

Assessing Cardiac Functionality by Means of Interpretable AI and Myocardial Strain

Marco S. Nobile^{*†}, Amalia Lupi[‡], Leone Bacciu^{*}, Matteo Grazioso^{*}, Chiara Gallese[§], Emilio Quaia[‡], Alessia Pepe[‡]

^{*}Department of Environmental Sciences, Informatics and Statistics

Ca' Foscari University of Venice, Italy

Email: marco.nobile@unive.it

[†]Bicocca Bioinformatics, Biostatistics and Bioimaging (B4) Research Center, Monza, Italy

[‡]Institute of Radiology, Department of Medicine, Padua University Hospital, Italy

Email: amalia.lupi@unipd.it

[§]Department of Law, University of Turin, Italy

Abstract—Cardiac Imaging is a powerful methodology for the accurate assessment of heart functionality. Among the possible approaches, Myocardial Strain assesses the functionality of the heart by tracking the movement and deformation of myocardium during the cardiac cycle. This information, that can be acquired also by means of Cardiac Magnetic Resonance, can pave the way to the development of predictive models using machine learning. In this work, we developed a predictive model of left ventricular ejection fraction, which is a measure of the heart's function to pump oxygen-rich blood to the body, trained using strain data. Specifically, we developed a fully interpretable model based on a rule-based Fuzzy Inference System, coupled with a novel methodology for the disambiguation of the rules. Our results show that the developed model is able to accurately estimate the ejection fraction, and can provide physicians with additional insights about the role of strain features.

Index Terms—Artificial Intelligence, Medical Image Analysis, Interpretability, Myocardial Strain.

I. INTRODUCTION

CARDIAC imaging is a fundamental technique for the assessment of heart functionality. The most important feature for the assessment of heart function is the Left Ventricular Ejection Fraction (LVEF) [1], that is, the ratio of blood that is pumped out of the left ventricle with each heart beat. The LVEF of an healthy individual usually ranges between 55% and 70% [2]. The most appropriate technique for the assessment of heart's state depends on the clinical situation.

The most widespread technique is echocardiography: widely available, non-expensive, not based on radiations, and it is not invasive leading to minimal patient discomfort [3]. Echo can give information about both systolic and diastolic heart function, ventricular function, along with pericardial and valvular functions. Regardless of its advantages, echo has limitations about how LVEF is assessed; in particular, it leads to a high inter-operator variability, and the outcome can change with the change of machine or algorithms [4]. Thus, better techniques are needed to mitigate the limitations mentioned above. Indeed, Cardiovascular Magnetic Resonance (CMR) is widely recognized as the gold standard for volumes, masses, and

function assessment. [5] Among the possibilities, Myocardial Strain (MS) denotes the change in myocardial fibre length over the cardiac cycle, and it used to assess cardiac muscle function. MS can be obtained by means of conventional techniques such as echocardiography and CMR, extending the clinical information collected using traditional techniques. In this context, Feature Tracking (FT) is a tissue tracking technique which quantifies both myocardial motion and deformation by means of MS, strain rate, torsion, and dyssynchrony. MS analysis can offer additional information with respect to established functional analyses. FT quantifies cardiovascular impairment, and characterizes myocardial function, through direct assessment of myocardial fiber deformation.

In this paper, we propose a novel approach to predict the LVEF of a patient given their quantitative information about myocardial strain and cardiac index. In order meet the novel legal requirements recently introduced by the European Union's legislation, in this paper we will adopt an interpretable Artificial Intelligence (AI) approach. The first legal aspect that we consider arises from Article 86 of the AI Act introduced the *right to explanation* of decision making related to high-risk systems [6], such as medical devices and thus any AI diagnostic software used on patients. This new right complements the previous Articles 13, 14 and 22 of the European General Data Protection Regulation (also known as GDPR), which prescribe the right of the data subject (in this case, the patient) to receive a meaningful explanation related to the use of their personal data in automated decision-making [7]. We stress the fact that, according to the European official guidelines [8], such explanations must be meaningful, and not merely formal. Stated otherwise, the patient should be able to understand the rationale behind the decision.

Article 14 of the AI Act explicitly incorporates the notion of interpretability within its requirements for human oversight of high-risk AI systems. Such article provides that users must be able to “correctly interpret the high-risk AI system's output, taking into account in particular the characteristics of the system and the interpretation tools and methods

available”, a mechanism that is supposed to facilitate an effective oversight. Within the legal framework, the provision denotes interpretability not merely as a technical feature, but rather as a normative safeguard designed to maintain a continuous and informed human presence in the operational loop. While interpretability is presented as one among several possible oversight measures, its inclusion is an evidence of the legislative intent to ensure that the output of high-risk AI applications is understandable. We argue that to comply with all legal obligations – including provisions regarding informed consent – the AI system must be interpretable, and not merely explainable [9], [10]. In this context, Cynthia Rudin has clearly conceptualized the difference between the two terms [11]: while explainability refers to a *post-hoc* model built to explain the output of a black-box model, interpretability refers to inherently interpretable models (*i.e.*, white boxes).

As a matter of fact, though, even the simplest of explanations related to medicine can be challenging to understand for patients, since they often lack any expertise and generally have insufficient background knowledge [12]. Thus, in this work, we assume that the interpretable model provides a clear motivation to the medical doctors who, in turn, will have the means to explain the rationale of the AI system to patients. In particular, we present an extended version of the pyFUME library for the generation of interpretable models based on Fuzzy Inference Systems. pyFUME adopts both unsupervised and supervised learning approaches to cluster the dataset and generate a rule for each cluster; each rule performs an accurate prediction of the target variable for that specific group. In this work we extended pyFUME to provide an additional level of explanation by performing a disambiguation of the rules and determine which rule has more impact to the prediction of a specific output. According to our results, the generated model is both accurate and easy to interpret, so that it can be trusted in the clinical environment and provide insights about the studied cohort.

This paper is structured as follows. We present Myocardial Strain and pyFUME in Section II. The new disambiguation approach, named RuDe, is introduced in Section III. The results of our tests are presented in Section IV. We conclude the paper with some remarks and future developments.

II. METHODS

A. Myocardial Strain

Cardiovascular Magnetic Resonance (CMR) is the primary imaging modality for non-invasive myocardial tissue characterization, and it is considered the gold standard test for quantifying cardiac function, volumes, and mass [5], [13]. For the morphological and functional evaluation, Steady State Free Precession (SSFP) sequences represent the state of the art [14]. Images are commonly acquired during brief periods of breath hold with EKG triggering to take cardiac motion into account. Standardized acquisition strategy to image the anatomical planes of the heart correctly is used. Since the long axis of the heart is usually not oriented with any of the standard axis of the body, it is necessary to perform a series

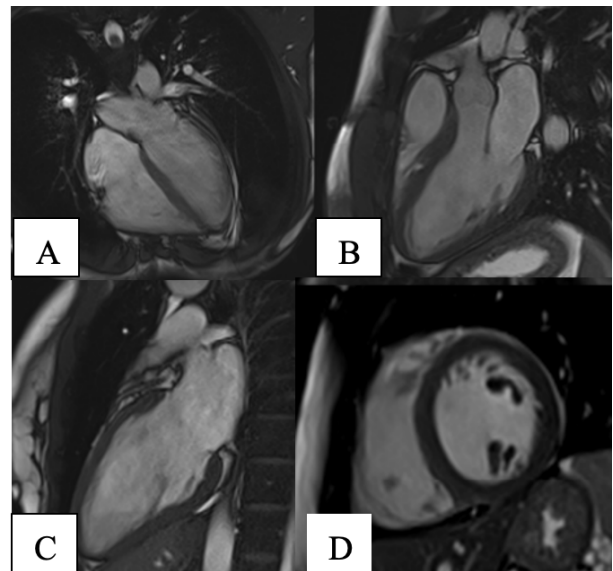


Fig. 1: SSFP-sequence projection for morpho-functional evaluation of the heart. A, B, C: 4-, 3- 2-chambers long axis view respectively. D: Short axis LV-RV outflow tract, basal, mid and apical left and right ventricle sections. All the images are taken in end-diastolic phase.

of angulations of the acquisition planes to obtain the 4-, 3-, 2-chambers and short axis view [15] (Figure 1).

Once all these projections are obtained, a short axis full coverage of the left ventricle, by acquiring a series of slices parallel to the atrioventricular plane from the base to the apex, is planned for the quantitative analysis of ejection fraction, ventricular mass, volumes, wall thickness, and to evaluate wall kinetics. Post-processing of SSFP cine images requires the contouring of the endocardial and epicardial borders of the left ventricle, in all the images of the end-diastolic and end-systolic phase. This task can be performed automatically using AI-based methods (*e.g.*, cvi42® software [16]), applying manual correction if necessary. Figure 2 shows an example of such borders contouring.

The need for in-depth characterization of heart function has led, in the last years, to the development of tissue tracking analysis, which provides a quantitative assessment of the global and regional kinetics of the myocardium, yielding additional information about heart function other than those obtained by the Ejection Fraction (EF) alone, a widespread diagnostic parameter that gives overall ventricular systolic function but does not provide information about regional myocardial kinetics and contractility nor allows evaluation of the diastole. Moreover, in many heart diseases, the EF is altered only late in the progression, leaving a diagnostic gap in asymptomatic or subclinical conditions [17]. Since its introduction in the late 1980s, tissue tracking has been extensively developed and is now gaining popularity due to research trends towards generating evidence of the adjunctive value of strain in early and differential diagnosis, risk stratification

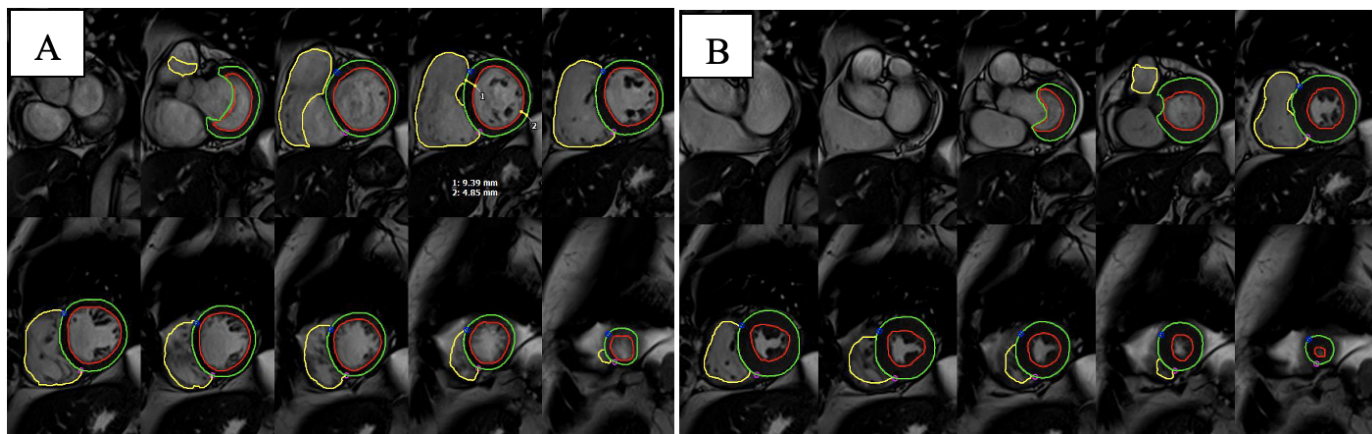


Fig. 2: SSFP-sequence short axis stack is shown in end-diastolic (A) and end-systolic phase (B). Left ventricular endocardial (red) and epicardial (green) and right ventricle endocardial (yellow) borders contouring is performed automatically by cvi42® software and is the base for volumes, mass and ejection fraction evaluations.

and prognosis determination of many heart diseases. Recent evidence has given strength to this technique so that it is starting to be included in CMR assessment. The determination of normative values both for healthy subjects and disease specific population is important to avoid misdiagnosis.

Myocardial strain is an index of length change between two points of the myocardium reflecting the myocardial deformation during contraction and relaxation. Strain is defined as follows:

$$\text{strain} = (L_1 - L_0) / L_0$$

where L_0 and L_1 represent the myocardial length at baseline (end-diastole) and at the end of the considered time interval (end-systole), respectively. As it is evident from the mathematical definition, negative values mean shortening, thinning and/or contraction whereas positive values mean lengthening, thickening and/or relaxation [18]. Myocardial points deformation sampled during the cardiac cycle defines curves, as shown in Figure 3.

The complex architectural structure of the left ventricle that is organized into three different layers – the subendocardial layer, in which fibers are oriented longitudinally from the base towards the apex, the mid-wall layer, where fibers are oriented circumferentially, and the subepicardial layer where fibers are oriented in a right-handed helicoidal fashion [19], [20] – induces the heart to deform, during systole and diastole, into three main different directions: longitudinal and circumferential shortening, radial thickening and torsion/rotation [21], as schematized in Figure 4. As a consequence, strain can be calculated for these directions of movement, defining global longitudinal (GLS), circumferential (GCS), radial (GRS) strain and torsion. Contraction of the sub-endocardial fibers accounts for longitudinal shortening, whereas contraction of the sub-epicardial fibers account for circumferential shortening. Both contribute to radial thickening.

Since myocardial involvement in heart diseases can be classified based on the predominantly involved layer into sub-

endocardial, mid-wall, sub-epicardial, or transmural dysfunction, the latter derived from the involvement of the entire cardiac wall [21], strain reduction could be found accordingly. Given these points, it is clear that strain could be an important

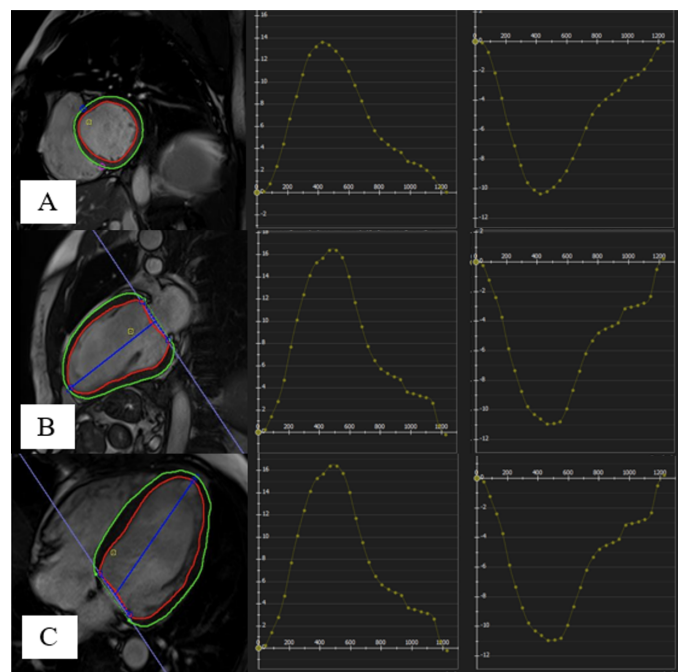


Fig. 3: Example of feature tracking analysis of the left ventricle. A: short axis slice with the epicardial and endocardial borders contoured with green and red respectively (on the left) and the radial and circumferential strain curves (on the right). B: 2-chambers long axis view (on the left) and the correspondent radial and longitudinal strain curves (on the right). C: 4-chambers long axis view (on the left) with the correspondent radial and longitudinal strain curves (on the right).

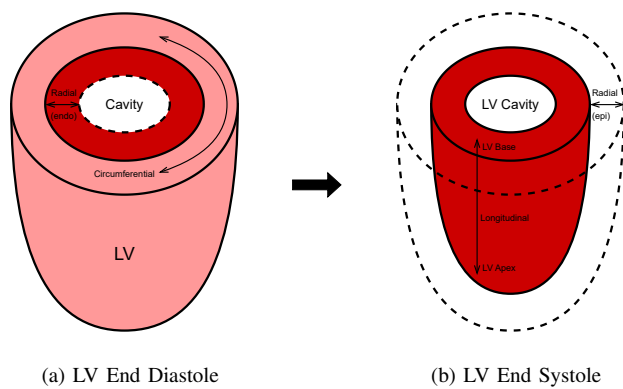


Fig. 4: Schematic image of the different directions of deformation of the Left Ventricle (LV).

tool to evaluate different types of myocardial dysfunction and eventually lead to an early diagnosis of progressive heart disease.

The technology of tissue tracking refers to methods for identifying a particular point in the heart and the capacity to follow it in order to evaluate the displacement of the same point during the cardiac cycle. This allows the evaluation of motion at a parametric level.

The first step is to define the end diastolic and end-systolic frames. The second step consists in defining the region of interest inside the cardiac wall. Segmentation is the third step, the critic one, that defines the set of points that are going to be tracked by the algorithm. At the end, the set of point is tracked, meaning followed during the cardiac cycle, and the strain curves are obtained [22].

Longitudinal strain is derived from 2-, 3- and 4-chambers long axis view; circumferential strain is derived from short axis images while radial strain is obtained from both 2-, 3- and 4-chambers long axis and short axis views.

FT was developed as a 2D technique which can be applied on standard SSFP cine sequences without the need to perform dedicated sequences and images acquisition. In addition to that, no complex post-processing analysis is required and therefore, many commercial software have implemented feature tracking-based tracking module in their packages. This possibility allows to evaluate strain onto pre-existent datasets to perform retrospective analysis without requiring new acquisitions [18]. For this reason, over the last years, FT has gained popularity and has experienced a great use in clinical setting.

The main limitation of FT is that 2D analysis suffers from through plane motion artifacts because of the difficulty to track the feature falling out of the plane on the consecutive frames. 3D analysis can overcome this limitation: many software packages offer this possibility, but experience of this analysis is still limited and not clinically validated, yet [23]. Global values of strain calculated with FT have shown to be more reproducible than segmental ones (which show also a significant inter-vendor variability in addition to inter/intra-observer variability), and to have a good concordance between vendors

and observers [24], especially with a standardized approach in post-processing using the same commercial software [25]. Inter- and intra-observer variability is ascribable to the absence of intrinsic markers in the tissue and the consequent need to draw the epicardial and endocardial contours. This task is performed automatically by some software, with possible manual correction in order to exclude the left ventricular outflow tract and the atrio-ventricular plan. This necessity obviously introduces a source of variability between different operators and the same operator in two different moments. Nevertheless, the systematic variability introduced by the different commercially available software packages, which has been ascribed to the differences between the algorithms used for tracking, highlights the need for a standardization of the acquisition of the SSFP cine images and analysis with the feature tracking module, that should be implemented. Finally, whether feature tracking should be included in the routine CMR protocol for clinical practice must be assessed on the basis of diagnosis, prognosis and clinical decision-making [18]. Several studies in patients with different heart diseases, such as hypertensive cardiomyopathy, ischemic and non-ischemic cardiomyopathy, heart failure, valvular disease and pulmonary hypertension, have evaluated FT additional value in terms of early diagnosis and prognosis. As for decision making, however, only limited clinical trial exist on the value of FT derived strain in clinical application and that's why the technique is not included in clinical algorithm yet [26].

Thus, further evidence from controlled clinical trials and prognostic data are needed to include strain evaluation into clinical decision trees for patients with heart disease.

In the last years, several studies analysed the impact of strain analysis in providing information of cardiac function in many heart diseases. In particular, the availability of FT led to its wide use in research settings and pointed out that it can provide more information on cardiac performance than ejection fraction alone, underlining its additional value in early diagnosis, differential diagnosis, and prognosis [18].

As far as early diagnosis is concerned, FT has proved to be able to detect subclinical myocardial impairment in patients with potential cardiomyopathy during their follow up. Other than this, feature tracking was able to detect early systolic and diastolic impairment in patients with preserved LVEF [27].

Feature tracking is certainly more useful to detect cardiac disfunction in patients with heart failure with preserved ejection function (HFpEF), than ejection fraction alone, as it can detect early diastolic disfunction [18].

Great limitation of FT derived strain for clinical practice is the lack of reference values coming from large cohorts of healthy and disease specific subjects. Several studies on healthy volunteers [28]–[31] have attempted to provide normal reference values for both global and segmental strain but results seem to have sources of potential variability introduced by different ethnicity of the population of study, MR scanner, acquisition sequences for kinetic imaging, and software for image post-processing. In addition to that, there is also a lack of normative values for disease specific conditions or

for specific population that could have different strain values compared to healthy subjects due to para-physiologic heart remodeling.

The aim of this study is to predict LVEF starting from strain analysis, through an interpretable machine learning approach in order to characterize patients function, and to assess the role of strain without the need for arbitrary thresholds.

B. Interpretable AI with pyFUME

Interpretable Artificial Intelligence is ever more a hallmark in fields where transparency and accountability of the decision are of the utmost importance. In high-stakes settings, models are not just supposed to generate correct predictions, but also to justify their internal logic. Fuzzy Inference Systems (FIS), under the form of Takagi–Sugeno fuzzy models, provide a natural way towards attaining this by encoding knowledge in the form of human-readable rules.

Interpretable AI is concerned with models whose internal logic can be directly examined and interpreted by human experts. FIS are rule-based and hence intrinsically interpretable, since each rule is equivalent to a local area of the input space with a corresponding output function. For a standard first-order Takagi–Sugeno model, the fuzzy rules are of the form:

$$R_j : \quad \text{IF } x_1 \text{ is } A_{j1} \text{ AND } \dots \text{ AND } x_N \text{ is } A_{jN} \\ \text{THEN } y_j = \mathbf{a}_j^\top \mathbf{x} + b_j \quad (1)$$

where:

- $j = 1, \dots, J$ denotes the cluster and rule number;
- $\mathbf{x} = (x_1, \dots, x_N)$ denotes the input vector;
- N is the number of input features (linguistic variables);
- A_{ji} is the fuzzy set for the i -th linguistic variable with respect to cluster j ;
- y_j is the consequent function for rule R_j .

The firing strength of a rule determines how much influence that rule has on the final output, given the current input, and it is calculated as:

$$\beta_j = \min(\mu_{A_{j1}}(x), \dots, \mu_{A_{jN}}(x)), \quad \text{for } j = 1, \dots, J$$

where:

- β_j is the firing strength of rule R_j ;
- $\mu_{A_{ji}}(x)$ is the membership function for the fuzzy set of feature i and cluster j ;

This transparent formulation enables practitioners to inspect directly the contribution of each input variable to the final decision, thus facilitating compliance with ethical guidelines and regulatory requirements, *e.g.*, the European AI Act.

pyFUME is a Python library that was created for automatic estimation of Takagi–Sugeno fuzzy models from data [32]. The library implements all essential steps for the data-driven generation of FIS:

a) *Data processing*: loading and normalization.

b) *Clustering*: partitioning the input-output space using clustering techniques such as classic Fuzzy C-Means [33] or a variant based on Fuzzy Self-Tuning Particle Swarm Optimization [34];

c) *Antecedent Estimation*: determining the membership functions that define the fuzzy sets associated with each input variable involves estimating the antecedent sets by fitting the convex envelope to the given membership values;

d) *Consequent Estimation*: estimating the parameters of the linear functions for each rule involves utilizing least-squares fitting techniques to determine the consequents, which can take the form of either zero-order ($y_j = c_j$) or first-order ($y_j = \mathbf{a}_j^\top \mathbf{x} + b_j$) linear models. The final output y^* is computed as a weighted combination of the individual rule outputs, as follow:

$$y^* = \frac{\sum_{j=1}^J \beta_j y_j}{\sum_{j=1}^J \beta_j}.$$

e) *Final Step*: Lastly, constructing an executable fuzzy model is performed using the *Simplful* library [35], which enables the design and assessment of the fuzzy rules in natural language. Although pyFUME supports many additional functionalities (*e.g.*, imputation, feature selection [36], automatic rules simplification [37], model size reduction [38]), they were not exploited in this work and will not be discussed further.

By integrating data-driven estimation and rule-based formulation, pyFUME closes the explanation-prediction gap. The clear design of fuzzy rules enables the inspection of single rules and their inherent membership functions, investigation of the input variable influence on the output via the linear consequent design, and generation of human-interpretable explanations of the decisions made by the model. These capabilities are an imperative in safety-critical fields for the validation of decisions. In a concise way, pyFUME is a framework for the integration of data-driven automated modeling and the principles of interpretability. By estimating Takagi–Sugeno fuzzy models through an open, step-by-step pipeline, pyFUME allows practitioners to develop AI systems that are both performant and explainable in order for state-of-the-art machine learning models to be trusted and interpretable in real-world, high-stakes applications.

One limitation of pyFUME – and FIS in general – is that it is not straightforward to determine which rule is responsible for the prediction of a given output, because it is ultimately dependent on the linear model of the consequent. Stated otherwise, it can be tricky to disambiguate *a priori* which rule is describing a given subset of data (*e.g.*, healthy patients vs patients affected by a disease). In this work we propose a novel algorithm to perform such disambiguation, described in the next section.

III. DISAMBIGUATION OF THE RULES WITH RuDe

The model produced by pyFUME consists of a set of rules in the form of Equation 1. In this manuscript we introduce a novel method for the automatic disambiguation of the role of the identified fuzzy rules, named RuDe.

RuDe operates as follows. For each sample in the dataset, we compute both the firing strength of the rule and the corresponding model prediction. To further clarify the contribution

of each rule in the decision process, we also determine and plot a smooth “contour” of the firing strengths. This upper bound is obtained by fitting a polynomial (up to a user-defined order) to the data points, with the additional constraint that the polynomial value is always greater than or equal to the observed firing strength at any given prediction.

The upper bound is derived by solving a linear optimization problem. The objective is to minimize the least squares error between the fitted polynomial and the actual firing strengths, subject to the constraint that the fitted polynomial does not fall below the observed firing strength for any data point. Formally, such optimization problem is defined as follows:

$$\min_{\mathbf{w}} \sum_{i=1}^n \left(\sum_{j=0}^d w_j p_i^j - s_i \right)^2$$

subject to:

$$\sum_{j=0}^d w_j p_i^j \geq s_i, \quad \forall i \in \{1, 2, \dots, n\}$$

where:

- n is the number of data points;
- d is the degree of the polynomial;
- w_j are the polynomial regression coefficients to be optimized;
- p_i are the model’s predictions;
- s_i are the corresponding firing strengths.

In a more compact notation, the problem can be written as:

$$\min_{\mathbf{w}} \|\mathbf{V}(n, d) \cdot \mathbf{w} - \mathbf{s}\|_2^2 \quad \text{subject to} \quad \mathbf{V}(n, d) \cdot \mathbf{w} \geq \mathbf{s}$$

where $\mathbf{V}(n, d)$ is the Vandermonde matrix constructed from the predictions, $\mathbf{w}$ is the column vector of polynomial coefficients, and $\mathbf{f}$ is the vector of firing strengths. To solve this optimization problem, we exploited the HIGHS solver in conjunction with the Python library `cvxpy` [40], [41]. According to our experiments, $d = 3$ yields the best results, which is hence used as default for RuDe.

IV. RESULTS

In this section we present and discuss an interpretable model, based on a FIS built with `pyFUME`, for the accurate prediction of LVEF exploiting strain data. We considered a data-set composed of 100 physically well-trained patients referred to the Institute of Radiology, Padua University Hospital, to perform CMR. The exams were performed with a 1.5T scanner. From the dataset, we extracted the data relative to Global Longitudinal Strain (GLS), Global Circumferential Strain (GCS), Long-axis Global Radial Strain (GRS LAX) and Short-axis Global Radial Strain (GRS SAX). These features were used to predict the left ventricle ejection fraction (LVEF), using an intrinsically interpretable model based on fuzzy logic.

All experiments that follow were performed on a workstation equipped with an Apple M2 SoC and 24 GB of unified memory. The main software packages used are: `Python` (3.11), `pyFUME` (0.3.4), `simplful` (2.12.0), `scipy` (1.15.1),

`numpy` (2.2.2), `pandas` (2.1.1), `scikit-learn` (1.3.0) and `cvxpy` (1.6.3). The model was automatically built with the `pyFUME` library, using Fuzzy C-Means clustering, three clusters, no merging thresholds, and degree 3 for the RuDe algorithm. The calculated RMSE of the model was 4.2.

Figure 5 shows the fuzzy sets for the four linguistic variables. There is a clear separation between the three clusters, with a smooth transition between the fuzzy sets.

Cluster 3 (green lines) suggests the following interpretation:

“Patients that have high peak longitudinal strain, high peak circumferential strain, low peak radial strain along the short axis, and low peak radial strain along the long axis”.

Now that we have a description of the prototypal patient associated to this cluster, we want to characterize the response of the AI system, that is, which kind of output this rule predicts. The RuDe algorithm can give a hint to disambiguate the role of Cluster 1 / Rule 1. The output of the algorithm is shown in Figure 6. The figure highlights that Rule 1 (green lines) is responsible for the prediction of a low value for the LVEF. Thus, the third cluster/rule describes, and predicts, patients characterized by reduced heart function. Interestingly, this conclusion is also supported by the normality ranges taken from the literature (see light gray areas in Figure 5): as a matter of fact, the membership to this fuzzy set/cluster tends to have a maximum two standard deviations far from the average normality value.

Let us now discuss the first rule. Figure 5 shows that Cluster 1 (blue lines) describes:

“Patients having medium peak longitudinal strain, medium peak circumferential strain, medium peak radial strain along the short axis, and medium peak radial strain along the long axis”.

If we compare these antecedents with the output of RuDe we see that they correspond high levels of LVEF, denoting a normal heart function. As a matter of fact, the peak of the membership functions for this group/cluster is very close to the average normality value across all linguistic variables.

Finally, let us consider the rule number two (orange lines), that is:

“Patients having very low peak longitudinal strain, very low peak circumferential strain, high peak radial strain along the short axis, and high peak radial strain along the long axis”.

This group of patients seems to be very peculiar. The fuzzy set have highest memberships within the normality ranges (Figure 5), but Figure 6 highlights that this group is characterized by a higher LVEF (orange lines, Rule 2). Indeed, this group is mainly composed of highly trained athletes practicing sports that remodel the myocardium. Further investigation will be necessary to fully understand and characterize this cluster.

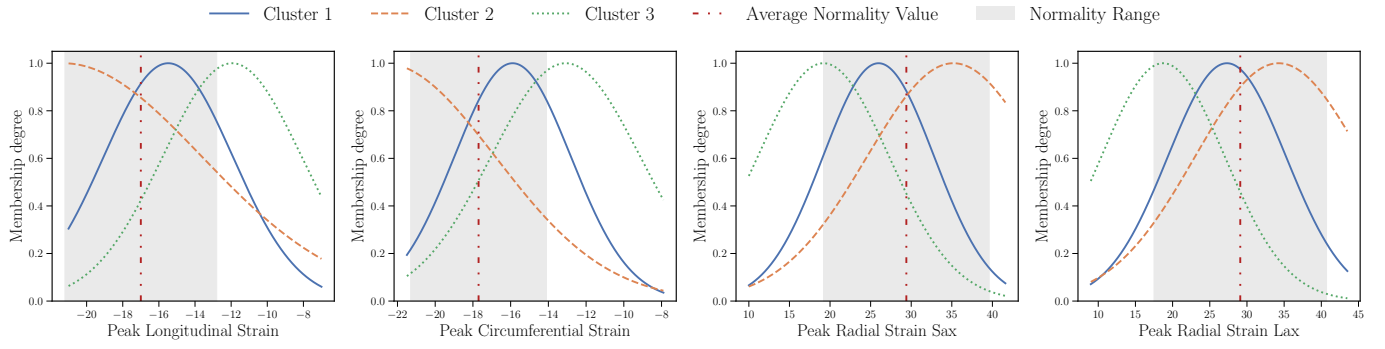


Fig. 5: Fuzzy sets produced by pyFUME for the first model. From left to right: global longitudinal strain, global circumferential strain, long and short axis global radial strain, cardiac index. The literature average normality values for the strain [39], exploited in the clinical settings, are denoted by the grey area

Thanks to its fuzziness, the FIS is able to smoothly transition between the three rules according to the characteristics of the patients (see Figure 6). It can be interesting to investigate the impact of the input features to the rule, its strength, and its predictions. Figure 7 summarizes the importance of the input features for the three rules as computed by pyFUME. The importances are calculated similarly to [42], that is, by normalizing the parameters of the linear regression models of the rules' consequents. If we write $\mathbf{X}_i$ for the i -th input row vector $\mathbf{x}$, and $\mathbf{y}$ for the output column vector, and $a_j^{(i)}$ is the i -th coefficient for the j -th rule, the corresponding variable importance is calculated as:

$$\tilde{a}_j^{(i)} = a_j^{(i)} \frac{\max \mathbf{X}_i - \min \mathbf{X}_i}{\max \mathbf{y} - \min \mathbf{y}}.$$

The results shown in the figure seems to suggest that:

- for Rule 1 and 2, the peak radial strain along the short axis is the most important feature. Interestingly, it is the least important feature for Rule 3 (that is, patients with higher cardiac functionality);
- in the case of Rule 3, the peak circumferential strain inversely affects the predicted LVEF, which is almost irrelevant for the predictions made by Rule 2.

As a final remark, the highest importance of peak radial strain sax in Figure 7 for Rule 1 can be explained by the fact that this variable is calculated using both the information about longitudinal and circumferential deformations, so that it is the most helpful to detect the nominal parameterizations.

Concerning the interpretation of the rules and the output of RuDe, the clinicians consider – in this case – the figure straightforward to interpret: the green, blue and orange lines describe which intervals are associated to rules 3, 1 and 2,

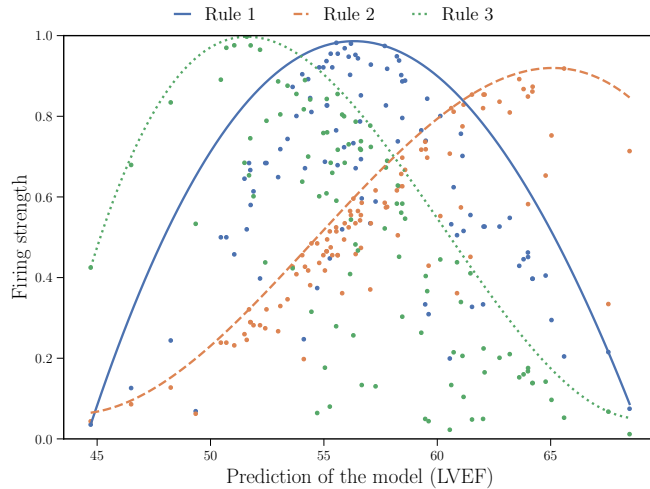


Fig. 6: Rules disambiguation provided by the RuDe algorithm. From visual inspection, it is clear that Rule 1 (blue line) is responsible for the prediction of lower LVEF, arguably corresponding to patients characterized by mildly reduced heart function. On the contrary, Rule 3 (orange line) is mainly activated to predict high values of LVEF.

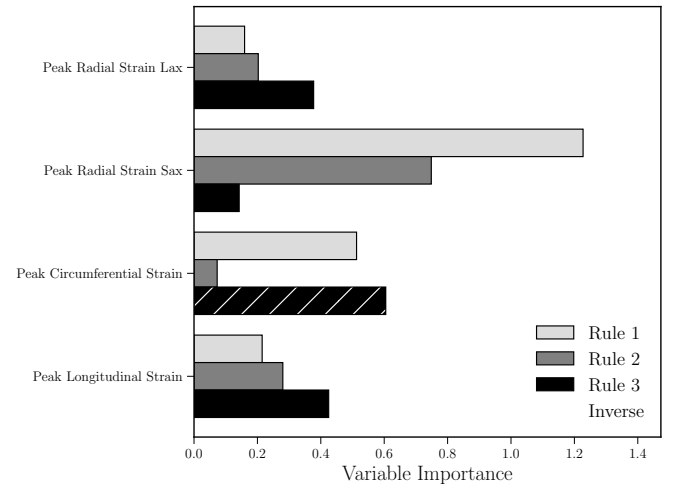


Fig. 7: Feature importance calculated from normalized rules' consequent parameters. The hatches on a bar denote a inverse importance of a given variable for a specific rule, that is, the variable has an inverse proportionality to the predicted output value for that rule.

respectively. The output also suggests that there are three clear and well-distinct sub-groups in the data, which corresponds to the expectations.

V. COMPARATIVE EVALUATION OF PREDICTIVE MODELS

Model	RMSE	MAPE	MAE
pyFUME	4.3130	0.0577	3.2438
Linear Regression	4.3454	0.0590	3.3318
MLP	8.1314	0.1140	6.5696
DT	5.9211	0.0839	4.7858
SVM	5.0030	0.0670	3.7990

TABLE I: Performance comparison of predictive models using average RMSE, MAPE and MAE on 5 folds.

Table I compares the performance of five predictive modeling strategies using five-fold cross-validation and traditional performance measures—Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). The results illustrated that the pyFUME model alone had the lowest mean RMSE (4.3130), MAPE (0.0577), and MAE (3.2438) than all models tested. These findings mean that pyFUME was not only the only model with accurate point estimates, it has greater consistency and stability with more minimized errors. The traditional models Linear Regression and Decision Tree (DT) as well as the more non-linear models such as Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) all had higher errors along each metric. The stronger evidence we can conclude is that it is clear pyFUME offers a more reliable and robust approach to predictive modeling.

VI. CONCLUSIONS

In this paper we discussed an interpretable AI system for the prediction of the left ventricle ejection fraction exploiting myocardial strain data. Our method extends the pyFUME library by adding a novel algorithm for the disambiguation of the rules' role named RuDe, which generates a plot that describes the relationship between the predicted output and the corresponding rule.

According to our results and to normal strain ranges taken from the literature, our prediction model allows to discriminate between patients with preserved and mildly reduced LVEF, confirming the correspondence between function and strain data, and identifying SAX GRS as the most important feature. These data can help to better understand myocardial deformation, making it suitable for the implementation of interpretable and trustable medical decision support systems. As future development, the application in different pathological conditions would allow to identify patients with preserved function but altered strain and *vice versa*, due to heart remodeling. Moreover, strain analysis carried out with a segmental approach should provide quantitative information also about regional myocardial injuries.

This work opens the way to trustworthy systems that go towards compliance to EU legal provisions and ethical principles. In fact, the inherent interpretability of our model helps

physicians understand the reasoning behind the output and the logic involved in the processing of health data, allowing for a better opportunity for patients to express an informed consent and exercise the rights enshrined in the GDPR and AI Act.

ACKNOWLEDGMENT

This work was partially funded by the European Union's Horizon Europe program with the DataCom Project (grant agreement no. 101108151). Views and opinions expressed are those of the author(s) only and do not reflect those of the EU or the European Commission.

REFERENCES

- [1] K. Kusunose, R. Zheng, H. Yamada, and M. Sata, "How to standardize the measurement of left ventricular ejection fraction," *Journal of medical ultrasonics*, vol. 49, no. 1, pp. 35–43, 2022.
- [2] A. M. Maceira, S. K. Prasad, M. Khan, and D. J. Pennell, "Normalized left ventricular systolic and diastolic function by steady state free precession cardiovascular magnetic resonance," *Journal of Cardiovascular Magnetic Resonance*, vol. 8, no. 3, pp. 417–426, 2006.
- [3] B. F. Bulten, A. M. Mavinkurve-Groothuis, L.-F. de Geus-Oei, A. F. de Haan, C. L. de Korte, L. Bellersen, H. W. van Laarhoven, and L. Kapusta, "Early myocardial deformation abnormalities in breast cancer survivors," *Breast cancer research and treatment*, vol. 146, pp. 127–135, 2014.
- [4] B. Brady, G. King, R. T. Murphy, and D. Walsh, "Myocardial strain: A clinical review," *Irish Journal of Medical Science (1971-)*, vol. 192, no. 4, pp. 1649–1656, 2023.
- [5] A. Pepe, F. Crimi, F. Vernuccio, G. Cabrelle, A. Lupi, C. Zanon, S. Gambato, A. Perazzolo, and E. Quaia, "Medical radiology: current progress," *Diagnostics*, vol. 13, no. 14, p. 2439, 2023.
- [6] E. Falletti and C. Gallese, "Credit scoring judicial review between the Court of Justice of the European Union and comparative case law," *MediaLaws*, 2024.
- [7] C. Gallese, "Legal aspects of AI in the biomedical field. the role of interpretable models," in *Big Data Analysis and Artificial Intelligence for Medical Sciences*, John Wiley & Sons, 2024.
- [8] A. . D. P. W. Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)*. European Commission, 2018.
- [9] C. Gallese, "The AI Act proposal: a new right to technical interpretability?," in *Internet Law and Digital Society: An International Overview*, Milano University Press, 2025.
- [10] E. Ballarin, L. Bortolussi, F. Cairoli, C. Gallese, L. Nenzi, and G. Saveri, "Reliable and Explainable AI in Trieste," in *Proceedings of Ital-IA 2023: 3rd National Conference on Artificial Intelligence*, 2023.
- [11] C. Rudin and J. Radin, "Why are we using black box models in AI when we don't need to? A lesson from an explainable AI competition," *Harvard Data Science Review*, vol. 1, no. 2, 2019.
- [12] H. Abbass, K. Crockett, J. Garibaldi, A. Gegov, U. Kaymak, and J. M. C. Sousa, "From explainable artificial intelligence (xAI) to understandable artificial intelligence (uAI)," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 9, pp. 4310–4314, 2024.
- [13] G. D. Aquaro, M. Perfetti, G. Camastra, L. Monti, S. Dellegrottaglie, C. Moro, A. Pepe, G. Todiere, C. Lanzillo, A. Scatteia, *et al.*, "Cardiac MR with late gadolinium enhancement in acute myocarditis with preserved systolic function: ITAMY study," *Journal of the American College of Cardiology*, vol. 70, no. 16, pp. 1977–1987, 2017.
- [14] S. Pujadas, G. P. Reddy, O. Weber, J. J. Lee, and C. B. Higgins, "MR imaging assessment of cardiac function," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 19, no. 6, pp. 789–799, 2004.
- [15] G. Pons-Llado, "Assessment of cardiac function by CMR," *European radiology*, vol. 15, pp. B23–32, 2005.
- [16] Circle Cardiovascular Imaging Inc., "cvi42@." <https://www.circlecvi.com>.
- [17] S. I. Mavrogeni, P. P. Sfikakis, L. Koutsogeorgopoulou, G. Markousis-Mavrogenis, T. Dimitroulas, G. Kolovou, and G. D. Kitas, "Cardiac tissue characterization and imaging in autoimmune rheumatic diseases," *JACC: Cardiovascular Imaging*, vol. 10, no. 11, pp. 1387–1396, 2017.

- [18] J. Xu, W. Yang, S. Zhao, and M. Lu, "State-of-the-art myocardial strain by CMR feature tracking: clinical applications and future perspectives," *European Radiology*, vol. 32, no. 8, pp. 5424–5435, 2022.
- [19] P. P. Sengupta and J. Narula, "Reclassifying heart failure: predominantly subendocardial, subepicardial, and transmural," *Heart failure clinics*, vol. 4, no. 3, pp. 379–382, 2008.
- [20] P. P. Sengupta, A. J. Tajik, K. Chandrasekaran, and B. K. Khandheria, "Twist mechanics of the left ventricle: principles and application," *JACC: Cardiovascular Imaging*, vol. 1, no. 3, pp. 366–376, 2008.
- [21] A. Scatteia, A. Baritussio, and C. Bucciarelli-Ducci, "Strain imaging using cardiac magnetic resonance," *Heart failure reviews*, vol. 22, pp. 465–476, 2017.
- [22] M. S. Amzulescu, M. De Craene, H. Langet, A. Pasquet, D. Vancraeynest, A.-C. Pouleur, J.-L. Vanoverschelde, and B. Gerber, "Myocardial strain imaging: review of general principles, validation, and sources of discrepancies," *European Heart Journal-Cardiovascular Imaging*, vol. 20, no. 6, pp. 605–619, 2019.
- [23] G. Pedrizzetti, P. Claus, P. J. Kilner, and E. Nagel, "Principles of cardiovascular magnetic resonance feature tracking and echocardiographic speckle tracking for informed clinical use," *Journal of cardiovascular magnetic resonance*, vol. 18, no. 1, p. 51, 2016.
- [24] M. Dobrovie, M. Barreiro-Pérez, D. Curione, R. Symons, P. Claus, J.-U. Voigt, and J. Bogaert, "Inter-vendor reproducibility and accuracy of segmental left ventricular strain measurements using CMR feature tracking," *European Radiology*, vol. 29, pp. 6846–6857, 2019.
- [25] R. Heinke, F. Pathan, M. Le, T. D'Angelo, L. Winau, C. Arendt, T. J. Vogl, A. Zeiher, E. Nagel, and V. O. Puntmann, "Towards standardized postprocessing of global longitudinal strain by feature tracking—OptiStrain CMR-FT study," *BMC cardiovascular disorders*, vol. 19, pp. 1–11, 2019.
- [26] T. Podlesnikar, G. Pizarro, R. Fernandez-Jimenez, J. M. Montero-Cabezas, J. Sanchez-Gonzalez, C. Bucciarelli-Ducci, N. Ajmone Marsan, Z. Fras, J. J. Bax, V. Fuster, *et al.*, "Effect of early metoprolol during ST-segment elevation myocardial infarction on left ventricular strain: feature-tracking cardiovascular magnetic resonance substudy from the METOCARD-CNIC trial," *JACC: Cardiovascular Imaging*, vol. 12, no. 7 Part 1, pp. 1188–1198, 2019.
- [27] H. Liu, J. Wang, Y. Pan, Y. Ge, Z. Guo, and S. Zhao, "Early and quantitative assessment of myocardial deformation in essential hypertension patients by using cardiovascular magnetic resonance feature tracking," *Scientific Reports*, vol. 10, no. 1, p. 3582, 2020.
- [28] D. Augustine, A. J. Lewandowski, M. Lazdam, A. Rai, J. Francis, S. Myerson, A. Noble, H. Becher, S. Neubauer, S. E. Petersen, *et al.*, "Global and regional left ventricular myocardial deformation measures by magnetic resonance feature tracking in healthy volunteers: comparison with tagging and relevance of gender," *Journal of Cardiovascular Magnetic Resonance*, vol. 15, no. 1, p. 8, 2013.
- [29] R. J. Taylor, W. E. Moody, F. Umar, N. C. Edwards, T. J. Taylor, B. Stegmann, J. N. Townend, K. N. Hor, R. P. Steeds, W. Mazur, *et al.*, "Myocardial strain measurement with feature-tracking cardiovascular magnetic resonance: normal values," *European Heart Journal-Cardiovascular Imaging*, vol. 16, no. 8, pp. 871–881, 2015.
- [30] F. Andre, H. Steen, P. Matheis, M. Westkott, K. Breuninger, Y. Sander, R. Kammerer, C. Galuschky, E. Giannitsis, G. Korosoglou, *et al.*, "Age-and gender-related normal left ventricular deformation assessed by cardiovascular magnetic resonance feature tracking," *Journal of Cardiovascular Magnetic Resonance*, vol. 17, no. 1, p. 25, 2015.
- [31] H. Liu, D. Yang, K. Wan, Y. Luo, J.-Y. Sun, T.-J. Zhang, W.-H. Li, A. Greiser, M.-P. Jolly, Q. Zhang, *et al.*, "Distribution pattern of left-ventricular myocardial strain analyzed by a cine MRI based deformation registration algorithm in healthy Chinese volunteers," *Scientific reports*, vol. 7, no. 1, p. 45314, 2017.
- [32] C. Fuchs, S. Spolaor, M. S. Nobile, and U. Kaymak, "pyFUME: a python package for fuzzy model estimation," in *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–8, 2020.
- [33] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & geosciences*, vol. 10, no. 2-3, pp. 191–203, 1984.
- [34] M. S. Nobile, P. Cazzaniga, D. Besozzi, R. Colombo, G. Mauri, and G. Pasi, "Fuzzy Self-Tuning PSO: A settings-free algorithm for global optimization," *Swarm and evolutionary computation*, vol. 39, pp. 70–85, 2018.
- [35] S. Spolaor, C. Fuchs, P. Cazzaniga, U. Kaymak, D. Besozzi, and M. S. Nobile, "Simpful: a user-friendly python library for fuzzy logic," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 1687–1698, 2020.
- [36] A. Tangherloni, P. Cazzaniga, N. Stranieri, F. M. Buffa, and M. S. Nobile, "A fast feature selection for interpretable modeling based on fuzzy inference systems," in *2024 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (IEEE CIBCB)*, pp. 1–8, IEEE, 2024.
- [37] C. Fuchs, S. Spolaor, M. S. Nobile, and U. Kaymak, "A graph theory approach to fuzzy rule base simplification," in *International conference on information processing and management of uncertainty in knowledge-based systems*, pp. 387–401, Springer, 2020.
- [38] C. Fuchs, U. Kaymak, and M. S. Nobile, "Building interpretable and parsimonious fuzzy models using a multi-objective approach," in *2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp. 1–8, IEEE, 2022.
- [39] C. Lim, E. Blaszczyk, L. Riazzy, S. Wiesemann, J. Schüler, F. von Knobelsdorff-Brenkenhoff, and J. Schulz-Menger, "Quantification of myocardial strain assessed by cardiovascular magnetic resonance feature tracking in healthy subjects—influence of segmentation and analysis software," *European radiology*, vol. 31, pp. 3962–3972, 2021.
- [40] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization," *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.
- [41] A. Agrawal, R. Verschuere, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems," *Journal of Control and Decision*, vol. 5, no. 1, pp. 42–60, 2018.
- [42] L. Bacciu, P. Cazzaniga, C. Gallese, C. Fuchs, U. Kaymak, D. M. Papetti, and M. S. Nobile, "Our fruitful relationship with sugeno inference from FUMOSO to pyFUME," in *Proceedings of the 20th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2024)*, (Lisbon, Portugal), 2024.