



Classification of Pathological Cases of Myocardial Infarction Using Convolutional Neural Network and Random Forest

Jixi Shi^{1,2}(✉), Zhihao Chen¹, and Raphaël Couturier¹

¹ FEMTO-ST Institute, UMR 6174 CNRS, Univ. Bourgogne Franche-Comté,
90000 Belfort, France

² ESIGELEC, UNIROUEN, Univ. Normandie, 76000 Rouen, France
jixi.shi@groupe-esigelec.org

Abstract. Myocardial infarction is one of the most common cardiovascular diseases. Clinical information and Delayed Enhancement cardiac MRI (DE-MRI) are crucial to diagnose the myocardial infarction. However, some discrepancies can occur between clinical characteristics and DE-MRI when the disease is diagnosed. In order to deal in an efficient way with the correlation between these data and to be able to automatically classify patients suffering from myocardial infarction, this paper proposes a mixed classification model that takes both the clinical characteristics and DE-MRI into account. In the mixed model, a 3D Convolutional Neural Network (CNN) encodes the MRI as the surface of infarction then the surface is fed with Random Forest and other clinical characteristics to make the final decision.

Keywords: Classification · MRI · Myocardial infarction · CNN · Random Forest

1 Introduction

Myocardial infarction has become one of the most common cardiovascular disease. A clinical diagnosis of physiological data followed by the Delayed Enhancement MRI (DE-MRI) exam [9] has become the clinical routine to identify myocardial infarction. Clinical diagnosis including, for example the troponin level from the blood test and the ST segment from the electrocardiography provides a quick evaluation result for the acute myocardial infarction. For patients likely to suffer from this pathology and whose clinical information reveals the possibility of myocardial infarction, the DE-MRI exam is then organized to get a more robust diagnosis.

This paper responds to the call of EMIDEC, a challenge of the workshop STACOM, co-organized with MICCAI 2020. One of the EMIDEC challenge's contest consists in classifying if a patient has myocardial infarction according to

the clinical physiological data and DE-MRI. To achieve this goal, a two-stage method is proposed which interprets the semantic information from both the physiological data and MRI to strengthen the classification robustness.

In the first stage of the 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 physiological features and takes part in the final classification through 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 methods followed by comparative results; discussion and conclusion.

2 Data

The dataset for the EMIDEC challenge [5] consists of 150 exams including DE-MRI associated with clinical physiological data. The training set consisting of 100 exams were publicly available for both the clinical data and the ground truth several weeks before the day of challenge. On the day of the challenge, 50 other cases were provided only with clinical data as the test set. The purpose of the data is to diagnose if a patient has acute myocardial infarction. In the entire dataset, 1/3 of the cases are normal and 2/3 of the cases are pathological. Challenge organizers chose to keep this unbalanced distribution to better reproduce the real clinical conditions of the cardiac emergency department.

The clinical physiological information includes 12 features. Each feature's value is either categorical, Boolean or floating point number. 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 the one on the X-Y plan. Manual segmentations were provided to challengers for the training set. Each segmentation consists of the normal and the pathological cardiac tissues including the cavity, myocardium, myocardial infarct and microvascular obstruction (MVO, or no-reflow, a subclass of the infarct). A patient that shows signs of infarction in its DE-MRI is considered to be pathologically positive.

3 Method

Each DE-MRI exam has multiple slices of 2D image but the clinical physiological information includes 12 pieces of 1D feature. In order to handle these data that have different dimensions and semantic information for the classification of the myocardial infarction, the proposal contains two stages as showed in Fig. 1: the encoding of DE-MRI then the classification on the fusion of encoded images and their paired clinical physiological features. The image encoding is realized by a 3D CNN and the classification of the myocardial infarction is done by Random Forest [8]. This conception aims at taking the advantage of the correlation between both types of data so that the classification result is more robust than on one single type of data.

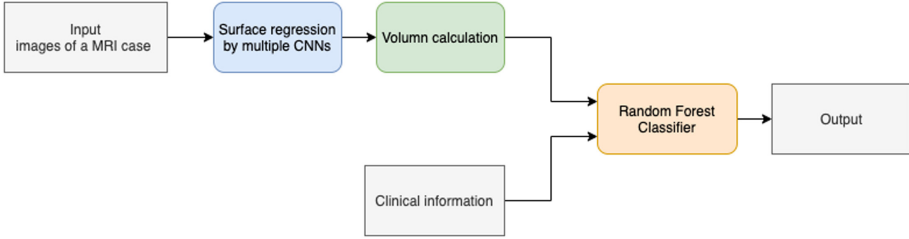


Fig. 1. Overview of our proposed architecture of prediction combining CNNs and Random Forest. A set of three-slice images come from the same MRI case are fed successively to the CNN. The output of the CNN is the regression of the predicted infarct surface of the middle slice in the input. Volume is calculated based on the predicted surfaces and the provided voxel spacing. Predicted volume and the clinical physiological information are concatenated as the input of the Random Forest classifier for the final classification.

3.1 Surface Regression by CNN

First, a 3D CNN is proposed to encode the MRI. The input of the 3D CNN is a three-layer MRI and the output is the predicted surface of the infarct. Since each MRI case can have different numbers of slices, to ensure that each CNN’s input has a fixed shape, the image preprocessing was firstly investigated. To match up each three-slice input, an optimized 3D CNN inspired by U-Net [6] was then proposed.

3.2 Image Preprocessing

In order to fully catch semantic information from the DE-MRI, the image preprocessing was executed on each MRI case. To learn the spatial information between adjacent slices and ensure a fixed-size of CNN’s inputs, three successive slices were taken as a single 3D input for the CNN. Assuming that an original MRI has N slices, the first and the last slice of the MRI were copied at the top and at the bottom side, hence N new 3D images were obtained and each 3D image was formed by three adjacent slices. Knowing that in the dataset the left ventricle myocardium is centred on the middle of each slice, to reduce the background’s size, a centre cropping of size $(96, 96)$ was performed on each slice. Therefore, each CNN’s input has the same shape of $(3, 96, 96)$. Figure 2 illustrates the way that the three-slice inputs were created. A 3D input could have had more slices. However, three adjacent slices were sufficient to provide enough spatial information. With more slices, more bottom and top slices should be copied, which was not efficient for the surface regression.

3.3 CNN with 3D Multi-kernel Convolution Block

The 3D convolution was added only at the first layer in our CNN since each input had only three slices. To expand the receptive field, multiple 3D convolution

kernels of the size $(3, 3, 3)$, $(3, 5, 5)$ and $(3, 7, 7)$ encoded the input image in parallel. This conception was inspired by Inception structures [7] and the objective was to flexibly extract features for objects of various sizes. For the 3D convolutions, the zero padding was performed only at 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 $(96, 96)$.

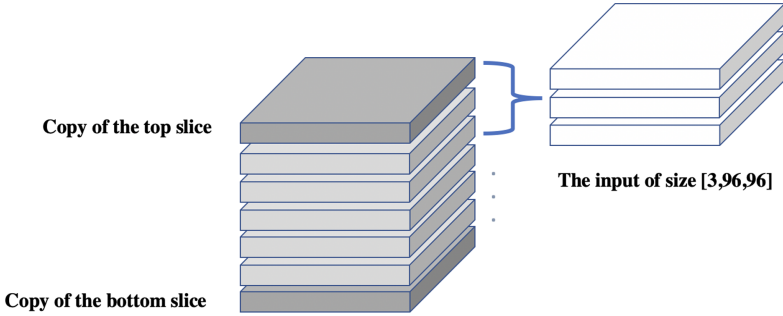


Fig. 2. Preparation of three-slice input for the 3D CNN. For each centre-cropped MRI case, firstly the top and the bottom slice are copied (darkgray). Secondly every three consecutive slices are chosen to form a 3D input for the CNN.

The 2D feature maps obtained by the 3D multi-kernel convolutional layer was then passed to residual modules which is similar to ResNet [3]. Each residual module performed 4 times of convolution + down-sampling + batch normalization [4] + ReLU activation on feature maps. To reinforce the semantic information interpretation, the Dense Atrous Convolution (DAC) block [2] was added at the last layer before the fully connected layers, motivated by the Inception-ResNet-V2 block [7] and atrous convolution. DAC had four cascade branches with the gradual increment of the number of atrous convolution, from 1 to 1, 3, and 5. Therefore, the network could extract high-level semantic information of different scales.

At the end of the CNN, the surface of the pathological tissue was predicted through the 2 fully connected layers as the CNN's final output. Smooth L1 loss was applied to penalize the error between the prediction and the ground truth (Fig. 3).

3.4 Volume Calculation

The output of the CNN is the predicted surface of the infarct in one MRI slice. In order to calculate the predicted volume of the infarct in one MRI case, the sum of the surfaces multiplying the pixel spacing (provided as the MRI meta data) was calculated. The predicted volume was used as an additional feature of the subsequent Random Forest model for the definitive classification of the myocardial infarction (Fig. 4).

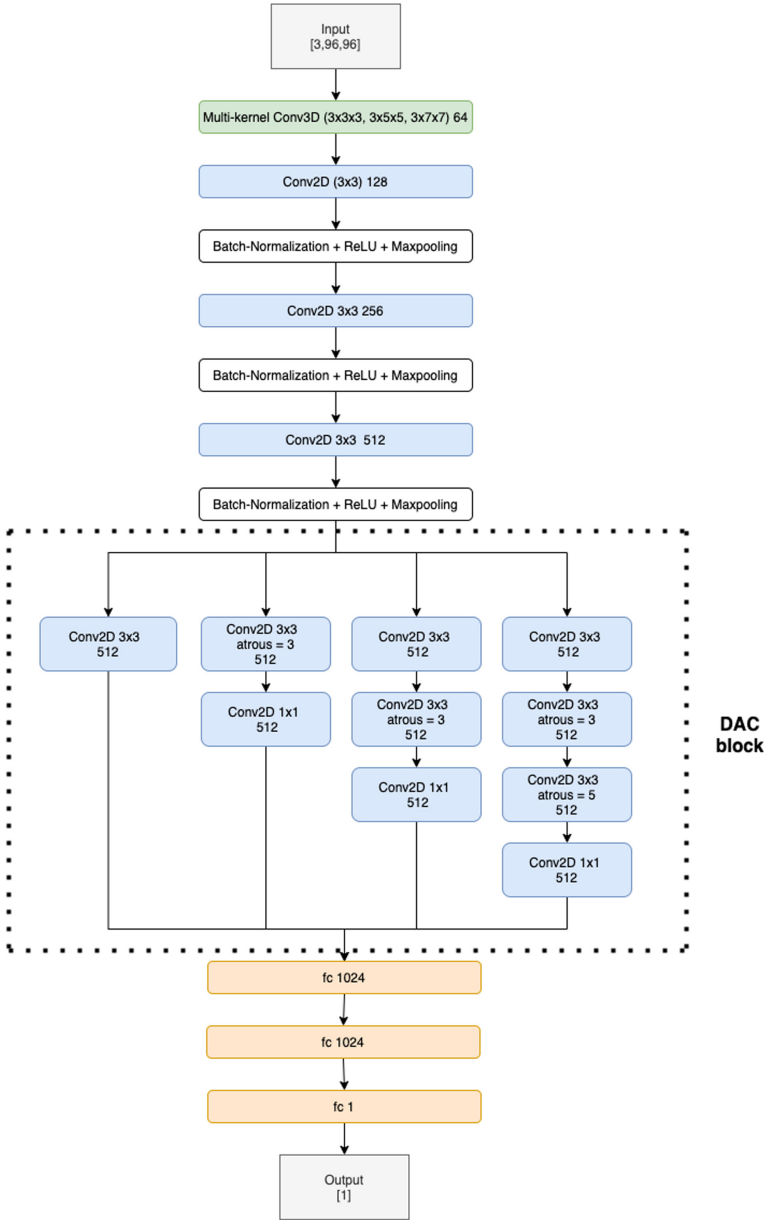


Fig. 3. The structure of the neural network. The 3D convolution is adopted only at the first layer. Before the fully connected layers, the DAC block enhances the cognition for both large and small areas.

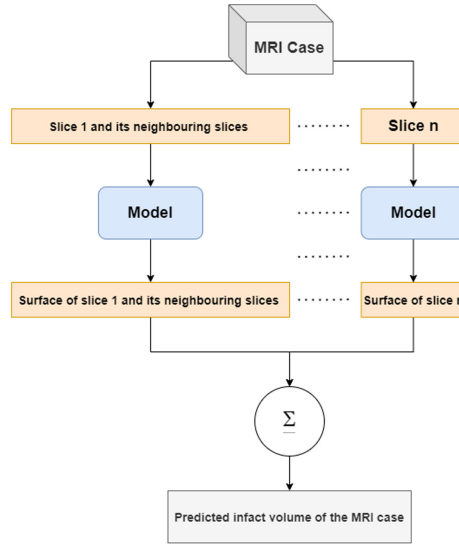


Fig. 4. The algorithm of volume calculation. The CNN predicts the surface of each slice. The predicted volume of a MRI case is calculated based on the predicted surfaces and the voxel spacing provided by the MRI meta data.

3.5 Random Forest Classifier

Random forest, developed by Breiman [1], is a classification algorithm that uses the 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.

As the predicted volume of infarct from the CNN was obtained during the first stage, at the second stage, the predicted volume was concatenated to the 12 clinical physiological features. The Random Forest was trained on these 13 features and the output was binary that indicated if the case was pathological or not.

4 Implementation Details

The CNN network was implemented in PyTorch. The CNN was trained 500 epochs and the predicted volume is the ensemble of multiple models’ prediction of different epochs. To show the advantage of the classification on merged MRI and clinical physiological information, the classifications made only on the CNN and only on Random Forest were also performed. The comparative tests used the same method from stage one or stage two and the data repartition of the cross-validation was identical to that of the above two-stage method, which can be considered as the baseline approaches.

5 Result

The proposed method and its comparative tests were evaluated on the dataset of MICCAI EMIDEC 2020 Challenge. A five-fold cross-validation was performed on the publicly available training set consisting of 100 annotated MRI exams and matched clinical physiological features.

Table 1 details the results of the cross-validation. The two-stage method achieved $95\% \pm 3\%$ accuracy, which is respectively 4% and 8% superior to the CNN only and Random Forest only approaches.

Table 1. Classification accuracy of three approaches

	Random Forest	CNN	Random Forest + CNN
Accuracy (%)	87	91	95
Standard deviation (%)	3	2	3

6 Conclusion

In this article, a two-stage machine learning framework for the myocardial infarction classification is proposed through the clinical physiological information and DE-MRI exam. A 3D CNN extracts the plane and spatial features from the images, and a Random Forest classifier combines the encoded image feature and physiological information to classify if the case is pathological or not. Our method shows a significant improvement of the classification accuracy on the training dataset compared to the single stage methods. Moreover, our two-stage approach could be applied for other types of disease diagnosis prediction containing significantly different data.

Acknowledgements. This work was partly supported by the ADVANCES project founded by ISITE-BFC project (number ANR-15-IDEX-0003) and by the EIPHI Graduate School (contract ANR-17-EURE-0002). We also thank the Mesocentre of Franche-Comté for the computing facilities.

References

1. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
2. Gu, Z., et al.: CE-Net: context encoder network for 2D medical image segmentation. *IEEE Trans. Med. Imaging* **38**(10), 2281–2292 (2019)
3. 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)
4. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. *Proc. Mach. Learn. Res.* **37**, 448–456 (2015)

5. Lalande, A., et al.: Emidec: a database usable for the automatic evaluation of myocardial infarction from delayed-enhancement cardiac MRI. *Data* **5**(4), 89 (2020)
6. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
7. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A.: Inception-v4, inception-ResNet and the impact of residual connections on learning. In: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, pp. 4278–4284 (2017)
8. Than, M.P., et al.: Machine learning to predict the likelihood of acute myocardial infarction. *Circulation* **140**(11), 899–909 (2019)
9. Vogel-Claussen, J., et al.: Delayed enhancement MR imaging: utility in myocardial assessment. *Radiographics* **26**, 795–810 (2006)