Edinburgh Research Explorer

Cirrhosis liver classification on B-mode ultrasound images by convolution neural networks with augmented images

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Cirrhosis liver classification on B-mode ultrasound images by convolution neural networks with augmented images

Yoshihiro Mitani, Robert B. Fisher, Yusuke Fujita, Yoshihiko Hamamoto, and Isao Sakaida

Abstract—In the medical imaging field, it is desirable to develop computer-aided diagnosis(CAD) systems. They are useful as a second opinion, and to objectively and quantitatively make diagnoses. In this study, we focus on liver ultrasound images. The cirrhosis liver is expected to progress to a liver cancer in the worst case. Therefore, we are investigating a CAD system to identify the cirrhosis liver sooner. In this paper, in order to classify cirrhosis or normal liver on regions of interest(ROIs) image from B-mode ultrasound images, we have proposed to use a convolution neural network(CNN). CNNs are one of promising techniques for medical image recognition. In a previous study, we tried to classify the cirrhosis liver using a Gabor features based method, a higher order local auto-correlation(HLAC) feature based approach and an improved version. However, the classification performance of our preliminary experimental results were poor. The average error rates were still over 40%. In order to more accurately classify the cirrhosis liver, we have explored the use of the CNNs. The experimental results show the effectiveness of the CNNs. Furthermore, by a data augmentation technique, the classification performance of the CNNs is improved.

Index Terms—Augumented images, B mode ultrasound images, Cirrhosis liver classification, Computer-aided diagnosis system, Convolution neural networks.

I. INTRODUCTION

In the medical imaging field, it is desirable to develop computer-aided diagnosis(CAD) [1] systems which can give a second opinion objectively and quantitatively. In this study, we focus on liver ultrasound images. Ultrasound images are widely used for diagnosis of liver cirrhosis [2]. The cirrhosis liver is expected to progress to liver cancer in the worst case. Therefore, we are investigating a CAD system to diagnose a cirrhosis liver sooner [3][4][5]. Fig. 1 shows B-mode ultrasound images. In the center, we can see the whole of the liver areas. Fig. 1 (a) is a normal liver, (b) is a cirrhosis liver. Cirrhosis is a kind of disease, abnormal situation. The cirrhosis liver becomes harder compared with the normal one. In this study we focus on regions of interest(ROIs) instead of the whole ultrasound image. We firstly consider examining small sized ROI images. Secondly from the experimental results of the ROI images, we try examining the whole sized ultrasound image. Fig. 2 shows ROI images. The ROI images are manually cut out from within the liver areas, by the physician. Fig. 2 (a) shows examples of normal liver. On the other hand, (b) shows cirrhosis examples. The size of the ROI images is the same, 32×32 pixels. This is a typical 2-class problem, normal or abnormal. In this paper, we have focused the classification of these ROI images. In a previous study [5], we explored cirrhosis liver classification using a Gabor features based method and a higher order local autocorrelation(HLAC) feature based approach. These features are expected to fit the texture recognition problem. Our ultrasound ROI images are considered to be a kind of texture. We want to investigate which feature is the best for classifying ROI images, because the better features are, the better the classification performance is. In a conventional approach, such as the Gabor and HLAC features approach, it is important to find better features. However, the experimental results of this conventional approach were poor, with even the best performance. The average error rate was still over 40%. Unfortunately we cannot easily classify the normal or cirrhosis. This might be a difficult pattern recognition problem.

In this paper, we have investigated the use of convolution neural networks(CNNs). CNNs are considered to be one of promising techniques in the image recognition field. CNNs originate from the artificial neural networks [6], and are highly suitable for images. The performance of CNNs has been improved by adding more layers and by deeply training to networks [7][8]. CNNs are successfully applied in the pattern recognition field, especially in the image recognition field. Recently, CNNs are reported to be used widely in the medical imaging field [9]. With CNNs, we don't need to consider what features or what classifier we should use. Though we are usually concerned about these kind of choices, CNNs can automatically connect between input and output. Although now we have the new issue of choosing the network model, network configuration, hyperparameters and training process. We are hoping that the deep nets will be able to improve performance in this difficult cirrhosis pattern recognition problem. However, CNNs need many training samples. There is also an over-training problem. This means that classification on the training samples is almost perfect, but test sample performance is low. From our preliminary experiments, we hypothesize that the CNN can memorize the class for every sample, as for ROIs images. The number of available samples is limited. It's known as the small sample size problem in the pattern recognition field [10][11]. Therefore, we also investigate using an augmentation of the ROI images such as the perspective transformation.

In this paper, in order to improve cirrhosis liver classification performance on ROI images from B mode ultrasound images, we propose to use a CNN. We further investigate an augmentation of the ROI images. The experimental results show the effectiveness of the CNNs. Furthermore, by using an augmentation technique, the classification performance of the CNNs is improved.

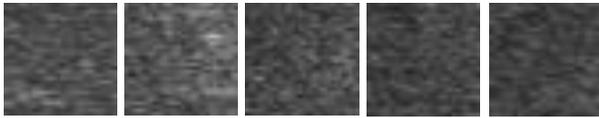This paper is organized as follows: In the second section,
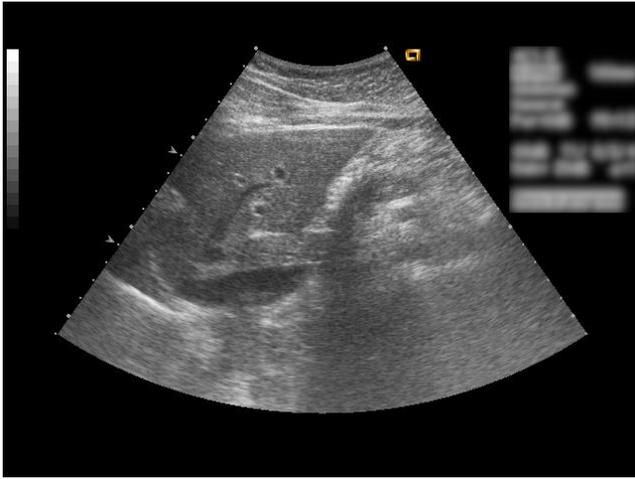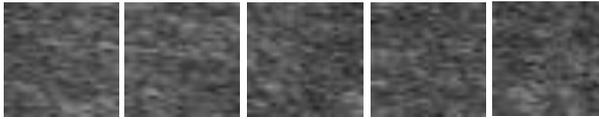
(a)Normal

(b)Cirrhosis

Figure 1 B-mode ultrasound images

(a)Normal

(b)Cirrhosis

Figure 2 ROI images

we discuss the previous work. The third section describes our methodology and the fourth is the experiments. The experimental results show the effectiveness of the CNNs compared to the conventional method.

II. PREVIOUS WORK

Submit your manuscript electronically for review. In order to classify the cirrhosis liver of B-mode ultrasound ROI images, the conventional approach focused on which features we should use. In a previous study [5], we compared the classification performance of Gabor features with that of HLAC features, because these features are preferable for texture images, like ultrasound images. For simplicity, we used a nearest neighbor classifier [12][13]. And we further investigated how the HLAC feature could be improved by using some image processing techniques [14]. The adaptive thresholding technique [15] was most effective. Table 1 shows the average error rates of the conventional method. The figures show average error rate and the 95% confidence interval. The HLAC feature* means the HLAC feature was improved with some image processing techniques. From the previous study, we found adaptive thresholding best for the classification. From the result, we could see the best performance of HLAC feature*, 44.1%. However, even the best average error rate was still over 40%, which is too low for

Table 1 The average error rates of the conventional method, with 95% confidence intervals.

Gabor feature	HLAC feature	HLAC feature*
48.3±0.6%	47.6±0.5%	**44.1±0.5%**

clinical use. Therefore, we have to consider another approach such as the CNNs.

III. METHODOLOGY

We describe the CNN and augmentation we used.

A. CNN architecture

First, we show the CNN architecture. The performance of CNNs depends on its network structure, learning method, the parameters to be determined, and so on. In preliminary experiments, we have decided these. We used the CNN shown in Fig. 3. The input of the CNN is the ROI image of size 32×32. Firstly, we convolve the ROI image by using 32 filters with a 3×3 filter size. And by 2×2 maxpooling, we reduce the ROI image size to a half-sized image, 16×16. Secondly, we repeatedly convolve and do max-pooling in the same manner. Then, we get 32 8×8 sized image. Thirdly, we flatten this image into $2,048(=32\times8\times8)$ -dimensional data. Finally, we make a fully connected artificial neural network. The network has one hidden layer. The number of the neurons also depends on the classification performance. For simplicity, we used 100. Then we used dropout. The rate of dropout is 0.5. The number of the outputs of the CNN is 2. This corresponds to a 2-class problem. Therefore, the structure of the fully connected artificial neural network is 2,048-100-2. All the activation functions are ReLU except for the output. In the output, we used softmax. The learning optimizer is adam. The epochs and batch size are 100 and 400, respectively.

B. Augmentation techniques

Here, we show an augmentation technique we used. In general, CNNs need a lot of samples to train properly. On the other hand, the number of the ROI images is limited in our study. Therefore, we generate additional samples by using the
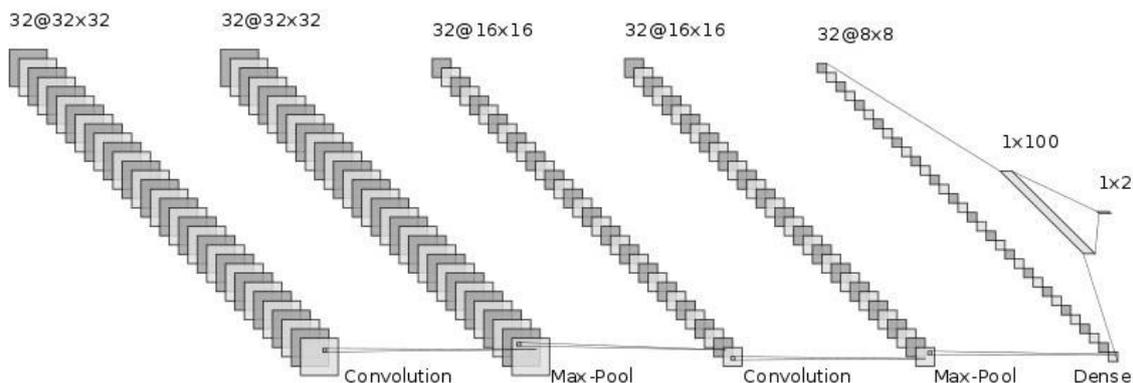
Figure 3 A network structure of CNN we used.

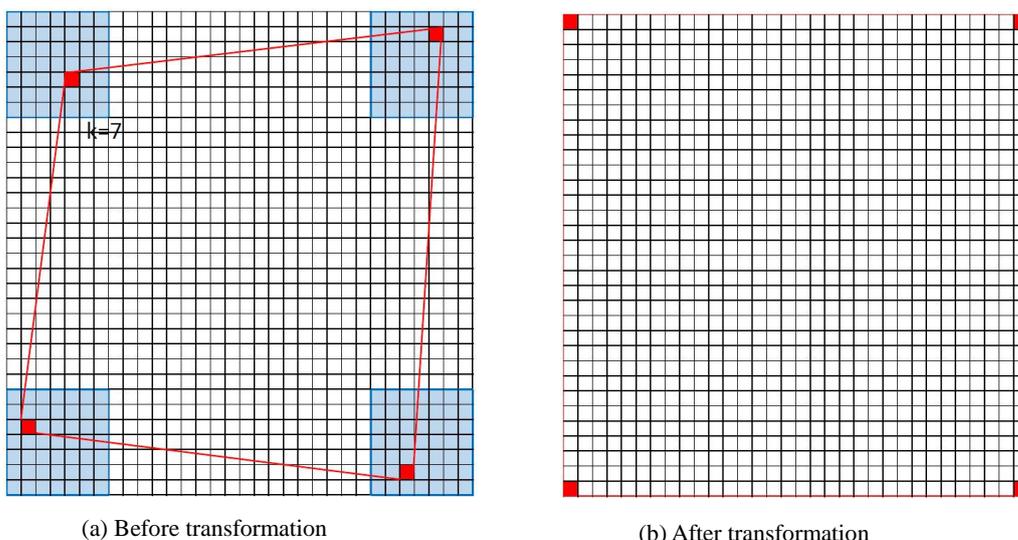(a) Before transformation

(b) After transformation

Figure 4 An illustration of a perspective transformation. The quadrilateral bounded by the left 4 red points is mapping by a homography of the right image.

augmentation technique. In this study, we have tried using a perspective transformation.

We describe a perspective transformation, which has some distortion. The newly generated image is usually a little smaller than an original image. Fig. 4 shows an illustration of a perspective transform. Fig. 4 (a) shows an illustration from before transformation. In the perspective transformation, we need at least 4 corresponding points between the before and after transformed images. In the experiments, we used 4 randomly generated points, such as in the light blue corners in Fig. 4 (a). In each of these areas, we choose a random point. In this case, we can see a selection box of size k = 7. In the experiments, we used k = 3, 5, 7, and 9. Every trial, we get an artificially generated new image. We can get an arbitrary number of images artificially generated. Given 400 training images, we could get 400×x training images. In the experiments, we used x = 8, 24, 48, and 128. Therefore, we used 3,200, 9,600, 19,200, and 51,200 training images.

IV. EXPERIMENTS

In the experiment, we used 500 available ROI images: 200 normal images and 300 cirrhosis images. This is a 2-class problem as mentioned previously. The gray level is 8 bits. The effectiveness of the CNNs is examined in terms of the error rate. The error rate is defined as a ratio of the number of test images misclassified to the number of all test images.

Error rate = #test images misclassified / # all test images ×100(%) (1)

For error rate estimation, the holdout method has been successfully used, because it maintains the statistical independence between the training and test images [16][12]. To evaluate the classification performance of the CNN, the average error rate was obtained by the holdout method. Fig. 5 shows the flow of the error rate estimation. First, we randomly divided the 500 available ROI images into 400 training ROI images and 100 test ROI images. The 400 training ROI images consist of 160 normal and 240 cirrhosis ROI images. The 100 test ROI images consist of 40 normal and 60 cirrhosis ROI images. Second, we augmented the training images as described above. Third, we train the CNN using the training images and compute the error rate using the test images. Finally, by 10 repetitions, the average error rate and 95% confidence interval were obtained. Fig. 5 illustrates the method used.

The purpose of experiment 1 is to investigate the classification performance of the CNN by using gray and binary ROI images in terms of the error rate. From the
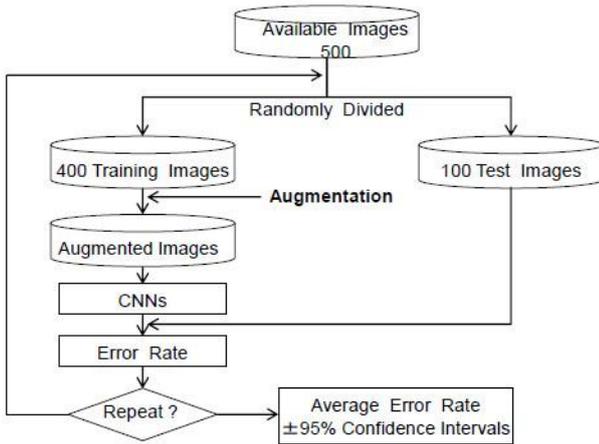
Figure 5 A flow of the error rate estimation.

Table 2 The average error rates of the CNN when using gray and binary ROI images.

Gray	Binary
34.4±3.5%	42.4±4.1%

Table 3 The average error rates of the CNN for each of the numbers of the augmentation.

3,200	9,600	19,200	51,200
35.1±2.1%	**31.9±2.5%**	34.9±3.2%	34.2±3.4%

previous study, we found that binary images were effective for classification. Therefore, we have conducted this experiment. The binarization method was the adaptive thresholding method [15]. Table2 shows the average error rate of the CNN by using the gray and binary ROI images. From Table 2, we see the gray ROI images are superior to the binary ones. Therefore, all remaining experiments were conducted using the gray ROI images.

The purpose of experiment 2 is to investigate effects of the augmentation by a perspective transformation. The image perspectively transformed has some distortion. But we can generate many samples by this technique. We can control the number of images artificially generated. In the experiments, the numbers of images augmented by the perspective transformation including original images are 3,200 (=400×8), 9,600 (=400×24), 19,200 (=400×48), and 51,200 (=400×128), respectively. From the experimental results, we chose k = 7 which determines the areas with 4 points randomly generated. Refer to the light blue areas in the Fig. 4. Table 3 shows the result of the average error rates of the CNN for each number of augmentations by the perspective transformation. From Table 3, we see the average error rate, 31.9%, is the lowest when the number of the images augmented is 9,600. The average error rate of the CNN with this augmentation outperforms that of the CNN without augmentation.

From all of the experimental results, we see the best average error rate of CNNs is 31.9%. This significantly outperforms the conventional method by HLAC features with the average error rate, 44.1%. It is significantly better in terms of t-test, $p \leq 0.01$.

V. CONCLUSION

In order to accurately classify the cirrhosis liver ultrasound ROI images, we have explored the use of CNNs. Furthermore, by an augmentation technique, the classification performance of the CNNs has been increased. The experimental results show that the classification performance of the CNN outperforms that of the conventional method, when using data augmentation by a perspective transformation technique. As we expected, the CNN has produced a dramatic improvement even with a difficult pattern recognition problem such ultrasound ROI image classification.

When applying CNNs to this problem, we could consider other network structures, parameter settings, other learning methods, and so on. In the future, we would get more real data and explore the whole ultrasound images such as Fig. 1 by using CNNs. Furthermore, we should try other augmentation techniques, such as by adding small Gaussian noise to images. There is a generative adversarial network(GAN) [17] which automatically generates fake samples. By using this technique, the fake images seem very similar to the real. Moreover, the deep convolutional generative adversarial networks(DCGANs) [18] was published recently. This technique is possibly preferable. In the future, we will investigate applying DCGANs to the cirrhosis liver classification problem.

REFERENCES

[1] Kunio Doi, "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Computerized Medical Imaging and Graphics*, vol. 31, pp. 198–211, 2007.

[2] A. Huber, L. Ebner, J. T. Heverhagen, and A. Christe, "State-of-the-art imaging of liver fibrosis and cirrhosis: A comprehensive review of current applications and future perspectives," *European Journal of Radiology Open*, vol. 2, pp. 90–100, 2015.

[3] Y. Fujita, T. Goto, Y. Hamamoto, Y. Mitani, M. Segawa, S. Terai, and I. Sakaida, "A liver cirrhosis detection method using probabilistic ROI combination," *Proc. of 11th International Conference on Quality Control by Artificial Vision*, pp. 81–85, 2013.

[4] Y. Fujita, Y. Mitani, Y. Hamamoto, M. Segawa, S. Terai, and I. Sakaida, "Training sample selection based on self-training for liver cirrhosis classification using ultrasound images," *Proc. of 13th International Conference on Quality Control by Artificial Vision*, 2017.

[5] K. Sasaki, Y. Mitani, Y. Fujita, Y. Hamamoto, and I. Sakaida, "A liver cirrhosis classification on b-mode ultrasound images by the use of higher order local autocorrelation features," *Proc. SPIE10225, 8th International Conference on Graphic and Image Processing*, 102250U, 2017.

[6] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams, "Learning representations by back propagating errors," *Nature*, vol. 323, no. 9, pp. 533–536, 1986.

[7] Geoffrey E. Hinton, Simon Osindero, and Yee-Whye Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[8] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning, *Nature*," vol. 521, no. 28, pp. 436–444, 2015.

[9] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M van der Laak, Bram van Ginneken, and Clara I. Sanchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017.

[10] S. J. Raudys and A. K. Jain, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners and open problem," *Proc. 10th International Conference on Pattern Recognition*, pp. 417–423, 1990.

[11] S. J. Raudys and A. K. Jain, "Small sample size problems in designing artificial neural networks," *Machine Intelligence and Pattern Recognition*, vol. 11, pp. 33–50, 1991.

[12] Richard O. Duda, Peter E. Hart, David G. Stork, "Pattern classification," Second Edition, Wiley Interscience, 2001.

[13] T. M. Cover and P. E. Hart, "Nearest neighbor pattern classification," *IEEE Trans.*, IT-13, pp. 21–27, 1967.

[14] J. C. Russ, "The image processing handbook," Third Edition, CRC Press, 1990.

[15] R. Fisher, S. Perkins, A. Walker, and E. Wolfart, "Adaptive thresholding," 2003, https://homepages. inf.ed.ac.uk/rbf/HIPR2/adpthrsh.htm , Last accessed on 2019-7-11.

[16] K. Fukunaga, "Introduction to statistical pattern recognition," Second Edition, Academic Press, 1990.

[17] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," *Proc. Advances in Neural Information Processing Systems*, vol. 27, 2014.

[18] Alec Radford, Luke Metz, and Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *Proc. 2016 International Conference on Learning Representations*, 2016

Yusuke Fujita is currently an Associate Professor at Yamaguchi University, Japan. His research interests include machine learning, image processing, pattern recognition and computer vision. Ph. D.

Yoshihiko Hamamoto is currently a Professor at Yamaguchi University, Japan. His research interests include pattern recognition. His research area is informatics, intelligent informatics. He is a member of IEEE. Ph. D.

Isao Sakaida is currently a medical doctor and a Professor at Yamaguchi University, Japan. His research interests include liver disease. His research area is clinical internal medicine, gastroenterology.

Yoshihiro Mitani is currently a Professor at National Institute of Technology, Ube College, Japan. His research interests include pattern recognition and image processing techniques. He is a member of IEEE. Ph. D.

Robert B. Fisher is currently a Professor at the University of Edinburgh, UK. His research interests include computer vision. The main topics of his recent research are: vision for human-robot interaction, and outdoor robotics. He is the leader of the TrimBot2020 gardening robot project.