Classification of pathological cases of myocardial infarction using Convolutional Neural Network and Random Forest

Jixi Shi[1,2], Zhihao Chen[2], and Raphaël Couturier[2]

[1] Normandie University, UNIROUEN, ESIGELEC, IRSEEM, 76000 Rouen, France
jixi.shi@groupe-esigelec.org
[2] FEMTO-ST Institute, UMR 6174 CNRS, Univ. Bourgogne Franche-Comté, 90000 Belfort, France

Abstract. Myocardial infarction is one of the most common cardiovascular diseases. Clinical information and Delayed Enhancement cardiac MRI (DE-MRI) [8] are crucial for pathological diagnosis of myocardial infarction. However, discordance can be present between clinical characteristics and DE-MRI for the disease diagnosis. To efficiently learn the correlation between these data the automatically classify if a patient has myocardial infarction, we propose a mixed classification model that take both the clinical characteristics and DE-MRI into account. In the mixed model, the 3D Convolutional Neural Network (CNN) encodes the MRI as the surface of infarction then the surface is feed to Random Forest with other clinical characteristic to make the final decision.

Keywords: MRI · myocardial infarction · CNN · random forest.

1   Introduction

Myocardial infarction has been one of the most common cardiovascular disease. A clinical diagnosis of physiology data followed by the Delayed Enhancement MRI (DE-MRI) has become one clinical routine to identify the myocardial infarction. Clinical diagnosis including the troponin level from the blood test and the ST segment from electrocardiography etc. provides a quick evaluation result for the acute myocardial infarction. For suspected patient that clinical information reveal the possibility of myocardial infarction, a DE-MRI test is organized to get a more robust diagnose.

EMIDEC is a challenge in the workshop STACOM, co-organized with MICCAI 2020. On account of the above-mentioned clinical application, EMIDEC proposed a challenge that classifies if a patient has myocardial infarction according to the clinical physiology data and DE-MRI. In response to the challenge, we propose a two-stage method that interprets the semantic information from both the physiology data and MRI to strengthen the classification robustness.

As the first stage of our proposal, a 3D Convolutional Neural Network (CNN) predicts the volume of the myocardial infarction from the DE-MRI. The predicted volume is then merged with the physiology features and makes the final classification by Random Forest in the second stage.

The paper is organized as follows: introduction of the data including the clinical features and the DE-MRI; presentation of our method followed by the comparative results; discussion and conclusion.

2 Data

The dataset for EMIDEC challenge [4] consists of 150 exams including DE-MRI associated with clinical physiology data. The 100 first exams are publicly available for both the clinical data and ground truth, the rest of 50 cases only provide with the clinical data during the challenge. The purpose of the data is to diagnose if a patient has acute myocardial infarction. In the entire dataset, 1/3 cases are normal and 2/3 cases are pathological. Challenge organizers keep this unbalance distribution to better reproduce the real clinical condition at the cardiac emergency department.

The clinical information includes 12 features. Each feature's value is either categorical, Boolean or float. DE-MRI exams include on average 7 slices per case. A DE-MRI study includes the entire left ventricle myocardium from apex to base. The pixel spacing between slices (axis Z) is larger than on the plan X-Y. Manual segmentations are provided to challengers for the training set. Each segmentation consists of the normal and the pathological cardiac tissues including cavity, myocardium, myocardial infarction and microvascular obstruction (MVO, or no-reflow, a subclass of infarction). A patient that contains infarction in its DE-MRI is thought pathologically positive.

3 Method

The DE-MRI is 3D array but the clinical information includes 12 pieces of 1D feature. In order to apply the data that have different dimension and semantic information for the classification of myocardial infarction, our proposal contains two stages: the encoding of DE-MRI and classification on the fusion of encoded images and their paired clinical features as showed in Figure. 1. The image encoding is realized by a 3D CNN and the final classification is done by Random Forest [7]. This conception aims to take the advantage of the correlation between the two types of data and ensure that the classification based on both the kinds of data is more robust than on one single type.

3.1 Surface Regression by CNN

We firstly propose a 3D CNN to encode the MRI. Since each MRI case could have different slice number, to ensure that each CNN's input has a fixed size, the data augmentation was firstly investigated. To match up each three-slice input, an optimized 3D CNN inspired by U-Net [5] was then proposed.
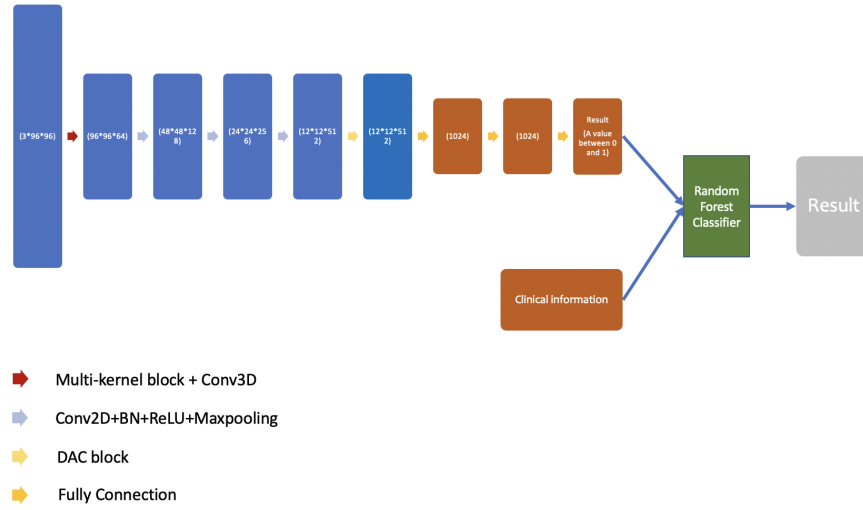
Fig. 1. Overview of our proposed architecture of prediction of combining the CNNs and random forest algorithm
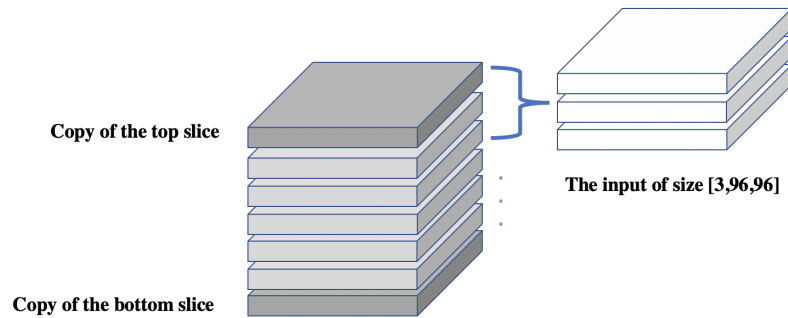
Fig. 2. Sub-sampling the image for the formed input of neural network

Data augmentation In order to fully catch semantic information from the DE-MRI, We made a data augmentation for the original MRI cases. To keep the information between adjacent slices and ensure the fixed-size of CNN's inputs, 3 successive slices were taken as a single 3D input. The first and the last slice were copied at the bottom and top side. Assuming that the original MRI has N slices, we got N new 3D images that each 3D image was formed by three adjacent slices in the augmented MRI. Known that the left ventricle myocardium is centred on the middle of each slice, to reduce the background's size, we did a centre cropping on each slice of size (96,96). Therefore, each CNN's input has the shape of (3,96,96). Figure. 2 illustrates the way that we created the three-slice inputs.

3D multi-kernel convolution block The 3D convolution was added only at the first layer in our CNN. We used a 3D convolution layer of multi-kernel to expand the receptive field. Multiple 3D convolution kernels whose sizes are (3,3,3), (3,5,5) and (3,7,7) encode the input image parallelly at the same layer. This conception is similar to Inception structures [6] and the objective is to flexibly extract features for objects with various sizes. For the 3D convolutions, padding was added only to the width and length dimension in light of the thickness of 3 at the dimension Z. Hence, the feature maps generated by each 3D kernel are in the same size of (96,96).

After we got the 2D feature maps, refferring to ResNet [2], we adopted the residual module on the 2D feature maps obtained by the previous 3D convolutions. Each residual module performed 4 times of convolution + down-sampling + batch normalization [3] + ReLU activation on feature maps. To reinforce the semantic information interpretation, the dense atrous convolution (DAC) block was added at the last layer before the fully connected layer, motivated by the Inception-ResNet-V2 block [] and atrous convolution []. DAC has four cascade branches with the gradual increment of the number of atrous convolution, from 1 to 1, 3, and 5. Therefore, the network can extract high-level semantic information from different scales.

At the end of the CNN, surface of the pathological tissue was obtained through the 2 fully connected layers as the CNN's final output. Surfaces of one MRI exam are cumulated as the regressive volume of pathological tissue.

3.2 Random Forest Classifier

The first stage we got a score of prediction of the surface of the pathological tissue, meanwhile, we expect to combine the clinical information and the score for precise classification, so we designed a random forest classifier as the second layer. Random forest is an algorithm for classification developed by Leo Breiman [1] that uses an ensemble of classification trees. Each of the classification trees is built using a bootstrap sample of the data, and at each split the candidate set of variables is a random subset of the variables. Thus, random forest uses both bagging (Bootstrap Aggregation), a successful approach for combining unstable

learners, and random variable selection for tree building. After the regression of first stage, we get a score of prediction of the images, we used the result of the first layer and the other 12 clinical information, in total 13 features as the inputs, to trained our random forest model, the output is set as 0 or 1 which indicate the case is normal one or pathological one.

4   Implementation Details

The CNN network is implemented by Pytorch. The image is preprocessed before network training. The MRI image is composed of multiple images. We divide the dataset into five parts, and select four part for training the model of regression, one part left for testing. We performed the experiment for five times with each time the different data for testing as a cross-validation. We set 500 epochs of training, and record the result of classification of each time. We also performed a test for single stage method of only using clinical information and only using the images for comparing.

5   Result

We evaluated our proposed framework on the MICCAI Emidec 2020 Challenge. We performed a five-fold cross-validation on the publicly available training set consisting of 100 annotated images to evaluate the method of proposed approach. The results of the five-fold cross-validation are as follows: the model of the random forest only using the pathological information represent a accuracy of $87\% \pm 3\%$, the model of the CNN only using the DE-MRI represent a accuracy of $91\% \pm 2\%$, and the model of random forest combined with the features of CNN represent a accuracy of $95\% \pm 3\%$

6   Conclusion

The classification of myocardial infarction and normal patients through pathological data can provide very effective reference information for diagnosis. In this article, we proposed a two-stage machine learning framework to make the classification through the clinical information and DE-MRI exams. We use CNN to extract the deep features from the images, and random forest to combine the image features and additional information. Our method shows a result of an accuracy improvement on the training datasets compared to the method of single stage method, the structure of CNN of the first stage can be applied to image data with information in depth. It is believed that the approach is a general one and can be applied to other type of disease diagnosis processing.

References

1. Breiman, L.: Random forests. Machine learning **45**(1), 5–32 (2001)

6

2. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
3. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
4. Lalande, A., Chen, Z., Decourselle, T., Qayyum, A., Pommier, T., Lorgis, L., de la Rosa, E., Cochet, A., Cottin, Y., Ginhac, D., Salomon, M., Couturier, R., Meriaudeau, F.: Emidec: A database usable for the automatic evaluation of myocardial infarction from delayed-enhancement cardiac mri. Data **5**(4) (2020)
5. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. vol. 9351, pp. 234–241 (10 2015)
6. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.: Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv preprint arXiv:1602.07261 (2016)
7. Than, M.P., Pickering, J.W., Sandoval, Y., Shah, A.S., Tsanas, A., Apple, F.S., Blankenberg, S., Cullen, L., Mueller, C., Neumann, J.T., et al.: Machine learning to predict the likelihood of acute myocardial infarction. Circulation **140**(11), 899–909 (2019)
8. Vogel-Claussen, J., Rochitte, C., Wu, K., Kamel, I., Foo, T., Lima, J., Bluemke, D.: Delayed enhancement mr imaging: Utility in myocardial assessment1. Radiographics **26**, 795–810 (05 2006). https://doi.org/10.1148/rg.263055047