Cardiac Pathology Prediction - Challenge report

IMA205, 2023.
Julián ALVAREZ DE GIORGI,
julian.alvarez@telecom-paris.fr,
Télécom Paris
Palaiseau, France
May 2023

***Abstract*—This report presents a pipeline for diagnosing a healthy control or four cardiac pathology's, Myocardial infarction, Dilated cardiomyopathy, Hypertrophic cardiomyopathy, and Abnormal right ventricle. The diagnosis is done from two MRI images of the end of systole and the end of diastole and the height and weight of the patient.**

I. Introduction

This is a report on the work carried out on Challenge 2023 on Cardiac Pathology Prediction of the IMA205 course of Télécom Paris. The goal of this challenge was to classify MRI images of the heart among five different diagnostic classes:

- Healthy controls
- Myocardial infarction
- Dilated cardiomyopathy
- Hypertrophic cardiomyopathy
- Abnormal right ventricle

The classification was performed on patient metadata, and by extracting features from masks of the left ventricle cavity (LV), right ventricle cavity (RV), and myocardium (MYO). The masks are segmented from the MRI images, in this report it is not detailed how to extract all these masks since the segmentation of RV and MYO was already available, however in works such as [1], it is detailed how to segment all the masks of interest.

The report is organized as follows:

- II Proposed model: The final model, features used, hyperparameters, and accuracy obtained are detailed.
- III Implementation: Comment on the implementation.
- IV Segmentation: The left ventricular segmentation process is detailed and results are analyzed.
- V Features extraction and analysis: The process of extraction and selection of features is detailed.
- VI Model training: Details of the training of the model.
- VII Conclusion.

II. Proposed model.

The model resulting from this work is based on the following 23 features:

- Height.
- Weight.
- LV volume at end of diastole.
- LV volume at end of systole.
- RV volume end of diastole.
- RV volume end of systole.
- MYO volume end of diastole.
- MYO volume end of systole.
- Ejection fraction of LV:

$$EF_{LV} = (Volume_{LV_{ED}} - Volume_{LV_{ES}})/Volume_{LV_{ED}}$$

- Ejection fraction of RV:

$$EF_{RV} = (Volume_{RV_{ED}} - Volume_{RV_{ES}})/Volume_{RV_{ED}}$$

- Ratio between RV and LV at end of diastole.
- Ratio between RV and LV at end of systole.
- Ratio between MYO and LV at end of diastole.
- Ratio between MYO and LV at end of systole.
- Displacement over all axis (x, y, and z) of the RV mass center.
- Displacement over all axis (x, y, and z) of the LV mass center.
- Displacement over all axis (x, y, and z) of the MYO mass center.

A Random forest classifier was trained to be able to diagnose a new patient, the hyperparameters used for the RF model are n_estimators = 150 and max_depth = 6 (the other parameters were left as default).

With this model, an accuracy of 85.714% was obtained in the testing set.

III. Implementation

All the code was done in Python, in the *IMA 205 Challenge 2023 - Cardiac Pathology Prediction.ipynb* file is shown step by step and comments on the whole process of building the model, from segmentation of the LV to training the model. Then, almost all the used functions are implemented in the *Library.py* file.

Several libraries were used for this purpose, among which the following are worth mentioning, *nibabel* [4], *nilearn* [5], *scipy.ndimage* [3] y *openCV* [2].

IV. Segmentation

In this section, the problem to solve was the LV segmentation. In the training dataset, we already had all the masks (LV, RV, and MYO), however in the test dataset, we only had the RV and MYO masks, and the LV was segmented as background, therefore it was necessary to implement a function to segment the LV.

Due to the anatomy of the heart and the arrangement of the MRI image slices, for each slice (setting z-direction) the LV is always within the MYO, therefore, for segmenting the LV a function called *segmentation_LV* was implemented in *Library.py*. For a better understanding of the layout of each region of interest, we can see the image below. 1.

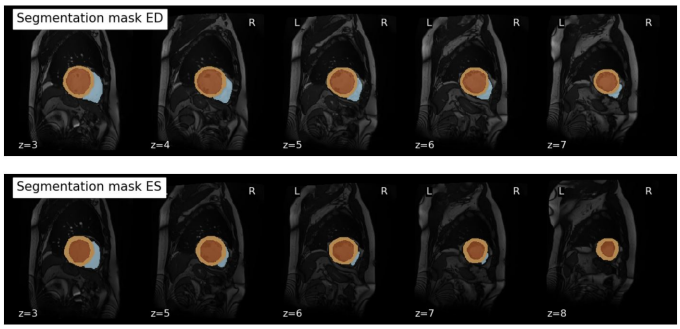

Figure 1: MRI images plus masks of different cuts

This function takes the dataset of all masks and segments LV for each patient. For this, first, the true background is identified and a value of 255 is assigned to it, this is achieved using the function *floodfill* of the OpenCV library [2]. The idea is to change to 255 all voxels that are 0 and are connected to the seed (0,0) (which always belongs to the true background). By implementation of *floodfill* it was necessary to do it for each slice because it works only for 2D images.

Finally, the part that remains as background (value 0) is taken and this corresponds to the LV.

In the following image 2, we can see the masks of any slice before segmenting the LV, after segmenting the LV, and the ground truth.

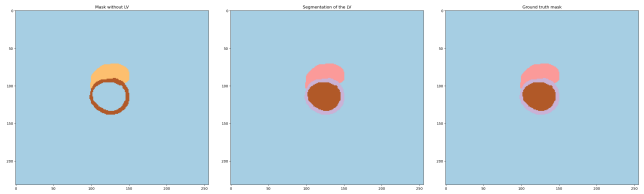

Figure 2: Maks of the segmentation process, without LV segmentation, with LV segmentation, and ground truth.

A. *Analysis of segmentation method results.*

To check that the method works well, the LV mask of the entire trainging dataset was set to background and the LV was segmented using the proposed method, then using the training dataset as ground truth the accuracy of the method was measured.

The Dice similarity coefficient was used as a metric to measure accuracy:

$$Dice = \frac{2(Area(S) \cap Area(GT)))}{Area(GT) + Area(S)} \quad (1)$$

Where:

- S is the segmented masks using the proposed method.
- GT is the ground truth mask segmentation.

The accuracy for all patients was 100%.

This analysis was performed on the notebook, *IMA 205 Challenge 2023 - Cardiac Pathology Prediction.ipynb*, and the results are shown there.

V. Features extraction and analysis.

In a first step, the features recommended by Khened, Mahendra and Alex, Varghese and Krishnamurthi, and Ganapathy in the article "Densely connected fully convolutional network for short-axis cardiac cine MR image segmentation and heart diagnosis using random forest" [1] were extracted, that is:

- Volume of the LV at the end of systole and the end of diastole phases.
- Volume of the RV at the end of systole and the end of diastole phases.
- Volume of the MYO at the end of systole and the end of diastole phases.
- Ejection fraction of the RV and LV.

Therefore, a function called *compute_volumes* was implemented. in *Libary.py*, to which the masks are passed as an argument and return the volumes of each mask. *compute_volumes*, counts for each patient the total number of voxels belonging to each class and multiplies them by the voxel volume (which varies in each MRI image).

Then, the function *compute_EF* in *Library.py* computes the ejection fraction.

With these features, the accuracy was 66,67%.

In order to improve this result, 4 features recommended in the paper [6] were added, the ratio between RV and LV volume at ED and ES, and the ratio between MYO and LV volume at ED and ES. These features were calculated directly in the notebook *IMA 205 Challenge 2023 - Cardiac Pathology Prediction.ipynb*

This improved the results to 86.67 percent.

To better understand this improvement, we can see the feature importance of these features 3 (was computed using cross-validation on the training set).

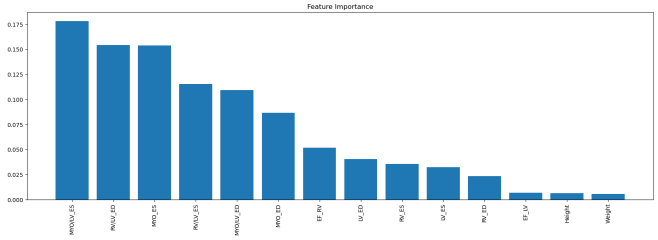

Figure 3: Feature Importance of the specified features

Then, in order to further improve the results, we tested adding the myocardium thickness, mentioned as important in several of the recommended bibliographies. For this, a function called *compute_thickness* was implemented in *Library.py*, which takes the widest thickness of each MRI image.

For this, the function *distance_transform_edt* was used from the *scipy.ndimage* [3] package, this function computes the distance from a foreground pixel to the nearest background pixel, therefore all the mask masks were set to the background (except MYO) and iterated for each slice of the MRI image and calculated the largest thickness.

Although, when performing the feature importance (see image 4) it appears as a relatively important feature, it did not improve the results. For this reason, it was not considered.

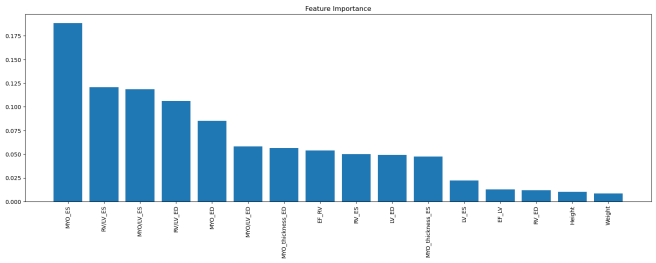

Figure 4: Feature Importance of the specified features

Later, it was considered that obtaining some displacement information between the diastole and systole of each of the sections of interest (LV, RV and MYO) could be interesting, so the center of mass of each of these in both stages (diastole and systole) was calculated and the displacement was calculated. Consequently, 9 more features were added, corresponding to the displacement in each coordinate (x,y and z) of each section (LV, RV and MYO).

This improved the results and an accuracy of 93,33% was obtained.

Once again, we performed a feature importance to see how important these new features are, see image 5.

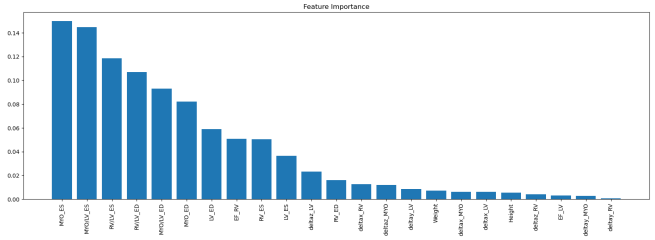

Figure 5: Feature Importance of the specified features

Observing that the results improved significantly and that these features were more important than some of the current ones (such as height and weight), it was decided to keep them.

We also tried adding the coordinates in each direction and for each region of interest but these did not add relevant information, as shown in the image below 6.

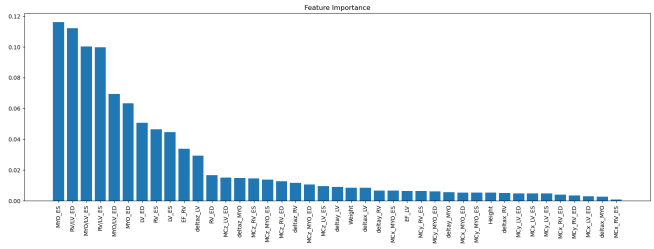

Figure 6: Feature Importance of the specified features

Finally, we show the box plots of all the features selected for the final model in the image 7.

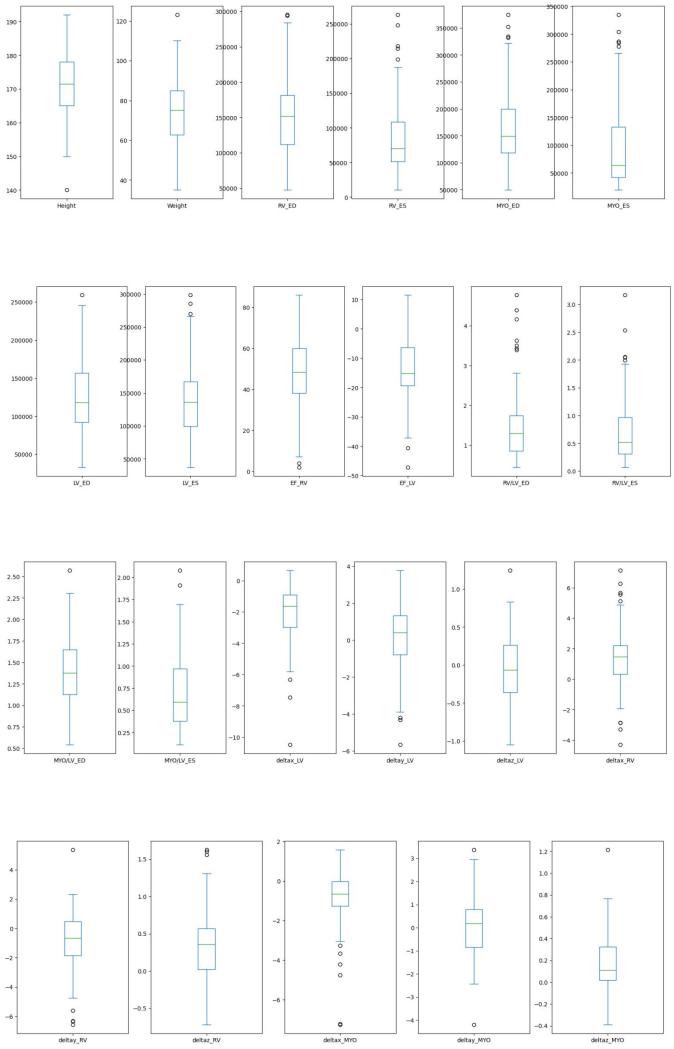

Figure 7: Box plot of all the features selected

From these box plots we see that there are some outliers but there are not many, these outliers may be the cases where the given feature is able to differentiate some class (heart disease). What surely allows us to affirm, is that we will need to make a standardization of the data to achieve that no feature is imposed before the others, if we want to test some other model besides random forest.

VI. Model training.

For the choice of the model, random forest was taken, since it was the one recommended by the available bibliography (such as [1]), and since it gave good results from the beginning.

First of all a standardization of the data was done, since the ranges were very different and some had negative values, and then a grid_searchCV was performed to find the best hyperparameters of n_estimators and max_depth. The results were as follows n_estimators = 150 and max_depth = 6.

As mentioned in the implementation section, the best parameters were such and such, yielding an accuracy of 85.714 percent.

VII. Conclusion

A method capable of diagnosing heart diseases with relatively high efficiency (85.714%) was implemented, putting into practice many concepts learned during the course, allowing one to learn how to adapt it to specific problems and enabling one to tangibly observe its capacity in real problem solving (as it was done in the PTs along the course).

Regarding the segmentation part, which was one of the most significant (not because of its complexity, but because of its importance for extracting all the features), the results obtained were very good (exactly the same as the ground truth when tested in the training dataset).

However, observing a posteriori and with the complete results of the predictions, there were decisions that were not necessarily the most appropriate, such as the fact of leaving out the thickness of the myocardium, although at the time of predictions, it did not improve the results, seeing the accuracy in the total test set it did, and the feature selection was giving us an indication that it was important, so the correct decision should have been to keep it.

Although other classification models such as Kernel-SVM, or applying PCA with all the mentioned features were tested, but not enough time was dedicated to achieve good results with these models, and for this reason, they were not mentioned in the report, dedicate more time to explore other models may also be interesting in the future.

In summary, although there were flaws, the overall result I think is good, and I am happy with it, I think I managed to identify the errors well and I know in which direction to continue if I want to improve the model even more.

References

[1] Mahendra Khened, Varghese Alex, and Ganapathy Krishnamurthi. "Densely connected fully convolutional network for short-axis cardiac cine MR image segmentation and heart diagnosis using random forest". In: *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers 8*. Springer. 2018, pp. 140–151.

[2] OPEN SOURCE COMPUTER VISION LIBRARY. *OpenCV*. URL: https://docs.opencv.org/4.6.0/.

[3] SCIPY LIBRARY. *SciPy ndimage*. URL: https://docs.scipy.org/doc/scipy/reference/ndimage.html.

[4] Chris Markiewicz Matthew Brett et al. *NiBabel*. URL: https://nipy.org/nibabel/.

[5] *Nilearn*. URL: https://nilearn.github.io/stable/index.html.

[6] Jelmer M Wolterink et al. "Automatic segmentation and disease classification using cardiac cine MR images". In: *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers 8*. Springer. 2018, pp. 101–110.