$$L = \arg\min_{G}\max_{D} \lambda_1 L_{adv}(G) + \lambda_2 L_{adv}(D) \tag{5}$$

λ_1 and λ_2 were coefficients.

Results

In our experiments, we used LiTS as our image dataset of liver CT images with tumors, which consisted of only 131 sequences. The size of LiTS was not big enough for the training of deep learning algorithms, such as liver tumor segmentation or classification. To enlarge the LiTS, we chose 4555 2D slices with tumors from 131 sequences of liver CT images. Then, all the images were normalized by using formula (6)

$$value_{normalized} = \frac{value_{original} - mean}{std} \quad (6)$$

$value_{original}$ and $value_{normalized}$ represented the original and normalized image pixels value, respectively. *mean* indicated the mean value of image pixels, and *std* indicated the standard deviation of image pixels. Moreover, we specially cut the tumor regions from the liver CT images and built a liver tumor dataset, then we augmented the tumor dataset by flipping, rotating, scaling the original tumor region so that we can create a liver tumor dataset of 50000 slices from the original 4555 slices, which were used as the mask attention map in our method.

The hardware and software configuration of our experiments were shown in Table 1. The quantitative evaluation metric used in our experiments was the peak signal to noise ratio (PSNR). There were four sections in our experiments, including training of our model, quantitative comparison between our method and other state-of-the-art methods, Turing test for synthesized images by radiologists, and the evaluation of the synthetic dataset for the medical image segmentation.

Table 1 Hardware and software configuration of our experiments

Item	Configuration
Operating system	Ubuntu 16.04
GPU	NVIDIA GeForce GTX 1080
CPU	Intel Core i5-7500 @3.4GHz
Software toolkit	Python 2.7、tensorflow 1.1、matlab 2016b
Disk	500 GB
GPU memory	8 GB
System memory	16 GB

3.1 Training of our model

The configurations of hyperparameters in our model during the training were shown in Table 2. The proposed MAGAN network was implemented by python 2.7 and TensorFlow 1.1 and trained on a NVIDIA GeForce GTX 1080 GPU using Adam optimizer with a learning rate 0.0002. It cost about ten hours for the whole procedure of the training.

Shown as Fig.5~Fig.9, and we can find that as the number of iterations increased, the performance of the synthesized CT liver images became better and better. After the first iteration of training (in Fig.5), the performance of the synthesized image from the generator network was terrible, for example, most pixels were black, the contour was blurring, intense chessboard effect. All these bad performances indicated

that the training had just started, and more iterations were needed. After ten iterations (in Fig.6), the whole image was more clear, the contour was more vivid, but the chessboard effect still existed. After one hundred iterations (in Fig.7), the performance of the synthesized image was much better and closer to the real image, more details can be visualized, human organs were vivid, the chessboard effect was weaker but still existed, whitened regions were not filled with tumor texture. After one thousand iterations (in Fig.8), the chessboard effect disappeared, all details of liver CT were restored, but only part of the whitened regions was filled with tumor texture. After ten thousand iterations (in Fig.9), it was hard to tell differences between synthesized image and real image, all details of liver CT was restored, the chessboard effect disappeared, all the whiten regions were filled with tumor texture.

Table 2 Hyperparameters of our model

parameter	value
initial learning rate	0.0002
Adam momentum	0.5
λ_1 in formula (5)	100
λ_2 in formula (5)	1
exponential decay	0.99
batch_size	1
epoch	10
dropout	0.5
frequency of saving loss value	100
frequency of saving model	500

The loss function of the generator network, discriminator network, and total network during the training were shown in Fig.10, Fig.11, Fig.12, respectively, and we can conclude that the loss functions decreased as the number of iterations increased, and became steady after about 10000 iterations, which indicated that our model performed well during the training.

Results of the synthesized image were shown in Fig.13, three liver tumor images with tumor masks were in the first row, which was used as inputs of our model, and can obtain the synthesized images in the second row. We compared the synthesized images and the real images and calculated the differences between them. The color image of the differences was shown in the fourth row. All these results showed that our method can synthesize liver CT images with tumors, and the synthesized images were almost identical to the real images.

To test the impact of the dilated convolution operators in the MAGAN, we replaced the dilated convolution operators with the regular convolution operators in the contracting path of the generator network and quantitatively compared PSNR of these two GAN networks. And we found that the network with regular convolution operators can provide PSNR of 59.66, while the MAGAN with dilated convolution operators can provide PSNR of 64.72, which indicated the effectiveness of the dilated convolution operators in our network.

To test the impact of the residual connections in the MAGAN, we removed the residual connections and quantitatively compared PSNR of these two GAN networks. And we found that the network without residual connections can provide PSNR of 55.23, while the MAGAN with residual connections can provide PSNR of 64.72, which indicated the effectiveness of the residual connections in our network.

Besides, we can also manually or automatically “add” tumor regions on the healthy liver CT images using our liver tumor dataset of 50000 slices, to create a diseased liver CT image, shown in Fig.14. The healthy

liver CT images were in the first row. In the second row, manually changing the pixel values of two regions to white color, which meant that these two regions were the selected tumor regions. Using our method, the results of the synthesized images were shown in the third row. All these results showed that our method can intelligently create liver CT images with tumors based on the healthy liver CT images, and the synthesized diseased images were almost identical to the real ones.

3.2 Quantitative comparison

In this section, we quantitatively compared our method with other seven state-of-the-art medical synthesis methods using the same dataset as ours: (1) atlas-based method[15]; (2) sparse representation (SR) based method; (3) structured random forest with ACM (SRF+)[16]; (4) manipulable object synthesis (MOS)[17]; (5) deep convolutional adversarial networks (DCAN) method[18]; (6) multi-conditional GAN(MC-GAN)[19];(7) mask embedding in conditional GAN (ME-cGAN)[20]. The first four methods were implemented by our group, and the source code of DCAN, MOS, ME-cGAN were downloaded from GitHub (www.github.com/ginobilinie/medSynthesis, www.github.com /HYOJINPARK/MC_GAN, and www.github.com/johnryh/Face_Embedding_GAN). The results of the quantitative comparison were shown in Table 3, which indicated that our method outperformed the other seven approaches and benefited from attention mechanism, dilated convolution operator, and residual connections.

Table 3 The quantitative comparison between our method and seven other approaches

	Method							
	Atlas[15]	SR	SRF+[16]	MOS[17]	DCAN[18]	MC-GAN[19]	ME-cGAN[20]	our method
mean PSNR(dB)	45.15	49.77	55.30	60.11	58.26	59.29	61.35	64.72

3.3 Turing test

To further verify the effectiveness of our method, we did the Turing test. Two experienced radiologists from Shengjing Hospital of China Medical University were asked to classify one hundred liver CT images into two types: real image or synthesized image. The radiologists were not aware of the answer to each image before the Turing test. The one hundred liver CT images consisted of fifty real CT images and fifty synthesized images. The results of the Turing test were shown in Table 4, radiologist #1 made correct judgments for 74% real image slices and 64% synthesized image slices, and radiologist #2 made correct judgments for 84% real image slices and 12% synthesized image slices. The radiologists made correct judgments for most of the real images and maybe psychologically influenced by the existence of a synthesized image, so they made some errors about the real images. Furthermore, the radiologists made difficult judgments for the synthesized images and can’t tell the obvious differences between the real images and the synthesized images. And according to radiologist #1, his most reliable evidence of telling the difference was the color of the tumor region was a little darker than the real ones, which was also the improvement we needed to do in the future. All these results of the Turing test indicated that our method can synthesize liver CT images with a tumor, which were almost identical to the real ones.

Table 4 The Turing test of our method

	real image (50 slices)		synthesized image (50 slices)	
	be judged as real images	be judged as synthesized images	be judged as real images	be judged as synthesized images
radiologist#1	37	13	18	32
radiologist#2	42	8	44	6

3.4 Evaluation of synthetic dataset for medical image segmentation

To evaluate the effectiveness of the synthetic dataset in the training of deep learning models, we used fully-connected network (FCN)[21] to perform the tumor segmentation task in the liver CT images and trained the FCN model using LiTS dataset (images from 131 subjects) and the new dataset obtained by our method (images from 131 real subjects and 865 synthetic subjects). And we used the Dice Index to quantitatively evaluate the performance of the segmentation results from the two trained FCN models. The FCN model trained by the LiTS dataset can provide a Dice value of 0.611 for the tumor segmentation, and the FCN model trained by a new dataset can provide a Dice value of 0.658 for the tumor segmentation. The result indicated that the synthesized liver CT images obtained by the proposed method can effectively enlarge the original dataset, and as the number of images in the dataset increased, the performance of the training of the deep learning model can become better, which resulted in the higher Dice value for the liver tumor segmentation.

Discussion

In the present study, we combined the attention mechanism and GAN model and proposed a novel CT image synthesis algorithm, which was MAGAN. As far as we know, the existing medical image synthesis methods mainly focused on the transformation of different modules or transformation between the different parameter of image acquisition, and our study was the first research of synthesizing the liver CT images with tumors in high resolution and enlarging the size of the medical image dataset.

Supposed that we had a dataset of chest CT images with lung nodules, whose size was one hundred. While we used this dataset for the training of deep learning, we may find that the trained model was not good enough due to the small size of the dataset. Under these circumstances, the proposed MAGAN can be used to synthesize thousands of chest CT images with lung nodules based on the original one hundred images. This kind of similar requirements from clinical researches and deep learning studies are very common. And the proposed method can meet the requirements.

From the quantitative comparison between the proposed method and the other seven state-of-the-art medical image synthesized methods, we can conclude that the proposed method outperforms the others, and the main reasons were the attention map, which mainly focused on the regions of interest in the medical images, such as liver tumors or lung nodules.

During the Turing test, two experienced radiologists can not clearly distinguish the synthesized liver CT images and the real liver CT images. We used the judgments of experts as the golden standard, and we may conclude that the synthesized liver CT images with tumors can be used as the real ones, and the size

of the training dataset of medical images can be enlarged from one hundred to thousands. Bigger the medical image dataset is, better the training performance can be.

Conclusions

In the present study, we proposed a method of synthesizing liver CT images with tumors based on mask attention generative adversarial networks. The experimental results showed that our method outperformed the other seven widely used approaches and can achieve 64.72db mean PSNR, and the Turing test indicated that even the experienced radiologists can't tell the differences between the synthesized images from our method and the real ones. All these results meant that using our method, we can build a huge medical image dataset to facilitate the diagnosis of computer-aided diagnosis and the training of deep learning.

Declarations

7.1 Ethics approval and consent to participate

Not applicable.

7.2 Consent for publication

Not applicable.

7.3 Availability of data and materials

All liver CT images used in our method were from a public liver CT dataset, which was Liver Tumor Segmentation (LiTS) [11, 12].

7.4 Competing interests

The authors declare that they have no competing interests.

7.5 Funding

This research was funded by National key research and development Project(2018YFB2003200), National Natural Science Foundation of China (61973058), and Fundamental Research Funds for the Central Universities (N2004020). The roles of the funders are medical imaging data provider and payment payer.

7.6 Authors' contributions

YL provided the original ideas. LM wrote the code and analyzed the data. JZ wrote the manuscript.

7.7 Acknowledgements

We appreciated Professor Jing Xiang from Cincinnati Children's Hospital Medical Center.

Abbreviations

MAGAN: mask attention generative adversarial network

PSNR: peak signal-to-noise ratio

GAN: generative adversarial networks

MR: magnetic resonance

PET: positron emission tomography

LiTS: Liver Tumor Segmentation

BN: batch normalization

SR: sparse representation

SRF+: structured random forest

MOS: manipulable object synthesis

DCAN: deep convolutional adversarial networks

MC-GAN: multi-conditional GAN

ME-cGAN: mask embedding in conditional GAN

FCN: fully-connected network

References

1. http://www.vlfeat.org/matconvnet/pretrained/.
2. Baur C, Albarqouni S, Navab N. Generating Highly Realistic Images of Skin Lesions with GANs, International Workshop on Computer-Assisted and Robotic Endoscopy:260–267, 2018.
3. Frid-Adar M, Diamant I, Klang E. GAN-based Synthetic Medical Image Augmentation for Increased CNN Performance in Liver Lesion Classification, Neurocomputing, 321:321–331, 2018.
4. Yang Q, Yan P, Zhang Y. Low Dose CT Image Denoising Using a Generative Adversarial Network with Wasserstein Distance and Perceptual Loss. IEEE Trans Med Imaging. 2018;37(6):1348–57.
5. Sun L, Wang J, Huang Y, Ding X, Greenspan H. J. Paisley. An Adversarial Learning Approach to Medical Image Synthesis for Lesion Detection, arXiv, https://arxiv.org/abs/1810.10850, 2019.
6. Nie D, Trullo R, Lian J. Medical Image Synthesis with Deep Convolutional Adversarial Networks. IEEE Trans Biomed Eng. 2018;65(12):2720–30.
7. Wolterink JM, Dinkla AM, Savenije MHF, Seevinck PR, Berg CAT, I. Išgum. Deep MR to CT Synthesis Using Unpaired Data, International Workshop on Simulation and Synthesis in Medical Imaging:14–23, 2017.

8. Jin CB, Jung W, Joo S. Deep CT to MR Synthesis Using Paired and Unpaired Data. Sensors. 2019;19(10):2361.
9. Costa P, Galdran A, Meyer MI, Abramoff MD, Niemeijer M, Mendonca AM, et al. End-to-End Adversarial Retinal Image Synthesis. IEEE Trans Med Imaging. 2018;37(3):781–91.
10. Isola P, Zhu J, Zhou T. Image-to-Image Translation with Conditional Adversarial Networks, IEEE Conference on Computer Vision and Pattern Recognition:5967–5976, 2017.
11. https://competitions.codalab.org/competitions/17094#learn_the_details-overview.
12. Bilic P, Christ PF, Vorontsov E. G. Chlebbus. The Liver Tumor Segmentation Benchmark (LiTS), arXiv, https://arxiv.org/abs/1901.04056, 2018.
13. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical Image Computing and Computer-Assisted Intervention (MICCAI), 9351:234–241, 2015.
14. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions, International Conference of Learning Representations (ICLR), 2016.
15. Vercauteren T. Diffeomorphic demons: Efficient non-parametric image registration, 45, 1(S61-S72), 2009.
16. Huynh T. Estimating ct image from mri data using structured random forest and auto-context model. IEEE Trans Med Imaging. 2016;35(1):174–83.
17. Siqi Liu E, Gibson S, Grbic Z, Xu AArindraA, Setio, et al. Decompose to manipulate: Manipulable Object Synthesis in 3D Medical Images with Structured Image Decomposition, publish online: arXiv:181201737, 2019.
18. Nie D, Trullo R, Lian J, Wang L. Medical Image Synthesis with Deep Convolutional Adversarial Networks. IEEE Trans Biomed Eng. 2018;65(12):2720–30.
19. Hyojin, Park. YoungJoon Yoo, Nojun Kwak. MC-GAN: Multi-conditional Generative Adversarial Network for Image Synthesis, Publish online: arXiv:180501123, 2018.
20. Ren Y, Zhu Z, Li Y, Lo J. Mask Embedding in conditional GAN for Guided Synthesis of High Resolution Images, Publish online: arXiv:190701710, 2019.
21. Ronneberger O. U-net: convolutional networks for biomedical image segmentation, International Conference of Medical Image Computing and Computer Assisted Intervention:234–241, 2015.

Figures

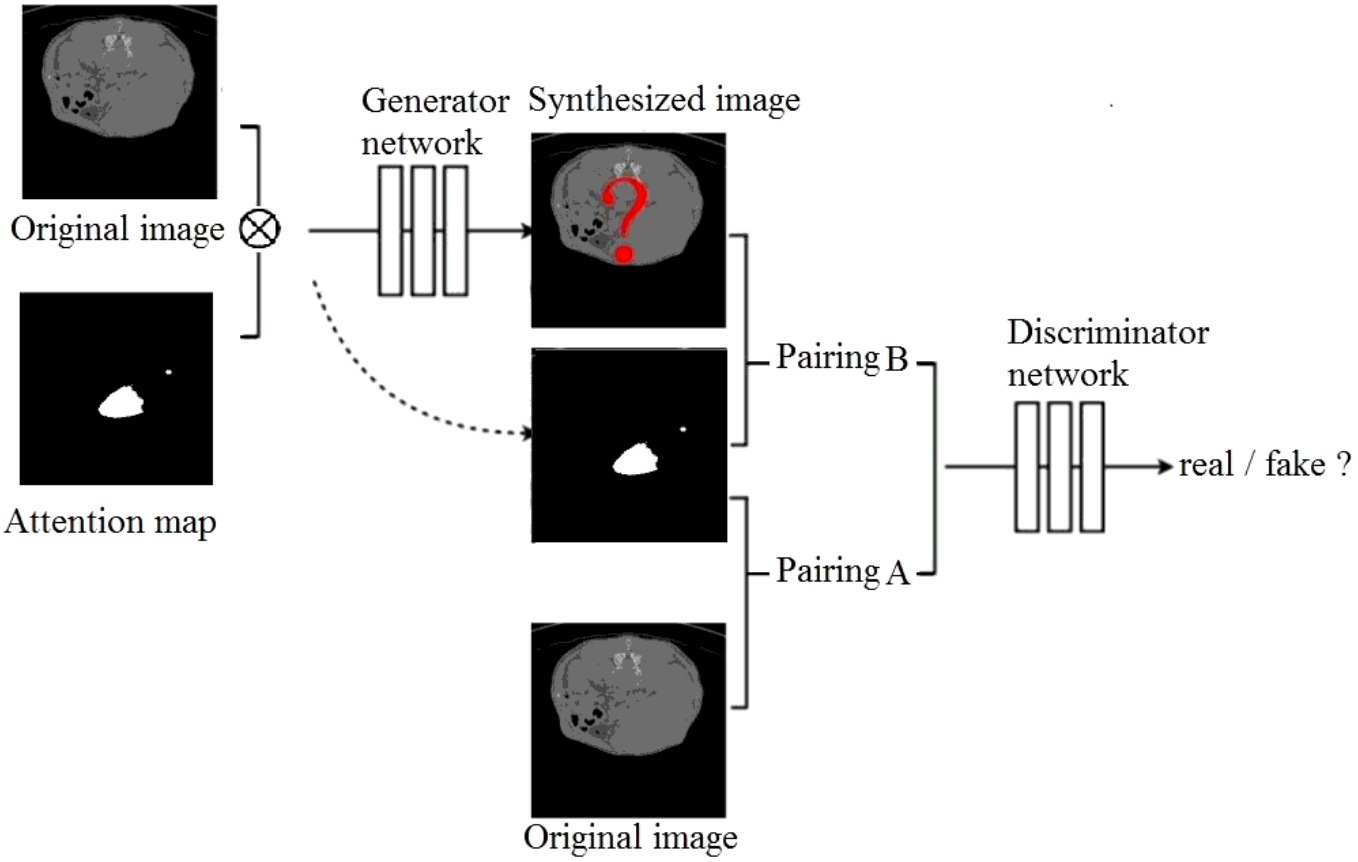


Figure 1

Framework of our model, ⊗ represented the matrix multiplication

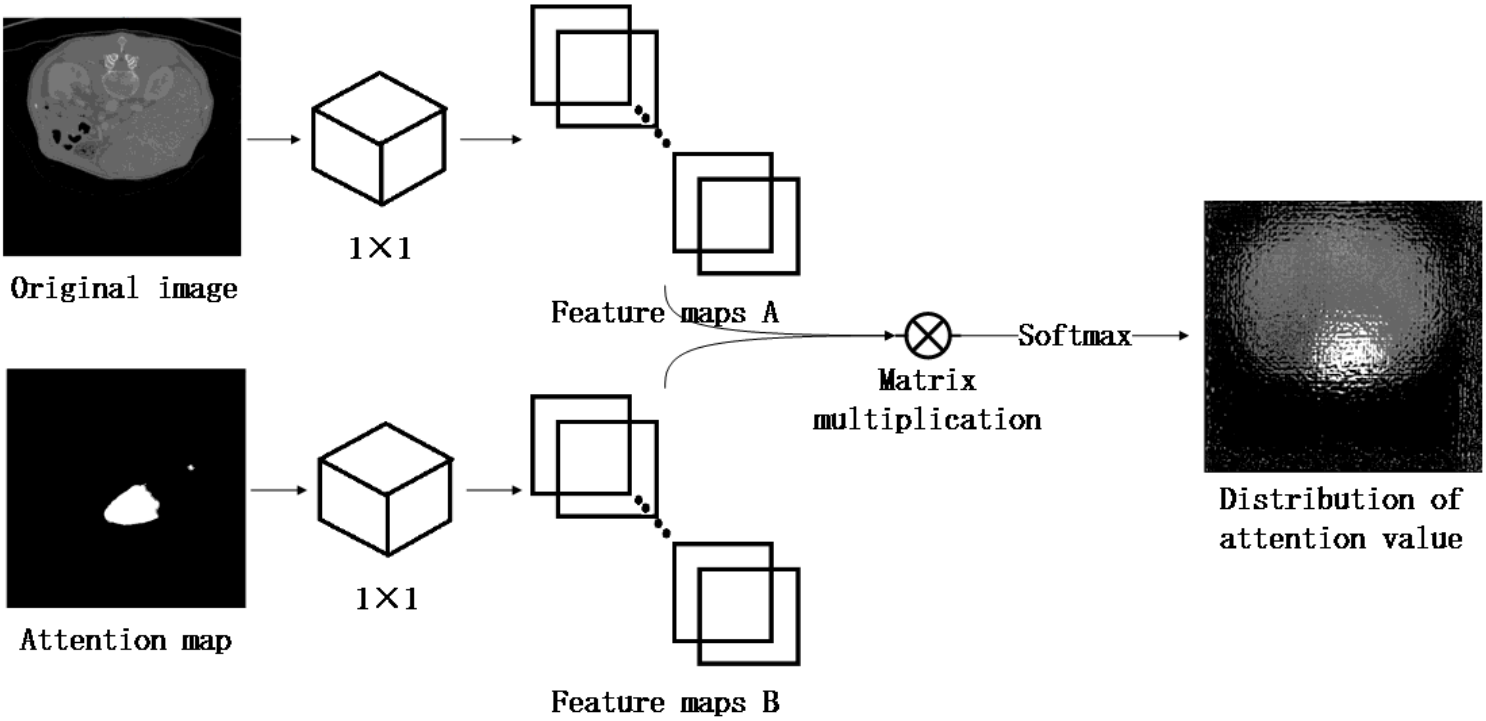


Figure 2

Framework of attention model

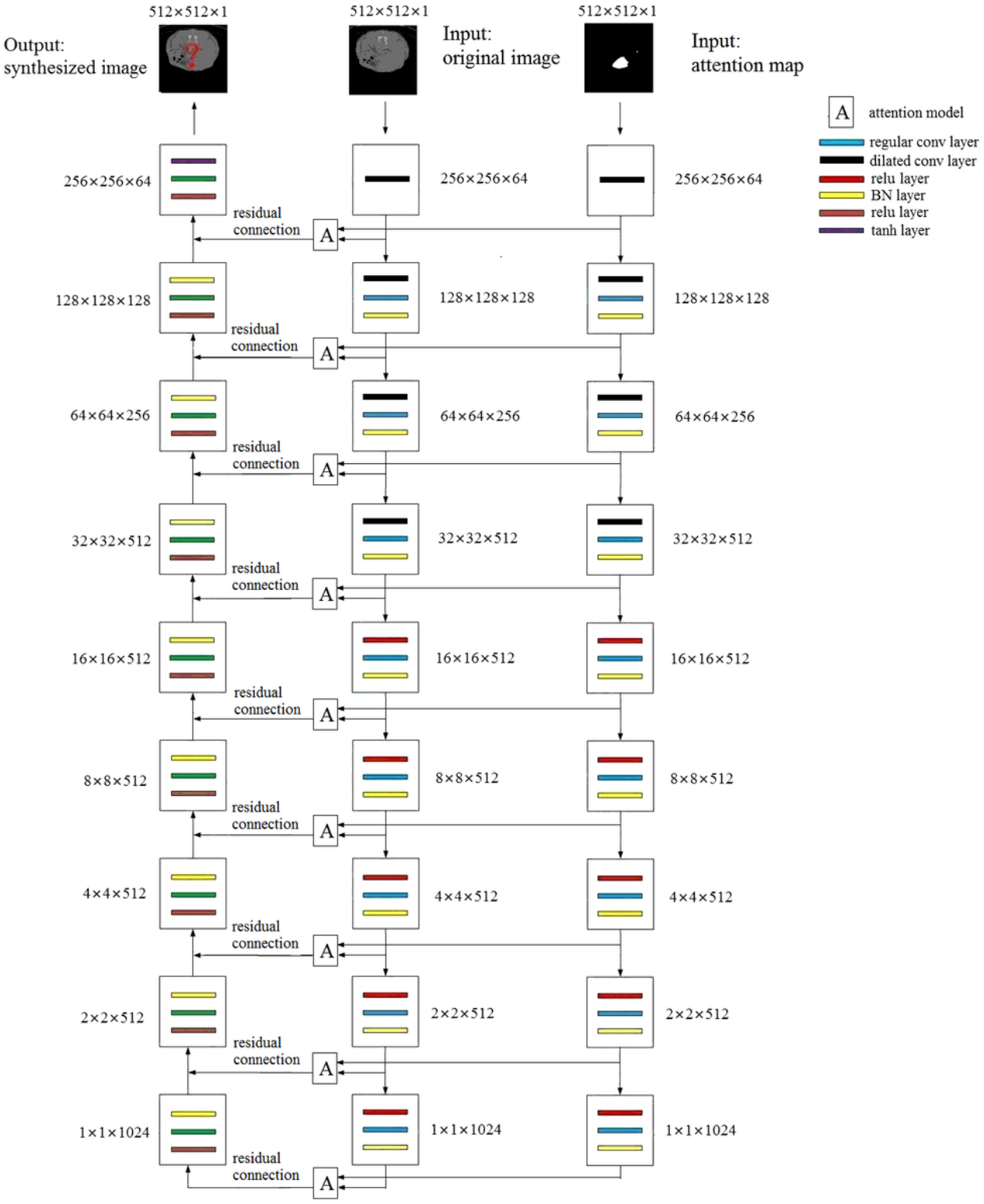


Figure 3

Framework of our generator network

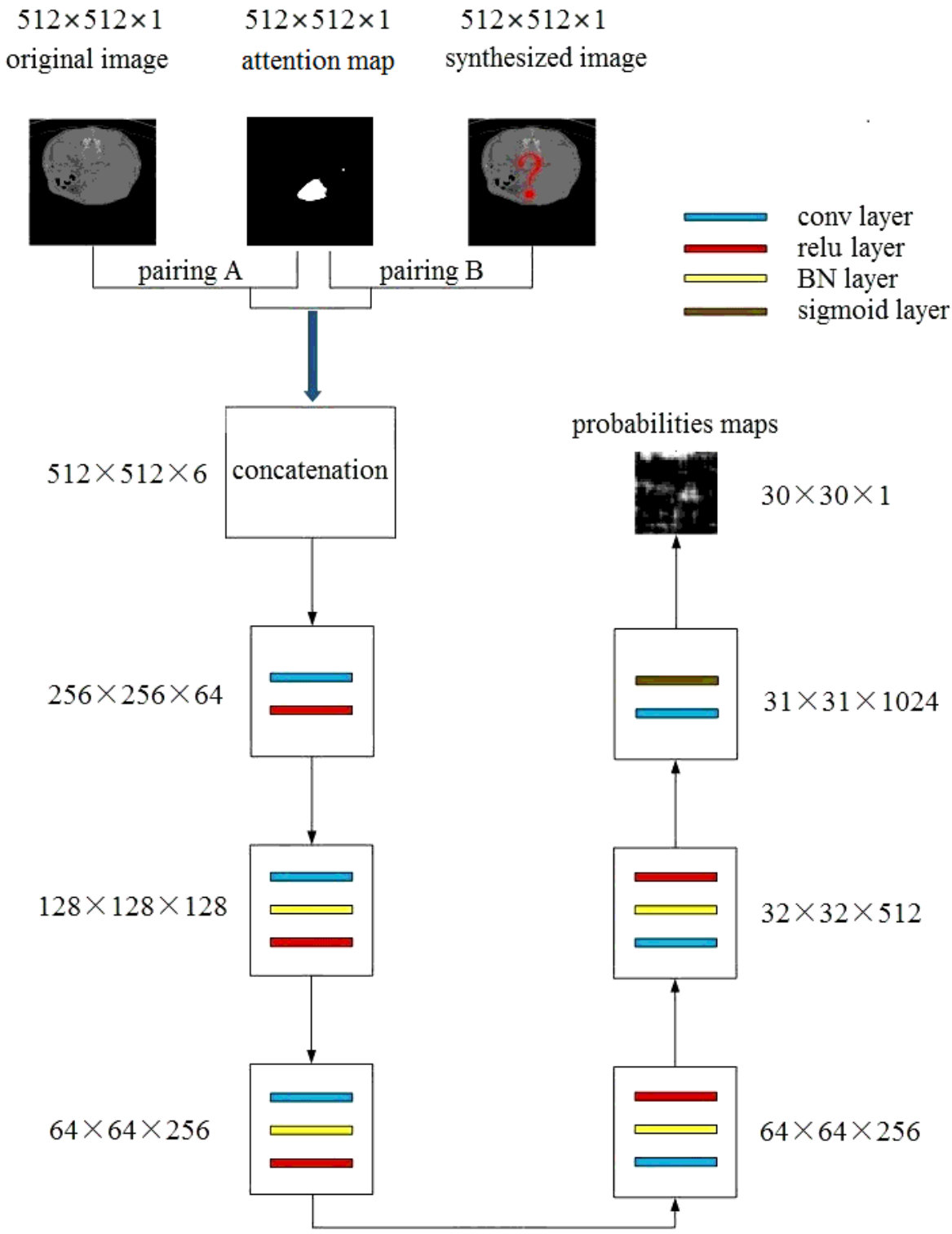


Figure 4

Framework of our discriminator network

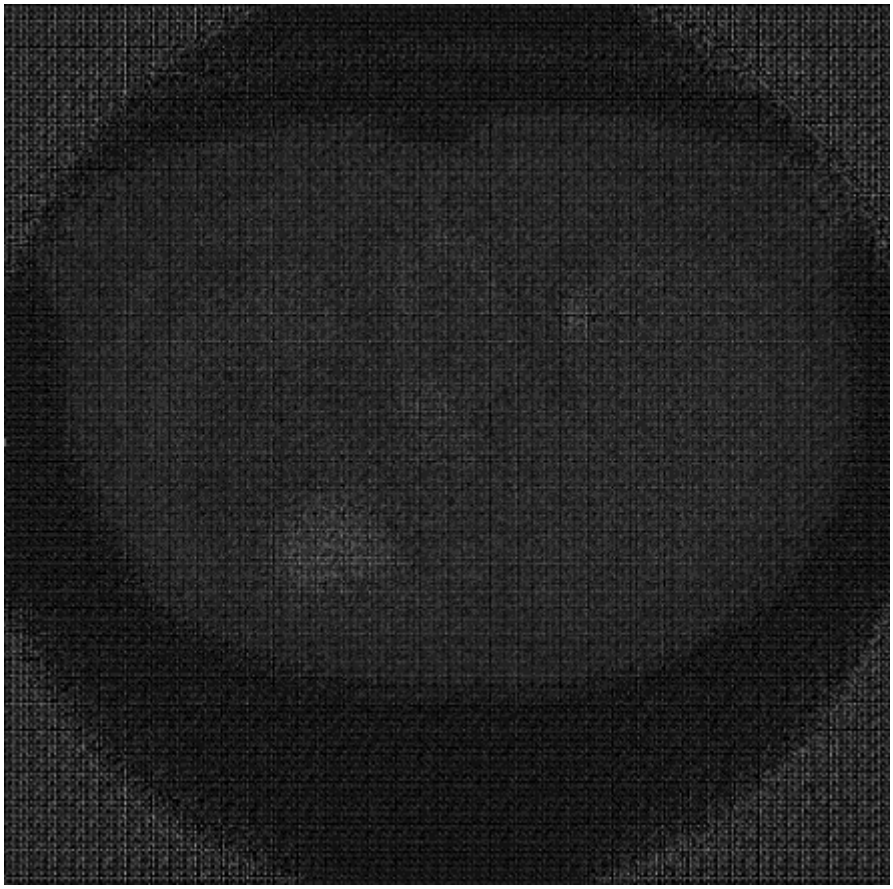


Figure 5

Synthesized image after one iteration of training

Figure 6

Synthesized image after ten iterations of training

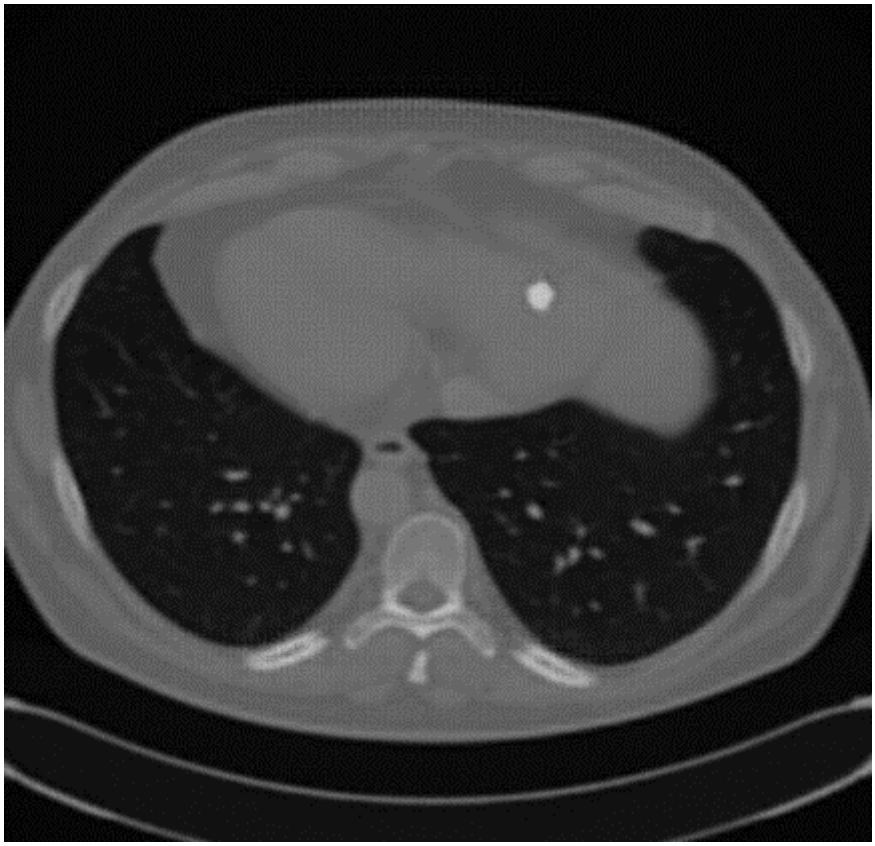
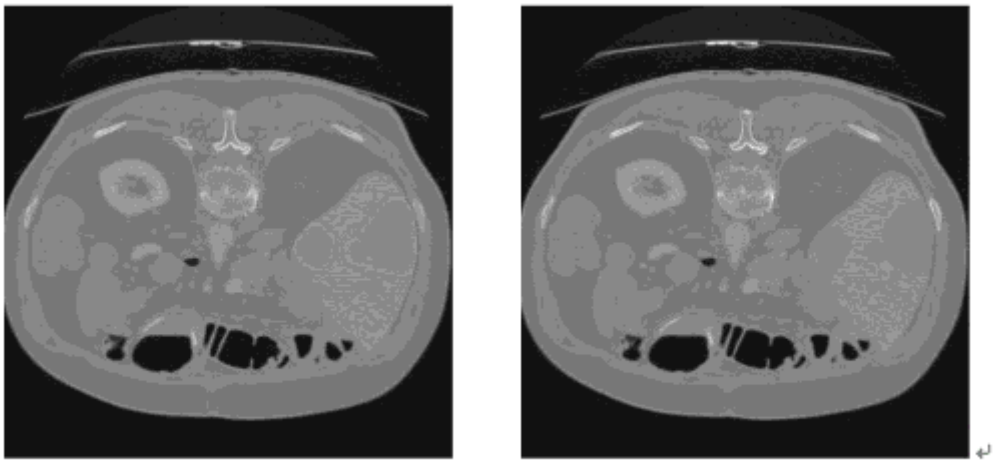


Figure 7

Synthesized image after one hundred iterations of training

Figure 8

Synthesized image after one thousand iterations of training



(a) real image (b) synthesized image

Figure 9

Synthesized image after ten thousands iterations of training

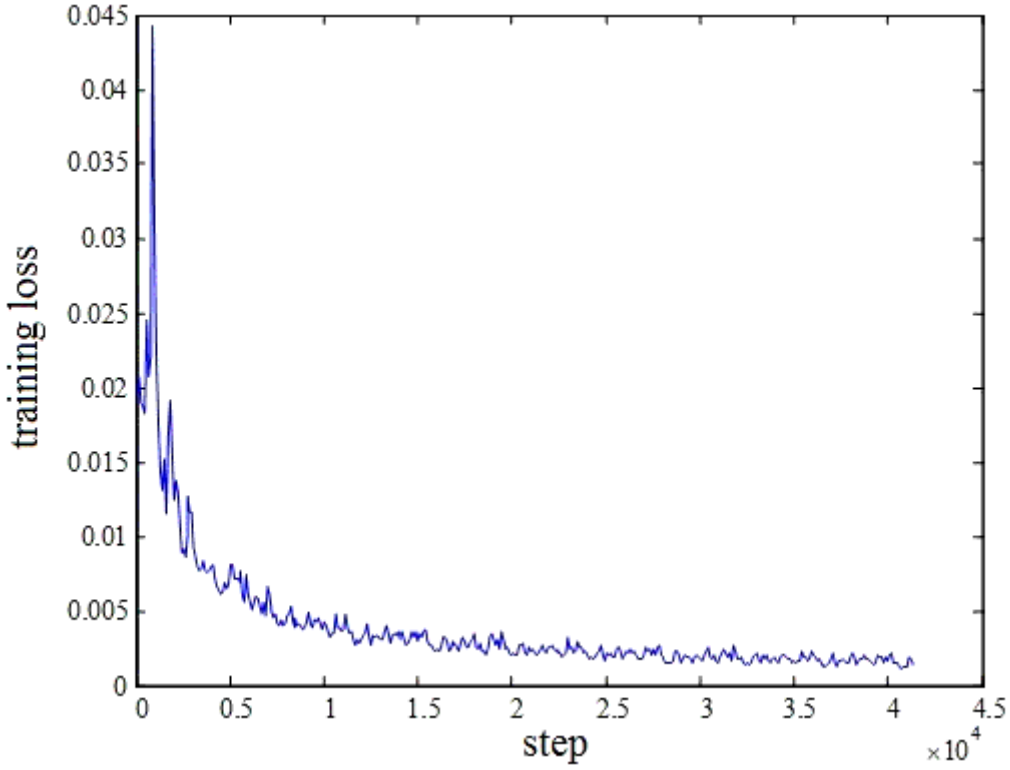


Figure 10

The loss function of the generator network during the training

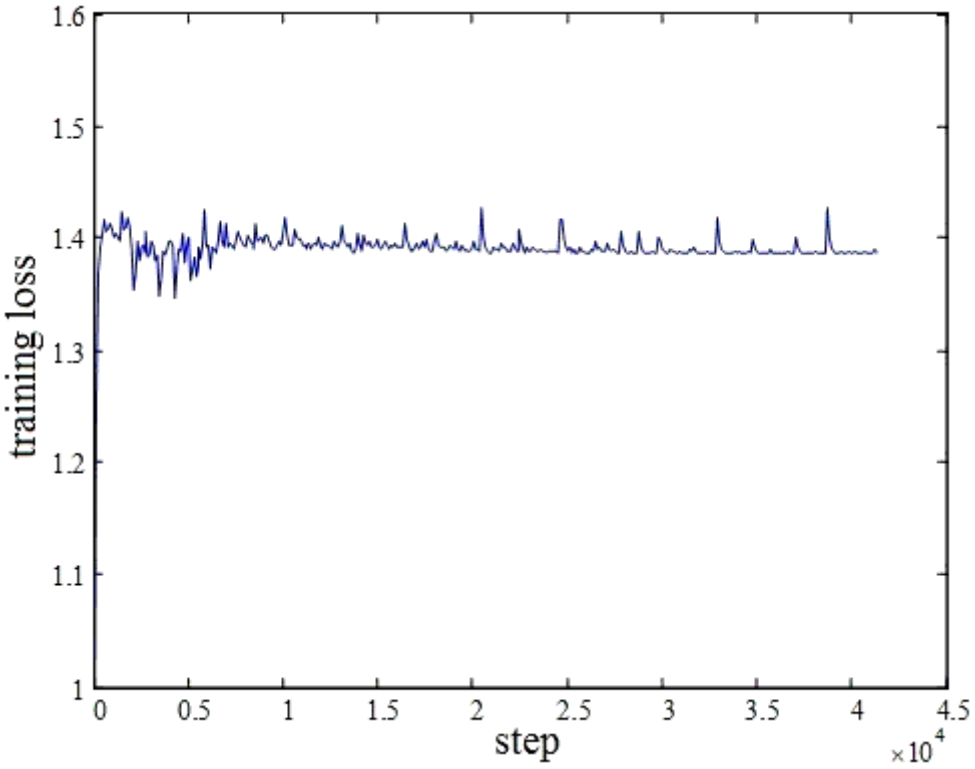


Figure 11

The loss function of the discriminator network during the training

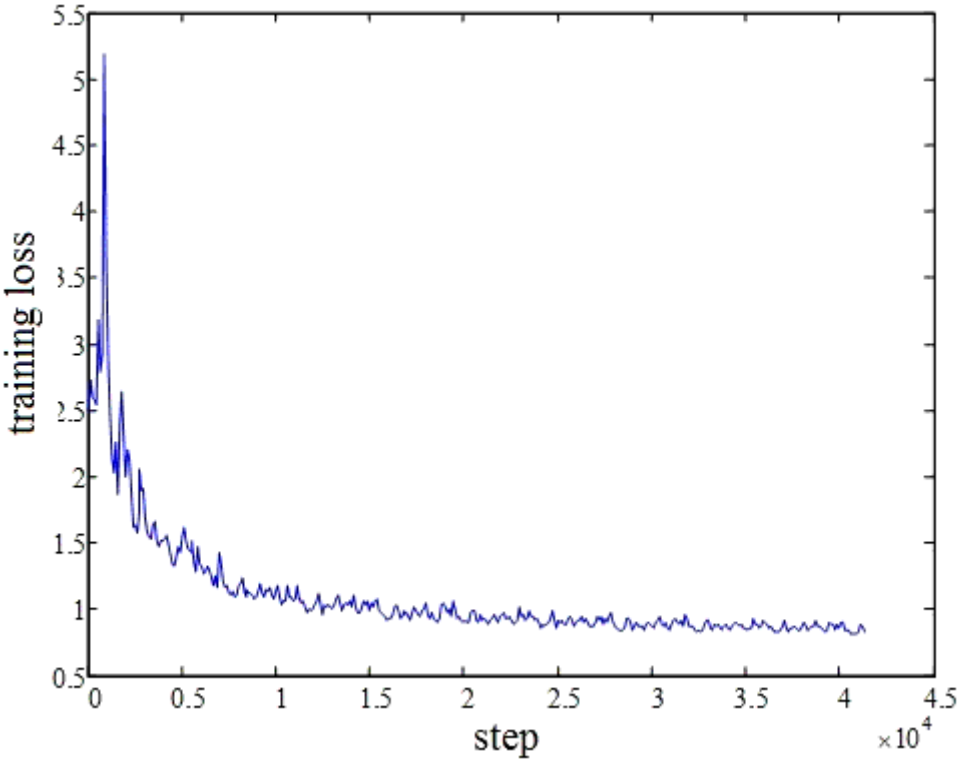


Figure 12

The loss function of the total network during the training

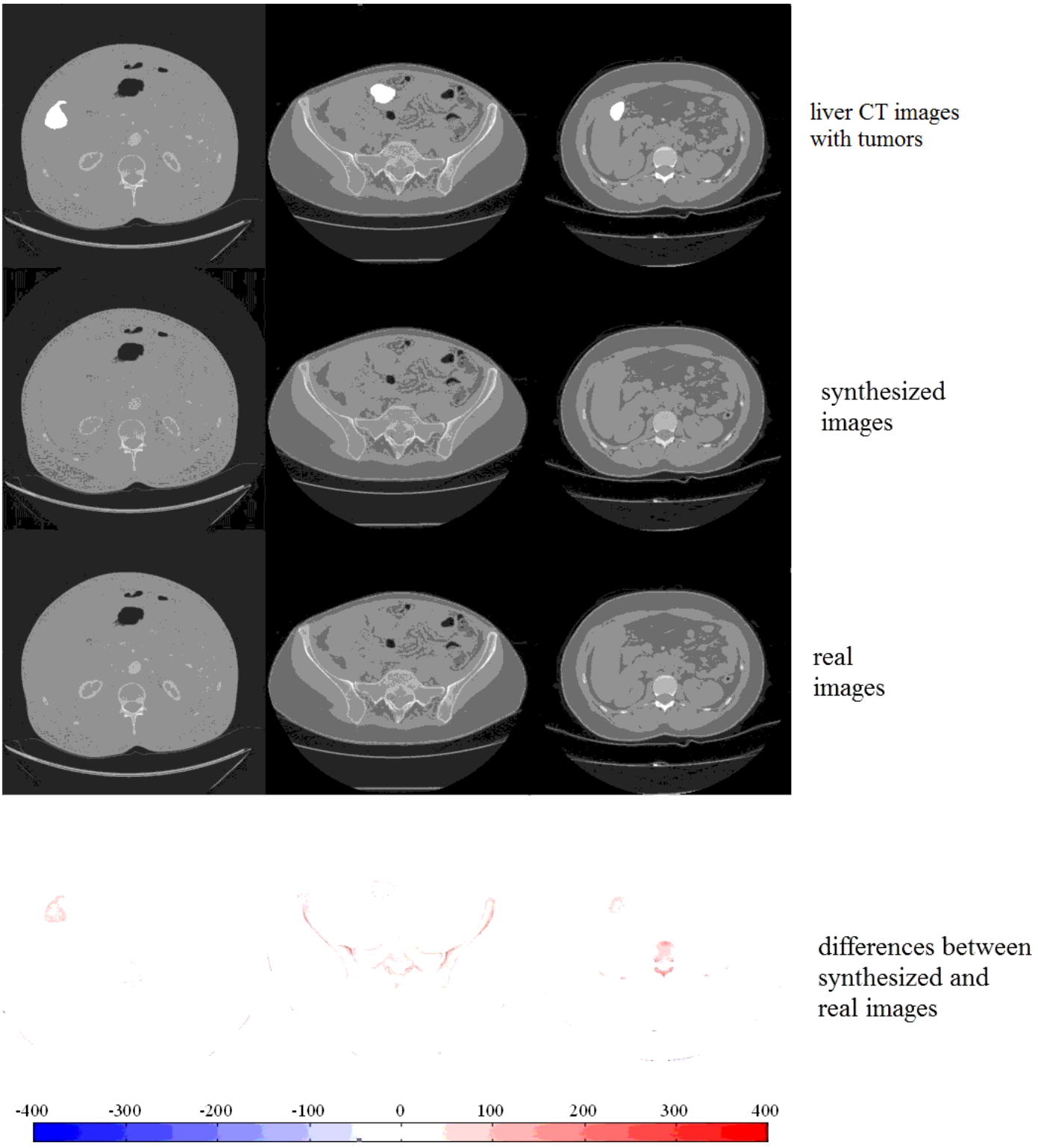


Figure 13

Results of the synthesized images, and the comparison between synthesized images and real images. The pixel values of the fourth rows are weak and low because the differences between real images and synthesized images were very small.

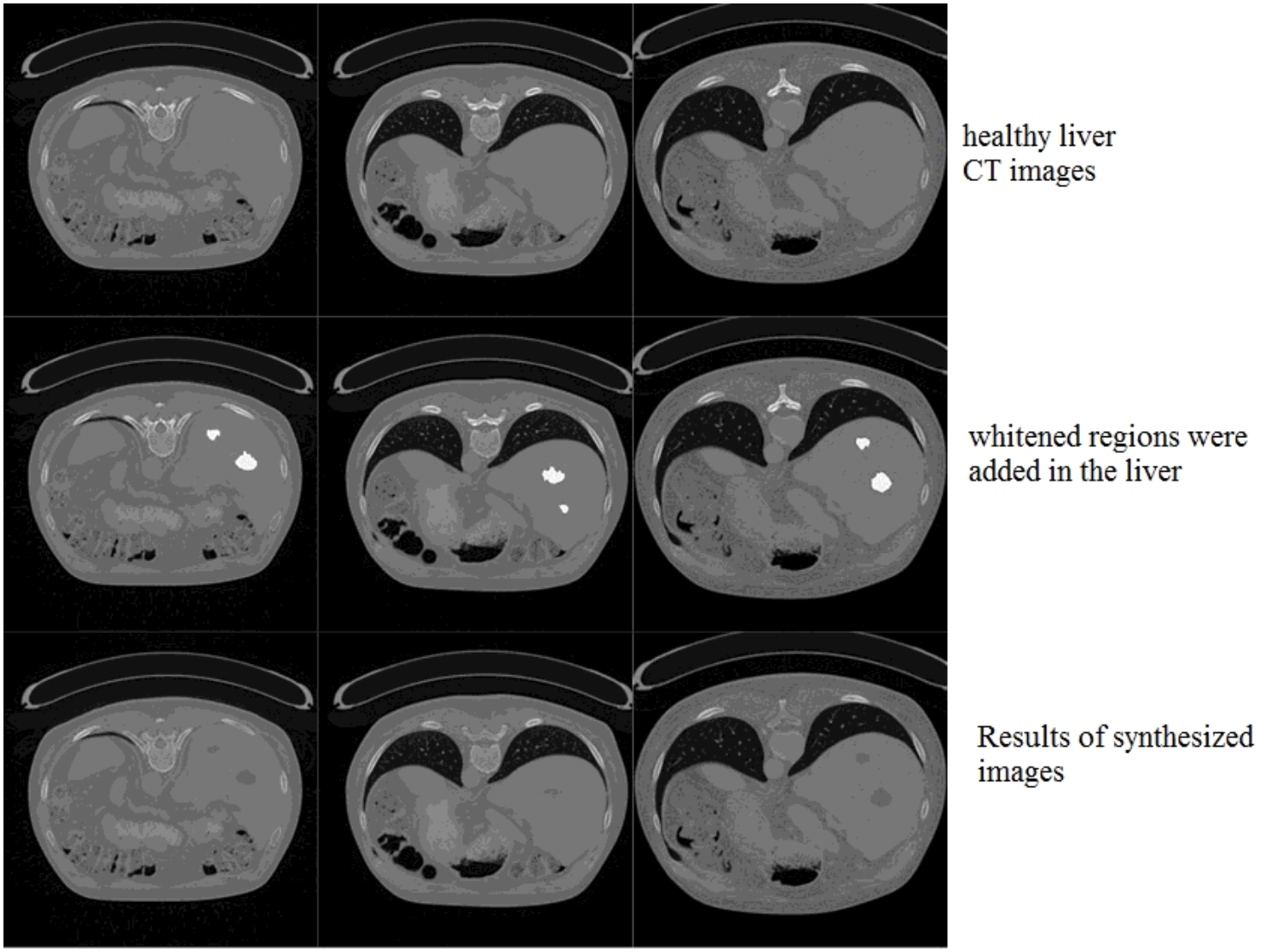


Figure 14

Adding tumor regions on the healthy liver CT images, and synthesizing diseased liver CT images using our method