

Received June 13, 2020, accepted June 30, 2020, date of publication July 10, 2020, date of current version July 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3008525

Dual-Enhanced Registration for Field of View Ultrasound Sonography

ZIJIA FAN¹, ZHONGYANG WANG¹, JUNCHANG XIN^{2,3}, (Member, IEEE),
ZHIQIONG WANG^{1,4,5}, (Member, IEEE), LU LIU¹, XIA ZHANG⁴, AND JIREN LIU⁴

¹College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110169, China

²School of Computer Science and Engineering, Northeastern University, Shenyang 110169, China

³Key Laboratory of Big Data Management and Analytics (Liaoning Province), Northeastern University, Shenyang 110169, China

⁴Institute of Intelligent Healthcare Technology, Neusoft Corporation, Ltd., Shenyang 110179, China

⁵Acoustics Science and Technology Laboratory, Harbin Engineering University, Harbin 150001, China

Corresponding author: Zhiqiong Wang (wangzq@bmie.neu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61472069 and Grant 61402089, in part by the China Postdoctoral Science Foundation under Grant 2019T120216 and Grant 2018M641705, in part by the Fundamental Research Funds for the Central Universities under Grant N2019007, Grant N180101028, Grant N180408019, and Grant N2024005-2, and in part by the CETC Joint Fund, the fund of Acoustics Science and Technology Laboratory.

ABSTRACT Extended Field of View Ultrasound Sonography (EFOV-US) uses the existing ultrasound images for image stitching, so as to display the shape and scope of organ occupation and the relationship with surrounding tissues comprehensively. However, there are still some problems in Extended Field of View Ultrasound Sonography, such as matching error and unstable quality of image stitching. In view of these problems, we propose Dual-enhanced EFOV-US method that overcomes the limitation and produces higher quality results. Firstly, the gray enhancement method is used to improve the image contrast and reduce the noise interference. Then the super-resolution method based on the generative adversarial network is used to improve the resolution of the ultrasonic image further and increase the number of feature point matching between stitching images. The high quality ultrasound wide-range image is gotten by stitching and fusing the double enhanced image. The experimental results show that the proposed method is effective and practical.


INDEX TERMS Extended field of view ultrasound sonography, gray enhancement, generative adversarial network, super-resolution, image registration.

I. INTRODUCTION

Extended Field of View Ultrasound Sonography (EFOV-US) [1], makes up for the limitation of scanning field of conventional real-time ultrasound diagnostic equipment [2]. In the examination of larger organs and lesion areas, the scope and morphology of lesion can be visualized, and the size of lesion areas can be obtained quantitatively [3], [4]. At present, some wide-range imaging methods have been implemented. Zheng *et al.* [5] select fewer blocks which are regarded as the most valid blocks based on the importance of image content in each block to reduce computational complexity. Zheng *et al.* [6] estimate the moving speed of the probe and ignore redundant image data by processed a smaller number of frames according to a frame interval, producing EFOV images in real-time. Yoo *et al.* [7] estimate

motion and position of acquired images accurately with respect to each other in an EFOV display format.

In the traditional ultrasonic extended field of view imaging technology, it is an effective method to obtain wide range images by image stitching using feature point matching [8]. The feature points in image stitching come from the processed image obtained by certain processing of the original image, and the matching of the original ultrasonic image is realized by matching the feature points and their matching relationship [9]. Based on feature points, processed image matching relationship is calculated to match original ultrasound image. The more feature points, the more accurate the image matching will be, and the better result will be [10]. However, the process of feature point extraction is affected by noise and the lack of image clarity. The stitching is often limited by the lack of clarity of the original image and the lack of accuracy of image matching, resulting in poor results [11], [12]. So we propose Dual-enhanced EFOV-US method which can reduce noise interference and enhance the

The associate editor coordinating the review of this manuscript and approving it for publication was Kumaradevan Punithakumar .

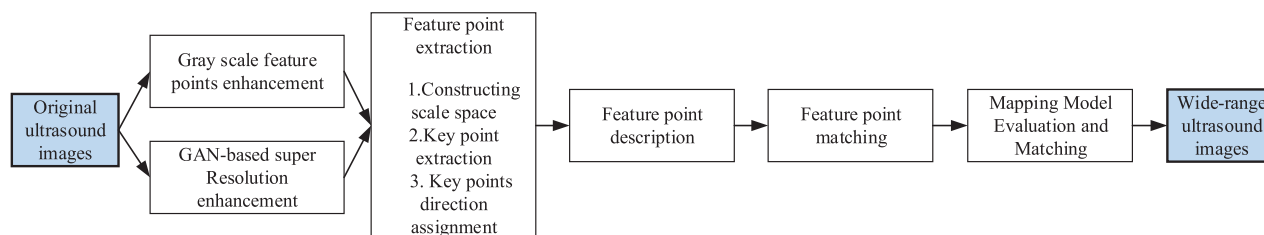


FIGURE 1. The process of dual-enhanced registration for field of view ultrasound images.

resolution of the image, thus improving the quality of the ultrasound wide-range image.

First of all, due to the limitation of ultrasonic image acquisition process [13], [14], there is a lot of noise interference in the image and insufficient image contrast, resulting in a large number of feature points being masked, which affects the result of feature point extraction. Among the methods of removing image noise, the denoising method based on gray transformation is widely used in ultrasound image denoising because of its good effect on ultrasound images [15], [16]. Gray enhancement updates the gray value of each pixel in the original image according to a certain mathematical transformation formula, so as to display more details, improve image contrast, highlight the characteristics of interest, and obtain images with less interference. Therefore, the ultrasonic image is processed by gray enhancement firstly.

Image resolution, as an important index of image evaluation [17], affects the extraction effect of feature points directly. Low resolution image will lead to some feature points which can not be displayed. In recent years, image super-resolution methods based on deep learning have been rapidly developed. Generative adversarial network (GAN) [18], as a rising star in deep learning, has been widely used in various scenarios. The main body of GAN is composed of two convolutional neural networks: generator and discriminator. The generator outputs a high-resolution image given a low resolution image, and the discriminator discriminate real and fake during training. Through training generator and discriminator against, the ability of generation and discrimination gradually improve until the training completed in this way, and generate an ultrasound image that is difficult to discriminate real and fake finally, so as to improve the image resolution greatly and make the details of the reconstructed image better. It is also widely used in medical images [19]. Therefore, we use the super-resolution method based on GAN to enhance the image so as to increase the number of feature points further.

As one of the most important processes, image registration affects the quality of ultrasound panoramic image directly. The registration method based on image features is widely used in panoramic stitching [20]. In many image registration algorithms, the registration method based on feature points which uses the local features of the image for matching, has the advantages of simple calculation, high robustness, stable results, etc. And it is widely used in ultrasound images [21].

Therefore, we use the registration method based on feature points for image registration.

As follows, the contribution of this work are:

- 1) The Dual-enhanced EFOV-US method which can improve the number of feature points effectively.
- 2) The gray transform enhancement method to reduce the noise interference in the ultrasonic image and to solve the problem that the feature is not obvious due to the weak contrast of the image.
- 3) The super-resolution of ultrasonic image realized by using generative adversarial network to enhance the image display effect and achieve more accurate calculation of feature points.

The rest of this paper is organized as follows. section II introduce the enhancement algorithm based on gray transformation, the image super-resolution algorithm based on generative adversarial network and the process of image registration. The effect of the method proposed in this paper is verified through several groups of experiments. in section III. Finally section IV draws conclusions.

II. METHOD

The dual-enhanced registration method for field of view ultrasound sonography includes three parts: Image feature point enhancement preprocessing, image super-resolution reconstruction, and image registration and fusion. The specific process is shown in Figure 1.

First, we use an ultrasound image enhancement method based on grayscale transformation to denoise the ultrasound image sequence. Then we use the algorithm based on GAN to process the image with super resolution. Next we select the scale-invariant feature transform (SIFT) [22], [23] method to perform feature point extraction and matching operations, then we perform mapping model evaluation to find the spatial transformation relationship between reference images. Finally, we perform geometric transformation through the spatial transformation relationship and the super-resolution processed image to form a wide-range image by image fusion and stitching. Section II-A introduces the method of feature point enhancement based on gray value transform. Section II-B introduces image super-resolution enhancement methods based on generative adversarial networks. Section II-C describes the specific method of matching and fusion of feature points between images.

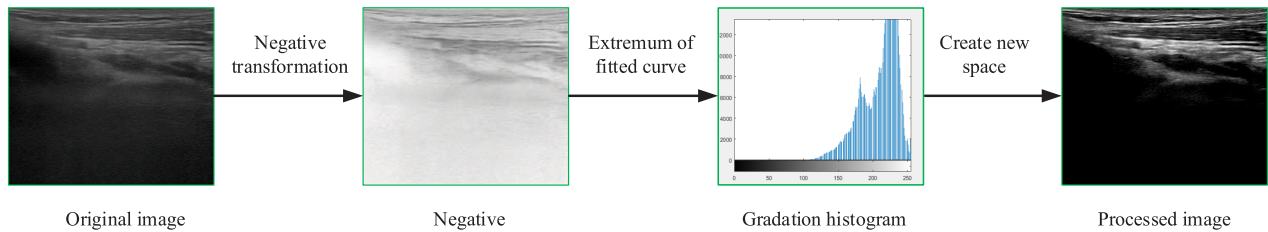


FIGURE 2. Feature point enhancement algorithm based on gray-scale transform.

A. ENHANCEMENT OF GRAY SCALE FEATURE POINTS

Due to its principle characteristics, the ultrasound image content is not obvious and has a lot of speckle-noise, which leads to the formation of interference echo signals. To reduce noise interference, Ayres *et al.* [24] makes improvements based on location information to obtain choroidal tumors images. Huang *et al.* [25] proposed a 2.5D ultrasound wide-field imaging technology, successfully applying EFOV-US technology to 2.5D ultrasound equipment.

In this paper, we propose the enhancement algorithm of ultrasonic image feature points based on gray-scale transformation to effectively solve the above problems from the perspective of retaining effective image information. The main purpose of this method is to delete useless information for feature extraction and improve the effectiveness of subsequent feature extraction and matching. The specific process is as follows: First, we acquire ultrasound images and then convert acquired image into grayscale images, using the grayscale inversion transform. Inverting the image in the gray value range $[0, L - 1]$ requires using $L - 1$ minus the image gray value and the formula is as follows.

$$s = L - 1 - r \quad (1)$$

where L is the gray range of the image, r is the gray value of the image, and s is the gray level after the inversion. Then we turn the image to negative for subsequent operations. Next we build a grayscale histogram for the image, fit a two-dimensional quadratic curve based on the specific grayscale histogram distribution, find the minimum point where the abscissa x is closest to 255, take its abscissa xt , and set it as threshold. Then we map pixels with gray range in $[0, xt]$ to $[0, 255]$ to remove unwanted noise for feature extraction. Next we perform the grayscale inversion transformation on the negative film, and restore the denoised negative film image to a normal image. Then we normalize the grayscale histogram to enhance the contrast of the image. Finally, the processed image is output. The detailed process is shown in Figure 2.

B. IMAGE SUPER-RESOLUTION BASED ON GENERATIVE ADVERSARIAL NETWORK

Image super-resolution (SR) aims to restore a high-resolution image (HR) from a low-resolution image (LR) or image sequence [26]. Compared with traditional super-resolution method based on interpolation or reconstruction, the method based on deep learning has excellent computing power

and good processing effect [27]. Since Dong *et al.* [28] propose the pioneering work of SRCNN, methods based on deep convolutional neural network (CNN) have been developed rapidly and gradually become mainstream. Choi and Kim [29] propose a new network selection unit called SelNet, selecting values from the feature map as channels between convolutional layers to improve performance and reduce computational consumption. Zhang *et al.* [30] propose a model based on simulated real image degradation (SRMD), which uses stitched low-resolution images and their degraded images for training to simulate real blurred images and obtains good results. The Laplacian pyramid super-resolution network (Lap-SRN) constructed by Lai *et al.* [31], use a pyramid structure to convert the residual image of each sub-network to the input image to generate the final image. Generative adversarial network (GAN) as a rising star in deep learning algorithms, operating through the confrontation of generators and discriminators, greatly improves training speed and generation details, and thus is widely used in images [32], [33]. Although these GAN based methods work well on natural images, they are limited in medical images. GAN based models pre-trained on natural images may synthesise unrealistic patterns in medical images which could affect the clinical interpretation and diagnosis. Therefore, we make improvements in this part, which is a big difference between Wang *et al.* [34] and our method. We use grayscale ultrasound images to train a pre-trained model. A pre-trained model close to the reality can improve super-resolution results. The image super-resolution method we propose in this paper mainly includes dataset enhancement, network construction and network training. First, we organize the collected ultrasound images, uniform image size and use downsampling degradation methods to build a near-realistic low-resolution image (LR) dataset. The vanilla GAN architecture may suffer from unstable training and collapse mode that can also affect results. Wasserstein GAN (WGAN) [35], [36] are proposed to replace the noncontinuous divergence with the Wasserstein-distance to overcome the problems of vanilla gan in convergence and training instability, which is another improvement over Wang *et al.* [34] method. In this work, G aims to generate SR images as realistic as possible to fool D and D aims to distinguish the SR images from real HR images which can be described as:

$$\hat{G}, \hat{D} = \min_G \max_{D \in \mathcal{D}} \mathbb{E}_{I_{HR}} [D(I_{HR})] - \mathbb{E}_{I_{LR}} [D(G(I_{LR}))] \quad (2)$$

where I_{LR} and I_{HR} are full size LR and HR images, and \mathbb{E} is the expectation of the D 's positive outputs (i.e. input is HR ground truth). We build a GAN-based convolutional neural network, sum the adversarial loss, pixel loss (MSE loss) and perceptual loss [37], [38] for calculation of the loss function. The main structure is shown in Figure 3.

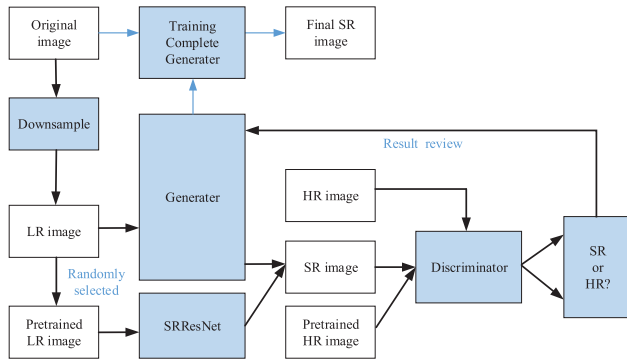


FIGURE 3. Ultrasound image super-resolution enhancement based on GAN.

The generator adversarial network (GAN) includes a generator G , a discriminator D and some assisted network. Pre-trained HR images are randomly selected from HR images while pre-trained LR images are selected from Multi spectral image database and degenerated. As a part of improvement [32], [39], we train D in advance to ensure the discriminating ability and training speed to prevent the generation gap caused by low discrimination ability. During the training, we put LR images training set into G for training, and put pre-trained LR images into SRResNet to avoid the mode collapse caused by the difference between the abilities of generators G and D in the early stage of training. We get the SR images of the two network respectively. We mix generated images with some HR images from training and put them into D for discrimination, and the judgment result is returned to G . Under continuous adversarial training in this way, SR images generated by the generator become more similar to the HR images, and the discriminator's ability to distinguish whether the image is SR or HR is stronger. The training is completed until the discriminator cannot judge the authenticity.

The specific neural network structure is shown in Figure 4. We employ the basic architecture of SRResNet [32], where most computation is done in the LR feature space. Specifically, we use two convolutional layers with small 3×3 kernels and 64 feature maps followed. We use parametricReLU [40] as the activation function, which introduces a learnable parameter to help adaptively learn some negative coefficient. Different from the generator network, the BN layers in residual block in SRResnet are removed to increase performance and reduce computational complexity. The SRResnet helps generator quickly keep up with the discriminator and prevent mode collapse and excessive training gap. The discriminator network contains eight convolutional layers with an increasing number of 3×3 filter kernels, from

64 to 512 kernels and finally returns a probability for sample classification.

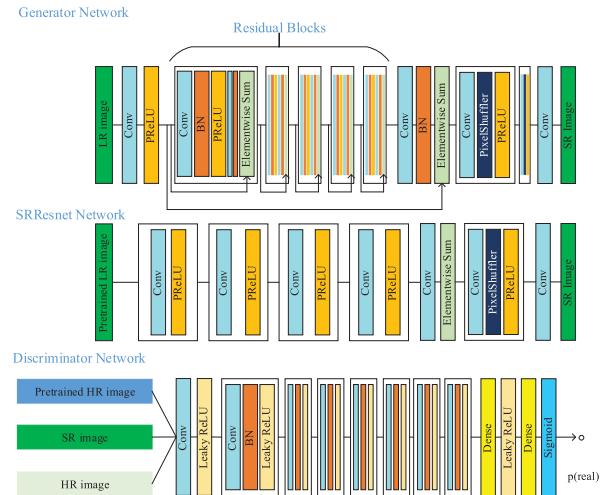


FIGURE 4. Generator, SRResnet, discriminator network.

C. IMAGE REGISTRATION

Image registration is a method of mapping one image to another by looking for a spatial transformation relationship, corresponding the points in the same position under the same spatial coordinate system in two images to achieve information fusion. Image registration is the core part of EFOV-US technology and the processing result directly affects the quality of the final output wide-range ultrasound image [41]. In this paper, we use feature point-based method, Scale-invariant feature transform (SIFT), for feature point extraction. The matching process includes three important steps: feature extraction, feature description, and feature matching and fusion. As shown in Table 1, according to the verification of the effects of Li [41], Routray *et al.* [42], and Juan and Gwon [43] on multiple image matching methods, SIFT has the largest number of successful matches both in scale-invariant and rotation-invariant. The stable processing of the gray-scale image and the accurate extraction of feature points by SIFT algorithm are suitable for our image registration.

1) FEATURE POINT EXTRACTION

Feature point extraction using SIFT is specifically divided into three steps: constructing scale space, spatial extreme point detection, and key point positioning and direction assignment.

a: CONSTRUCTING SCALE SPACE

We convolve the original image with a two-dimensional Gaussian function at different scales can generate multiscale Gaussian space, and the mathematical formula is expressed as follows.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (3)$$

TABLE 1. Correct matches of feature detection algorithms.

Method	Harris	FAST	BRISK	MSER	SIFT	SURF	ORB
Correct matches (scale invariance)	13	0	5	10	121	62	16
Correct matches (rotation invariance)	28	20	13	27	279	81	40

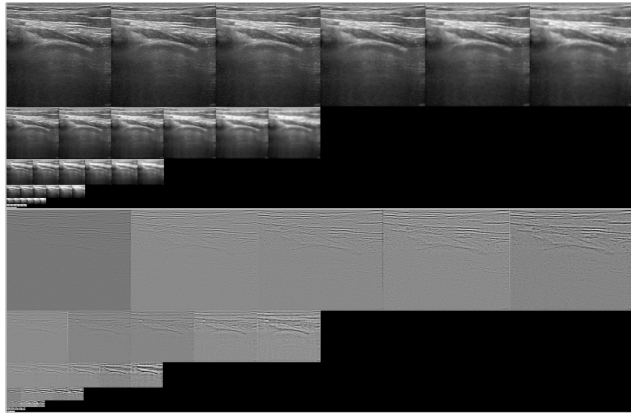


FIGURE 5. An example of Gauss pyramid and Gauss difference pyramid.

where $L(x, y, \sigma)$ represents multiscale Gaussian space, $G(x, y, \sigma)$ represents Gaussian kernel function, $I(x, y)$ represents image matrix, $*$ represents convolution operator, and σ represents scale space factor.

First, we take the input image for downsampling and Gaussian blur at different scales and establish multiple sets of multi-scale space sequences to form the image Gaussian pyramid. Then we subtract adjacent images in each set of scale-space sequences, forming a difference of Gaussian pyramid(DOG) [44].

The mathematical expression is as follows.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (4)$$

Which leads to:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (5)$$

b: KEY POINT EXTRACTION

key point is the local extreme point of DOG. To form a local three-dimensional space, we compare the pixel at the core position with 8 neighboring points of same surrounding scale and 9*2 different scale points in upper and lower neighboring points. Then we set the maximum or minimum point in each local stereo space as key point, fit DOG function in scale space and remove key points with poor stability.

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (6)$$

$X = (x, y, \sigma)^T$ and the offset of extreme points is shown as follows.

$$\hat{X} = -\frac{\partial^2 D}{\partial X^2} \frac{\partial D}{\partial X} \quad (7)$$

$\hat{X} = (x, y, \sigma)^T$ represents the offset from the interpolation center. When the offset is greater than 0.5, which means that the interpolation center has shifted, we need to change the current keypoint position and find the principal curvature through Hessian matrix to eliminate unstable edge response points.

c: KEY POINTS DIRECTION ASSIGNMENT

In Gaussian pyramid image, we calculate the gradient modulus and direction of all pixels in a circular area with a radius of 3σ around each key point, use which as a reference to assign a reference direction for each key point. The magnitude and direction of the gradient are calculated respectively, using formula:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (8)$$

$$\theta(x, y) = \tanh^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (9)$$

$\theta(x, y)$ represents the gradient direction angle of the feature point (x, y) , $m(x, y)$ represents the gradient modulus of the feature point (x, y) , and $L(x, y)$ is the pixel value of the feature point (x, y) in the Gaussian pyramid. We set feature point as the center, and keep the main direction at zero degrees. Then we use the gradient histogram to count the above information. We divide the gradient direction into 36 columns, and each column is divided by a span of 10 angles as abscissa, gradient magnitude as ordinate. The largest gradient amplitude is taken as the main direction.

2) FEATURE POINT DESCRIPTION

To perform feature matching, a feature vector needs to be defined as the feature descriptor of each feature point as a unique "label" for each feature point. We choose Gradient location and orientation histogram (GLOH) [45], [46] as feature descriptor, as shown in the Figure 6. The experimental results of Mikolajczyk and Schmid [47] prove the advantages of choosing GLOH: GLOH has a higher successful matching rate, and compared with PCA-SIFT and standard SIFT descriptors, GLOH has better results on edge features. GLOH also performs better with smooth images and blurred images, and has less dependence on sample ultrasound images. GLOH descriptors use a logarithmic polar hierarchy to replace the 4-quadrant traditional descriptor proposed by Lowe [48]. Take a radius of 6, 11, 15 in space, and divide it into 8 intervals in angle (except for the middle area),

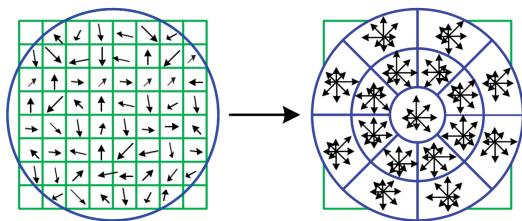


FIGURE 6. GLOH features descriptor.

and we can get 136 (17 * 8) dimensional vector as the final feature vector.

3) FEATURE POINT MATCHING

We calculate the Euclidean distance between the feature vectors corresponding to each pair of feature points, using which to determine the correspondence. The closer the Euclidean distance between two point feature vectors is, the greater the chance of successful matching is. The formula for calculating the Euclidean distance in n-dimensional space is as follows.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{10}$$

The specific method of feature point matching is as follows. First, we select the feature point B in the image to be matched. We find the feature point with the smallest Euclidean distance from B in the reference image and record it as B' . Then we find the nearest feature point C and the second nearest point D to point B' in the reference image. Next we calculate the ratio of the distance between points B and C to the distance between points C and D . The specific formula is as follows.

$$\frac{d(B, C)}{d(B, D)} < Threshold \tag{11}$$

When the ratio is less than the given threshold (According to the experimental conclusion of Lowe [48] and the comparison of different thresholds we make in Table 2), where $Threshold = 0.9$, the match succeeds, otherwise it fails. An example of the feature point matching effect is shown in Figure 7.

TABLE 2. Threshold value selection.

Threshold	0.2	0.4	0.6	0.8	0.9	1.0
Correct matches	34	68	73	98	121	101

4) MAPPING MODEL EVALUATION AND MATCHING

After feature matching, we select random sample consensus (RANSAC) [49] to evaluate the mapping model, and calculate the spatial geometric transformation relationship between the image to be matched and the reference image. We get the transformation relationship between adjacent frames in two steps. First, we use a recursive formula to find the transformation relationship between each frame image and the reference

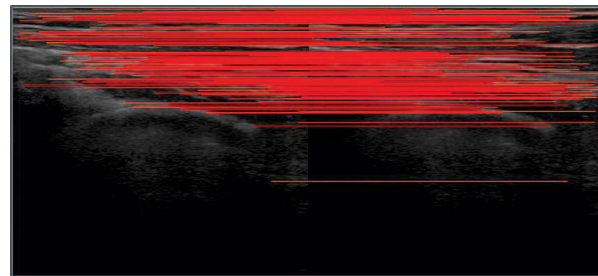


FIGURE 7. Feature point matching effect.

image. We assume that the first frame of the image sequence is the reference image, and T_i is the spatial transformation matrix between the $i(i > 1)$ frame image and the $i + 1$ frame image. According to the following formula,

$$T = T_1 \times T_2 \cdots T_{i-1} \times T_i \tag{12}$$

We calculate the spatial transformation matrix between the $i + 1$ frame image and the reference image. According to the transformation relationship, we map all pixel points (x, y) in the $i + 1$ frame image to the corresponding (x', y') position in the reference image. Then we use the weight average method for fusion and stitching on all images to form an ultrasound image with a wider range.

III. EXPERIMENTS

A. DATA AND TRAINING DETAILS

The training images of our GAN-based super-resolution network come from Multi spectral image database, which contains 420 grayscale images. The pre-trained HR images are randomly selected from database while the pre-trained LR images are randomly selected and gotten bicubic interpolation degradation. The ultrasound image datasets used in experiments for test are obtained clinically from signal processing laboratory. The database contains 84 images of common carotid artery (CCA) of ten volunteers (mean age 27.5 ± 3.5 years) with different weight (mean weight 76.5 ± 9.7 kg). All images are taken by the specialists with five-year experience with scanning. Taking the average number of the feature points of the volunteers as result, the training loss curve of the generator is shown in the figure below. All the implementation uses Python 3.5 with TensorFlow and TensorLayer, which are now widely used in solving various medical image analysis problems. All the experiments are performed with one NVIDIA TITAN X Pascal GPU. All neural networks are trained and tested on the GPU. All experiments use the same initial learning rate of 10^{-4} , which decays to 10^{-5} at the midpoint of the training. The model is trained for 200 epochs to establish a fair comparison. The specific generator loss function curve is shown in figure 8.

B. EXPERIMENT DESIGN AND RESULTS

In this part, we compare dual-enhanced registration method for the field of view ultrasound sonography proposed in

TABLE 3. Vertical conoral experiment evaluation index.

Algorithm	Average feature points	Standard deviation(σ)	Average increased number	Increase ratio	Average number of successful matches
Reference image	309	10.1	-	-	108
Image to be matched	329	10.6	-	-	108
Reference image (Median filter)	586	13.2	277	90%	140
Image to be matched (Median filter)	589	12.8	260	79%	140
Reference image (Gaussian filter)	853	14.4	544	176%	174
Image to be matched (Gaussian filter)	904	13.7	575	175%	174
Reference image (Gray transform denoised)	1203	16.5	894	289%	367
Image to be matched (Gray transform denoised)	1232	17.0	943	274%	367

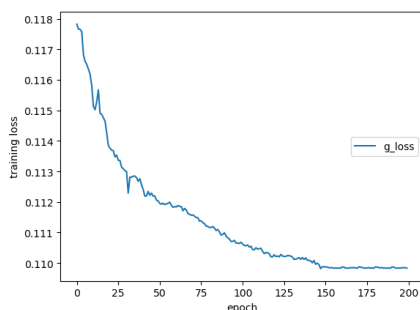


FIGURE 8. Training loss of generator.

this paper with some other EFOV-US based algorithms. We divide the evaluation methods of EFOV into two categories. (1) Number measurement using quantitative differences to measure algorithm performance. For example, Yerli and Eksioglu [50] determine the performance of EFOV by detecting the number of lesions. (2) Distance measurement use the offset in the distance to measure the performance of the algorithm, such as [51], [52], Silbernagel *et al.* [53] who analyze with the intraclass correlation coefficient (ICC) and compare an injured tendon with the same subject's healthy tendon. Our method is the same as the above method in comparison with the gold standard. Compared to the previous methods, the feature point evaluation criteria we use are more intuitive and quantified, avoiding distance calculations and complex markings to reduce calculation complexity, reduce the calculation time to 3 seconds and thus achieve better results. We also use several evaluation indicators to evaluate and compare our method with several other classic EFOV methods mentioned above to judge the pros and cons.

1) FEATURE POINT ENHANCEMENT ALGORITHM EXPERIMENT

In this set of experiments, we first perform feature point extraction and matching operations on reference images and images to be matched. Then we evaluate the effect of denoised methods on image feature point extraction and feature matching by comparing the number of feature points and the number of successfully matched feature points.

As shown in Table 3, after the denoising enhancement processing the number of feature points of reference image increases from the original 309 to 1203, which is higher than median filter of 586 and Gaussian filter of 853, an increase of 289%. The number of feature points of the image to be

matched increases from the original 329 to 1232, an increase of 274%. The above results show that our method performs better than several other traditional filtering methods in the number of feature points extracted. Median filtering is beneficial to preserve the sharpness of the edges but it will wash away the texture in the uniform medium area. Gaussian filter is effective for noises that follow a normal distribution, but it does not work well for other noises. In particular, compared with other methods, the number of successfully matched feature points by our method increases more, reaching 367. The number of feature points that are successfully matched will effectively improve the accuracy of the mapping model evaluation and the quality of image matching.

2) ULTRASOUND IMAGE SUPER-RESOLUTION EXPERIMENT

In this part, we compare images processed with a super-resolution algorithm with original images, in the number of feature points extracted and the effect of successful matching to evaluate the superiority of the algorithm [54]. The reference group performs feature point extraction and feature matching on the reference image and the image to be matched without super-resolution processing. The control group uses super-resolution processing based on Bicubic interpolation (BIC) and GAN (ours), then we perform feature point extraction and feature matching respectively. Finally we compare and analyze the number of successfully matched feature points number and the matching effects.

From the results in Figure 9 and Table 4, after denoising and super-resolution processing based on bicubic interpolation, the number of feature points in the reference image has increased directly from the original 1203 to 3582, an increase of 197.8%, while the method based on bicubic interpolation is 1784. Besides, the number of feature points of the image to be matched increases from the original 1232 to 3373, an increase of 173.8%. Bicubic interpolation method increases 48.4% to 1829. And the number of feature points of the reference image and the image to be successfully matched increases from the original 367 to 496, an increase of nearly 35%. The specific matching results are shown in Figure 10. As mentioned earlier, feature points which successfully matched will directly affect the accuracy and quality of subsequent image matching. This experiment can show that the super-resolution enhancement method of ultrasonic image based on the generative adversarial network proposed in this paper is helpful to increase the number of feature points and the number of feature points that are successfully matched. With

TABLE 4. Super resolution experimental evaluation index.

Algorithm	Average feature points	Standard deviation(σ)	Average increased number	Increase ratio	Average number of successful matches
Reference image (Gray transform denoised)	1203	16.5	-	-	367
Image to be matched (Gray transform denoised)	1232	17.0	-	-	367
Reference image(Denoised+bicubic interpolation)	1784	17.3	581	48.2%	479
Image to be matched(Denoised+bicubic interpolation)	1829	21.2	597	48.4%	479
Reference image(Denoised+GAN-based)	3582	19.5	2379	197.8%	496
Image to be matched (Denoised+GAN-based)	3373	20.5	2379	173.8%	496

TABLE 5. Image matching quality evaluation.

Algorithm	Information Entropy	Average gradient	Cross mutual information	Standard Deviation
Common EFOV-US	6.4510	2.2001	1.2853	37
Rapid Image Registration EFOV-US	4.2410	2.3056	1.3245	44
EFOV-US with two dimensional array probe	5.4020	3.2160	1.1438	45
Real-time EFOV-US (PC)	6.1211	3.1059	1.3144	50
EFOV-US (BIC)	6.4410	2.9005	1.2923	52
DRFOV-US (ours)	6.4287	5.6258	1.3788	92

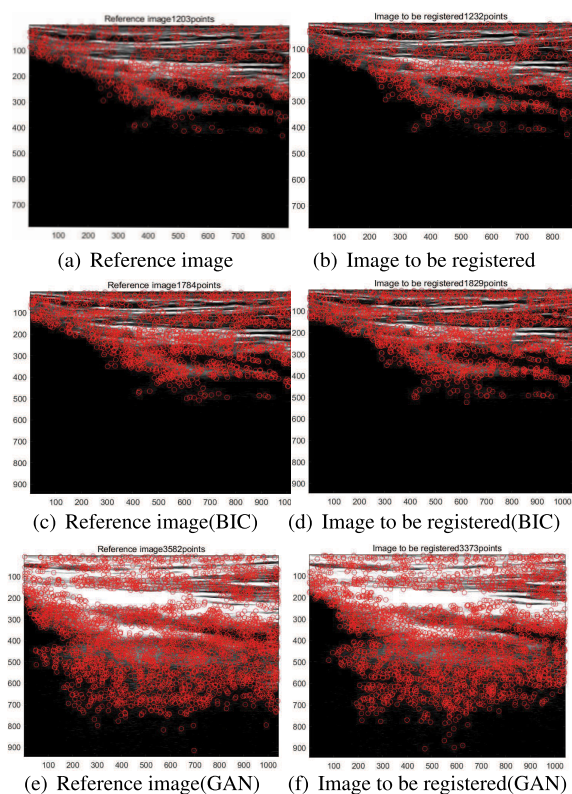


FIGURE 9. The contrast results of ultrasonic image super-resolution experiment, (a)(b) is original, (c)(d) is bicubic interpolation and (e)(f) is GAN-based method(ours).

the same feature point extraction method such as SIFT, our method captures more effective successful matching feature points as extra feature points. Compared with the BIC method, our method has slight improvement in the number of successfully matched feature points. However, the effective matching feature points extracted by different methods are not inclusive. Our method has a small increase in quantity, but the extra successful matching feature points play a key role in the quality of wide-range ultrasound images, which is clearly reflected in the following quality evaluation indicators experiment.

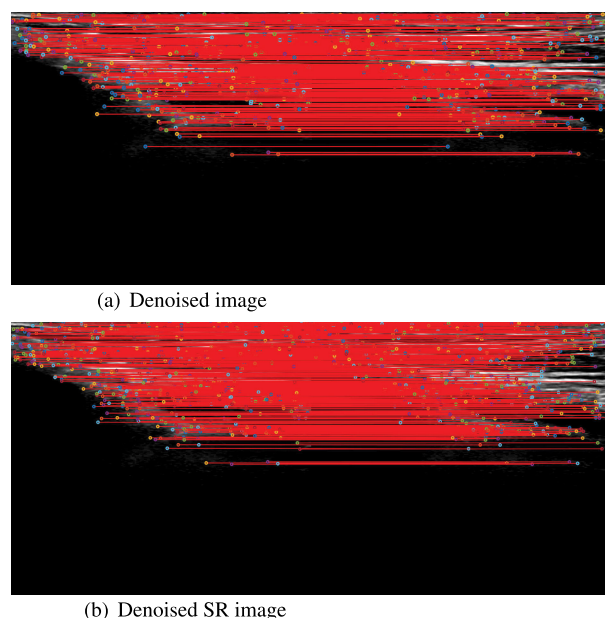


FIGURE 10. Matching results of image feature points processed by super resolution.

3) ULTRASOUND WIDE-RANGE IMAGE MATCHING EXPERIMENT

We use the following evaluation indicators for comparison.

a: INFORMATION ENTROPY

The stable value means that the fused image has no extra information. The specific definition is as follows, where $p(i)$ represents the probability value of gray value i

$$IE = - \sum_{i=0}^L p(i) \log_2 p(i) \quad (13)$$

Our method is as stable as the images generated by other methods.

b: AVERAGE GRADIENT

Average Gradient can objectively reflect the clarity of the image. If the value of average gradient is larger, the resulting

$$AG = \frac{\sum_{i=1}^{M-1} \sum_{j=1}^{N-1} [(f(i, j) - f(i + 1, j))^2 + (f(i, j) - f(i, j + 1))^2] / 2}{(M - 1) \times (N - 1)} \quad (14)$$

fusion image will be clearer and the visual effect will be better. The specific definition is as follows. M and N represent the length and width of the image respectively, $f(i, j)$ represents the pixel value at position (i, j) in (14), as shown at the top of this page.

Our method has a nearly double improvement in the average gradient compared with other methods, which means better visual effect and clearer image.

c: CROSS MUTUAL INFORMATION

Cross mutual information represents the interactive information of the fused image F and original images A and B . The larger the value of cross mutual information is, the more source image information the fused image contains. Our method is slightly improved compared to other methods.

d: STANDARD DEVIATION

The standard deviation index reflects the degree of dispersion between the pixel value and the average of the image. The larger the standard deviation is, the better the image quality and the richer details contain in the image. The specific definition is as follows.

$$SD = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (f(i, j)^2 - \bar{f})^2}{M \times N}} \quad (15)$$

M and N represent the length and width of the image respectively, $f(i, j)$ represents the pixel value at position (i, j) in image, and \bar{f} is specifically defined as follows.

$$\bar{f} = \frac{\sum_{i=1}^M \sum_{j=1}^N f(i, j)}{M \times N} \quad (16)$$

Our method performs nearly doubled in average standard deviation compared with other methods, which means better wide-range image quality and richer details.

According to the result in Table5, we can find that our DRFOV-US method based on the dual enhancement has first-class image quality and stability. Our method is also richer in image details and retains more information with higher matching accuracy than other models. Particularly, the results of the GAN-enhanced and BIC-enhanced ultrasound images are shown in Figure 11, in the case where other parts are consistent, our GAN-based method has a significant improvement in detail and clarity compared to method based on BIC. Doctors can compare wide-range images obtained by these algorithms. Our improved DRFOV-US algorithm obtains better result, clearer texture, less noise interference than several other EFOV-US methods. The higher quality wide-angle images are more helpful for doctors' observation and diagnosis.

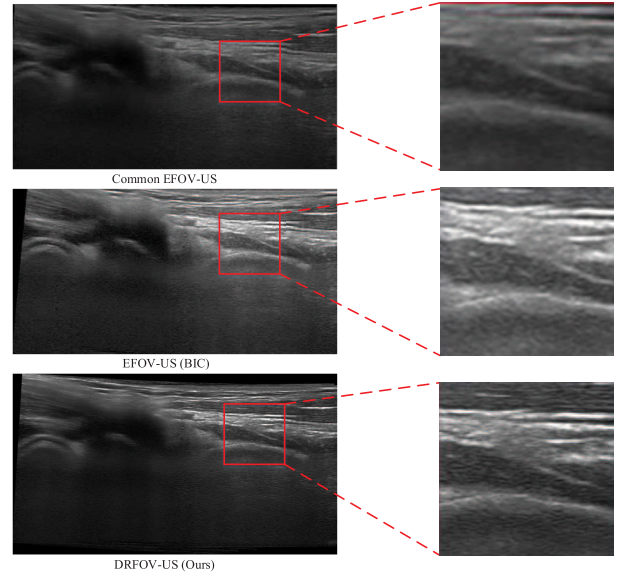


FIGURE 11. Final ultrasound wide-range imaging result.

IV. CONCLUSION

In this paper, we present a dual-enhanced registration method for the field of view ultrasound sonography. We first use feature point enhancement algorithm based on gray-scale transform to reduce noise in image preprocessing, and then conduct ultrasound image super-resolution enhancement based on generative adversarial network. We combine these two innovations to further improve the effect of feature point extraction, the image matching, the fusion accuracy, and therefore get high quality wide-range ultrasound images. Experiments show that the method proposed in this paper has achieved good results, greatly improved the quality of the final wide-range image, and can effectively help doctors.

REFERENCES

- [1] S. H. Kim, B. I. Choi, K. W. Kim, K. H. Lee, and J. K. Han, "Extended field-of-view sonography: Advantages in abdominal applications," *Ultrasound Med. Biol.*, vol. 29, no. 5, pp. S163-S164, 2003.
- [2] B. D. Fornage, E. N. Atkinson, L. F. Nock, and P. H. Jones, "US with extended field of view: Phantom-tested accuracy of distance measurements," *Radiology*, vol. 214, no. 2, pp. 579-584, Feb. 2000.
- [3] A. N. Adkins, P. W. Franks, and W. M. Murray, "Demonstration of extended field-of-view ultrasound's potential to increase the pool of muscles for which *in vivo* fascicle length is measurable," *J. Biomech.*, vol. 63, pp. 179-185, Oct. 2017.
- [4] L. Kwan, K. Nishihara, A. Cheung, C. D'Amico, A. Hart, N. Keshwani, and S. Mathur, "Reliability and feasibility of extended field of view ultrasound imaging techniques for measuring the limb muscle cross-sectional area," *Physiotherapy Canada*, vol. 72, no. 2, pp. 149-157, 2019.
- [5] S. Zheng, Q. Huang, L. Jin, and G. Wei, "Real-time extended-field-of-view ultrasound based on a standard PC," *Appl. Acoust.*, vol. 73, no. 4, pp. 423-432, Apr. 2012.

- [6] S. Zheng, Q. Huang, M. Lu, L. Jin, and T. Wang, "Rapid image registration for Extended-Field-of-View ultrasound," in *Proc. 4th Int. Conf. Bioinf. Biomed. Eng.*, Jun. 2010, pp. 1–4.
- [7] Y. Yoo, J. Jago, J.-M. Jong, R. R. Entekin, M. Anderson, and L. J. Olsson, "Extended field of view ultrasonic imaging with a two dimensional array probe," U.S. Patent 8 539 838 B2, Sep. 24, 2013.
- [8] E. Dyer, U. Zeeshan Ijaz, R. Housden, R. Prager, A. Gee, and G. Treece, "A clinical system for three-dimensional extended-field-of-view ultrasound," *Brit. J. Radiol.*, vol. 85, no. 1018, pp. e919–e924, Oct. 2012.
- [9] L. Kong, Q. Huang, M. Lu, S. Zheng, L. Jin, and S. Chen, "Accurate image registration using SIFT for Extended-Field-of-View sonography," in *Proc. 4th Int. Conf. Bioinf. Biomed. Eng.*, Jun. 2010, pp. 1–4.
- [10] D. Ni, Y. P. Chui, Y. Qu, X. Yang, J. Qin, T.-T. Wong, S. S. H. Ho, and P. A. Heng, "Reconstruction of volumetric ultrasound panorama based on improved 3D SIFT," *Computerized Med. Imag. Graph.*, vol. 33, no. 7, pp. 559–566, Oct. 2009.
- [11] C. Che, T. S. Mathai, and J. Galeotti, "Ultrasound registration: A review," *Methods*, vol. 115, pp. 128–143, Feb. 2017.
- [12] N. I. Tanaka, M. Ogawa, A. Yoshiko, R. Ando, and H. Akima, "Reliability of size and echo intensity of abdominal skeletal muscles using extended field-of-view ultrasound imaging," *Eur. J. Appl. Physiol.*, vol. 117, no. 11, pp. 2263–2270, Nov. 2017.
- [13] E. Shugard, N. Mistry, J. Cheung, J. Pouliot, and J. Chen, "SU-F-207-09: Evaluating the dosimetric accuracy of extended field-of-view CT reconstructions using clinical data with real patient geometries," *Med. Phys.*, vol. 42, no. 6Part26, pp. 3543–3544, Jun. 2015.
- [14] E. F. Brouwer, S. B. Myhrvold, J. Š. Benth, and S. E. Hoelsbrekken, "Ultrasound measurements of Achilles tendon length using skin markings are more reliable than extended-field-of-view imaging," *Knee Surg., Sports Traumatol., Arthroscopy*, vol. 26, no. 7, pp. 2088–2094, Jul. 2018.
- [15] C. Lu, C. Merrill, A. Medellin, K. Novak, and S. R. Wilson, "Bowel ultrasound state of the art: Grayscale and Doppler ultrasound, contrast enhancement, and elastography in crohn disease," *J. Ultrasound Med.*, vol. 38, no. 2, pp. 271–288, Feb. 2019.
- [16] J. Zhou, W. Zheng, L. Cao, M. Liu, F. Han, and A. Li, "Antiangiogenic tumor treatment: Noninvasive monitoring with contrast pulse sequence imaging for contrast-enhanced grayscale ultrasound," *Academic Radiol.*, vol. 17, no. 5, pp. 646–651, 2010.
- [17] S. Anwar, S. Khan, and N. Barnes, "A deep journey into super-resolution: A survey," 2019, *arXiv:1904.07523*. [Online]. Available: <http://arxiv.org/abs/1904.07523>
- [18] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [19] N. Varish and A. K. Pal, "A novel image retrieval scheme using gray level co-occurrence matrix descriptors of discrete cosine transform based residual image," *Int. J. Speech Technol.*, vol. 48, no. 9, pp. 2930–2953, Sep. 2018.
- [20] G. Song, J. Han, Y. Zhao, Z. Wang, and H. Du, "A review on medical image registration as an optimization problem," *Current Med. Imag. Rev.*, vol. 13, no. 3, pp. 274–283, Jul. 2017.
- [21] W. Venderink, M. de Rooij, J. P. M. Sedelaar, H. J. Huisman, and J. J. Fütterer, "Elastic versus rigid image registration in magnetic resonance imaging–transrectal ultrasound fusion prostate biopsy: A systematic review and meta-analysis," *Eur. Urol. Focus*, vol. 4, no. 2, pp. 219–227, Mar. 2018.
- [22] T. Lindeberg, "Scale invariant feature transform," *Scholarpedia*, vol. 7, no. 5, p. 10491, 2012.
- [23] K. D. Lakshmi and V. Vaithyanathan, "Image registration techniques based on the scale invariant feature transform," *IETE Tech. Rev.*, vol. 34, no. 1, pp. 22–29, Jan. 2017.
- [24] B. Ayres, A. Stacey, J. Grant, T. McClendon, and H. Demirci, "Comparative study of clinical, ultrasonographic, conventional imaging, and Ultra-Wide-Field fundus for measurements of the longest basal diameter of choroidal tumors," *Ophthalmic Surgery, Lasers Imag. Retina*, vol. 48, no. 6, pp. 459–464, Jun. 2017.
- [25] Q. Huang, Z. Zeng, and X. Li, "2.5-D extended field-of-view ultrasound," *IEEE Trans. Med. Imag.*, vol. 37, no. 4, pp. 851–859, Apr. 2018.
- [26] Z. Wang, J. Chen, and S. C. H. Hoi, "Deep learning for image super-resolution: A survey," 2019, *arXiv:1902.06068*. [Online]. Available: <http://arxiv.org/abs/1902.06068>
- [27] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 286–301.
- [28] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis. Zürich, Switzerland: Springer*, 2014, pp. 184–199.
- [29] J.-S. Choi and M. Kim, "A deep convolutional neural network with selection units for super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 154–160.
- [30] K. Zhang, W. Zuo, and L. Zhang, "Learning a single convolutional super-resolution network for multiple degradations," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3262–3271.
- [31] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep Laplacian pyramid networks for fast and accurate super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 624–632.
- [32] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4681–4690.
- [33] J. Zhu, G. Yang, and P. Lio, "How can we make GAN perform better in single medical image super-resolution? A lesion focused multi-scale approach," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1669–1673.
- [34] R. Wang, Z. Fang, J. Gu, Y. Guo, S. Zhou, Y. Wang, C. Chang, and J. Yu, "High-resolution image reconstruction for portable ultrasound imaging devices," *EURASIP J. Adv. Signal Process.*, vol. 2019, no. 1, pp. 1–12, Dec. 2019.
- [35] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, *arXiv:1701.07875*. [Online]. Available: <http://arxiv.org/abs/1701.07875>
- [36] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein GANs," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5767–5777.
- [37] X. Yi, E. Walia, and P. Babyn, "Generative adversarial network in medical imaging: A review," *Med. Image Anal.*, vol. 58, Dec. 2019, Art. no. 101552.
- [38] A. Yadav, S. Shah, Z. Xu, D. Jacobs, and T. Goldstein, "Stabilizing adversarial nets with prediction methods," 2017, *arXiv:1705.07364*. [Online]. Available: <http://arxiv.org/abs/1705.07364>
- [39] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis. Amsterdam, The Netherlands: Springer*, 2016, pp. 630–645.
- [40] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034.
- [41] S. Li, "A review of feature detection and match algorithms for localization and mapping," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 231, Jul. 2017, Art. no. 012003.
- [42] S. Routray, A. K. Ray, and C. Mishra, "Analysis of various image feature extraction methods against noisy image: SIFT, SURF and HOG," in *Proc. 2nd Int. Conf. Electr., Comput. Commun. Technol. (ICECCT)*, Feb. 2017, pp. 1–5.
- [43] L. Juan and L. Gwon, "A comparison of SIFT, PCA-SIFT and SURF," *Int. J. Signal Process., Image Process. Pattern Recognit.*, vol. 8, no. 3, pp. 169–176, 2007.
- [44] L. Kou, L. Zhang, K. Zhang, J. Sun, Q. Han, and Z. Jin, "A multi-focus image fusion method via region mosaicking on Laplacian pyramids," *PLoS ONE*, vol. 13, no. 5, May 2018, Art. no. e0191085.
- [45] Y. Cui, N. Hasler, T. Thormählen, and H.-P. Seidel, "Scale invariant feature transform with irregular orientation histogram binning," in *Proc. Int. Conf. Image Anal. Recognit. Berlin, Germany: Springer*, 2009, pp. 258–267.
- [46] N. K. Karn and F. Jiang, "Improved GLOH approach for one-shot learning human gesture recognition," in *Proc. Chin. Conf. Biometric Recognit. Chengdu, China: Springer*, 2016, pp. 441–452.
- [47] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.
- [48] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [49] R. Raguram, J.-M. Frahm, and M. Pollefeys, "A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus," in *Proc. Eur. Conf. Comput. Vis. Berlin, Germany: Springer*, 2008, pp. 500–513.
- [50] H. Yerli and S. Y. Eksioğlu, "Extended Field-of-View sonography: Evaluation of the superficial lesions," *Can. Assoc. Radiologists J.*, vol. 60, no. 1, pp. 35–39, Feb. 2009.

- [51] Y.-K. Bae, J.-W. Lee, and S. Hong, "Effects of image distortion and Hounsfield unit variations on radiation treatment plans: An extended field-of-view reconstruction in a large bore CT scanner," *Sci. Rep.*, vol. 10, no. 1, 2020, Art. no. 473.
- [52] C. Sahinis, E. Kellis, N. Galanis, K. Dafkou, and A. Ellinoudis, "Intra- and inter-muscular differences in the cross-sectional area of the quadriceps muscles assessed by extended field-of-view ultrasonography," *Med. Ultrasonography*, vol. 22, no. 2, p. 152, May 2020.
- [53] K. G. Silbernagel, K. Shelley, S. Powell, and S. Varrecchia, "Extended field of view ultrasound imaging to evaluate achilles tendon length and thickness: A reliability and validity study," *Muscle Ligaments Tendons J.*, vol. 6, no. 1, p. 104, 2016.
- [54] M. S. M. Sajjadi, B. Scholkopf, and M. Hirsch, "EnhanceNet: Single image super-resolution through automated texture synthesis," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4491–4500.



ZIJIA FAN received the bachelor's degree in computer science and technology from Northeastern University, in 2016, where he is currently pursuing the M.D. degree with the College of Medicine and Biological Information Engineering. His main research interests include image super resolution, deep learning, and medical assisted diagnosis.



ZHONGYANG WANG received the B.Sc. and M.Sc. degrees in biomedical engineering from Northeastern University, in 2012 and 2015, respectively, where he is currently pursuing the Ph.D. degree with the College of Medicine and Biological Information Engineering. His main research interests include machine learning, digital image processing, and bioinformatics, brain functional network research, data mining, and data analysis.



interests include big data, uncertain data, brain functional network research, and bioinformatics.



ZHIQIONG WANG (Member, IEEE) received the M.Sc. and Ph.D. degrees in computer science and technology from Northeastern University, China, in 2008 and 2014, respectively. She visited the National University of Singapore, in 2010, and The Chinese University of Hong Kong, in 2013, as an Academic Visitor. She is currently an Associate Professor with the College of Medicine and Biological Information Engineering, Northeastern University. She has published more than 60 articles. Her main research interests include computer aided diagnosis, medicine information, big health data analysis, cloud computing, and machine learning.



LU LIU received the bachelor's degree in computer science and technology from the Shenyang University of Technology, in 2015, and the M.Sc. degree in biomedical engineering from Northeastern University, in 2019. His main research interests include image stitching and image registration.



XIA ZHANG used to be a Teaching Assistant and a Lecturer with Northeastern University. In 2014, she joined Neusoft Group Corporation, Ltd., as the Director of the Department of Database and the Director of software product. From 1999 to 2004, she served as the Director of the Software Technology Center and Neusoft Group Corporation, Ltd. In 2004, she served as the Vice President. She is currently the Senior Vice President, the Chief Technology Officer, and the Chief Knowledge Officer at Neusoft Group Corporation, Ltd. Her research interests include software engineering, databases, and service-oriented smart applications.



JIREN LIU graduated from Northeastern University, China, in 1980. He went to the National Institute of Standards and Technology, USA, for further education, in 1986. He is currently the President of Neusoft Group Corporation, Ltd., and a Professor with Northeastern University. His research interests include medical image processing and analysis, computer network technology, and the embedded software technology.

...