4IM05 Report: Kaggle Challenge

Prisha Chandiramani

May 2025

1 Introduction

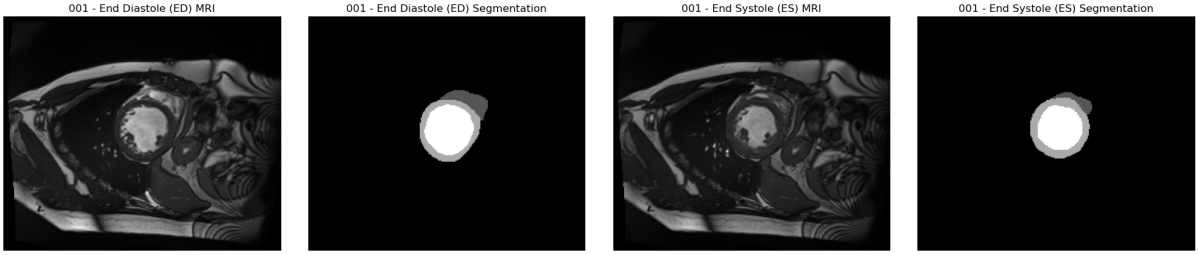
Segmentation of the heart in cardiac MRIs is clinically used to quantify cardiac function as MRI images of the heart at different time points contain very useful information for heart disease diagnosis. The challenge was to implement a method for segmentation of cardiac MRI and cardiac disease classification using machine learning methods. Patients can be classified as Healthy (HC), with Myocardial infarction (MI), with dilated cardiomyopathy (DCM), with hypertrophic cardiomyopathy (HCM) or an abnormal right ventricle (ARV). In order to do that, I was provided with a dataset of 150 subjects with their MRI images and, when available, their corresponding segmentations and metadata (subject height and weight). The data was randomly split into a training-validation set of 100 subjects and a test set of 50 subjects. For each subject, two images are provided at two different time points along the cardiac cycle: one image at end diastole (ED) and one image at end systole (ES). For the training-validation set MRI images, segmentations of the left ventricle (LV), right ventricle (RV) and myocardium (M) were given whereas for the test set MRI images, only segmentations of the right ventricle and myocardium were given. Therefore, as preprocessing of the test data, segmentation of the left ventricle was necessary. I used a Random Forest Model for classification.

The code related to the whole pipeline is in the jupyter notebook **final_notebook.ipynb**. The code related to the different segmentation methods used and the computation of the Dice Scores for each method is in the notebook **hand_segmentation.ipynb**.

2 Dataset Description

2.1 Train-Validation dataset

The train-validation dataset contains data for 100 subjects. For each subject, we have access to MRI images with the segmentation of the left ventricle (LV), right ventricle (RV) and myocardium (M) at end diastole (ED) and end systole (ES) (cf. Figure 1).



(a) Image from ED MRI of Patient 1 (b) Image from ED Segmentation of Patient 1 (c) Image from ES MRI of Patient 1 (d) Image from ES Segmentation of Patient 1

Figure 1: Example of images for the train-validation dataset

Alongside the MRI images, for each patient metadata of height and weight is provided (cf. Table 1).

Id	Category	Height	Weight
1	2	184.0	95.0
2	2	160.0	70.0
3	2	165.0	77.0

Table 1: The metadata provided for the first three patients

There are 20 patients per class.

2.2 Classes

The classification task aims to classify patients into 5 different categories:

0. Healthy Controls (HC)

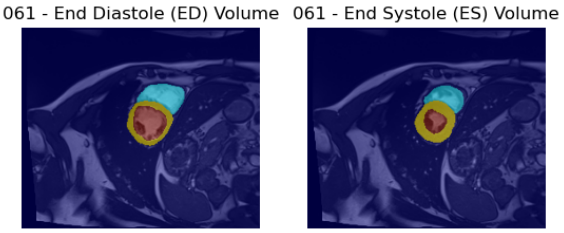
1. Myocardial Infarction (MI)

2. Dilated Cardiomyopathy (DCM)

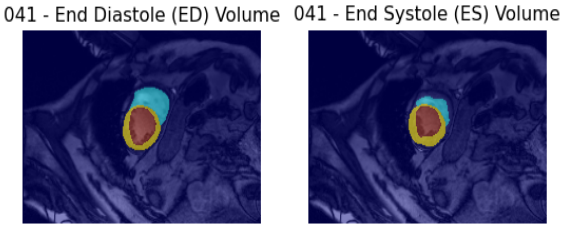
3. Hypertrophic Cardiomyopathy (HCM)

4. Abnormal Right Ventricle (ARV)

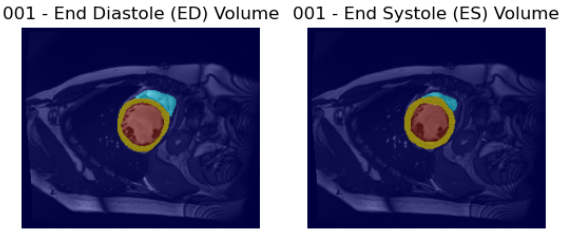
MRI images with segmentation for each class can be seen on Figure 2.



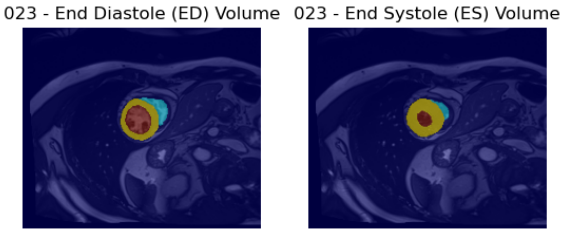
(a) MRI segmentation of a healthy patient



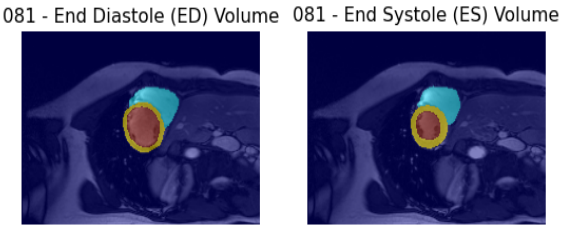
(b) MRI segmentation of a patient with myocardial infarction



(c) MRI segmentation of a patient with dilated cardiomyopathy



(d) MRI segmentation of a patient with hypertrophic cardiomyopathy

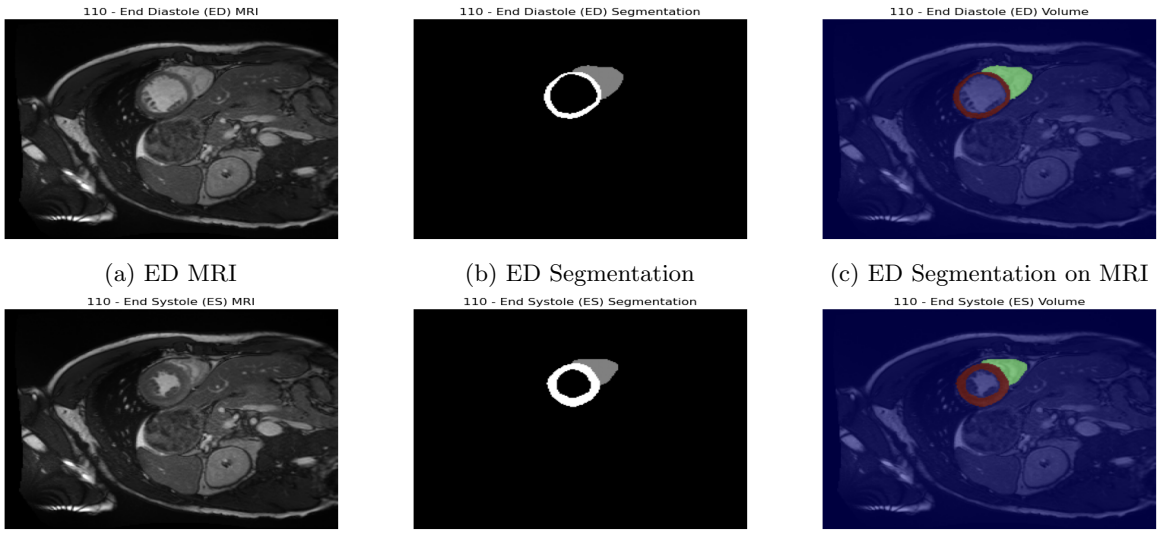


(e) MRI segmentation of a patient with an abnormal right ventricle

Figure 2: MRI images with segmentation for each class

2.3 Test dataset

The test dataset contains data for 50 subjects. For each subject, we have access to MRI images with the segmentation of the right ventricle (RV) and myocardium (M) at end diastole (ED) and end systole (ES). Examples of these images provided can be seen in Figure 3.



(a) ED MRI (b) ED Segmentation (c) ED Segmentation on MRI

(d) ES MRI (e) ES Segmentation (f) ES Segmentation on MRI

Figure 3: MRI segmentation provided for a patient in the test dataset

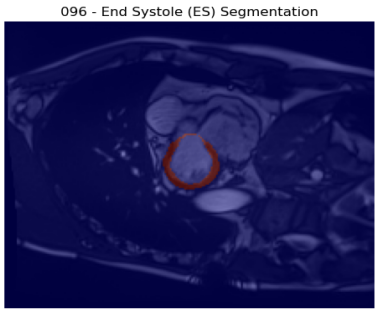
Metadata of height and weight is also provided for each patient. The left ventricle (LV) segmentation is not provided. Therefore, we have to find a good method to segment the left ventricle.

3 Segmentation

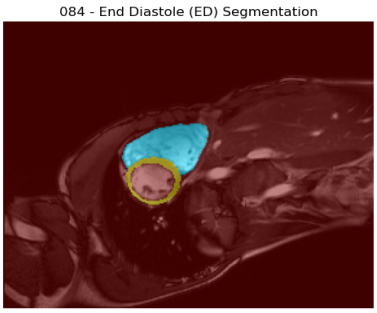
One main challenge of this Kaggle project is to do an accurate segmentation of the left ventricle for MRI images in the test dataset.
The main idea used to compute this segmentation is that the left ventricle is delimited by the myocardium. So, if we are able to fill the inside of the myocardium, we have the segmentation of the left ventricle. I decided to use morphology operations in order to compute the left ventricle's segmentation.

3.1 First method

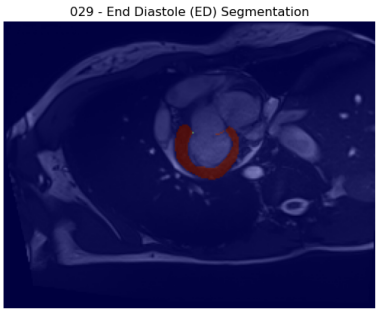
I started to use the FloodFill function from OpenCV. This function fills a connected region of an image with a specific color, starting from a seed point. I used as seed point the center of mass of the myocardium and then used the FloodFill function to fill the mask from the seed point until it reaches all borders of the myocardium. This enabled me to find the left ventricle mask. However, a few problems arised from this method. Sometimes the myocardium contour in the mask was too thin and other times the myocardium mask was open. Therefore, the FloodFill function did not give satisfactory results for the segmentation. Some images illustrating the problems faced can be seen in Figure 4.



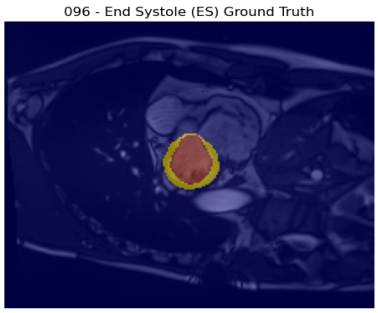
(a) Left Ventricle not filled



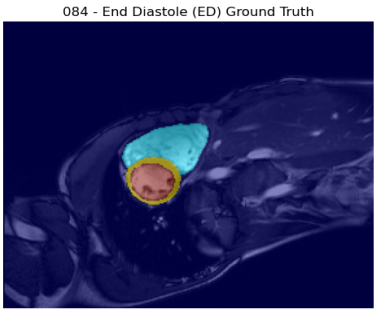
(b) Whole Background filled



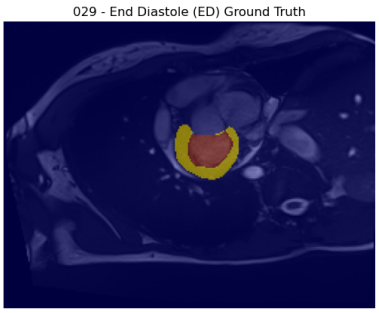
(c) Open Myocardium



(d) Ground Truth



(e) Ground Truth



(f) Ground Truth

Figure 4: Examples of wrong segmentations compared to the right one

To solve the first problem, I tried to use morphological closing on my binary mask, in order to have thicker boundaries. However, it wasn't a success, there were still a few images for which it did not perform correctly. The Mean Dice Scores for this method can be seen in Table 2.

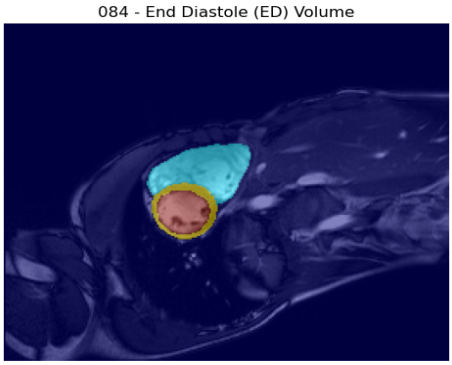
	HC	MI	DCM	HCM	ARV	**All categories**
ED	0.78	0.73	0.58	0.53	0.75	**0.67**
ES	0.78	0.73	0.58	0.54	0.76	**0.68**
Total	**0.78**	**0.73**	**0.58**	**0.54**	**0.76**	**0.68**

Table 2: The Dice Scores for each category at End Diastole and End Systole times

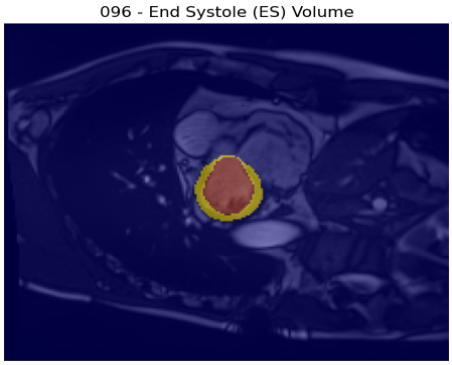
By looking at this table, we can see that the segmentation is not very accurate especially for categories 2 and 3. Therefore I decided to use another function that was easier to use and was more robust to solve this problem: the binary fill holes function.

3.2 Second method

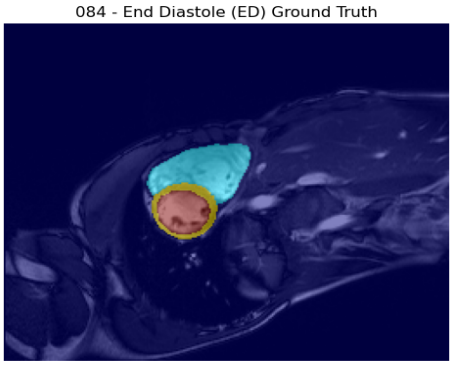
The binary fill holes function finds holes and fills them but is prone to the same problem: if the myocardium ring is not smooth, the left ventricle mask cannot be found. However, the binary closing operation, with a radius of 8, works well to solve the problem of "thin myocardium ring". An example of this improvement can be seen in Figure 5.



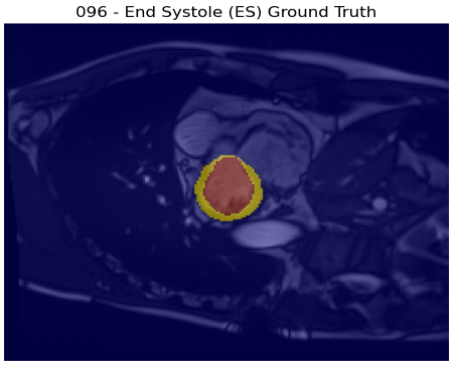
(a) Segmentation with binary_fill_holes



(b) Segmentation with binary_fill_holes



(c) Ground Truth



(d) Ground Truth

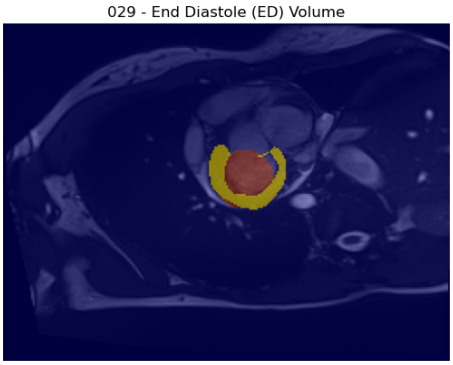
Figure 5: Improvements with the binary_fill_holes function

The Dice Scores given by this method can be seen in Table 3.

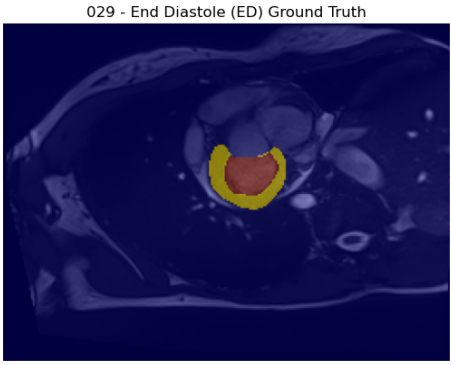
	HC	MI	DCM	HCM	ARV	**All categories**
ED	1	1	0.99	0.99	1	**0.99**
ES	1	1	0.99	0.99	1	**0.99**
Total	**1**	**1**	**0.99**	**0.99**	**1**	**0.99**

Table 3: The Dice Scores for each category at End Diastole and End Systole times

I, then, had to solve the second problem: compute the left ventricle mask when the myocardium is open. In order to do that, I had to first check if the myocardium in the image was closed or not. So I decided to create a bounding box around the myocardium in each image and use the fill holes function. If the filled mask is the same as the original one (not filled), it means that no holes were found and that the myocardium was open. In that case, I drew a circle that had for center, the center of mass of the myocardium and for radius, the mean of all distances from the center of mass to points on the myocardium. This filled circle was defined as the left ventricle mask. In figure 6 is shown an example of the segmentation computed when the myocardium is open.



(a) Segmentation with a bounding box



(b) Ground Truth Segmentation

Figure 6: Segmentation computed when the myocardium is open

With this method, I got the Dice Scores shown in Table 4.

	HC	MI	DCM	HCM	ARV	**All categories**
ED	0.99	1	0.99	0.99	1	**0.99**
ES	0.99	1	0.99	0.99	1	**0.99**
Total	**0.99**	**1**	**0.99**	**0.99**	**1**	**0.99**

Table 4: The Dice Scores for each category at End Diastole and End Systole times

3.3 Third method

I tried to used a third method to compute the segmentation of the left ventricle. It envolved using the openCV find contours and draw contours functions in order to draw the contour of the myocardium from which I would deduce the left ventricle mask delimited by the contour. When computing the dice scores, I would get a very good score of 99% on average. However, when computing the error in the computation of the left ventricle volume, I would get an error of 99% which I didn't understand and found quite strange. Therefore, I decided to use the second method for the segmentation of the left ventricle.

Now that the segmentation works well. The next step is feature extraction to train the machine learning model.

4 Feature Extraction

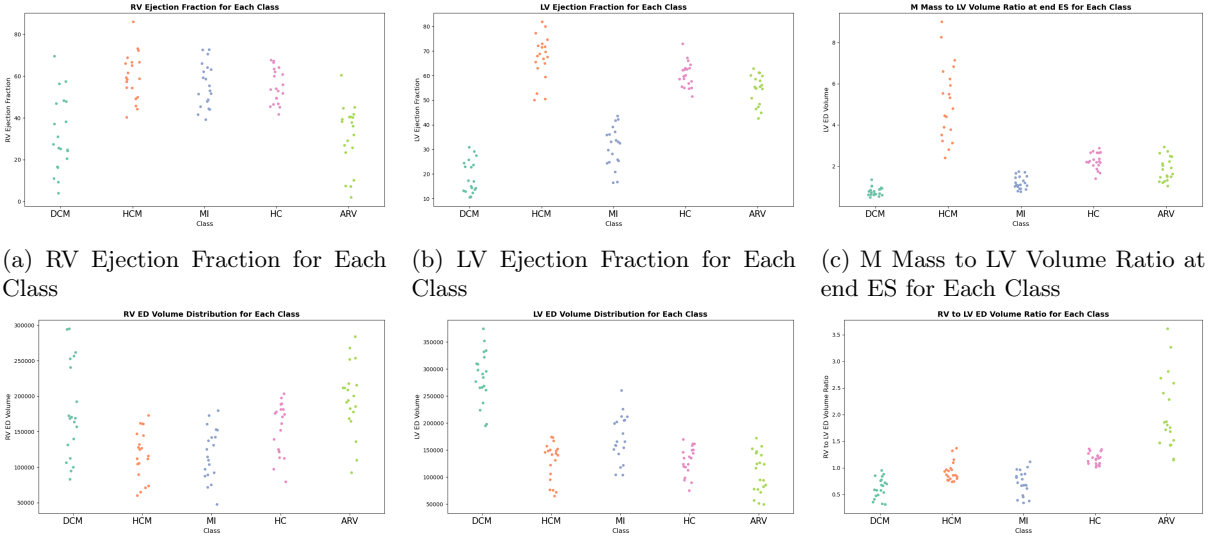
4.1 Feature Choice

The choice of the features used was based on the different papers provided. I used patient-based characteristics, **height** and **weight** as well as image-based characteristics. Quantification of volumetric changes in the heart during the cardiac cycle is essential for diagnosis and monitoring of cardiac diseases. This is why I extracted the following volumes, mass and ejection fractions: **volume of the left ventricle at end diastole and end systole**, calculated with an error of 0.26% at ED and 0.30% at ES after segmentation of the left ventricle, **volume of the right ventricle at end diastole and end systole**, **volume of the myocardium at end diastole and end systole**, **mass of the myocardium at end diastole**, **left ventricle ejection fraction** and **right ventricle ejection fraction**. These parameters are calculated routinely to diagnose a subject as healthy or diseased and were used in the paper "Densely Connected Fully Conventional Network for Short-Axis Cardiac Cine MR Image Segmentation and Heart Diagnosis Using Random Forest" to train a Random Forest model. I, thus, first trained my model with these. I got an accuracy of 93% on the public leaderboard. So I decided to add features. According to the paper"Automatic Segmentation and Disease Classification Using Cardiac Cine MR Images", the three most important features were the left ventricular ejection fraction, the ratio between the right and left ventricular volume at ED, and the ration between the myocardial and left ventricular volume at ES. This is why I decided to add as features the **volume ratio between right ventricle and left**

ventricle at end diastole and end systole and the **volume ratio between myocardium and left ventricle at end diastole and end systole**. I also added the **BSA** (Body Surface Area), computed with the Mosteller formula as stated in the paper "Automatic Cardiac Disease Assessment on cine-MRI via Time-Series Segmentation and Domain Specific Features", the **BMI** (Body Mass Index) and the **Normalized volumes and mass above by the BSA**. These features aim to reflect diagnostic clinical procedures.

4.2 Feature Distribution

According to the paper "Automatic Segmentation and Disease Classification Using Cardiac Cine MR Images": right ventricular abnormality patients generally have a large RV to LV volume ratio compared with healthy patients. Patients with dilated cardiomyopathy and heart failure with infarction have a reduced LV Ejection Fraction, while this value is higher for healthy patients and patients with hypertrophic cardiomyopathy. The myocardial volume is relatively small compared with the LV volume in patients with dilated caridomyopathy, indicating thinning of the myocardium, but large in patients with hypertrophic cardiomyopathy. And this information is verified by plotting these features for all classes, as shown in Figure 7.



(a) RV Ejection Fraction for Each Class

(b) LV Ejection Fraction for Each Class

(c) M Mass to LV Volume Ratio at end ES for Each Class

(d) RV ED Volume Distribution for Each Class

(e) LV ED Volume Distribution for Each Class

(f) RV to LV Volume Ratio for Each Class

Figure 7: Plots of Features' Distributions for Each Class

From these plots, we have a lot of information to discriminate each class. The Ejection fractions and the Myocardium Mass to Left Ventricle Volume Ratio are useful to detect the MI and HC classes ((a),(b),(c), Figure 7). The Myocardium Mass to Left Ventricle Volume ratio at end systole is very useful to detect the HCM class ((c) Figure 7). The Left Ventricle Volume Distribution ((e) Figure 7) is useful to detect the DCM class. The Right Ventricle to Left Ventricle Volume ratio ((f) Figure 7) is useful to detect the ARV class.
Therefore, the features chosen above are relevant for the classification task.

5 Classification model

For the classification task, I used a machine learning model. After reading the papers, there were two possibilities: Random Forest or a Multi Layer Perceptron. I decided to first train a Random Forest Model as I only had access to a small training set for which Random Forest would be more accurate and stable.

5.1 Random Forest Model

In the paper "Densely Connected Fully Convolutional Network for Short-Axis Cardiac Cine MR Image Segmentation and Hearth Diagnosis Using Random Forest", they used a 100 trees Random Forest model and in the paper "Automatic Cardiac Disease Assessment on cine-MRI via Time-Series Segmentation and Domain Specific Features" a 1000 trees, full depth Random Forest model. I decided to use the GridSearchCV function from scikit-learn to find the optimal hyperparameters for my model. To tune the model, I used a 5-fold cross-validation and the accuracy evaluation metric. I played on the following hyperparameters: number of trees, maximum depth, minimum number of samples required to split an internal node, minimum number of samples required to be a leaf node, number of features to consider when looking for the best split. The optimal model was found having an accuracy of 95% with 50 trees, 2 samples minimum to split a node, 1 sample minimum to be a leaf node, full depth and 'sqrt' function to determine the maximum number of features to consider when looking for the best split. On the public leaderboard, I got a 100% accuracy and on the private leaderboard an 85% accuracy.

5.2 Multi Layer Perceptron Model

I tried to train a Multi Layer Perceptron Model, to see if it allowed me to get a better accuracy. Inspired from the paper "Automatic Cardiac Disease Assessment on cine-MRI via Time-Series Segmentation and Domain Specific Features", my MLP's architecture consists of four hidden layers, each containing 32 units, followed by batch normalization, leaky ReLU nonlinearity and a Gaussian noise layer ($\sigma = 0.1$). I used the Cross Entropy Loss and the Adam Optimizer with a 0.001 learning rate. After training this model, I got a 97% accuracy but when submitting it to Kaggle challenge, It did not go above 86% of accuracy in the public leaderboard. I think that using so little data to train a Multi Layer Perceptron lead to overfitting and it was not the best model for this classification task.

The best results I got were with the Random Forest Model which is coherent with the size of training dataset we had at hand.

6 Features Importance Analysis

For the Random Forest Model, we get the feature importance analysis shown in Figure 8.

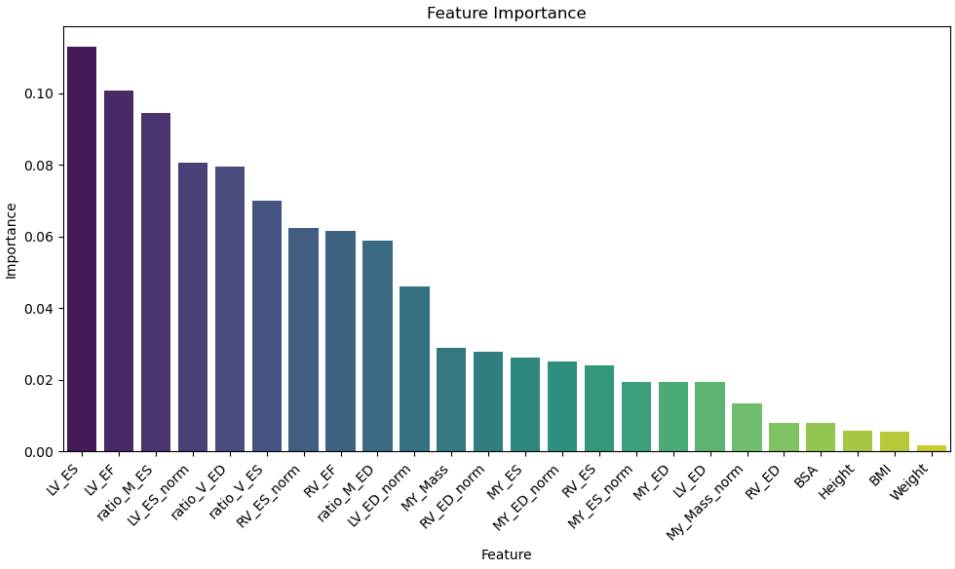
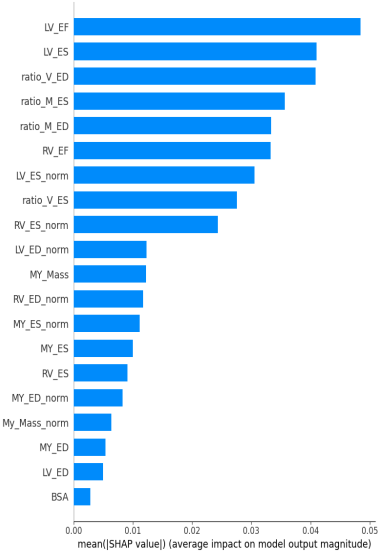


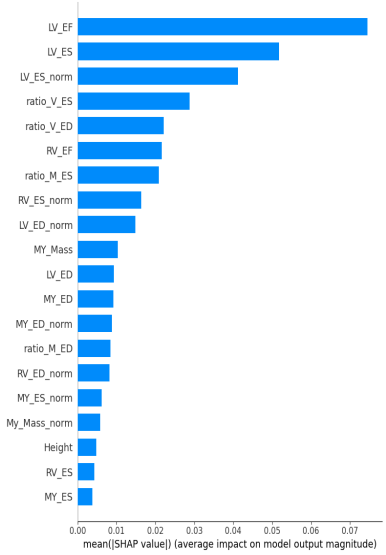
Figure 8: Feature Importance Analysis for the Random Forest Model

We can deduce from this analysis that the most important features for this classification task are the volume of the left ventricle at the end of the systole, the left ventricle ejection fraction, the ratio between the Mass of the Myocardium and the volume of the left ventricle at the end of the systole, the ratio of the volume of the right ventricle and the left ventricle at the end of the diastole and the systole and the right ventricle ejection fraction. We can say that the most important features are the ratios and the

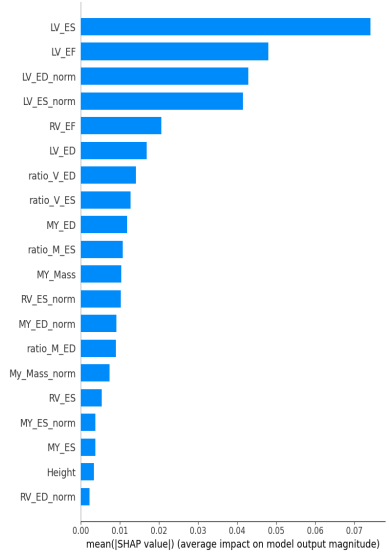
ejection fractions. This coincides with what was said in the different papers so our algorithm works as expected.
We can check the importance of each feature in our model for each class. These are plotted in Figure 9.



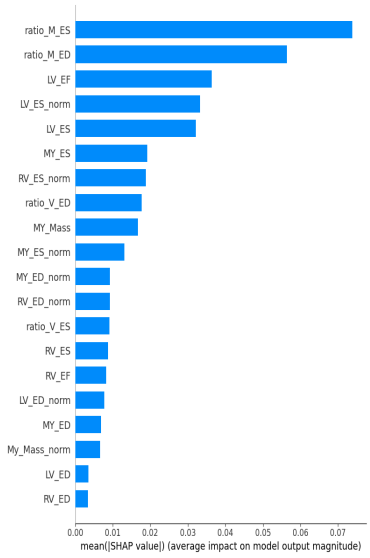
(a) Feature Importance Analysis for Healthy Controls (class 0)



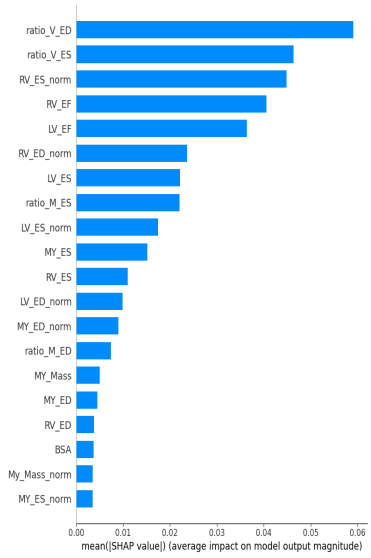
(b) Feature Importance Analysis for Myocardial Infarction (class 1)



(c) Feature Importance Analysis for Dilated Cardiomyopathy (class 2)



(d) Feature Importance Analysis for Hypertrophic Cardiomyopathy (class 3)



(e) Feature Importance Analysis for Abnormal Right Ventricle (class 4)

Figure 9: Feature Importance Analysis for Each Class

On Figure 9, we can see that the left ventricle ejection fraction is very important to determine both healthy patients and patients with Myocardial Infarction. The volume of the left ventricle at the end of the systole is also important. Moreover, the ratio between volumes of right and left ventricle at the end of the diastole is very important to classify correctly healthy patients. The volume of the left ventricle is the most important feature to classify patients with dilated cardiomyopathy. The left ventricle ejection fraction is also important for this class. The most important features for Hypertrophic Cardiomyopathy are the ratios between the mass of the myocardium and the volume of the left ventricle at end systole and end diastole and the ones to classify abnormal right ventricle patients are the ratios between the volume of the right ventricle and the left ventricle at end diastole and end systole. This analysis coincides with what is shown Figure 8.
We can also look at the correlation matrix Figure 10.

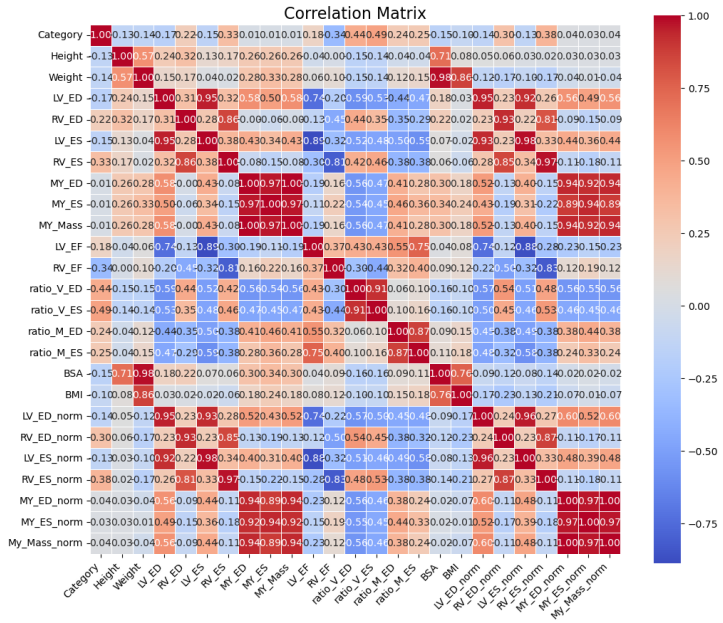


Figure 10: Correlation Matrix

We can see that the normed features are very correlated to the ones that are not normed.
Since many features are used by the model, from all the information above we can reduce the number of features by removing normed features and the BMI and we get the same accuracy. When training the Random Forest Model, I also tried to normalize the features using the MinMaxScaler from the library scikit-learn, but it did not improve the accuracy of the model.

7 Conclusion

The aim of the challenge was to implement a method for the segmentation of cardiac MRI and cardiac disease classification. After testing several techniques, the best segmentation method for the left ventricle found was the one using the morphology operation binary fill holes which computed the segmentation with a 99% Dice Score. It took 3.9 seconds to compute the segmentation of all masks in the test dataset. A Random Forest model was chosen to classify the data. After carefully choosing the input features of the model and fine tuning the hyperparameters, I was able to get 100% accuracy on the public leaderboard and 85% accuracy on the private leaderboard of the Kaggle challenge.

References

[1] Khened, M., Alex, V., Krishnamurthi, G. (2018). Densely Connected Fully Convolutional Network for Short-Axis Cardiac Cine MR Image Segmentation and Heart Diagnosis Using Random Forest. In: Pop, M., et al. Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges. STACOM 2017. Lecture Notes in Computer Science(), vol 10663. Springer, Cham. https://doi.org/10.1007/978-3-319-75541-0_15

[2] Isensee, F., Jaeger, P.F., Full, P.M., Wolf, I., Engelhardt, S., Maier-Hein, K.H. (2018). Automatic Cardiac Disease Assessment on cine-MRI via Time-Series Segmentation and Domain Specific Features. In: Pop, M., et al. Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges. STACOM 2017. Lecture Notes in Computer Science(), vol 10663. Springer, Cham. https://doi.org/10.1007/978-3-319-75541-0_13

[3] Wolterink, J.M., Leiner, T., Viergever, M.A., Išgum, I. (2018). Automatic Segmentation and Disease Classification Using Cardiac Cine MR Images. In: Pop, M., et al. Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges. STACOM 2017. Lecture Notes in Computer Science(), vol 10663. Springer, Cham. https://doi.org/10.1007/978-3-319-75541-0_11