Virtual Reality-Assisted Deep Learning for Support Vector Machine-based Analysis of Mass Sports Fitness Medical Images

Qin Yijun[1*]

Henan Polytechnic Institute, Nanyang, China

Corresponding author: Qin Yijun, hngyqyj2021@163.com

Abstract. This study aims to investigate the application of a deep learning method of support vector machine in mass sports fitness medical images. In accordance with the inclusion and exclusion criteria, 120 infant patients who were hospitalized in the neonatal intensive care unit (NICU) from January to October of 2019 were selected and randomly divided into three groups (40 cases in each group), namely, colostrum combined with sodium bicarbonate nursing (experimental group), colostrum (control group I), and sodium bicarbonate (control group II) groups. The primary outcomes measured included incidence rates of VAP and oral infection, positive rate of pathogenic bacteria after sputum culture, mechanical ventilation time, and length of stay (LOS). Taking the GAN with better reconstruction effect as the main body, the residual dense block is introduced into the generator of the GAN, so that the generator can still extract different levels of features of LR images even when it is deeper and wider to achieve feature fusion. The loss function composed of perception loss and confrontation loss is used to replace the MSE loss used only in the past. The main task of this paper is the research of medical image super-resolution reconstruction technology based on deep learning. This paper introduces the research background and significance of image super-resolution reconstruction algorithm, and conducts some research on image super-resolution reconstruction technology. In the aspect of network structure design, common network structures are analyzed. Combined with the particularity of super-resolution reconstruction task, SRGAN is taken as the basic framework and improved. In the aspect of loss function design, content loss function and generation loss function are used to ensure the restoration of high-resolution images while producing texture details that are more consistent with the human eyes.

Keywords: Super resolution; Medical image; Residual dense block; Retinal fundus image; Virtual Reality
DOI: https://doi.org/10.14733/cadaps.2023.S14.208-2015

1 INTRODUCTION

Medical image means that doctors use medical imaging equipment to obtain images of some organs or tissues in patients without invading the human body, so as to achieve the purpose of diagnosis or medical research on patients' conditions1-3. Medical images are an important tool for judging the existence of diseases and analyzing experimental results4-6. Higher quality and precision images can provide more details for doctors to improve the diagnostic accuracy in pathological research, making it easier for them to identify the precise location of lesions and understand the details of the affected areas7. virtual reality (VR) technology further enhances the image viewing experience, providing a realistic and immersive environment for medical professionals to analyze high-resolution medical images. At present, ultrasonic imaging (US), magnetic resonance imaging (MRI) and X-ray equipment are the most widely used medical imaging equipment in medical institutions8. As an important index to evaluate the quality of medical images, image resolution is very important in medical research9. The resolution of medical images generated by current medical imaging equipment is low, and accurate medical judgment cannot be separated from high-resolution images10. The most important factor affecting the diagnosis effect is the resolution of medical images. High resolution medical images can help experts analyze patients' conditions more accurately. Using super resolution reconstruction technology to improve the resolution of medical images can greatly improve the quality of diagnosis, while saving materials and funds required for equipment upgrading.

2 MATERIALS AND METHODS

2.1 Generate Countermeasure Network

The GAN framework is usually composed of generator G and discriminator D. The basic idea of GAN is antagonistic learning: train G to generate images with rich details, and train D to distinguish whether a given image is real or generated. At the end of the training, D will be a very good classifier, which can separate the real and generated images. G can learn to create solutions that are highly similar to the real images, so it is difficult to use D for classification. Through adversarial learning, the generator and discriminator have been improved. The confrontation process between D and G is:

$$\min_{G}\max_{D} V(D,G) = E_{x\sim Pdata(x)}\left[\log D(x)\right] + E_{z\sim P_z(z)}\left[\log\left(1-D\left(G(z)\right)\right)\right] \tag{1}$$

2.2 Residual Dense Block

With the introduction of VGG network, network depth is considered to be crucial. Many computer vision problems have obtained satisfactory results from deeper networks. Depth convolution neural network has made a series of breakthroughs in image processing. The depth network integrates different levels of image features and classifiers in an end-to-end multi-layer way, and can enrich features by adding deep depth. However, most of the super-resolution models based on convolutional neural networks do not fully utilize the hierarchical features of the original LR image, so their performance is low. The combination of residual dense block (RDB) with residual learning and dense block can solve this problem. As shown in Figure 1, residual dense block consists of dense connection layer, local feature Fusion, LFF) layer and residual learning. The input data of the module, that is, the feature map, passes through a series of dense connection layers composed of 3×3 convolution layers and ReLU, fully learns and extracts high-frequency information, and then carries out dimension reduction through 1×1 feature fusion layer to realize feature fusion.

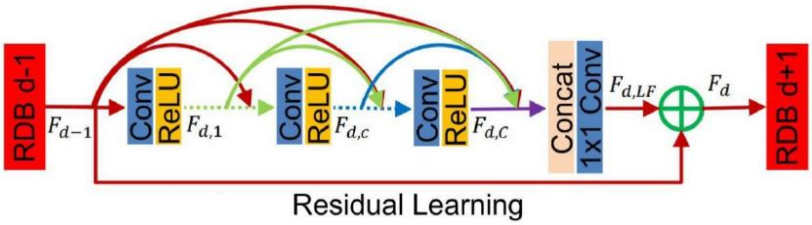


Figure 1: Residual dense block.

Then, the final output is obtained by combining the output result at this time with the output result of the last RDB through residual learning. Specifically:

(1) Transfer the output of the previous layer of RDB to each convolution layer of the current RDB. Set Fd1 and Fd as the input and output of layer d RDB, respectively. Both have G0 characteristic graphs. Then the output of layer c convolution layer of layer d RDB can be expressed as:

$$F_{d,c} = \sigma\left(W_{d,c}\left[F_{d-1}, F_{d,1} \cdots, F_{d,c-1}\right]\right) \tag{2}$$

(2) Using the idea of dense network for reference, the output of the former RDB and all convolution layers of the current RDB are adaptively fused by local feature fusion. Because a large number of repeated feature maps are input into each convolution layer at the back, the last feature map of the network is redundant, which will take up a large amount of video memory and slow down the operation, and it is difficult to train a dense network that is too deep. Therefore, 1×1 local fusion layer H_{LFF}^{d} is introduced to reduce the number of feature maps, adaptively fuse and retain the previously learned features. This process is expressed as:

$$F_{d,LF} = H_{LFF}^{d}\left(\left[F_{d-1}, F_{d,1}, \cdots, F_{d,c}, \cdots, F_{d,C}\right]\right) \tag{3}$$

(3) RDB introduces residual learning to further improve the information flow. The residual unit structure is shown in Figure 2:

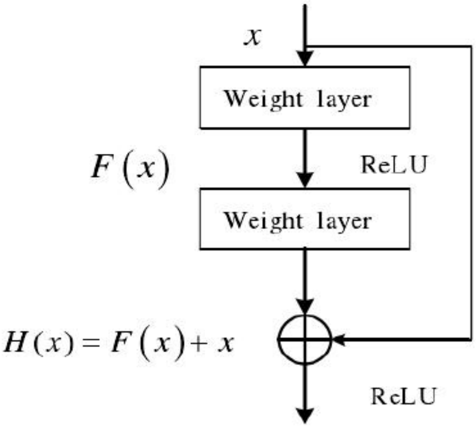


Figure 2: Residual unit structure.

2.3 GAN Based on Residual Dense Block

How to define the loss function lSR is critical to the performance of the generator. Although lSR is usually based on MSE modeling, through the improvement of Johnson and Bruna's research, this paper uses a new loss function based on perceptual loss, which is the weighted sum of content loss l_X^{SR} and confrontation loss l_{Gen}^{SR}.

(1) Content loss: The difference between SR image and HR image is expressed by the loss function based on mean square error (MSE):

$$l_{MSE}^{SR} = \frac{1}{r^2WH}\sum_{x=1}^{rW}\sum_{y=1}^{rH}\left(I_{x,y}^{HR} - G_{\theta_G}\left(I^{LR}\right)_{x,y}\right)^2 \tag{4}$$

MSE is the optimization target widely used by many image super-resolution algorithms at present. Although the reconstructed image can obtain high PSNR, it lacks high-frequency texture information, and the perceptual visual effect under high magnification is not good.

(2) Confrontation loss: In addition to the content loss described above, we also added GAN's confrontation loss to the perceived loss. By cheating the discriminator network, the countermeasure loss can get a solution that is more inclined to the natural image manifold.

$$l_{Gen}^{SR} = \sum_{n=1}^{N} -\log D_{\theta_D}\left(G_{\theta_G}\left(I^{LR}\right)\right) \tag{5}$$

(3) To sum up, the losses of the whole generator are:

$$G_{loss} = MSE + VGG_{loss} + Dis_{loss} = l_{MSE}^{SR} + l_{VGG}^{SR} + l_{Gen}^{SR} \tag{6}$$

In order to make the three losses within the same comparable scale range, the above formula is modified in this paper:

$$G_{loss} = l_{MSE}^{SR} + 2\times10^{-6} l_{VGG}^{SR} + 10^{-3} l_{Gen}^{SR} \tag{7}$$

Using mini-batch training mode, the batch size is 16, and each HR image of each batch of mini-batch is randomly cropped by 128×128 to obtain sub-images. Using MATLAB to Bicubic down-sample HR sub-image, LR image is obtained. The LR and its corresponding HR sub-images of this mini-batch are input into the generator and discriminator for training, and the iterative network parameters are continuously updated. The initial learning rate is set at 10-4, and 6×104 iterations are carried out. After 3×104 iterations, the learning rate is reduced to 10-5. Adam optimization algorithm is used in the training process, β1=0.9, β2=0.999.

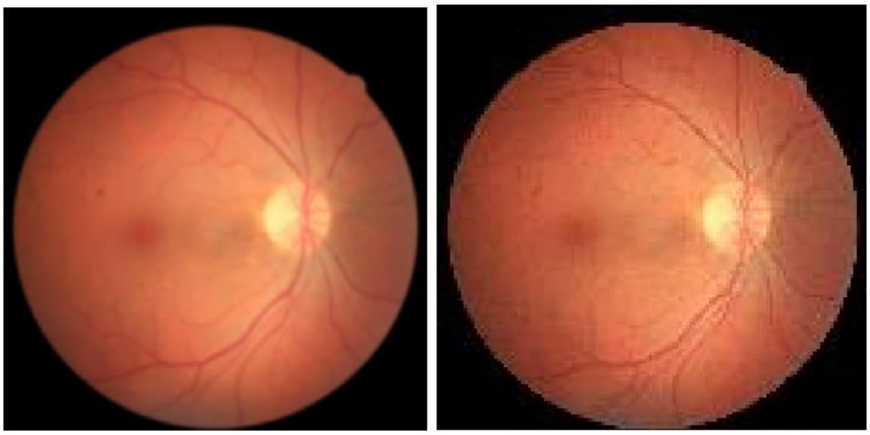
3 RESULTS

3.1 SR of Retinal Fundus Image

EyePACS is a free screening platform for retinopathy in the United States, which is used by more than 600 organizations. It is the largest screening program in the United States outside the vertically integrated system. Take 6,000 retinal fundus images from EyePACS, 5,000 for training and 1,000 for testing. First, the original image of 1536×1024 is clipped, and the extra black edge outside the retina is removed, and the image is adjusted to HR image of 1024×1024. The LR images with sizes of 512×512, 256×256 and 128×128 are obtained by downsampling image scaling factors R of 2, 4 and 8, respectively. The indicators of this quantitative evaluation are PSNR, SSIM and S3 (sharpness measurement).

It can be seen from Figure 3 that the image reconstructed by bicubic is blurred, and there are still jagged mosaics at the edge. And only the aorta can be clearly seen, SRCNN can blur some microvessels, and the image quality clarity after SRGAN reconstruction is obviously improved, but the texture details at the edge of the blood vessels are still blurred. In this paper, the sharpness of the algorithm is obviously improved, and it has a better performance in restoring details, which is closer to the original image.

HR Ours SRGAN



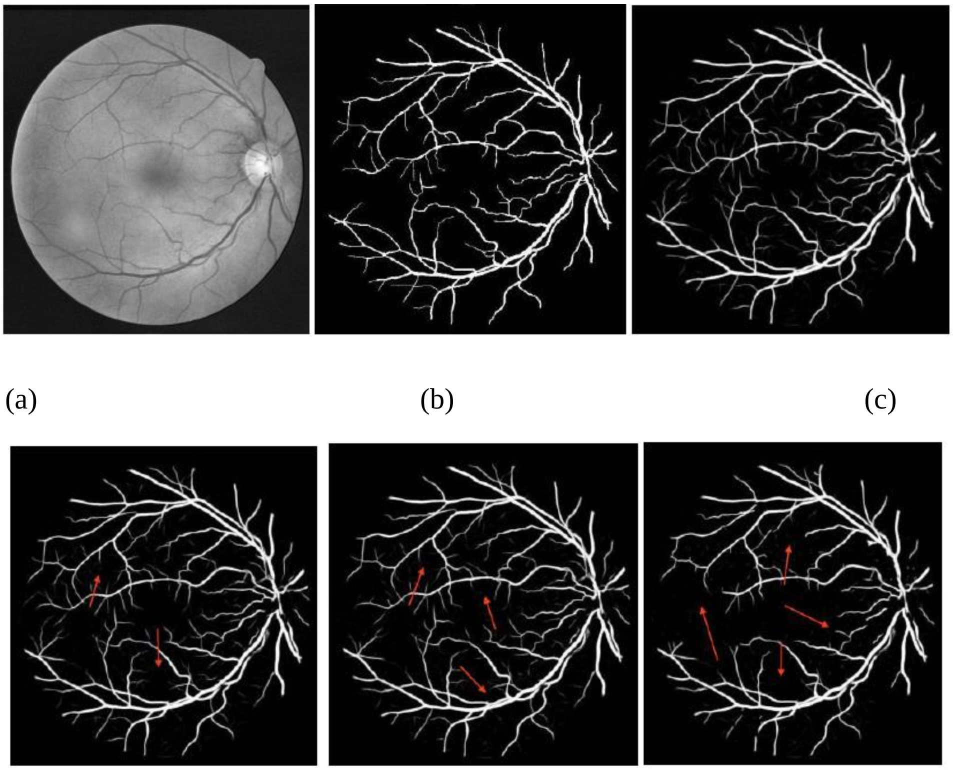
SRCNN bicubic

Figure 3: Results of LR image reconstruction of retinal fundus oculi with r=8.

4 DISCUSSION

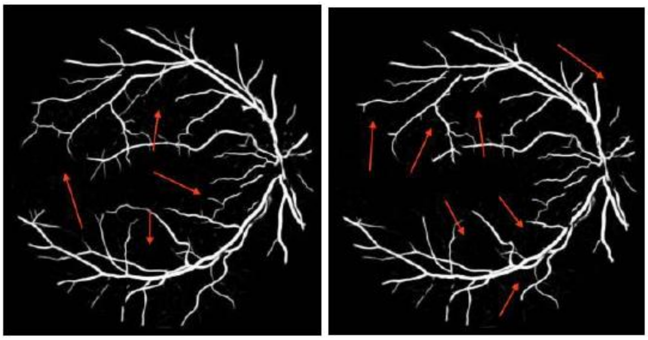
The image acquisition of remote ophthalmology is completed by a low-quality camera. Even if high-resolution devices are used, images transmitted on the network still need to be compressed from high resolution to low resolution. In this case, the image quality received by clinicians cannot meet their requirements for disease analysis. One of the main purposes of super-resolution reconstruction of retinal images is to better map image analysis, such as accurate pathological detection and marker segmentation. Through segmentation of SR images reconstructed by different algorithms, the reconstruction algorithms are compared and analyzed from the objective segmentation index and subjective evaluation.

The DRIVE dataset is used for training and segmentation on the U-net partition network. The DRIVE dataset consists of 40 fundus images with manually labeled fundus vessels, 20 of which are training sets and 20 of which are test sets. The original HR image is sampled down by the multiples of 2, 4 and 8 respectively to obtain LR image, and LR image is input into the trained super resolution model to generate different SR images. Different U-Net architectures are trained with HR, LR, and SR images. The segmentation result of HR image is the upper limit of the best super-resolution reconstruction performance, and the LR image segmentation result is used as the baseline to compare the performance of super-resolution reconstruction algorithm. If the reconstruction algorithm is successful, the segmentation performance of SR image should be better than that of LR image, and the segmentation result should be close to that of the original HR image.



(a) (b) (c)

(d) (e) (f)



(g) (h)

Figure 4: R=8 retinal blood vessel segmentation result.

Figure 4 (a) shows the retinal fundus image after preprocessing, and (b) shows an example of manual segmentation of retinal fundus vascular image. (c) To use the original HR image to train the U-Net, (d), (e), (f) and (g) are the segmentation results of the super-resolution images generated by the algorithm, SRGAN, SRCNN and bicubic after training respectively, and (h) is the segmentation results of LR images with a scaling factor of 8. The regions that cannot be accurately segmented by different methods are indicated by red arrows. Obviously, the segmentation result of SR image restored by this algorithm is the closest to that of HR image.

5 CONCLUSION

The main task of this paper is the research of medical image super-resolution reconstruction technology based on deep learning. This paper introduces the research background and significance of image super-resolution reconstruction algorithm, and introduces the research status of image super-resolution reconstruction algorithm at home and abroad. Some researches on image super-resolution reconstruction technology are carried out, including traditional image super-resolution reconstruction technology based on interpolation, reconstruction and learning, and image super-resolution technology based on depth learning. In the aspect of network structure design, common network structures are analyzed. Combining the particularity of super resolution reconstruction task, SRGAN is taken as the basic framework and improved: residual dense blocks are used to replace residual blocks as the basic blocks of the generator, and the discriminator still uses VGG networks. The antagonism characteristic of GAN enables us to obtain a reconstructed image more in line with human perception. The residual dense block can learn the high-frequency details of medical images. This algorithm enables medical images to be reconstructed effectively. In the aspect of loss function design, content loss function and generation loss function are used to ensure the restoration of high resolution images, while producing texture details that are more consistent with human perception.

Qin Yijun, https://orcid.org/0009-0000-8061-5555

REFERENCES

[1] Alemdar, M.: Abdominal epilepsia partialis continua in a patient with astrocytoma treated with Lacosamide - value of repetitive EEG recordings, Eur Rev Med Pharmacol Sci, 25, 2021, 6277-6282.

[2] Chen, F.; Zhao, X.: Systematic family therapy for the psychological support of the public under the COVID-19 epidemic, International Journal of Psychiatry, 47(3), 2020, 4-11.

[3] Gao, S.; Wu, L.H.; Chen, Z.G.: Functional movement screening is related to the risk of military training injury of special service personnel, Southeast Natl Def Med, 23, 2021, 3-17.

[4] Kannan, S.; Shaik, S.A.; Sheeza, A.: Monkeypox: epidemiology, mode of transmission, clinical features, genetic clades and molecular properties, Eur Rev Med Pharmacol Sci, 26, 2022, 5983-5990.

[5] Lin, Y; Li Q; Jiang, A.: Research Progress on psychological support of tumor patients through social media, Journal of Nursing, 36(17), 2021, 3-11.

[6] Mao, W.: Analysis of sports injury characteristics of College Aerobics athletes based on functional action screening, Liaoning Sports Sci Tech, 44, 2022, 5-10.

[7] Mukdes, H.; Cheng J.; Zhang L.: Effects of intensive psychological support intervention on adverse reactions, psychological resilience and quality of life of patients with ovarian cancer undergoing chemotherapy, Cancer Progress, 18(3), 2020, 4-9.

[8] Rinaldi, V.E.; Alonzo, R.D.; Di, C.G.; Verrotti, A.: COVID-19 and abdominal pain: a pediatric case report and a point of view in pediatric emergency medicine, Eur Rev Med Pharmacol Sci, 25, 2021, 7115-7126.

[9] Wu, S.Y.; Jia, Y.: Application and research of functional movement screening to prevent injuries and injuries of female gymnasts, Fujian Sports Sci Tech, 41, 2022, 5-12.

[10] Zhang, B.; Dong, Y.: Analysis of the value of language communication combined with psychological support in urodynamic examination of patients with BPH. Basic medical forum 2020; 24: 3488-3489[3] Gao S, Wu LH, Chen ZG. Functional movement screening is related to the risk of military training injury of special service personnel, Southeast Natl Def Med 2021; 23, 3-17.