Cardiac Pathology Prediction

Challenge IMA 205

Khalil BRAHAM

7 mai 2023

Table des matières

1 Problematic

Cardiac pathologies are among the leading causes of death worldwide, and early identification of these conditions is crucial for patient management and therapy decision-making. Several approaches have been proposed for automatic diagnosis from cardiac magnetic resonance imaging (CMRI), including non-invasive computer-aided diagnosis (CAD). The goal of this challenge is to classify MRI images of the heart among five different diagnostic classes, including healthy controls, myocardial infarction, dilated cardiomyopathy, hypertrophic cardiomyopathy, and abnormal right ventricle. In this report, we will describe our approach for feature extraction and classification algorithms used to accurately classify MRI images of the heart.

2 Data

The dataset provided for this challenge includes MRI images of the heart for 150 subjects, along with corresponding segmentations and metadata. The dataset has already been split into a training-validation set of 100 subjects and a test set of 50 subjects. We only have the classification made by clinicians of the training-validation set, and the goal of the challenge is to estimate the correct class of each subject in the test set. For each subject, two MRI images are provided : one at end diastole and one at end systole, each containing the heart and adjacent structures. The segmentation maps provided for the training-validation set consist of three substructures : left ventricle cavity, right ventricle cavity, and myocardium. The left ventricle cavity label is missing in the test set, and we may need to perform left ventricle segmentation as an intermediate step.

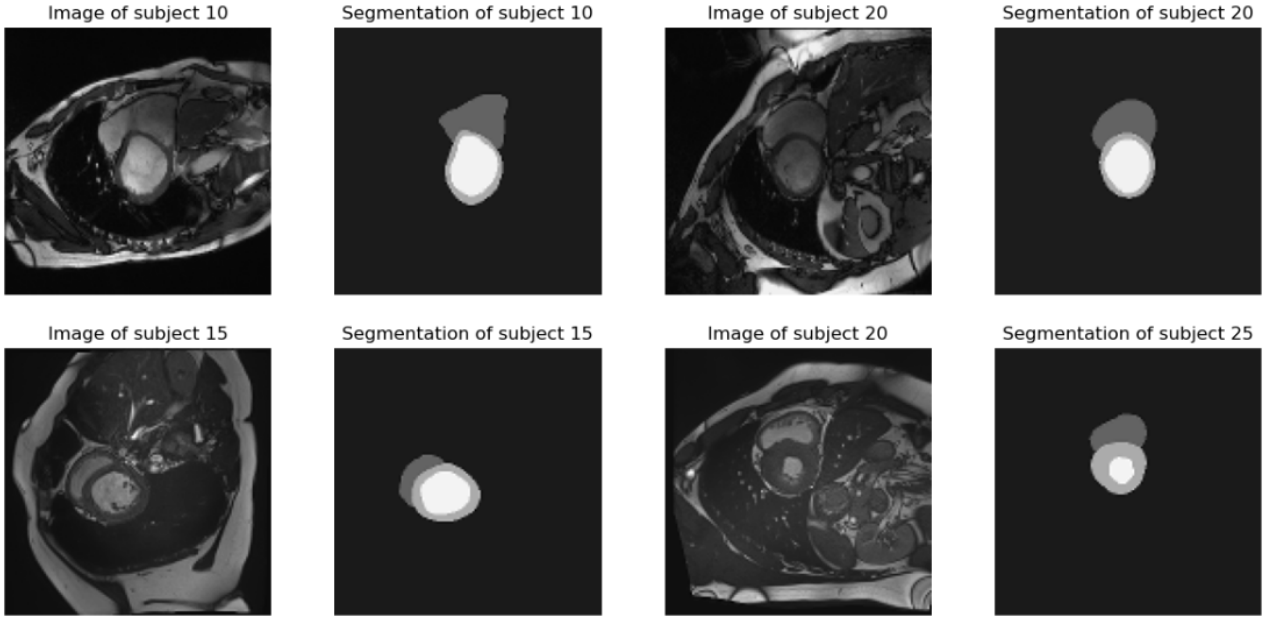


FIGURE 1 – Visualisation of exemples

3 Left Ventricle Segmentation

In the context of the cardiac MRI classification challenge, the left ventricle (LV) cavity segmentation is an essential step in extracting features that are relevant for classification. The LV cavity is a structure of interest as it allows the calculation of parameters such as the LV volume, mass, and ejection fraction, which are commonly used in clinical diagnosis and patient management.

3.1 Loss function

The dice coefficient is a widely used evaluation metric in medical image segmentation tasks. It is a similarity measure that ranges from 0 to 1, where 1 indicates a perfect overlap between the predicted and true segmentations. The dice coefficient can be interpreted as the harmonic mean of precision and recall, making it a balanced metric for imbalanced datasets. In the case of the left ventricle segmentation problem, where the left ventricle region is relatively small compared to the whole image, the dice coefficient can be especially useful because it can penalize false positives and false negatives equally.

In addition to being a good evaluation metric, the dice coefficient can also be used as a loss function to optimize the segmentation model. The dice loss is defined as 1 minus the dice coefficient and is commonly used in conjunction with cross-entropy loss. The combination of dice loss and cross-entropy loss has been shown to improve segmentation accuracy in medical imaging tasks compared to using cross-entropy loss alone. In the combined loss function, the hyperparameter alpha controls the weight given to the cross-entropy loss relative to the dice loss. By setting alpha to a value between 0 and 1, we can balance the importance of both loss functions and improve the overall segmentation performance.

Overall, the dice coefficient and its associated dice loss are well-suited for the left ventricle segmentation problem because they provide a balanced metric for evaluating segmentation accuracy and a loss function that can be used to optimize the segmentation model.

3.2 Optimizers

In addition to the loss function, the optimization algorithm or optimizer used in the model training process can also have a significant impact on the performance of the segmentation model.

The Adam optimizer is used with default settings to minimize the combined loss. Adam is a popular optimization algorithm that is well-suited for deep learning models, and its adaptive learning rate helps to efficiently converge to the optimal solution.

The EarlyStopping callback monitors the validation loss during training and stops training if the loss does not decrease for a specified number of epochs (in this case, 10).

The ModelCheckpoint callback saves the best model weights based on the validation loss, which can be used later for inference or further training.

The LearningRateScheduler callback adjusts the learning rate of the optimizer during training. The scheduler function defines a learning rate schedule that reduces the learning rate by a factor of 0.95 after the first 10 epochs of training. This helps the model to converge more slowly and to avoid getting stuck in local minima.

3.3 U-Net

3.3.1 Preview

One of the most popular neural network architectures for semantic segmentation tasks such as the LV cavity segmentation is the U-Net. U-Net is an encoder-decoder architecture that uses skip connections to combine feature maps from the encoder and decoder paths, allowing the model to recover fine-grained details while maintaining contextual information. U-Net has shown excellent performance in various segmentation tasks, including medical image analysis. Therefore, it was a natural choice for the LV segmentation task in the cardiac MRI classification challenge.

3.3.2 Architecture

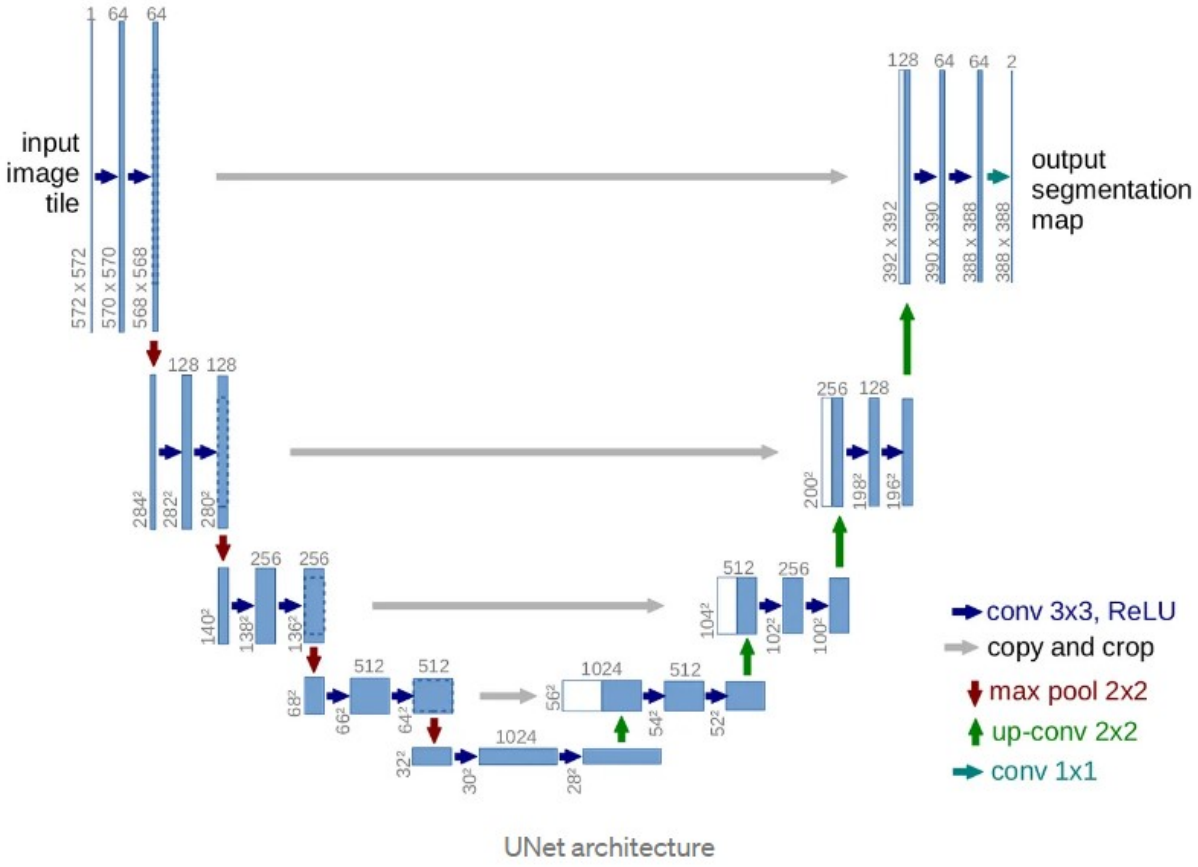


FIGURE 2 – UNet architecture

Contracting Path The contracting path comprises multiple layers of Convolutional Neural Network (CNN), followed by Batch Normalization and Rectified Linear Unit (ReLU) activation. The input layer accepts an MRI image of shape (128, 128, 1) by default. This path consists of four sets of convolutional blocks, where each block contains two Conv2D layers with the same number of filters, followed by MaxPooling2D for downsampling. The number of filters in the convolutional layers doubles after each MaxPooling2D layer.

Bottom Layer At the bottom of the U-Net, a fifth convolutional block with 512 filters is placed, followed by a Dropout layer with a dropout rate of 0.5.

Expanding Path The expanding path comprises a series of Conv2DTranspose layers for upsampling, followed by Concatenate layers to merge the upsampling layers with their corresponding residual connections from the contracting path. Subsequently, the convblock function is applied to generate new convolutional layers with batch normalization and ReLU activation. The expanding path consists of four sets of these layers.

Output Layer The output layer utilizes a Conv2D layer with filters equal to the number of classes (default : 4) and a softmax activation function for class probabilities.

Residual Connections The model uses residual connections to allow gradients to propagate more efficiently through the network. These connections involve adding the output of the original input layer to the output of the second Conv2D layer in each convblock.

3.3.3 Results

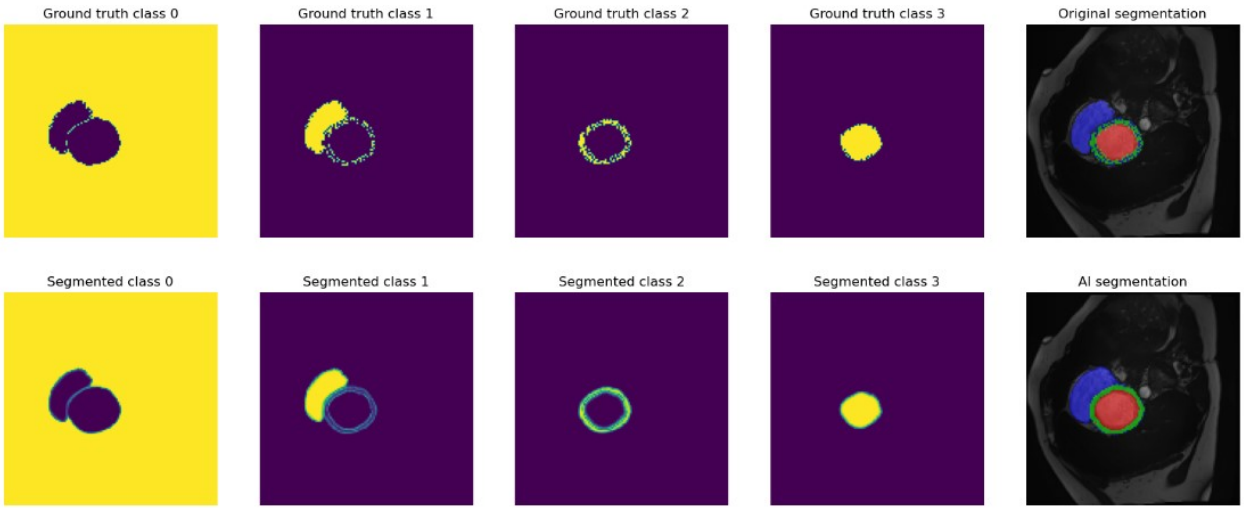


FIGURE 3 – UNet segmentation results

3.4 SegNet

3.4.1 Preview

The U-Net architecture has some limitations, such as its tendency to lose spatial information due to pooling operations in the encoder path that can be very important in our case. To address these limitations, SegNet was proposed as an alternative architecture for semantic segmentation in medical images. SegNet uses a similar encoder-decoder structure as U-Net, but instead of using skip connections, it uses max-pooling indices to upsample the feature maps in the decoder path. This approach allows SegNet to maintain the spatial information lost during the pooling operations in the encoder path, while also reducing the number of parameters and the risk of overfitting.

3.4.2 Architecture

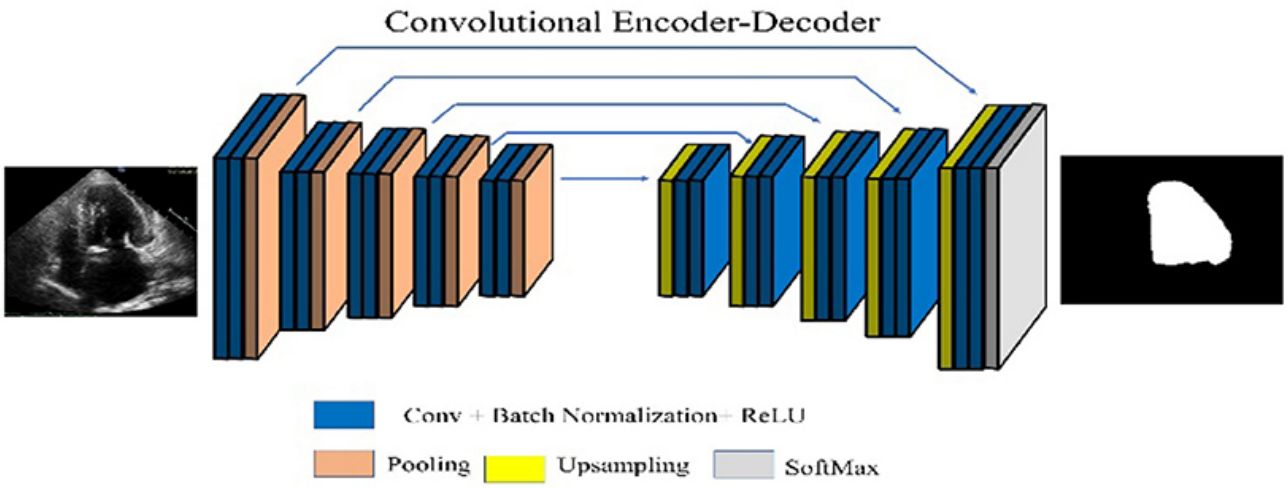


FIGURE 4 – Segnet architecture

Encoding Path The encoding path consists of a series of Conv2D layers, BatchNormalization, Activation (ReLU) layers, and MaxPooling2D layers. This path helps in extracting hierarchical features from the input MRI image. The encoding path has five sets of convolutional layers, with each set followed by a MaxPooling2D layer to reduce the spatial dimensions.

Fully Connected Layers After the encoding path, there are two fully connected (Dense) layers with 1024 neurons each and ReLU activation. These layers help in creating higher-level features from the extracted features in the encoding path.

Decoding Path The decoding path comprises a series of UpSampling2D layers, Conv2DTranspose layers, BatchNormalization, and Activation (ReLU) layers. The UpSampling2D layers increase the spatial dimensions of the feature maps, while Conv2DTranspose layers help in reconstructing the segmented image. The decoding path mirrors the encoding path, with five sets of Conv2DTranspose layers, each followed by an UpSampling2D layer.

Output Layer The output layer is a Conv2DTranspose layer with 4 filters, followed by a BatchNormalization layer and a softmax activation function. This layer provides the probability distribution of the four classes for each pixel in the input MRI image. The output shape is reshaped to match the desired segmented image size (128,128,4).

3.4.3 Results

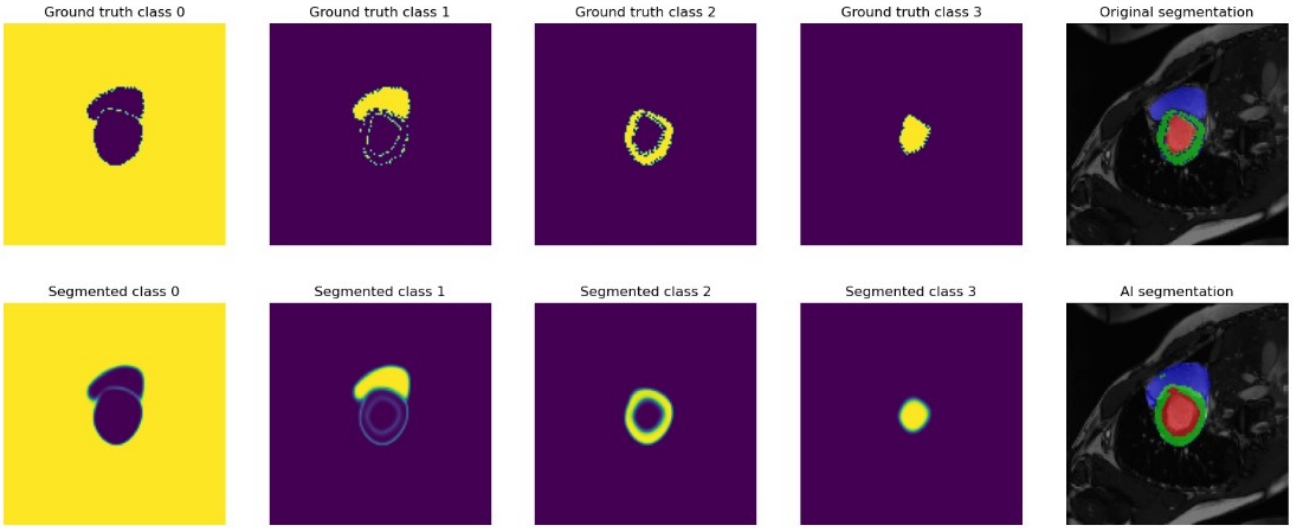


FIGURE 5 – Segnet segmentation results

3.5 SASNet

3.5.1 Preview

Even with these improvements, SegNet can still struggle to capture the fine details in the LV cavity segmentation, which can affect the accuracy of the subsequent feature extraction and classification steps. To overcome this limitation, the Segment Anything SegNet (SASNet) architecture was proposed. SASNet is a modified version of SegNet that includes dilated convolutional layers in the decoder path, which allows the model to capture more detailed information at different scales. Additionally, SASNet includes residual connections between the encoder and decoder paths, which helps the model recover the spatial information lost during the pooling operations while maintaining a low number of parameters.

3.5.2 Architecture

The model uses a modified version of the SegNet architecture, which includes the Segment Anything Model (SAM) blocks. These SAM blocks consist of two parallel paths of Conv2D layers followed by Batch Normalization. The outputs of the two paths are added together and passed through a ReLU activation function to obtain the final output of the SAM block.

The SegNet-SAM model consists of an encoding and decoding layer, both of which use the SAM blocks instead of the traditional Conv2D or Conv2DTranspose layers. The encoding layer consists of a series of SAM blocks and MaxPooling2D layers, while the decoding layer consists of UpSampling2D layers and SAM blocks. The output layer is a Conv2D layer followed by Batch Normalization and a Softmax activation function to provide class probabilities for each pixel.

3.5.3 Results

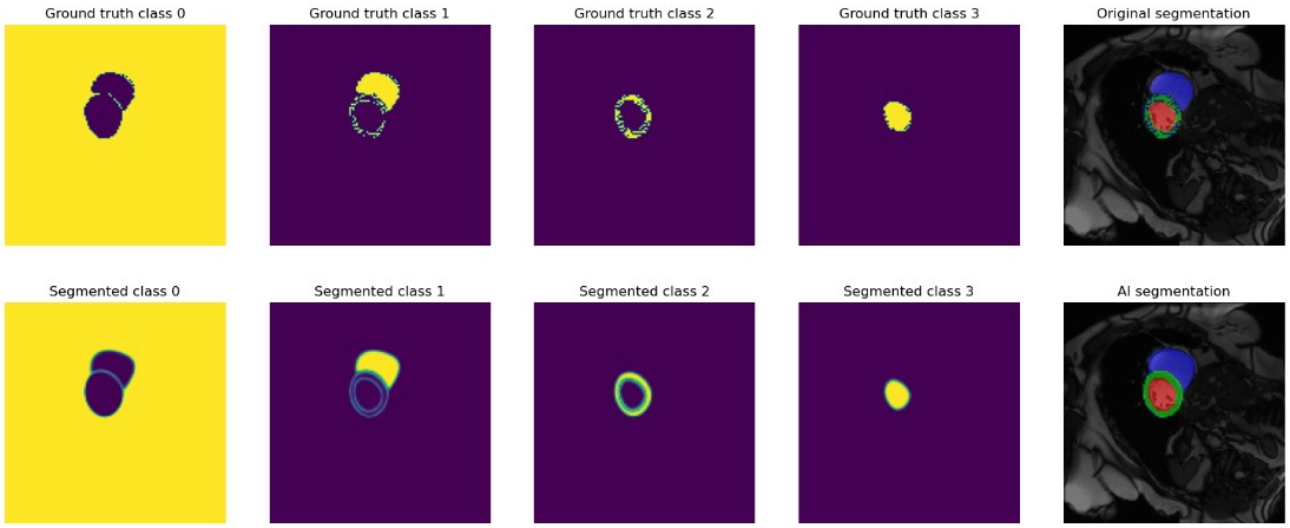


FIGURE 6 – SASnet segmentation results

3.6 Comparative study

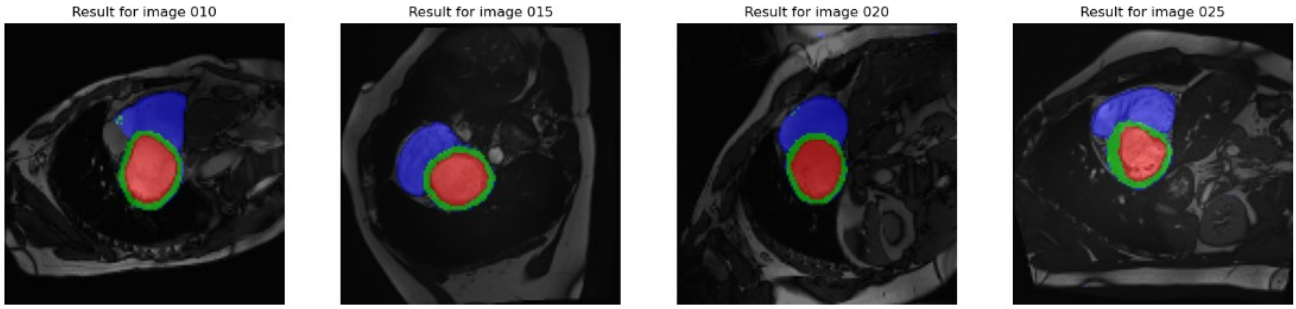


FIGURE 7 – Unet segmentation results

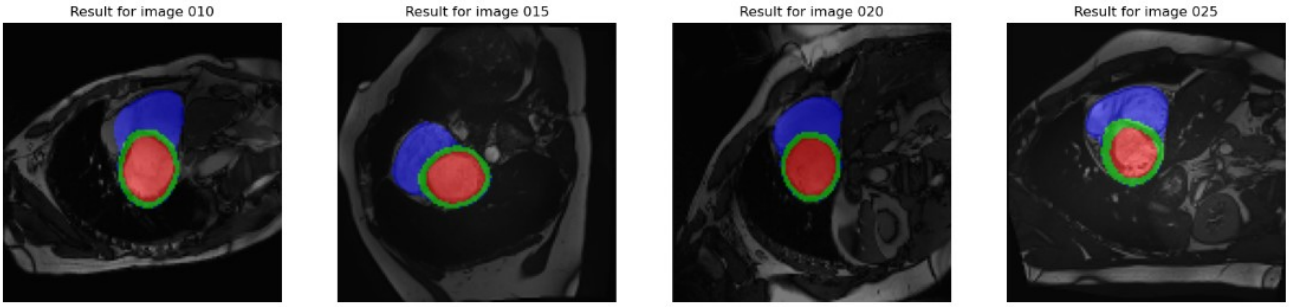


FIGURE 8 – Segnet segmentation results

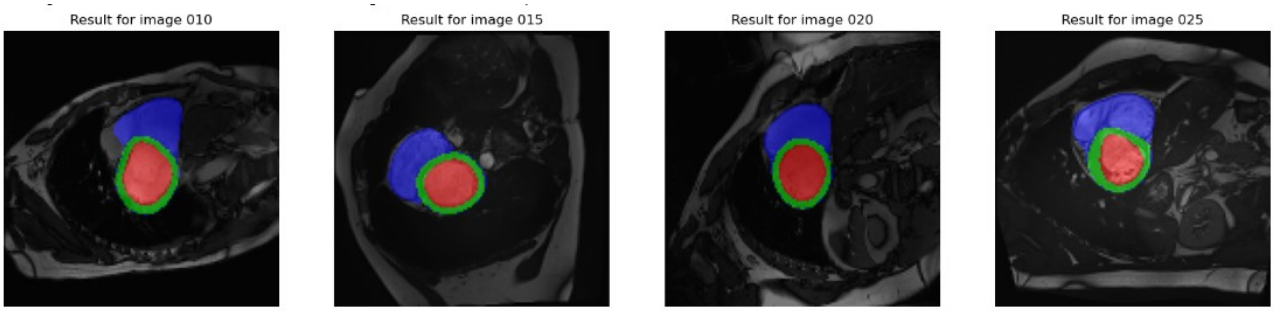


FIGURE 9 – SASnet segmentation results

The SegNet-SAM model has proven to be the top performer in the segmentation and classification of MRI cardiac images when compared to both U-Net and the standard SegNet model. The superior performance of SegNet-SAM can be attributed to the incorporation of SAM blocks into the SegNet architecture. This modification has enhanced the model's ability to process complex medical image data, leading to more accurate and reliable segmentation results.

Segment Anything Model (SAM) blocks serve as an innovative component that improves the overall network architecture. SAM blocks consist of two parallel paths of Conv2D layers followed by Batch Normalization. The outputs of these paths are then combined using an Add layer and passed through a ReLU activation function. This unique structure enables the model to extract richer features and capture complex patterns from the input data, resulting in a more accurate representation of the MRI cardiac images.

In addition, the SegNet-SAM model benefits from its ability to adapt to varying input sizes and process high-resolution images, which is a crucial factor when dealing with intricate details present in medical images. This flexibility allows the model to maintain high segmentation performance even when faced with diverse and challenging image datasets.

Furthermore, the advanced structure of the SAM blocks provides the SegNet-SAM model with an increased capacity to model contextual information. This advantage enables the model to better differentiate between various tissue types and anatomical structures, leading to improved segmentation accuracy and reliability.

The combination of the SegNet architecture and the innovative SAM blocks has proven to be highly effective in the segmentation and classification of MRI cardiac images. With its advanced feature extraction capabilities and ability to handle complex medical image data, the SegNet-SAM model shows great promise for use in a wide range of medical image processing tasks. This model's superior performance highlights its potential as a powerful tool for assisting medical professionals in the analysis and diagnosis of various cardiac conditions based on MRI images.

3.7 Summary

In summary, we used U-Net as a starting point for the LV cavity segmentation because of its excellent performance in semantic segmentation tasks. However, we moved to SegNet to address its limitations in preserving spatial information, and finally, we adopted SASNet to improve the accuracy of the LV cavity segmentation by capturing more detailed information at different scales. The advantages of each model were their ability to preserve spatial information, avoid overfitting,

and capture fine-grained details at different scales, respectively. The ameliorations of SASNet were the inclusion of dilated convolutions and residual connections to improve the model's accuracy while maintaining a low number of parameters.

4 Feature Extraction

4.1 Manually extracted features

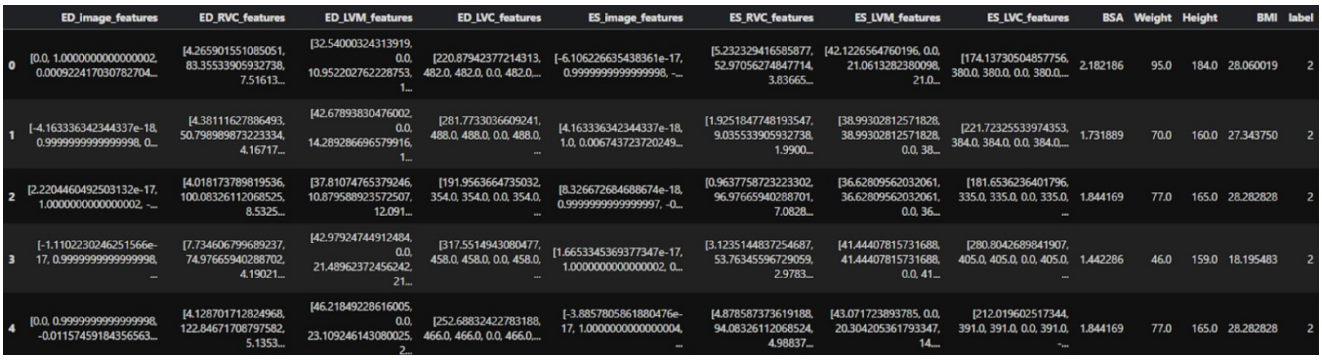
The following features have been extracted from cardiac magnetic resonance (CMR) images :

4.1.1 Intensity-based features :

— Mean intensity : This feature provides an estimate of the overall brightness of the image. In cardiac pathology, changes in the mean intensity can indicate abnormalities in tissue composition, such as edema or fibrosis.
— Variance of intensity : This feature measures the spread of intensities in the image. It can provide information about the heterogeneity of tissue composition and the presence of regions with different signal intensities, which can be indicative of pathology.
— Gradient magnitude : This feature measures the sharpness of edges in the image. In cardiac pathology, changes in the gradient magnitude can indicate the presence of boundaries between different tissues or structures, which can be relevant for identifying pathologies such as myocardial infarction or fibrosis.

4.1.2 Texture-based features (GLCM) :

— Contrast : This feature quantifies the difference in intensity between adjacent pixels. In cardiac pathology, changes in contrast can indicate the presence of regions with varying signal intensities, such as areas of fibrosis or scar tissue.
— Dissimilarity : This feature measures the difference in texture between adjacent pixels. It can provide information about the heterogeneity of tissue composition, which can be indicative of pathology.
— Homogeneity : This feature measures the uniformity of the distribution of intensities in the image. In cardiac pathology, changes in homogeneity can indicate the presence of regions with varying signal intensities, such as areas of fibrosis or scar tissue.
— Energy : This feature provides a measure of the overall complexity of the texture in the image. In cardiac pathology, changes in energy can indicate the presence of regions with varying signal intensities, such as areas of fibrosis or scar tissue.
— Correlation : This feature measures the linear relationship between the intensities of adjacent pixels. In cardiac pathology, changes in correlation can indicate the presence of regions with different signal intensities, such as areas of fibrosis or scar tissue.



	ED_image_features	ED_RVC_features	ED_LVM_features	ED_LVC_features	ES_image_features	ES_RVC_features	ES_LVM_features	ES_LVC_features	BSA	Weight	Height	BMI	label
0	[0.0, 1.0000000000000002, 0.0009224170307827 04...	[4.265901551085051, 83.35533905932738, 7.51613...	[32.540003243139 19, 0.0, 10.952202762228753, 1...	[220.87942377214313, 482.0, 482.0, 0.0, 482.0....	[-6.1062266354383 61e-17, 0.9999999999999998, -...	[5.232329416585877, 52.97056274847714, 3.83665...	[42.1226564760196, 0.0, 21.06132823809 98, 21.0...	[174.13730504857756, 380.0, 380.0, 0.0, 380.0....	2.182186	95.0	184.0	28.060019	2
1	[-4.1633363423443 37e-18, 0.9999999999999998, 0...	[4.38111627886493, 50.798989873223334, 4.16717...	[42.678938304760 02, 0.0, 14.289286696579916, 1...	[281.7733036609241, 488.0, 488.0, 0.0, 488.0, ...	[4.1633363423443 37e-18, 1.0, 0.0067437237202 49...	[1.9251847748193547, 9.035533905932738, 1.9900...	[38.99302812571828, 38.99302812571828, 0.0, 38...	[221.72325533974353, 384.0, 384.0, 0.0, 384.0....	1.731889	70.0	160.0	27.343750	2
2	[2.2204460492503132e-17, 1.0000000000000002, -...	[4.018173789819536, 100.08326112068525, 8.5325...	[37.810747653792 46, 10.879588923572507, 12.091...	[191.9563664735032, 354.0, 354.0, 0.0, 354.0, ...	[8.3266726846886 74e-18, 0.9999999999999997, -0...	[0.9637758723223302, 96.97665940288701, 7.0828...	[36.62809562032061, 36.62809562032061, 0.0, 36...	[181.6536236401796, 335.0, 335.0, 0.0, 335.0, ...	1.844169	77.0	165.0	28.282828	2
3	[-1.1102230246251 566e-17, 0.9999999999999998, ...	[7.734606799689237, 74.97665940288702, 4.19021...	[42.979247449124 84, 0.0, 21.48962372456242, 21...	[317.5514943080477, 458.0, 458.0, 0.0, 458.0, ...	[1.6653345369377 347e-17, 1.0000000000000002, 0...	[3.1235144837254687, 53.76345596729059, 2.9783...	[41.44407815731688, 41.44407815731688, 0.0, 41...	[280.8042689841907, 405.0, 405.0, 0.0, 405.0, ...	1.442286	46.0	159.0	18.195483	2
4	[0.0, 0.9999999999999998, -0.011574591843565 63...	[4.128701712824968, 122.84671708797582, 5.1353...	[46.218492286160 05, 0.0, 23.109246143080025, 2...	[252.68832422783188, 466.0, 466.0, 0.0, 466.0....	[-3.8857805861880 476e-17, 1.0000000000000004, ...	[4.878587373619188, 94.08326112068524, 4.98837...	[43.071723893785, 0.0, 20.304205361793347, 14....	[212.019602517344, 391.0, 391.0, 0.0, 391.0, -...	1.844169	77.0	165.0	28.282828	2

FIGURE 10 – Manually extracted features

These features are relevant to cardiac pathology prediction because they provide information about the tissue composition and texture of the heart, which can be indicative of various pathologies. For example, changes in intensity and texture can indicate the presence of fibrosis, edema, or scar tissue, which are common features of myocardial infarction and heart failure. The extraction of these features can therefore aid in the identification and diagnosis of such pathologies, and can ultimately contribute to the development of more accurate and effective predictive models for cardiac disease.

In addition to the previously mentioned features, for each of the three classes (RVC, LVM, LVC), we calculate features based on the segmented regions. The common features that were calculated previously, such as circularities, circumferences, and volumes, are also used here.

For RVC, we calculate features such as mean circularity, maximum circumference, mean circumference, sum of volumes normalized by body surface area, maximum volume, kurtosis of volumes, skewness of volumes, and standard deviation of volumes. These features can provide insights into the shape and size of the right ventricle, which can be used for predicting cardiac pathology.

For LVM, we calculate features such as maximum thickness, mean thickness, mean circularity, maximum circumference, mean circumference, sum of volumes normalized by body surface area, maximum volume, volume range, kurtosis of volumes, skewness of volumes, and standard deviation of volumes. These features provide insights into the size and shape of the left ventricle mass, which can also be used to predict cardiac pathology.

For LVC, we calculate features such as sum of volumes normalized by body surface area, maximum volume, minimum volume, volume range, median volume, kurtosis of volumes, skewness of volumes, and standard deviation of volumes. These features can provide insights into the size and shape of the left ventricle cavity, which can also be used to predict cardiac pathology.

Overall, these features provide a comprehensive understanding of the shape, size, and structure of the heart, which can be used for predicting and diagnosing cardiac pathology.

4.2 Pyradiomics features

This feature extraction process utilizes PyRadiomics library to extract a set of relevant features from medical images that can be used for predicting cardiac pathology. The process takes as input

two image files, an image and its corresponding binary mask, which identifies the region of interest (ROI) to be analyzed.

	features_1	features_2	features_3	features_4	features_5	features_6	bsa	h	w	label
0	[140075.68359375, 0.16941521415537886, 23188.2...	[163723.7548828125, 0.31480151411269575, 50213...	[295147.705078125, 0.0868170583771158, 25466.4...	[59902.9541015625, 0.22783956509562964, 13236....	[193963.623046875, 0.25014029416920666, 48069.8...	[226675.4150390625, 0.09736921560255554, 21903....	2.182186	95.0	184.0	2
1	[95133.0184936523 4, 0.18716492005537064, 17488...	[160655.5461883545, 0.319557779985523, 50508.7...	[264732.6946258545, 0.09421187511278188, 24782...	[29068.422317504883, 0.3800477866166557, 10022....	[192976.3793945312 5, 0.24461403387295766, 4661...	[187487.9360198974 6, 0.10711962395275 88, 19922...	1.731889	70.0	160.0	2
2	[192187.5, 0.1339136133154044 4, 25474.42763716...	[190722.65625, 0.28741248281156634, 53274.6917...	[275976.5625, 0.09029196885424089, 24733.66550...	[175195.3125, 0.14530476947217966, 25094.28031...	[201757.8125, 0.25825613453944923, 51230.88814...	[240332.03125, 0.09510756495619356, 22664.2844...	1.844169	77.0	165.0	2
3	[105771.1124420166, 0.18778594150856254, 19476...	[169057.6076507568 4, 0.29804492736494 34, 49135...	[261751.3179779052 7, 0.09277115987658159, 2413...	[84291.6488647461, 0.2263766654462115, 18737.1...	[174613.8095855713, 0.28261200413396 88, 48117....	[226584.6252441406 2, 0.09841061896101107, 2215...	1.442286	46.0	159.0	2
4	[170592.4072265625, 0.15832202832425576, 26650...	[200492.7978515625, 0.27682947912841 86, 54768...	[291061.6149902344, 0.09183037706332679, 26574...	[74217.041015625, 0.223025373572, 16632.290962...	[233263.0920410156 2, 0.21965400644580946, 5058...	[224052.7038574218 8, 0.09887542776287726, 2200...	1.844169	77.0	165.0	2

FIGURE 11 – Pyradiomics features

The PyRadiomics library is used to extract a set of features from the ROI. The feature extraction process is initiated by creating a PyRadiomics feature extractor object, and the parameters of this object are set to define the desired feature extraction settings. The feature extraction process is performed separately for three cardiac classes : Myocardium (MYO), Right Ventricle (RV), and Left Ventricle (LV).

Each cardiac class has its own binary mask that is created by thresholding the original binary mask using specific threshold values. After obtaining the binary masks for each cardiac class, the feature extraction process is executed for each mask and the corresponding image using the PyRadiomics library.

The extracted features are stored in separate arrays for each cardiac class, and the arrays are returned at the end of the feature extraction process. These extracted features can be utilized to train and test machine learning models for predicting cardiac pathology.

The features extracted using this method include a combination of first-order, shape, texture, and gray-level co-occurrence matrix (GLCM) features. First-order features capture the distribution of pixel intensities within the ROI, while shape features describe the geometric characteristics of the ROI. Texture features represent the spatial arrangement of pixel intensities within the ROI, and GLCM features capture the spatial distribution of intensity pairs within the ROI.

Overall, the feature extraction process described here is a valuable tool for extracting a comprehensive set of features from medical images that can be used for cardiac pathology prediction.

4.3 Features selection

4.3.1 Variable selection

Variable selection aims to identify the most relevant subset of features for a given problem. It can be done using various techniques, such as statistical tests or machine learning algorithms. The goal is to identify a small set of features that can accurately represent the data and lead to better model performance.

4.3.2 Principal component analysis

PCA is a dimensionality reduction technique that can be used to reduce the number of features in a dataset while preserving as much of the original information as possible. It works by finding a

new set of variables (principal components) that are linear combinations of the original variables and explain the maximum possible amount of variance in the data.

4.3.3 Discussion

In the context of our cardiac pathology prediction problem, both variable selection and PCA could be relevant. With 100+ features, it is possible that some of them may not be relevant for the prediction task and could even introduce noise and decrease the accuracy of our models.

However, it's worth noting that PCA has its limitations. It assumes that the data is linearly related, which may not always be the case in our problems. Also, the interpretability of the resulting principal components may not always be straightforward, especially when the original features have complex relationships.

These approaches did not show any improvements or significant results, we may assume that thay will add a generalization to our final model that can not be seen with our size-limited dataset.

4.4 Data augmentation

In order to generalize our model, since we have a small dataset, we apply a light data augmentation techniques on-the-fly to efficiently feed the input data volumes into our model. On a random basis, the data is rotated between 10 to +10∘, and shifted by 0.1 on the x and y axis. This ensures slight robustness and variability in training the models.

5 Classification Algorithms

We evaluated several machine learning algorithms to classify the MRI images into the five diagnostic classes. These algorithms included random forests, support vector machines, boosting models and neural networks. We used the scikit-learn library to implement these algorithms in Python 3. We also experimented with different hyperparameters for each algorithm to optimize their performance.

5.1 LightGBM

LightGBM is a gradient boosting framework that uses tree-based learning algorithms. It is known for its high efficiency and low memory usage. For our problematic, LightGBM is a good choice because it can handle a large number of features and is well-suited for classification tasks. It can also handle missing data, which is important in medical datasets where missing data can be common. One potential downside of LightGBM is that it can be sensitive to overfitting, so it's important to use proper regularization techniques to avoid this.

5.2 MLP : Multilayer perceptron

MLP is a type of artificial neural network that is commonly used for classification tasks. It consists of multiple layers of nodes, with each node applying a nonlinear activation function to its inputs. MLPs can handle large datasets and are able to model complex relationships between inputs and outputs. For our problematic, MLPs could be a good choice for modeling the relationship between the extracted features and the classification labels. In order to reduce the variance, we tried a boosting approach by looping on 100 MLP models where we generated the probabilities

instead of the results of each test subject. We averaged the probabilities to assign the subject to the class with the highest average probability.

5.3 Random Forest

Random Forest is a type of ensemble learning method that uses decision trees. It creates multiple decision trees and then aggregates the results to make a final prediction. In our context, Random Forest can be a good choice because it can handle high-dimensional data and can model complex relationships between features and labels. It is also less prone to overfitting than some other models, due to its use of multiple decision trees. However, it can be slower and less efficient than some other models.

5.4 XGBoost

XGBoost is another gradient boosting framework that uses tree-based learning algorithms. It is similar to LightGBM in many ways, but is known for its scalability and efficiency. XGBoost can be a good choice because it can handle large datasets with many features and is well-suited for classification tasks. It also has built-in regularization techniques to help prevent overfitting.

5.5 Support vector machine

SVM is a supervised learning algorithm that can be used for classification and regression tasks. It works by finding the best hyperplane that separates the data points into their respective classes.

In our case, it seems like a good choice because it can handle high-dimensional feature spaces and is less prone to overfitting compared to other classifiers. Additionally, SVM is known to work well with small to medium-sized datasets, which is our case.

Furthermore, SVM is a versatile algorithm that can be applied to different types of data, including both linear and nonlinear data. SVM can use different types of kernel functions (e.g., linear, polynomial, radial basis function) to map the input data into a higher-dimensional space where a hyperplane can separate the data points into their respective classes. This makes SVM an excellent candidate for classification problems with complex decision boundaries.

Overall, SVM is a powerful and flexible algorithm that is well suited for classification problems with high-dimensional data such as the cardiac pathology prediction problem that we are working on.

5.6 Autoencoder Teacher-Student SVM : Self supervised model

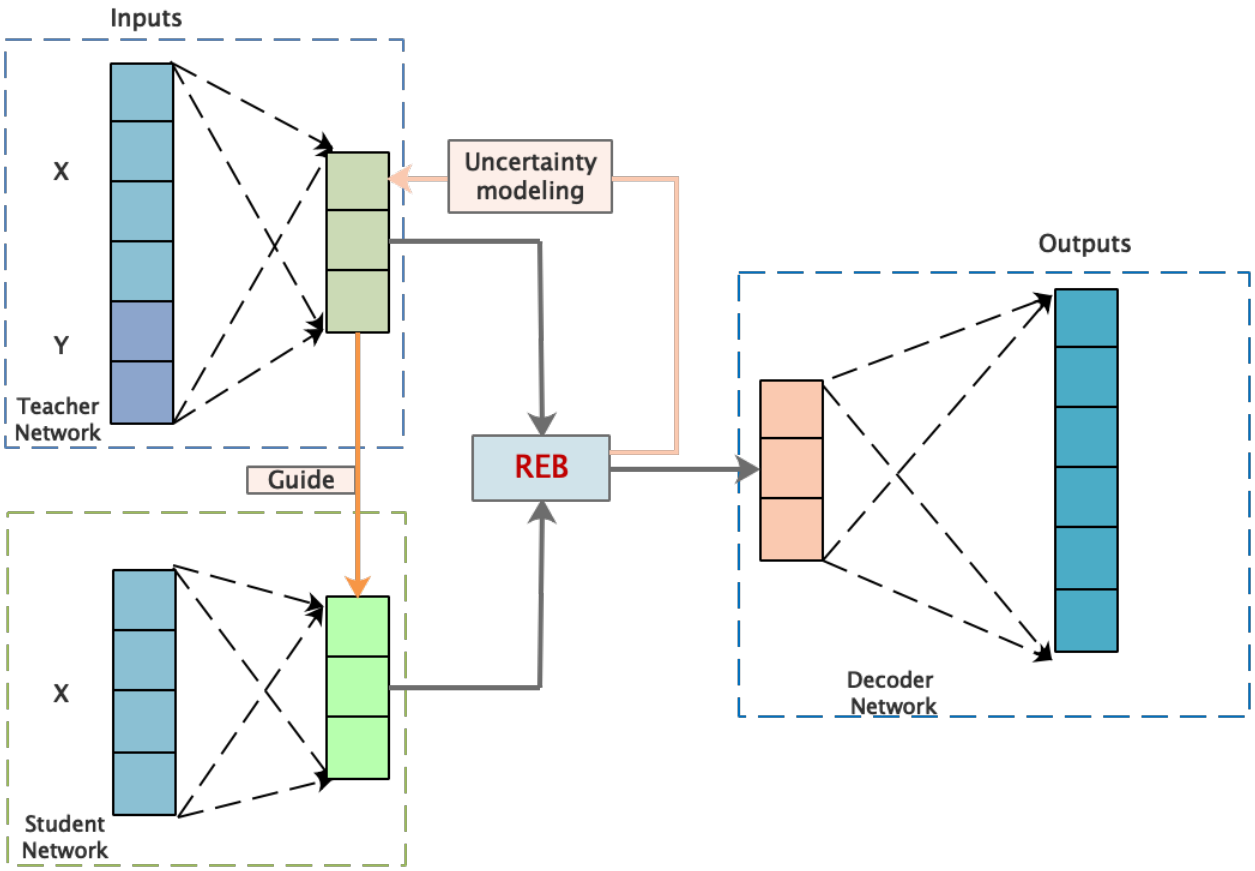


FIGURE 12 – Autoencoder Teacher-Student architecture

The autoencoder is a state of the art neural network that is trained to compress the high-dimensional MRI scan data into a low-dimensional feature representation. This feature representation is then fed into the teacher network, which is a neural network that is trained to predict the cardiac pathology labels using the compressed features as inputs. The teacher network serves as a guide for the SVM classifier, which is the final component of the model. The SVM classifier is trained using the compressed features and the corresponding cardiac pathology labels.

The Autoencoder Teacher-Student SVM model is relevant to the cardiac pathology prediction problem because it can extract meaningful features from the MRI scans and use these features to accurately predict the presence or absence of cardiac pathologies. The autoencoder component of the model helps to reduce the dimensionality of the input data and extract important features that are relevant to the prediction task. The teacher network helps to guide the training process and ensure that the features extracted by the autoencoder are informative for the prediction task. Finally, the SVM classifier uses the compressed features to make the actual predictions.

One advantage of the Autoencoder Teacher-Student SVM model is that it can be used to make predictions on new MRI scans that were not included in the training data. This is possible because the autoencoder has learned to extract general features from the MRI scans that are relevant to the cardiac pathology prediction task. Additionally, the model can be adapted to incorporate

new data and update the feature representation as new data becomes available. This makes the model a useful tool for clinicians who need to make predictions on new patients and monitor the progression of cardiac pathologies over time.

6 Results

Model	Fold 1		Fold 2		Fold 3		Fold 4		Fold 5		Overall mean	
	TA	VA	TA	VA	TA	VA	TA	VA	TA	VA	TA	VA
LightGBM	0.91	0.80	0.95	0.85	0.89	0.75	0.90	0.65	0.9	0.70	0.90	0.75
SVM	0.96	0.85	0.92	0.80	0.96	0.80	0.96	0.90	0.95	0.70	0.95	0.81
Self-supervised AE-SVM	0.96	0.90	0.90	0.90	0.95	0.80	0.95	0.90	0.90	0.75	0.92	0.85
Random forest classifier	0.90	0.80	0.90	0.90	0.90	0.65	0.90	0.80	0.9	0.75	0.90	0.78
XGBoost	0.90	0.90	0.90	0.90	0.90	0.80	0.90	0.80	0.90	0.75	0.90	0.83
MLP	0.90	0.80	0.90	0.90	0.90	0.85	0.90	0.89	0.9	0.75	0.90	0.78

TA: Train accuracy
VA: Validation accuracy

FIGURE 13 – Classification results

Looking at the overall mean column, it appears that the SVM model achieved the highest validation accuracy (0.81), followed by the self-supervised AE-SVM model with a validation accuracy of 0.85. The other models achieved lower validation accuracies ranging from 0.65 to 0.78.

The SVM model's high performance may be attributed to its ability to find the optimal hyperplane that separates the different classes in the feature space. On the other hand, the self-supervised AE-SVM model, which combines the unsupervised learning capabilities of the autoencoder and the supervised learning capabilities of the SVM, performed relatively well with an average validation accuracy of 0.85.

Overall, the results suggest that the self-supervised AE-SVM model, which is a combination of unsupervised and supervised learning, is a promising approach for this problem, and that it is worth exploring further. However, the SVM model remains the best performer in terms of validation accuracy, suggesting that it may be the best choice for this particular classification problem.

7 Conclusion

With great excitement, we present our pioneering approach for automatic diagnosis from cardiac magnetic resonance imaging. Our journey was a rich, variant, and complex pipeline that produced outstanding results.

Firstly, we implemented the state-of-the-art "Segment Anything Segnet," a neural network that surpassed classic segmentation models in its precision and specificity to our particular challenge. Witnessing its performance was a highlight of our work.

Secondly, we embarked on a research expedition to identify pathologies and determine possible features manually. We scoured through research papers and basic algorithms to arrive at a set of features extracted from MRI scans with the Pyradiomics package. We meticulously examined and extracted the most significant features using Principal Component Analysis (PCA) and feature importance analysis.

Lastly, we brought everything together by implementing a diverse range of classification models, each expertly tuned and modified to optimize our results.

All of these steps represent the culmination of our theoretical and practical knowledge acquired during the IMA 205 class. We thoroughly enjoyed the experience of using our newfound knowledge to tackle this fascinating and challenging objective.