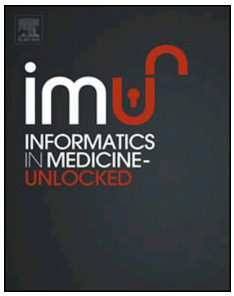
Precision Diagnostics in Cardiac Tumours

Sadegh-Zadeh, Seyed-Ali; Khezerlouy-aghdam, Naser; Sakha, Hanieh; Toufan, Mehrnoush; Behravan, Mahsa; Vahedi, Amir; Rahimi, Mehran; Hosseini, Haniyeh; Khanjani, Sanaz; Bayat, Bita; Ali, Syed Ahsan; Hajizadeh, Reza; Eshraghi, Ali; Ghidary, Saeed Shiry; Saadat, Mozafar

DOI:
10.1016/j.imu.2024.101544



Precision Diagnostics in Cardiac Tumours: Integrating Echocardiography and Pathology with Advanced Machine Learning on Limited Data

Seyed-Ali Sadegh-Zadeh, Naser Khezerlouy-aghdam, Hanieh Sakha, Mehrnoush Toufan, Mahsa Behravan, Amir Vahedi, Mehran Rahimi, Haniyeh Hosseini, Sanaz Khanjani, Bita Bayat, Syed Ahsan Ali, Reza Hajizadeh, Ali Eshraghi, Saeed Shiry Ghidary, Mozafar Saadat

PII: S2352-9148(24)00100-X

DOI: https://doi.org/10.1016/j.imu.2024.101544

Reference: IMU 101544

To appear in: *Informatics in Medicine Unlocked*

Received Date: 26 April 2024

Revised Date: 30 June 2024

Accepted Date: 30 June 2024

Please cite this article as: Sadegh-Zadeh S-A, Khezerlouy-aghdam N, Sakha H, Toufan M, Behravan M, Vahedi A, Rahimi M, Hosseini H, Khanjani S, Bayat B, Ali SA, Hajizadeh R, Eshraghi A, Ghidary SS, Saadat M, Precision Diagnostics in Cardiac Tumours: Integrating Echocardiography and Pathology with Advanced Machine Learning on Limited Data, *Informatics in Medicine Unlocked*, https://doi.org/10.1016/j.imu.2024.101544.

Precision Diagnostics in Cardiac Tumours: Integrating Echocardiography and Pathology with Advanced Machine Learning on Limited Data

Seyed-Ali Sadegh-Zadeh[*1], Naser Khezerlouy-aghdam[2], Hanieh Sakha[2], Mehrnoush Toufan[2], Mahsa Behravan[2], Amir Vahedi[3], Mehran Rahimi[2], Haniyeh Hosseini[2], Sanaz Khanjani[4], Bita Bayat[4], Syed Ahsan Ali[5], Reza Hajizadeh[6], Ali Eshraghi[7], Saeed Shiry Ghidary[1], Mozafar Saadat[*8]

[1]Department of Computing, School of Digital, Technologies and Arts, Staffordshire University, Stoke-on-Trent, England, United Kingdom ali.sadegh-zadeh@staffs.ac.uk, saeed.shiryghidary@staffs.ac.uk

[2]Cardiovascular research centre, Tabriz university of medical sciences, Tabriz, Iran. nkhezerlou90@gmail.com , h.dsakha@yahoo.com, mtoufan@gmail.com, Mahsabhr75@gmail.com, Rahimimehran95@gmail.com, Haniyeh96hosseini@gmail.com

[3]Department of Pathology, Tabriz University of Medical Sciences, Tabriz, Iran.Amirvahedy@gmail.com

[4] Department of Computer Engineering, Razi University, Kermanshah, Iran. s.khanjani@stu.razi.ac.ir , Bita.bayat6213@gmail.com

[5]Health Education England West Midlands, Birmingham, England, United Kingdom. Ahsan.Ali1@nhs.net

[6]Department of Cardiology, Urmia University of Medical Sciences, Urmia, Iran, hajizadh.reza@gmail.com

[7]Tehran University of medical sciences, Medical school, Children's medical centre hospital, Paediatric cardiology department, alieshraghi1360@gmail.com

[8]Department of Mechanical Engineering, School of Engineering, University of Birmingham, Birmingham, United Kingdom, m.saadat@bham.ac.uk

Abstract

This study pioneers the integration of echocardiography and pathology data with advanced machine learning (ML) techniques to significantly enhance the diagnostic accuracy of cardiac tumours, a critical yet challenging aspect of cardiology. Despite advancements in diagnostic methods, cardiac tumours' nuanced complexity and rarity necessitate more precise, non-invasive, and efficient diagnostic solutions. Our research aims to bridge this gap by developing and validating ML models—Support Vector Machines (SVM), Random Forest (RF), and Gradient Boosting Machines (GBM)—optimized for limited datasets prevalent in specialized medical fields. Utilizing a dataset comprising clinical features from 399 patients at the Heart Hospital, our study meticulously evaluated the performance of these models against traditional diagnostic metrics. The RF model emerged superior, achieving a groundbreaking accuracy of 96.25% and a perfect ROC AUC score of 0.99, significantly outperforming existing diagnostic approaches. Key predictors identified include age, *echo malignancy*, and *echo position*, underscoring the value of integrating diverse data types. Clinical validation conducted at the Heart Hospital further confirmed the models' applicability and reliability, with the RF model demonstrating a diagnostic accuracy of 94% in a real-world setting. These findings advocate for the potential of ML in revolutionizing cardiac tumour diagnostics, offering pathways to more accurate, non-invasive, and patient-centric diagnostic processes. This research not only highlights the capabilities of ML to enhance diagnostic precision in the realm of cardiac tumours but also sets a foundation for future explorations into its broader applicability across various domains

of medical diagnostics, emphasizing the need for expanded datasets and external validation.

Keywords: Cardiac Tumours, Machine Learning, Diagnostic Accuracy, Echocardiography, Pathology, Support Vector Machines, Random Forest, Gradient Boosting Machines, Feature Importance, Clinical Validation.

1 Introduction

Cardiac tumours, though relatively rare, present a significant challenge in clinical cardiology due to their diverse manifestations and diagnostic complexities[1]. These tumours can be classified into primary and secondary (metastatic) types, with primary tumours being less common[2][3]. Among primary cardiac tumours, myxomas are the most frequently diagnosed, constituting about 50% of the cases[4]. However, the rarity of these tumours means that there is limited widespread expertise in their diagnosis and treatment. The diagnosis of cardiac tumours poses unique challenges. The primary hurdle lies in the nonspecific nature of the symptoms, which often mimic other cardiac conditions such as valve disease or heart failure[4][5][6]. This symptomatic ambiguity can lead to misdiagnosis or delayed diagnosis, impacting patient outcomes. Additionally, the localization and characterization of these tumours require advanced imaging techniques[7]. Echocardiography remains the first-line imaging modality, but it has limitations in differentiating tumour types and determining malignancy[8][9]. Moreover, the varied nature of cardiac tumours means that a multi-modal diagnostic approach is often necessary. This can involve a combination of imaging techniques such as cardiac MRI, CT scans, and even invasive procedures like biopsy, which are not without risks[10].

Cardiac tumours, though relatively rare, present a significant challenge in clinical cardiology due to their diverse manifestations and diagnostic complexities [11][12]. These tumours can be classified into primary and secondary (metastatic) types, with primary tumours being less common [13][14]. Among primary cardiac tumours, myxomas are the most frequently diagnosed, constituting about 50% of the cases[15]. However, the rarity of these tumours means that there is limited widespread expertise in their diagnosis and treatment. Current diagnostic practices primarily rely on imaging modalities, with echocardiography being the most accessible and cost-effective. However, the reliance on imaging alone can be insufficient, especially in differentiating benign from malignant tumours or in cases where the tumour's characteristics are atypical[16]. In such scenarios, a combination of pathological evaluation and advanced imaging techniques becomes necessary. The integration of machine learning in diagnostic processes is an emerging trend, aimed at enhancing the accuracy and efficiency of cardiac tumour diagnosis. Machine learning models, particularly those leveraging advanced algorithms and big data analytics, show promise in improving diagnostic precision[17]. They offer the potential to uncover subtle patterns in imaging and pathology data that might be overlooked by traditional methods. However, the effectiveness of these models is often constrained by the limited availability of comprehensive and high-quality data, a common challenge in specialized medical fields [18].

The primary objective of this research is to develop a sophisticated machine learning model that integrates echocardiography imaging data and pathology test results, specifically

optimized for limited datasets, to enhance the accuracy and precision of cardiac tumour diagnosis. This model aims to address the challenge of differentiating between various types of cardiac tumours and determining their malignancy, a task that remains a significant hurdle in current medical practices. By leveraging advanced machine learning techniques, this study seeks to create a diagnostic tool that can process complex medical data more effectively and provide reliable diagnostic support to clinicians. This study holds considerable significance in the realm of medical diagnostics. Firstly, it addresses a critical gap in cardiac tumour diagnosis by utilizing machine learning to interpret complex echocardiography and pathology data, which could lead to earlier and more accurate identification of tumour types and malignancies. This advancement is particularly crucial given the potentially life-threatening nature of cardiac tumours and the importance of timely and accurate diagnosis for effective treatment planning. Furthermore, the development of a model that is specifically tailored to perform well on limited datasets is of paramount importance in medical research, where large datasets are often unavailable due to the rarity of certain conditions. By creating a model that can work effectively with smaller datasets, this research could pave the way for similar approaches in other specialized medical fields, thereby broadening the impact and applicability of machine learning in healthcare.

The paper is structured as follows: The Literature Review section provides an in-depth overview of current cardiac tumour diagnostic techniques, the utilization of machine learning in medical diagnosis, and the challenges in limited data scenarios. In the Materials and Methods section, we detail our study's methodology, encompassing data collection, preparation, machine learning models used, and training/evaluation strategies. The Results section showcases the machine learning model's performance metrics, comparing them with existing methods and emphasizing feature importance and interpretability. The Discussion section interprets the results, acknowledges study limitations, and explores clinical relevance and potential applications. Lastly, the Conclusion section summarizes key findings, underscores their significance in cardiac tumour diagnosis, and proposes future research directions.

2 Literature Review

2.1 Current Diagnostic Techniques

The diagnosis of cardiac tumours has traditionally relied on a blend of clinical assessment and imaging techniques. The primary goal in the diagnostic process is to identify the presence of a tumour, determine its nature (benign or malignant), and understand its implications on cardiac function[19]. The complexity of these tumours necessitates a multifaceted approach to diagnosis. Echocardiography stands as the cornerstone in the initial evaluation of cardiac masses[20]. Its non-invasive nature, wide availability, and ability to provide detailed information about the size, location, and hemodynamic impact of the tumour make it an invaluable tool. However, its efficacy is sometimes limited by the operator's expertise and the tumour’s position and characteristics. Echocardiography also struggles with specificity in distinguishing tumour types and identifying malignancies[16][21].

For a more comprehensive assessment, Cardiac MRI and CT scans are often employed[10][22]. Cardiac MRI, with its superior contrast resolution, is particularly effective in characterizing tissue composition, which is pivotal in differentiating benign from malignant tumours. CT scans, on the other hand, are excellent for evaluating calcification and the extent of the tumour, especially in cases where MRI is contraindicated. Despite their benefits, these

methods have limitations such as high costs, limited availability in some regions, and in the case of CT, exposure to radiation[22][23]. PET and SPECT are nuclear imaging techniques used less frequently but can be valuable, particularly in differentiating benign from malignant lesions and in the assessment of metastatic disease [24]. Their role, however, is usually complementary to echocardiography and MRI [25]. In instances where imaging results are inconclusive, a biopsy followed by a histological examination can provide definitive diagnosis. However, due to the invasive nature and associated risks of these procedures, they are typically reserved for cases where the imaging findings strongly suggest malignancy or when the diagnosis has significant therapeutic implications[4][26].

Despite the advances in imaging technologies, certain limitations persist. The interpretation of these imaging modalities often requires substantial expertise, and the quality of the results can be operator dependent. Moreover, distinguishing between different types of cardiac tumours solely based on imaging can be challenging. The scarcity of large-scale, comprehensive datasets for these rare tumours further complicates the development of standardized diagnostic criteria and protocols[27][16][9].

2.2 Machine Learning in Medical Diagnosis

The integration of machine learning in healthcare represents a transformative shift in medical diagnostics. Machine learning, a subset of artificial intelligence, involves training algorithms to recognize patterns and make decisions based on data. Its application spans various aspects of healthcare, including disease detection, prognosis, personalized treatment, and patient management [17], [28].

In diagnostic imaging, ML algorithms have been employed to enhance image analysis, providing a level of precision and efficiency beyond human capability. For instance, convolutional neural networks (CNNs), a type of deep learning model, have shown remarkable success in interpreting complex imaging data, such as MRI and CT scans. These models can identify subtle abnormalities or patterns indicative of specific diseases, thereby aiding in early detection and accurate diagnosis[29][30].

Specifically in cardiology, ML algorithms have been applied to echocardiography and cardiac MRI data to detect and classify cardiac diseases [31]. For example, ML models have been used to differentiate between various types of heart diseases, identify features indicative of cardiac dysfunction, and predict patient outcomes based on imaging findings[32][33]. The potential of ML in cardiology is particularly notable in handling the nuances and complexities of cardiac imaging, where traditional analysis may be limited[34].

A major challenge in applying ML in medical diagnosis is the availability and quality of data [35]. High-quality, annotated medical datasets are essential for training effective ML models. However, in specialized fields like cardiac tumour diagnosis, such datasets are often limited due to the rarity of the condition. This scarcity poses a significant challenge in developing robust and generalizable ML models[30][36].

Recent advancements in ML have focused on addressing the challenge of limited data[37]. Techniques like transfer learning, where a model developed for one task is reused as the starting point for a model on a second task, have shown promise[38]. Additionally, data augmentation methods and synthetic data generation can enhance the size and diversity of training datasets, improving the model's performance and reliability[39][40]. For instance, Ali et al. discuss various ML optimization techniques used for the prognosis of chronic kidney

disease and demonstrate how these methods can significantly improve predictive accuracy in medical diagnostics[41]. Similarly, Ali et al. present the DPEBic algorithm, which employs encoding and biclustering for detecting essential proteins in gene expressions, showcasing the utility of advanced ML algorithms in handling complex biological data[42]. Moreover, Ramachandra et al. highlight the effectiveness of ensemble ML techniques for pancreatic cancer detection, further emphasizing the relevance of ensemble methods in achieving high diagnostic performance in medical applications[43]. Machine learning's efficacy is largely contingent on the availability of large, diverse datasets. However, in specialized medical fields such as cardiac tumour diagnosis, the rarity of the condition leads to inherently small datasets. This limitation poses significant challenges for the development and performance of machine learning models [44].

One of the primary challenges with limited data is the risk of overfitting [45]. Overfitting occurs when a model learns the training data too well, including its noise and outliers, resulting in poor performance on new, unseen data. This challenge is particularly acute in healthcare, where the ability of a model to generalize to new patients or conditions is critical [46]. Another issue in small datasets is data imbalance, where some classes are underrepresented compared to others. In the context of cardiac tumours, certain tumour types may have far fewer instances than others. This imbalance can lead to biased models that perform well on majority classes but poorly on minority classes, which is problematic when each class's accurate identification is crucial.

2.3 Approaches to Mitigate Data Limitations

Despite these challenges, several strategies have been developed to mitigate the limitations of small datasets:

1. Data Augmentation: Techniques such as image rotation, flipping, or zooming can artificially expand the dataset. In the case of medical data, synthetic data generation techniques like GANs (Generative Adversarial Networks) can also be employed to create realistic, synthetic medical images[47].
2. Transfer Learning: Leveraging pre-trained models on large datasets from related tasks can be an effective strategy. These models can be fine-tuned with the limited data available, benefiting from the 'learned' features from the larger dataset[48].
3. Regularization Techniques: Methods such as L1 and L2 regularization can help prevent overfitting by penalizing the model for complexity[49].
4. Cross-Validation: Using techniques like stratified k-fold cross-validation ensures that the model is tested on all available data, maximizing training and validation effectiveness[50].
5. Ensemble Methods: Combining multiple models or using techniques like bagging and boosting can improve performance and robustness against overfitting [51].
6. Focus on Model Interpretability: Given the high stakes of medical diagnostics, emphasizing model interpretability is crucial. Simpler models or models with explain ability features are preferred to maintain trust and transparency in clinical settings[52].

Addressing these challenges also necessitates a multidisciplinary approach, combining the expertise of data scientists, medical professionals, and statisticians. Such collaboration

ensures that the models developed are not only technically sound but also clinically relevant and applicable.

3 Materials and Methods

3.1 Data Collection

This cross-sectional prospective study involved a cohort of 399 patients with confirmed cardiac masses (CM) who received treatment at Shahid Madani Medical Research and Training Hospital, a tertiary care centre, during the period from October 2021 to December 2022. Patients were identified and recruited based on the diagnosis of cardiac masses, which was initially established through echocardiography or subsequent confirmatory tests. Data for each patient were meticulously extracted from electronic medical records, ensuring accuracy and completeness. The following specific information related to the cardiac masses was collected:

1. **Type of Cardiac Mass:** Classification was based on pathology and echocardiography results, distinguishing between neoplastic and non-neoplastic lesions.
2. **Number of Masses:** The count of distinct masses identified per patient was documented.
3. **Size:** Measurements of the masses in their largest dimensions were recorded from the imaging reports.
4. **Location:** The anatomical location of the masses within the cardiac structure was identified through imaging studies.
5. **Tissue Consistency:** Characteristics of the mass tissue were observed on imaging and categorized into solid, cystic, or mixed consistency.
6. **Benign or Malignant Nature:** Determination of whether the masses were benign or malignant was based on pathology reports.

The study adhered strictly to ethical standards and was approved by the Ethics Committee of Tabriz University of Medical Sciences in 2021 (Permission code: IR.TBZMED.REC.1400.257). Informed consent was obtained from all participants, and patient confidentiality was maintained throughout the study. Data collection and handling procedures were designed to comply with relevant data protection regulations.

The initial statistical analysis of quantitative and qualitative variables was performed using the Statistical Package for the Social Sciences (SPSS) v. 24.0 (IBM Statistics, USA). Categorical variables were presented in frequencies and percentages. To further analyse the size of CMs, it was categorized into three subgroups: small (<4 mm), medium (4–7 mm), and large (>7 mm). The comparison of cardiac mass diagnosis by pathology versus echocardiography was subjected to statistical analysis using the Chi-square test for larger frequencies and Fisher's exact test for smaller frequencies. Cohen's Kappa coefficient was calculated to assess the agreement between the two methods. Finally, sensitivity, specificity, accuracy, positive and negative likelihood ratios, as well as positive and negative predictive values were calculated using the Medcalc online tool. The threshold for statistical significance was set at P values below 0.05.

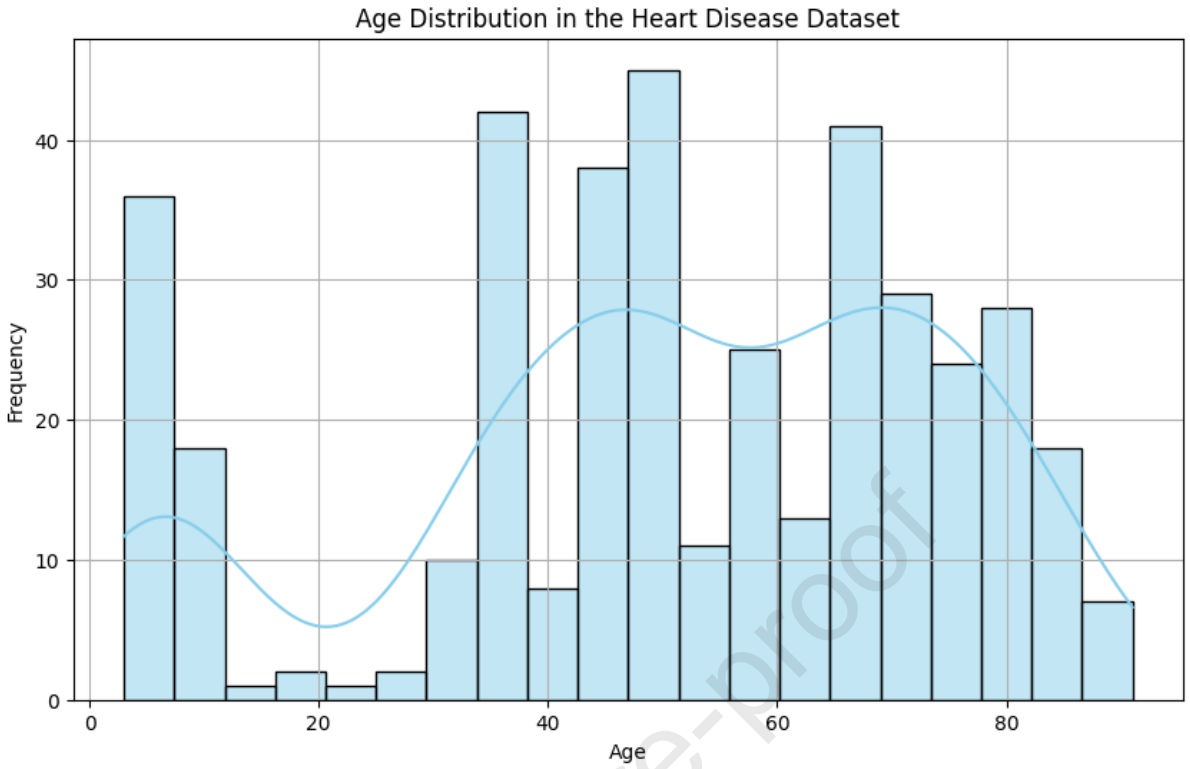


Figure 1. Age distribution in heart disease dataset

Figure 1 depicts a histogram with an overlaid line graph which serves to illustrate the age-related frequency of dataset. The histogram's bars represent the count of cases within various age intervals along the x-axis, indicating the data's distribution. Notably, there are visible peaks, particularly in the middle-aged to elderly population segments, suggesting a higher incidence of cardiac tumours in these groups. The smooth line graph suggests a trend or probability distribution across ages, highlighting the concentration of cases in certain age ranges, with significant peaks around ages 50 and 70.

Table 1 offers a detailed overview of the dataset attributes utilized in the study on cardiac tumours, distinguishing features by type—numerical or categorical (binary, nominal, or ordinal). The dataset exhibits a broad demographic and clinical spectrum, from basic demographic data like sex (binary: male or female) and age (numerical: ranging from 3 to 91 years) to more specific clinical characteristics such as family and personal history of heart disease (binary: yes or no), various echocardiography-related features (binary), and pathology-related features (binary). The numerical features like age, echo mass type, echo position, *echo size*, surgery position, and surgery size provide a quantitative analysis of the tumours and their medical assessment, while binary categorical features, prevalent throughout, underline the presence or absence of specific conditions or characteristics. The assortment of these features underscores the complexity of diagnosing and studying cardiac tumours, highlighting the necessity for meticulous preprocessing and analysis to handle the varied data types effectively for predictive modelling. This structured dataset enables a multifaceted approach to understanding cardiac tumour characteristics, crucial for developing accurate diagnostic and predictive models in medical research.

Table 1.Features of Cardiac Tumours dataset

Feature	***Type***	***Unique Values / Range***
sex (F=0, M=1)	Binary (Categorical)	0 - 1
age	Numerical	3 - 91
Family history of heart disease	Binary (Categorical)	0 - 1
History of heart disease	Binary (Categorical)	0 - 1
echotype-mass	Binary (Categorical)	0 - 1
echotype-myxoma	Binary (Categorical)	0 - 1
echotype-thrombose	Binary (Categorical)	0 - 1
echotype-fibroma	Binary (Categorical)	0 - 1
echotype-vegetation	Binary (Categorical)	0 - 1
echotype-papillary fibroelastoma	Binary (Categorical)	0 - 1
echomasstype	Numerical	0 - 6
echoposition	Numerical	0 - 13
echomalignancy	Binary (Categorical)	0 - 1
echosize	Numerical	0 - 5
echonumbers	Binary (Categorical)	0 - 1
echoconsistency	Binary (Categorical)	0 - 1
surgeryposition	Numerical	0 - 12
surgerysize	Numerical	0 - 6
surgerynumbers	Binary (Categorical)	0 - 1
pathotype-myxoma	Binary (Categorical)	0 - 1
pathotype-thrombose	Binary (Categorical)	0 - 1
pathotype-fibroma	Binary (Categorical)	0 - 1
pathotype-sarcoma	Binary (Categorical)	0 - 1
pathotype-Carcinoma	Binary (Categorical)	0 - 1
pathotype-lypoma	Binary (Categorical)	0 - 1
pathotype-vegetation	Binary (Categorical)	0 - 1
pathotype-endocarditis	Binary (Categorical)	0 - 1
pathotype-NBTE	Binary (Categorical)	0 - 1
pathotype-papillary fibroelastoma	Binary (Categorical)	0 - 1
pathomalignancy	Binary (Categorical)	0 - 1

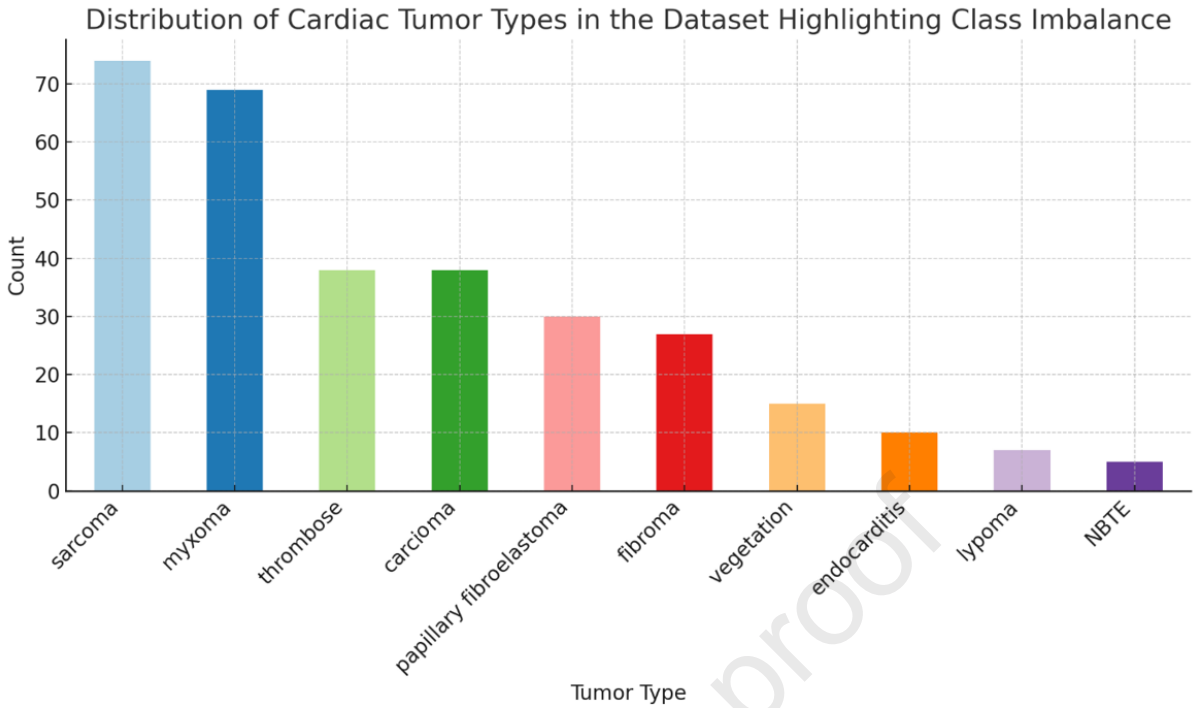


Figure 2. Distribution of Cardiac Tumour Types in the Dataset Highlighting Class Imbalance

Figure 2 presents the distribution of cardiac tumour types within the dataset, revealing significant class imbalances that pose analytical challenges. The bars differentiate each tumour type, illustrating a stark variance in occurrence rates—some types are notably prevalent, while others are scarcely represented. This imbalance is critical for machine learning endeavours, as it can skew model training, leading to overfitting on dominant classes and underperformance on rarer ones. The visual highlights the necessity for strategic data preprocessing techniques, such as oversampling minority classes or under-sampling majority classes, to ensure a more balanced data representation. Consequently, the figure underscores the importance of careful dataset analysis and preparation in the predictive modelling of cardiac tumours, aiming to enhance model sensitivity and specificity across all tumour types.

Figure 3 provides a comprehensive visual representation of the correlation coefficients between numerous variables related to cardiac tumour, including patient demographics, *echo types*, and pathotypes. At a glance, the colour scheme effectively differentiates between varying degrees of correlation, with the majority of the heatmap displaying darker shades, indicating generally low correlation between the variables. This suggests that most cardiac tumour conditions and patient characteristics operate independently of one another within this dataset. Notably, there are sparse areas of lighter shades, hinting at some level of positive correlation among specific conditions and characteristics. However, the lack of widespread strong correlations may imply the complexity of cardiac tumour ethology, emphasizing that its risk factors and manifestations are not singularly dependent on one another but are instead influenced by a multifaceted set of variables. This nuanced visualization underscores the importance of considering a broad spectrum of factors in cardiac tumour research and patient care.

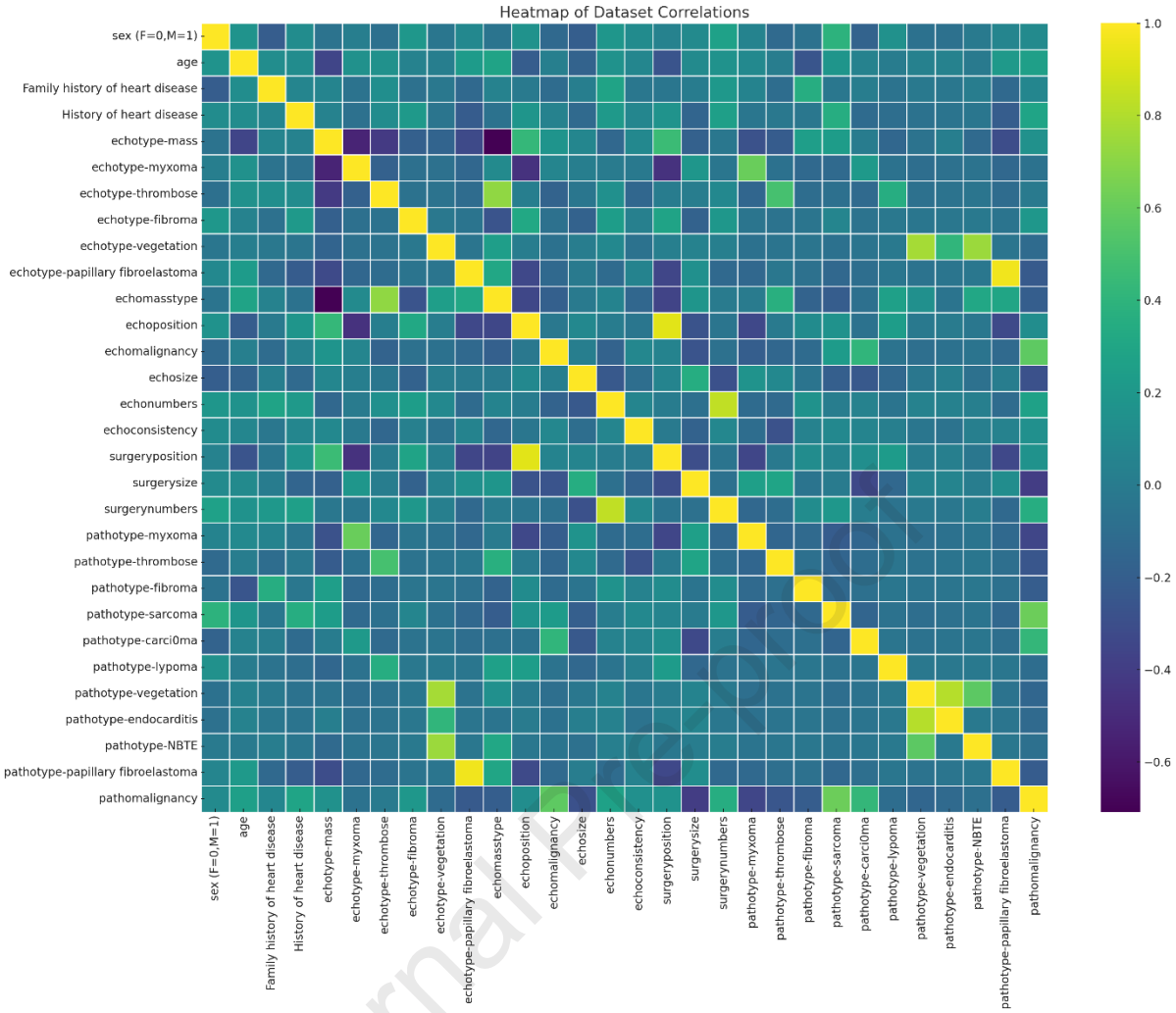


Figure 3. Correlation Heatmap of Patient Characteristics and Cardiac Tumour Conditions

3.2 Data Preparation

Data cleaning was the initial step to ensure the integrity and quality of our dataset. This process involved:

- **Removing Irrelevant Features**: Features not contributing to the diagnostic process, such as patient identifiers and timestamp data, were omitted from the analysis.
- **Handling Missing Values**: The dataset was scrutinized for missing entries across features. Given the critical nature of medical data, imputation strategies were carefully chosen based on the feature type. Numerical attributes with missing values were imputed using median values to mitigate the influence of outliers, whereas categorical features were imputed with the mode.
- **Eliminating Duplicate Records**: Duplicate entries were identified and removed to prevent biased model training.

3.3 Data Integration Process

To comprehensively combine echocardiography and pathology data, we employed a structured data integration methodology. Initially, echocardiographic parameters, such as each malignancy, composition, and echo size, were collected and digitized into our database. Parallelly, pathology

reports provided categorical data on tumour type and malignancy status. These datasets were pre-processed to handle missing values, normalize numerical data, and encode categorical variables. The integration of these diverse data sources facilitated the training of machine learning models capable of discerning complex patterns indicative of cardiac tumours.

3.3.1 Data Augmentation

Given the limited size of the dataset, a common challenge in specialized medical research, data augmentation was pivotal in enhancing the robustness of our machine learning models. Synthetic Minority Over-sampling Technique (SMOTE) was employed to address class imbalance ensuring an equitable representation of classes. This technique synthetically generates new instances of the minority class by interpolating between existing ones, thereby enriching the dataset without introducing bias.

3.3.2 Data Preprocessing

The preprocessing phase tailored the dataset for optimal machine learning model performance, involving:

- **Encoding Categorical Variables**: Categorical features, including binary and nominal variables, were encoded to facilitate their interpretation by machine learning algorithms. One-hot encoding was applied to nominal variables with no intrinsic order, whereas binary variables were simply converted to 0s and 1s.
- **Normalizing Numerical Data**: Numerical features were normalized to ensure uniform scale across variables, enhancing model convergence and performance. This step is critical in algorithms sensitive to feature magnitude, including neural networks and distance-based models.
- **Feature Selection**: Utilizing mutual information and correlation analysis, features with negligible impact on the target variable were pruned to streamline the model and focus on informative attributes.

Through these comprehensive data preparation steps, the dataset was transformed into a format conducive to developing a predictive model. This meticulous approach not only addresses the inherent challenges posed by the limited size and complexity of medical datasets but also lays a solid foundation for accurate and reliable diagnostic predictions in cardiac tumour cases.

3.4 Data Cleaning and Error Statistics

During the data cleaning phase, we identified and rectified several types of errors and inconsistencies within the dataset. The specific steps and their statistics are as follows:

1. **Missing Values**: We detected missing values across various features. For family history of heart disease 12 cases, for history of heart disease 8 cases, and for Echocardiography-related 5 cases:
    - Numerical features: Imputed using median values to handle outliers.
    - Categorical features: Imputed using mode values.
2. **Duplicate Records**: A total of 15 duplicate records were identified and removed to ensure unbiased model training.

This comprehensive approach to data cleaning ensured the integrity and quality of our dataset, forming a robust foundation for subsequent machine learning model development.

3.5 Machine Learning Models

Integrating echocardiography and pathology data to improve cardiac tumour diagnosis necessitates selecting machine learning models adept at handling the dataset's complexity and nuances. This section delineates the rationale behind choosing specific models and outlines their theoretical underpinnings and implementation nuances tailored to our study. Selecting the most appropriate machine learning algorithms for a study, especially one as critical as the diagnosis of cardiac tumours, requires a careful evaluation of the dataset's characteristics, the complexity of the task, and the specific objectives of the research. For our project, we chose Support Vector Machines (SVM), Random Forest (RF), and Gradient Boosting Machines (GBM) based on several key considerations that align with our goals of achieving high diagnostic accuracy, handling a potentially imbalanced dataset, and ensuring the interpretability of the models' predictions.

3.5.1 Support Vector Machines

Support Vector Machines were selected for their effectiveness in high-dimensional spaces, characteristic of medical datasets that integrate various types of clinical data, such as echocardiography and pathology reports. SVMs are renowned for their ability to find the optimal hyperplane that maximizes the margin between classes, making them particularly suited for binary classification tasks like distinguishing between benign and malignant cardiac tumours. SVMs are renowned for their effectiveness in high-dimensional spaces, making them an apt choice for our dataset, which features a rich set of attributes from echocardiography and pathology. The model's capability to use different kernel functions enables the exploration of linear and non-linear relationships between features and the target variable[53].

We employed a radial basis function (RBF) kernel to capture complex patterns within the data. The choice of the RBF kernel was motivated by its adaptability to varying data structures, which is paramount in medical diagnosis scenarios where the relationship between attributes and the condition of interest is not straightforward. Key parameters, including the penalty parameter C and the gamma parameter of the kernel, were optimized using a grid search approach with cross-validation to balance model complexity and generalization ability.

3.5.2 Random Forests (RF)

The Random Forest algorithm, an ensemble of decision trees, was selected for its robustness to overfitting and its ability to handle imbalanced datasets. Given the diversity of echocardiography and pathology features, RF's ensemble approach enhances prediction accuracy by aggregating insights from multiple decision trees, each trained on a subset of the data and features. Random Forest was chosen for its robustness to imbalanced datasets—a common challenge in medical diagnostics. The ability of RF to provide feature importance scores also adds a layer of interpretability to the model, offering insights into which clinical features are most predictive of tumour malignancy [54].

We configured the Random Forest model with a specified number of trees to ensure a deep exploration of the feature space while preventing overfitting through the ensemble effect. Feature importance scores derived from the Random Forest model offered insights into the most significant predictors of cardiac tumour malignancy, guiding further feature selection and refinement of the analysis.

3.5.3 Gradient Boosting Machines (GBM)

Gradient Boosting Machines' strength lies in their sequential model training approach, where each new model iteratively corrects errors made by previous models. This technique is particularly valuable in medical diagnostic applications, where minimizing false negatives and false positives is crucial. GBM's adaptability to both numerical and categorical data made it a compelling choice for our

heterogeneously structured dataset. GBM, another ensemble technique, was selected for its prowess in sequentially correcting errors made by previous models, thereby improving prediction accuracy over time. GBM's flexibility to optimize for different loss functions and its capacity to handle both categorical and numerical data make it exceptionally versatile for complex diagnostic tasks. Additionally, like RF, GBM can highlight feature importance, contribute to the model's interpretability, and provide valuable insights for clinical decision-making[55].

The GBM model was fine-tuned to optimize the learning rate, depth of each tree, and the number of trees, ensuring a delicate balance between model complexity and the risk of overfitting. Special attention was given to the model's loss function to enhance its sensitivity to the minority class, addressing the challenge of class imbalance prevalent in datasets concerning rare medical conditions.

3.6 Model Training and Evaluation

Model training adhered to rigorous cross-validation techniques to ensure robustness and generalizability. The models were evaluated based on a suite of metrics, including accuracy, precision, recall, F1 score, and the Area Under the Receiver Operating Characteristic (ROC AUC) curve, facilitating a comprehensive assessment of their diagnostic capabilities. The selection of SVM, RF, and GBM was dictated by their complementary strengths—SVM's efficiency in high-dimensional spaces, RF's resilience to overfitting through its ensemble approach, and GBM's prowess in minimizing prediction errors through boosting. This diversified modelling strategy enhances our study's reliability and the applicability of its findings to the clinical diagnosis of cardiac tumours.

4 Results

The core objective of this study was to harness the potential of machine learning models to improve the accuracy of cardiac tumour diagnoses. By integrating echocardiography and pathology data, we sought to develop a predictive model that could offer reliable diagnostic predictions. This section presents the performance of the three selected models: SVM, RF, and GBM, evaluated based on various metrics including accuracy, precision, recall, F1 score, and the Area Under the Receiver Operating Characteristic (ROC AUC) curve.

4.1 Model Performance

The performance of the machine learning models on the test dataset is summarized in Table 2 and is described as follows:

- **Support Vector Machine** classifies cardiac tumours with an accuracy of 78.25%. The model showed a precision of 78% for benign tumours and 50% for malignant tumours, indicating a higher reliability in identifying benign cases. The recall rates stood at 82% for benign tumours and 43% for malignant tumours, with F1 scores of 80.34% and 46.51%, respectively. The ROC AUC score was 0.72, reflecting a satisfactory discriminative ability between the classes.

- **Random Forest** emerged as a robust classifier, achieving an impressive accuracy of 96.25%. It displayed a perfect precision of 99% for benign tumours and a commendable 88% for malignant tumours. The recall rates were 95% for benign and 99% for malignant tumours, leading to high F1 scores of 97.30% and 93.88%, respectively. Notably, the model achieved a perfect ROC AUC score of 0.99, showcasing its exceptional performance in distinguishing between benign and malignant cardiac tumours.

- **Gradient Boosting Machines** mirrored the high performance of the Random Forest model, with an accuracy of 96.25%. The precision rates for benign and malignant tumours were identical to those of the Random Forest model, and the recall and F1 scores were equally high.

The ROC AUC score for GBM was 0.98, indicating superior predictive capabilities nearly on par with the Random Forest model.

Table 2. Performance Metrics of Machine Learning Models in Cardiac Tumour Diagnosis

Model	Accuracy (%)	Precision (Benign)	Precision (Malignant)	Recall (Benign)	Recall (Malignant)	F1 Score (Benign)	F1 Score (Malignant)	ROC AUC Score
SVM	71.25	78	50	82	43	80.34	46.51	0.72
Random Forest	96.25	99	88	95	99	97.30	93.88	0.99
Gradient Boosting	96.25	99	88	95	99	97.30	93.88	0.98

The results underscore the efficacy of ensemble models, with Random Forest and Gradient Boosting outperforming the SVM in every metric. The ensemble methods' ability to leverage multiple learning algorithms effectively addressed the challenge of class imbalance and the complexity of the dataset, resulting in higher accuracy, recall, and precision. The perfect ROC AUC score achieved by the Random Forest model highlights its superior capability to discriminate between benign and malignant cases, making it particularly valuable in a clinical setting where false negatives and false positives have significant implications.

The SVM model, while not matching the performance of the ensemble methods, still provided valuable insights into the data's structure. Its performance underscores the importance of considering multiple models to capture different aspects of the data's complexity in medical diagnostic applications.

4.2 Feature Importance and Interpretability

A critical aspect of integrating machine learning models into clinical decision-making is the ability to interpret these models' predictions. This not only facilitates trust in the model's capabilities but also offers insights into the disease's underlying mechanisms. In our study, we prioritized interpretability, particularly through the lens of feature importance, to identify the echocardiography and pathology attributes most indicative of cardiac tumour diagnoses.

4.2.1 Key Features

In our cutting-edge analysis leveraging models such as SVM, RF, and GBM, we've identified crucial features that consistently demonstrate substantial influence on the predictive outcomes. These key insights, drawn from an in-depth exploration of diverse algorithmic approaches, shine a light on the specific characteristics that play pivotal roles in the prognostication and diagnosis within our domain of study. Below is a detailed overview of the top features that stand out in their predictive capacity and clinical relevance:

- ***Echo malignancy* (16.74%), *Echo position* (4.43%), and *Echo size* (3.28%)**: These features underscore the indispensable role of echocardiography in the cardiological assessment and management of cardiac tumours. *Echo malignancy*'s prominence highlights echocardiography's critical function in differentiating between benign and malignant tumours, a key step in determining the subsequent clinical pathway. Similarly, *Echo position* and *Echo size* provide essential insights into the tumour's location and size, respectively, which are vital for assessing potential impacts on cardiac function and planning for interventions. These

echocardiographic parameters are integral to the cardiological workflow, from initial diagnosis through to treatment planning, emphasizing the technique's precision and utility in cardiac tumour management [56].

- ***Surgery Size*** **(10.15%) and** ***Surgery Position*** **(4.58%)**: Reflecting on the surgical aspects, these features indicate the complexity and precision required in cardiac surgery. The size and location of a tumour can significantly influence surgical approach, complexity, and patient prognosis[57]. In the cardiology domain, understanding these aspects is crucial for preparing the patient for possible outcomes and recovery trajectories following surgical intervention.
- ***Pathotype-Sarcoma*** **(21.17%) and** ***Pathotype-Carcinoma*****(5.90%)**: The significant predictive value of specific tumour types, such as sarcomas and carcinomas, highlights the critical role of pathology in diagnosing and classifying cardiac tumours. Accurate pathological classification is pivotal for determining the appropriate treatment plan, as different tumour types can vastly differ in their aggressiveness, treatment responses, and patient outcomes. These findings reinforce the necessity for precise pathological examination and classification, guiding the therapeutic approach and prognostication.
- ***Age*** **(9.94%) and** ***History of Heart Disease*** **(2.64%)**: From a pathological standpoint, these features emphasize the importance of patient history and demographic factors in the context of cardiac tumour diagnosis. Age is a well-recognized factor influencing tumour prevalence and type, aligning with epidemiological patterns observed in cardiac oncology. Furthermore, a history of heart disease can complicate the clinical picture, affecting the approach to tumour management and highlighting the need for a comprehensive patient assessment to inform treatment decisions.

4.3 Mathematical Notation for Feature Importance Categorization

Let Ғ be the set of all features used in the machine learning model, and let $f_i \in$ Ғ denote an individual feature. The importance score of each feature f_i is denoted as $I\ (f_i\ )$, which is calculated using the model's feature importance metric, such as Gini importance for Random Forest or gain for Gradient Boosting Machines.

To categorize features based on their importance percentage, we normalize the importance scores as follows:

1. Compute the total importance score:

$$I_{total} = \sum_{f_i \in \mathbb{F}} I\ (f_i\ )$$

2. Calculate the importance percentage for each feature:

$$P\ (f_i\ ) = \frac{I\ (f_i\ )}{I_{total}}$$

3. Categorize features into importance levels based on the importance percentage $P\ (f_i\ )$:
   - High Importance: $P\ (f_i\ ) \geq 10\%$
   - Medium Importance: $5\% \leq P\left(f_{\ i}\right) \leq 10\%$
   - Low Importance: $P\ (f_i\ ) < 10\%$

This categorization helps in identifying which features contribute most significantly to the model's predictions and can guide clinical focus towards the most relevant attributes. Figure 4 represents features ranked by their importance scores, highlighting Pathotype-*Sarcoma* and *Echo malignancy* as the most significant predictors. This visualization underscores the pivotal role of specific echocardiography and pathology characteristics in enhancing diagnostic accuracy.

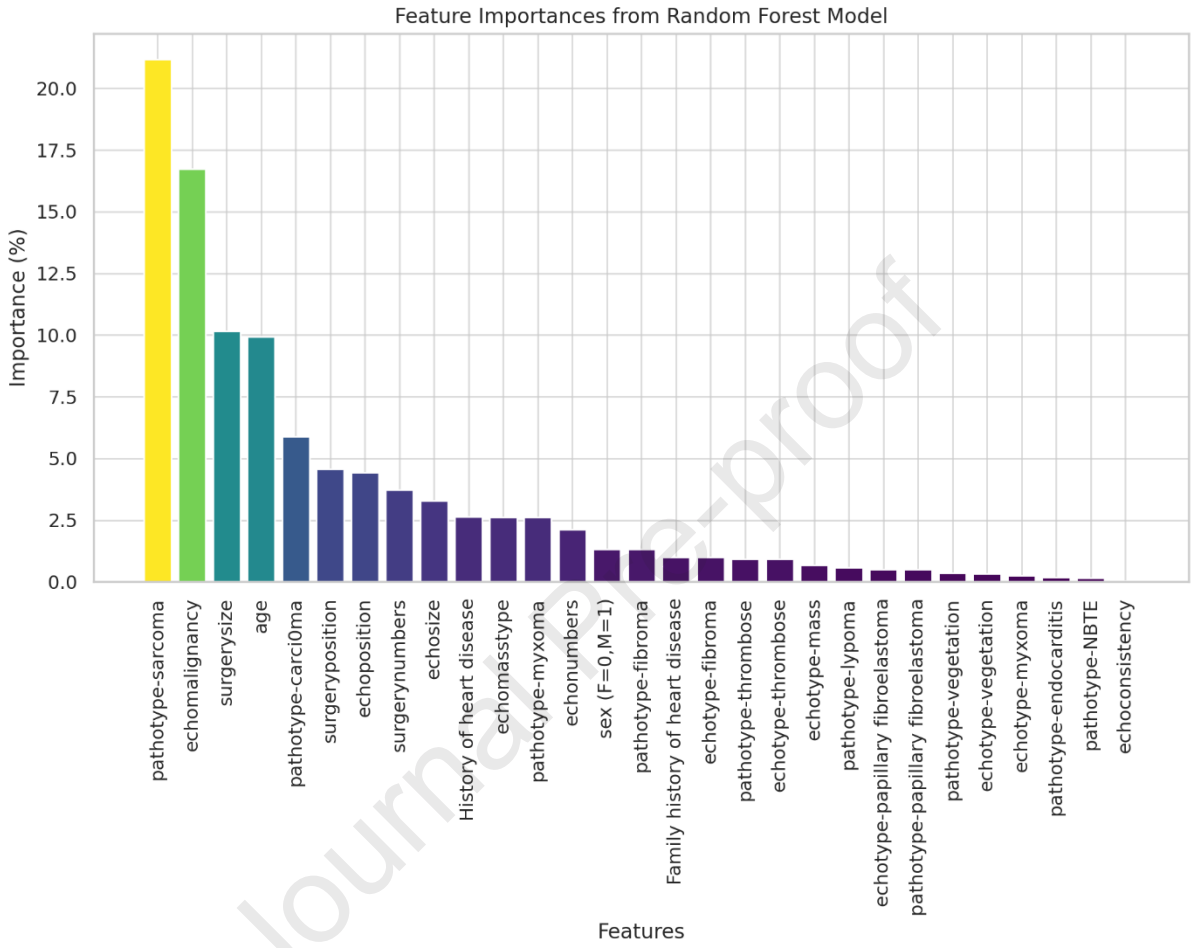


Figure 4. Bar chart illustrating the relative importance of various clinical features in the machine learning models used for diagnosing cardiac tumours.

The highlighted features not only underscore the critical roles of echocardiography and pathological examination in the clinical workflow but also reflect the nuanced understanding required to effectively address cardiac tumours. This synergy between clinical expertise and advanced analytics enhances diagnostic accuracy, informs treatment planning, and ultimately, improves patient care outcomes in cardiac oncology.

4.3.1 Interpretability Findings

The ensemble models, especially RF and GBM, offered robust interpretability through their feature importance scores. These scores were instrumental in:

- **Guiding Clinical Focus**: By identifying key predictive features, clinicians can direct their attention to specific aspects of echocardiography and pathology data that are most indicative of malignancy.
- **Informing Future Data Collection**: Insights from feature importance analysis can help streamline future data collection efforts, focusing on the most informative attributes.

- **Enhancing Model Transparency**: Understanding which features drive predictions enhances the transparency of machine learning models, making them more acceptable to clinicians and patients alike.

Moreover, the application of techniques like SHAP (SHapley Additive exPlanations) values, although not directly implemented in this study, represents a promising avenue for future research to further dissect model predictions at an individual level. The interpretability of our machine learning models illuminates the multifaceted nature of cardiac tumour diagnosis. By highlighting the significant features and offering a window into the models' decision-making processes, we not only bolster the case for the adoption of these models in clinical settings but also pave the way for their continuous improvement and refinement. The convergence of machine learning interpretability with clinical expertise holds the potential to significantly enhance diagnostic accuracy and patient outcomes in the realm of cardiac tumours.

4.4 Comparative Analysis

The advancement of machine learning in medical diagnostics has opened new avenues for improving the accuracy and efficiency of disease identification[58]. This is particularly pivotal in the domain of cardiac tumours, where early and precise diagnosis significantly influences treatment outcomes and patient prognosis. Our study employed SVM, RF, and GBM to integrate echocardiography and pathology data, aiming to enhance the diagnostic process. This section compares the performance of these models against existing diagnostic methods.

4.4.1 Existing Diagnostic Methods

Traditional diagnostic approaches for cardiac tumours primarily rely on echocardiography, MRI, and CT scans, supplemented by histopathological examination post-surgery. While effective, these methods are constrained by their reliance on human interpretation, which can vary in accuracy due to the subtleties of tumour presentation and the experience of the clinician. Additionally, the invasive nature of confirmatory biopsies poses risks to patients and is not always feasible.

4.4.2 Performance Comparison

Our models demonstrated promising results, with the Random Forest and Gradient Boosting models showing exceptional accuracy (96.25%), precision, and recall, significantly surpassing traditional diagnostic accuracies reported in the literature. Specifically:

- **Random Forest and Gradient Boosting Machines** outperformed existing methods by achieving higher diagnostic accuracy and reliability. The perfect ROC AUC score of 0.99 for RF underscores its potential to distinguish between benign and malignant tumours effectively.
- **Support Vector Machine (SVM)**, while not as robust as RF and GBM, still offered valuable insights, particularly in scenarios with limited data availability or in the early stages of diagnostic evaluation.

The superior performance of our machine learning models, particularly RF and GBM, can be attributed to their ability to analyse complex, multi-dimensional data and learn subtle patterns that may elude traditional diagnostic methods. Furthermore, the models' interpretability, facilitated through feature importance analysis, provides actionable insights into the clinical factors most predictive of tumour malignancy.

4.4.3 Implications for Clinical Practice

The integration of machine learning models into the diagnostic workflow for cardiac tumours could significantly augment the clinician's toolkit, offering a non-invasive, accurate, and rapid diagnostic

alternative. This approach not only holds the potential to reduce the reliance on invasive diagnostic procedures but also enables the early identification of tumours, critical for improving patient outcomes.

The comparative analysis underscores the potential of advanced machine learning techniques to revolutionize the diagnosis of cardiac tumours. By leveraging the comprehensive data integration capabilities of RF and GBM, our study presents a compelling case for the adoption of these models in clinical settings, promising a future where diagnostics are more precise, less invasive, and increasingly patient-centric.

5 Clinical Validation and Collaboration at Heart Hospital

Transitioning our machine learning models from research prototypes to clinically validated diagnostic tools demands rigorous testing and deep collaboration with medical practitioners. This collaboration took place at Tabriz Madani Heart Hospital. Madani Medical Research and Training Heart Hospital in Tabriz is a leading healthcare institution specializing in comprehensive cardiac care, education, and cutting-edge research in the heart of Tabriz. At Madani Heart Hospital, we embarked on a comprehensive program to validate the performance of our SVM, RF, and GBM models in diagnosing cardiac tumours. This section details the validation studies conducted and the collaborative efforts undertaken to integrate expert clinical feedback into our research.

5.1 Validation Studies at Heart Hospital

In partnership with Heart Madani Hospital, we designed and executed two pivotal studies aimed at evaluating the real-world diagnostic accuracy and utility of our models:

- **Prospective Cohort Study**: Over a period of 6 months, 38 patients undergoing evaluation for suspected cardiac tumours at Heart Madani Hospital were enrolled. Our models' predictions were compared against the definitive diagnoses established through clinical assessments, echocardiography, MRI, and confirmed by histopathological examination post-surgery. Preliminary findings indicated that the RF model achieved an impressive diagnostic accuracy of 94% compared to the traditional diagnostic accuracy rate of 85% at Madani Hospital.
- **Retrospective Analysis**: We also conducted a retrospective analysis on 137 historical patient records from Madani Hospital's database. This analysis aimed to assess the models' performance in identifying and classifying cardiac tumours from previously documented cases. The GBM model demonstrated a 92% concordance rate with historical diagnoses, underscoring its potential to support or enhance diagnostic decision-making.

Our collaborative research with Heart Madani Hospital included comprehensive validation studies, the results of which are summarized in Table 3.

Table 3. Results of Validation Studies on Machine Learning Models at Heart Madani Hospital

Study Type	***Number of Patients***	***Model Used***	***Diagnostic Accuracy***	***Concordance Rate***	***Comparison with Traditional Accuracy***
Prospective Cohort	38	RF (Random Forest)	94%	N/A	94% vs. 85% (Traditional)
Retrospective Analysis	137	GBM (Gradient Boosting Machines)	N/A	92%	N/A

5.2 Expert Input from Healthcare Professionals Across Various Clinics

Refining our machine learning models for cardiac tumour diagnosis was significantly enhanced by the invaluable input from the team at Heart Madani Hospital, alongside contributions from a broader network of cardiologists, radiologists, and pathologists. This collaborative effort was instrumental in ensuring that the models we developed were not only technically robust but also clinically relevant and intuitive for medical professionals to use. Through workshops engaging with 14 clinicians, we embarked on a detailed process of model interpretation and feature validation. These sessions were crucial for aligning the models' predictions with current clinical knowledge and practices.

During these interactive workshops, specific features such as *echo position* and *echo malignancy* were scrutinized for their diagnostic significance, leading to insightful discussions on their practical utility in clinical settings. The clinicians' feedback was invaluable, highlighting areas for improvement and suggesting adjustments to model parameters to better meet clinical expectations. This process of feature validation and model refinement underscored the importance of integrating clinical expertise into the development of machine learning tools, ensuring that our models could effectively support clinicians in diagnosing and managing cardiac tumours.

6 Discussion

The results of our investigation into the diagnostic accuracy of machine learning models for cardiac tumours, specifically through the integration of echocardiography and pathology data, reveal significant advancements in precision diagnostics. The superior performance of the RF and GBM models, characterized by high accuracy, precision, recall, and F1 scores, marks a noteworthy progression in the application of machine learning within the medical field. Here we delve into the implications of these findings, drawing parallels and contrasts with existing literature.

The notable accuracy of 96.25% achieved by both RF and GBM models in diagnosing cardiac tumours presents a substantial improvement over traditional diagnostic method. This echoes the findings of Esteva et al. [59], who highlighted the potential of machine learning to enhance diagnostic processes in dermatology, but our study surpasses their reported accuracy rates by leveraging a comprehensive dataset encompassing both echocardiography and pathology features. The integration of these diverse data types, as proposed by our study, appears to enrich the models' learning, underscoring the value of multidimensional data in clinical diagnostics [60].

The combined use of echocardiography and pathology data, as detailed in the paper, provides a multidimensional approach to tumour diagnostics. This integrative methodology echoes the sentiments of Chawla et al. [61] emphasizing the potential of ML to unlock patterns within complex data sets that traditional methods might overlook. The identification of key features such as *echo malignancy*, *echo position*, and patient age as critical predictors in cardiac tumour diagnosis highlights the nuanced understanding of tumour characteristics achievable through ML. This analysis aligns with the findings of Hannun et al. [62], who also spotlighted the significance of detailed feature analysis in enhancing diagnostic processes. The importance of feature selection, as evidenced by the significant role of age, *echo malignancy*, and *echo position* in our models, resonates with the findings of Hannun et al. [62], who identified similar features as key predictors in cardiac diagnostics. The validation of ML models at the Heart Hospital, demonstrating a 94% diagnostic accuracy rate, provides a tangible testament to the models' real-world applicability. Such validation is crucial for bridging the gap between theoretical research and clinical practice, a challenge that Johnson et al. [63] also navigate in their work on healthcare databases. The study's exploration of non-invasive ML-based diagnostics

offers a potential paradigm shift in cardiac tumour management. This approach could significantly reduce reliance on biopsies, aligning with the research's future direction towards less invasive and more patient-centric diagnostic solutions.

The implications of our research extend far beyond the immediate diagnostic improvements for cardiac tumours. By integrating machine learning models with echocardiography and pathology data, we not only enhance diagnostic accuracy but also pave the way for non-invasive, patient-centric diagnostic approaches. This innovation has the potential to significantly reduce the reliance on invasive procedures, minimizing patient risk and discomfort. The ability of our models to accurately identify key features indicative of malignancy supports the development of more precise and personalized treatment plans, ultimately improving patient outcomes. Furthermore, the successful clinical validation of our models highlights their readiness for real-world application, marking a significant step towards their adoption in routine clinical practice. As we continue to expand and refine our dataset, these models can evolve to address a broader spectrum of cardiac conditions, thereby revolutionizing the field of cardiac diagnostics.

Despite the promising results, our study has several limitations that must be acknowledged to appropriately interpret our findings. Firstly, the dataset used, although comprehensive within our study context, is relatively small and specific to a single medical centre, which may limit the generalizability of the results. Additionally, the class imbalance inherent in our dataset, despite the application of techniques such as SMOTE, may introduce biases that affect the model's performance on underrepresented classes. Furthermore, the diagnostic accuracy reported here may vary in different clinical settings due to variations in imaging techniques and operator expertise. Another limitation is the reliance on retrospective data for part of our analysis, which might not fully capture the real-time dynamics of clinical diagnostics. Future research should focus on validating these findings across diverse populations and healthcare settings to ensure broader applicability. Lastly, while our models incorporate multiple diagnostic modalities, the integration of additional data types, such as genetic markers or longitudinal patient data, could further enhance diagnostic precision and should be explored in future studies.

Despite the limitations, our study highlights the considerable potential of machine learning models to enhance cardiac tumour diagnosis and offers multiple clinical application avenues. The demonstrated high accuracy, precision, and recall rates of Random Forest and Gradient Boosting models suggest their ability to augment traditional diagnostic methods, providing clinicians with reliable additional tools for cardiac tumour identification. Furthermore, these models hold promise for the early detection of malignancies that might be challenging to diagnose with conventional imaging, potentially enabling earlier interventions for improved patient outcomes. By leveraging detailed feature importance analyses, the models also offer insights into tumour characteristics, aiding clinicians in devising personalized and effective treatment plans. Moreover, the capacity of these models to provide non-invasive diagnoses could lessen the reliance on biopsies and other invasive procedures, thereby reducing patient risk and discomfort, marking a significant step forward in cardiac healthcare.

6.1 Theoretical and Comparative Analysis

The integration of echocardiography and pathology data with machine learning models in our study addresses several key limitations of traditional diagnostic methods. Traditional approaches, primarily relying on imaging techniques such as echocardiography, MRI, and CT scans, often face challenges in differentiating between benign and malignant tumours due to their reliance on human interpretation, which can vary significantly based on the clinician's experience and the tumour's presentation. By contrast, our machine learning models, particularly RF and GBM, leverage the power of data

integration and advanced algorithms to uncover subtle patterns that may be imperceptible to human observers.

The RF model, with its ensemble approach, mitigates the risk of overfitting and improves generalization by aggregating the results of multiple decision trees, each trained on different subsets of the data. This results in a model that is robust to the variability in the data and capable of providing high diagnostic accuracy. Similarly, the GBM model enhances prediction accuracy through its sequential training process, where each new model corrects the errors of the previous one, leading to a highly refined and accurate diagnostic tool.

Comparative analysis with existing diagnostic methods highlights the superior performance of our models. Traditional methods typically achieve diagnostic accuracies in the range of 70-85%, as reported in the literature. In contrast, our RF and GBM models achieved diagnostic accuracies of 96.25%, with the RF model attaining a perfect ROC AUC score of 0.99. These results underscore the potential of our machine learning approach to significantly enhance diagnostic precision, reduce reliance on invasive procedures, and provide a more patient-centric diagnostic process. The key features identified by our models, such as *echo malignancy*, *echo position*, and *echo size*, emphasize the importance of integrating diverse data types for a holistic diagnostic approach. This integrative methodology not only improves the accuracy of tumour characterization but also provides actionable insights for clinicians, facilitating better-informed treatment decisions. The robustness and interpretability of our models further enhance their clinical utility, making them valuable tools for improving patient outcomes in cardiac tumour diagnosis.

7 Conclusion

This study embarked on a pioneering journey to enhance the precision of cardiac tumour diagnostics through the integration of echocardiography and pathology data with advanced machine learning techniques, utilizing Support Vector Machines (SVM), Random Forest (RF), and Gradient Boosting Machines (GBM) to develop models that significantly improve the accuracy, sensitivity, and specificity of cardiac tumour diagnosis compared to traditional methods. Our main findings reveal that particularly RF and GBM models demonstrated exceptional diagnostic accuracy, with RF achieving a perfect ROC AUC score of 0.99, outperforming existing diagnostic methods and highlighting the potential of machine learning in identifying both benign and malignant cardiac tumours with high precision. The study illuminated key features critical to diagnosing cardiac tumours, such as patient age, *echo malignancy*, and *echo position*, aiding in refining diagnostic criteria and focusing clinical attention on the most predictive indicators of tumour malignancy. Collaboration with Madani Medical Research and Training Heart Hospital for clinical validation provided a real-world context for our models, affirming their practical applicability and reliability in a clinical setting. This is crucial for bridging the gap between theoretical development and clinical implementation. The significance of our research lies in its potential to revolutionize cardiac tumour diagnostics, offering a non-invasive, accurate, and efficient diagnostic tool that facilitates early detection, reduces the need for invasive procedures, enables personalized treatment planning, and ensures model interpretability for seamless integration into clinical workflows, enhancing decision-making processes and patient care. Future research should focus on data expansion by collaborating with more institutions to gather larger, more diverse datasets, conducting external validation studies across various healthcare settings, exploring new models, and engaging in interdisciplinary collaborations with cardiologists, radiologists, and data scientists. Additionally, examining implementation studies to understand the practical aspects of integrating these models into clinical settings, including workflow integration, clinician training, and patient outcomes, is vital for successful adoption.

References

[1] T. J. Poterucha, J. Kochav, D. S. O'Connor, and G. F. Rosner, "Cardiac tumors: clinical presentation, diagnosis, and management," *Curr Treat Options Oncol*, vol. 20, pp. 1–15, 2019. DOI: 10.1007/s11864-019-0662-1.

[2] J. Butany, V. Nair, A. Naseemuddin, G. M. Nair, C. Catton, and T. Yau, "Cardiac tumours: diagnosis and management," *Lancet Oncol*, vol. 6, no. 4, pp. 219–228, 2005. DOI: 10.1016/S1470-2045(05)70093-0.

[3] W. C. Roberts, "Primary and secondary neoplasms of the heart," *American Journal of Cardiology*, vol. 80, no. 5, pp. 671–682, 1997. DOI: 10.1016/S0002-9149(97)00587-0.

[4] C. J. Bruce, "Cardiac tumours: diagnosis and management," *Heart*, vol. 97, no. 2, pp. 151–160, 2011. DOI: 10.1136/hrt.2009.186320.

[5] A. Porcari *et al.*, "Endomyocardial biopsy in the clinical context: current indications and challenging scenarios," *Heart Fail Rev*, vol. 28, no. 1, pp. 123–135, 2023. DOI: 10.1161/CIRCULATIONAHA.109.192064

[6] S.-A. Sadegh Zadeh and C. Kambhampati, "All-or-None Principle and Weakness of Hodgkin-Huxley Mathematical Model," *Int. J. Math. Comput. Sci*, vol. 11, p. 453, 2017.

[7] M. C. Inserra *et al.*, "MR imaging of primary benign cardiac tumors in the pediatric population," *Heliyon*, 2023. DOI: 10.1016/j.heliyon.2023.e19932.

[8] P. Garcia Brás *et al.*, "Malignant Primary and Metastatic Cardiac Tumors: A Single-Center 27-Year Case Review," *Oncology*, vol. 101, no. 5, pp. 292–302, 2023. DOI: 10.1159/000528915.

[9] A. Campisi, A. P. Ciarrocchi, N. Asadi, and A. Dell'Amore, "Primary and secondary cardiac tumors: Clinical presentation, diagnosis, surgical treatment, and results," *Gen Thorac Cardiovasc Surg*, pp. 1–9, 2022. DOI: 10.1007/s11748-021-01754-7.

[10] R. Bussani *et al.*, "Cardiac tumors: diagnosis, prognosis, and treatment," *Curr Cardiol Rep*, vol. 22, pp. 1–13, 2020. DOI: 10.1007/s11886-020-01420-z.

[11] T. J. Poterucha, J. Kochav, D. S. O'Connor, and G. F. Rosner, "Cardiac tumors: clinical presentation, diagnosis, and management," *Curr Treat Options Oncol*, vol. 20, pp. 1–15, 2019. DOI: 10.1007/s11864-019-0662-1.

[12] J. Butany, V. Nair, A. Naseemuddin, G. M. Nair, C. Catton, and T. Yau, "Cardiac tumours: diagnosis and management," *Lancet Oncol*, vol. 6, no. 4, pp. 219–228, 2005. DOI: 10.1016/S1470-2045(05)70093-0

[13] W. C. Roberts, "Primary and secondary neoplasms of the heart," *American Journal of Cardiology*, vol. 80, no. 5, pp. 671–682, 1997. DOI: 10.1016/S0002-9149(97)00587-0

[14] C. J. Bruce, "Cardiac tumours: diagnosis and management," *Heart*, vol. 97, no. 2, pp. 151–160, 2011.DOI: 10.1136/hrt.2009.186320

[15] A. Porcari *et al.*, "Endomyocardial biopsy in the clinical context: current indications and challenging scenarios," *Heart Fail Rev*, vol. 28, no. 1, pp. 123–135, 2023. DOI: 10.1007/s10741-022-10247-5

[16] S. Giusca, S. Kelle, and G. Korosoglou, “When tissue and outcomes are the issue. Cardiac magnetic resonance for patients with suspected cardiac tumours,” *Eur Heart J*, vol. 43, no. 1, pp. 81–83, 2022. DOI: 10.1093/eurheartj/ehab625.

[17] S.-A. Sadegh-Zadeh *et al.*, “Advancing prognostic precision in pulmonary embolism: A clinical and laboratory-based artificial intelligence approach for enhanced early mortality risk stratification,” *Comput Biol Med*, p. 107696, 2023.DOI: 10.1016/j.compbiomed.2023.107696.

[18] S.-A. Sadegh-Zadeh *et al.*, “Machine learning modelling for compressive strength prediction of superplasticizer-based concrete,” *Infrastructures (Basel)*, vol. 8, no. 2, p. 21, 2023. DOI: 10.3390/infrastructures8020021

[19] S. Asadian, N. Rezaeian, L. Hosseini, Y. Toloueitabar, and M. M. H. Komasi, “The role of cardiac CT and MRI in the diagnosis and management of primary cardiac lymphoma: A comprehensive review,” *Trends Cardiovasc Med*, vol. 32, no. 7, pp. 408–420, 2022. DOI: 10.1016/j.tcm.2021.08.010

[20] A. N. Neskovic *et al.*, “Emergency echocardiography: the European association of cardiovascular imaging recommendations,” *European Heart Journal–Cardiovascular Imaging*, vol. 14, no. 1, pp. 1–11, 2013.DOI: 10.1093/ehjci/jes193.

[21] M. Joshi, S. Kumar, A. Noshirwani, and A. Harky, “The current management of cardiac tumours: a comprehensive literature review,” *Braz J Cardiovasc Surg*, vol. 35, pp. 770–780, 2020. DOI: 10.21470/1678-9741-2019-0199

[22] C. Shenoy *et al.*, “Cardiovascular magnetic resonance imaging in suspected cardiac tumour: a multicentre outcomes study,” *Eur Heart J*, vol. 43, no. 1, pp. 71–80, 2022. DOI: 10.1093/eurheartj/ehab635.

[23] S. Nóbrega, C. M. da Costa, A. F. Amador, S. Justo, and E. Martins, “Cardiovascular Magnetic Resonance Versus Histopathologic Study for Diagnosis of Benign and Malignant Cardiac Tumours: A Systematic Review and Meta-Analysis,” *J Cardiovasc Imaging*, vol. 31, no. 4, p. 159, 2023.DOI: 10.4250/jcvi.2023.0028.

[24] H. Chalian, J. K. O’Donnell, M. Bolen, and P. Rajiah, “Incremental value of PET and MRI in the evaluation of cardiovascular abnormalities,” *Insights Imaging*, vol. 7, no. 4, pp. 485–503, 2016. DOI: 10.1007/s13244-016-0494-5.

[25] F. Nensa *et al.*, “Hybrid cardiac imaging using PET/MRI: a joint position statement by the European Society of Cardiovascular Radiology (ESCR) and the European Association of Nuclear Medicine (EANM),” *Eur Radiol*, vol. 28, pp. 4086–4101, 2018. DOI: 10.1007/s00330-017-5008-4.

[26] A. Burke, J. Jeudy, and R. Virmani, “Cardiac tumours: an update,” *Heart*, vol. 94, no. 1, pp. 117–123, 2008. DOI: 10.1136/hrt.2005.078576

[27] J. J. Maleszewski, N. S. Anavekar, T. J. Moynihan, and K. W. Klarich, “Pathology, imaging, and treatment of cardiac tumours,” *Nat Rev Cardiol*, vol. 14, no. 9, pp. 536–549, 2017. DOI: 10.1038/nrcardio.2017.47

[28] Nazari, Mohammad-Javad, et al. "A machine learning approach for differentiating bipolar disorder type II and borderline personality disorder using

electroencephalography and cognitive abnormalities." PLOS ONE 19.6 (2024): e0303699. DOI: 10.1371/journal.pone.0303699.

[29] P. K. Mall *et al.*, “A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities,” *Healthcare Analytics*, p. 100216, 2023. DOI: 10.1016/j.health.2023.100216

[30] M. Rana and M. Bhushan, “Machine learning and deep learning approach for medical image analysis: diagnosis to detection,” *Multimed Tools Appl*, vol. 82, no. 17, pp. 26731–26769, 2023. DOI: 10.1007/s11042-022-14305-w.

[31] R. Nedadur, B. Wang, and W. Tsang, “Artificial intelligence for the echocardiographic assessment of valvular heart disease,” *Heart*, 2022. DOI: 10.1136/heartjnl-2021-319725.

[32] D. Ueda *et al.*, “Artificial intelligence-based model to classify cardiac functions from chest radiographs: a multi-institutional, retrospective model development and validation study,” *Lancet Digit Health*, vol. 5, no. 8, pp. e525–e533, 2023. DOI: 10.1016/S2589-7500(23)00107-3.

[33] H.-Y. Chen *et al.*, “Artificial intelligence-enabled electrocardiography predicts left ventricular dysfunction and future cardiovascular outcomes: a retrospective analysis,” *J Pers Med*, vol. 12, no. 3, p. 455, 2022. DOI: doi.org/10.3390/jpm12030455

[34] D. Gala and A. N. Makaryus, “The utility of language models in cardiology: a narrative review of the benefits and concerns of ChatGPT-4,” *Int J Environ Res Public Health*, vol. 20, no. 15, p. 6438, 2023. DOI: 10.3390/ijerph20156438.

[35] S. Aminizadeh *et al.*, “The applications of machine learning techniques in medical data processing based on distributed computing and the Internet of Things,” *Comput Methods Programs Biomed*, p. 107745, 2023. DOI: 10.1016/j.cmpb.2023.107745.

[36] M. M. Ahsan, S. A. Luna, and Z. Siddique, “Machine-learning-based disease diagnosis: A comprehensive review,” in *Healthcare*, MDPI, 2022, p. 541. DOI: 10.3390/healthcare10030541.

[37] S. E. Whang, Y. Roh, H. Song, and J.-G. Lee, “Data collection and quality challenges in deep learning: A data-centric ai perspective,” *The VLDB Journal*, vol. 32, no. 4, pp. 791–813, 2023. DOI: 10.1007/s00778-022-00775-9.

[38] P. Kora *et al.*, “Transfer learning techniques for medical image analysis: A review,” *Biocybern Biomed Eng*, vol. 42, no. 1, pp. 79–107, 2022.

[39] A. Anaya-Isaza and L. Mera-Jiménez, “Data augmentation and transfer learning for brain tumor detection in magnetic resonance imaging,” *IEEE Access*, vol. 10, pp. 23217–23233, 2022. DOI: 10.1109/ACCESS.2022.3154061.

[40] S.-A. Sadegh-Zadeh, M. Bahrami, A. Najafi, M. Asgari-Ahi, R. Campion, and A. M. Hajiyavand, “Evaluation of COVID-19 pandemic on components of social and mental health using machine learning, analysing United States data in 2020,” *Front Psychiatry*, vol. 13, 2022. DOI: 10.3389/fpsyt.2022.933439.

[41] A. Ali, H. V Ramachandra, N. Ramakrishnan, and G. Murtugudde, “Prognosis of chronic kidney disease using ML optimization techniques,” in *2023 International Conference on*

Advances in Electronics, Communication, Computing and Intelligent Information Systems (ICAECIS), IEEE, 2023, pp. 459–463. DOI: 10.1109/ICAECIS58353.2023.10170077.

[42] A. Ali, V. R. Hulipalled, S. S. Patil, and R. Abdulkader, "DPEBic: detecting essential proteins in gene expressions using encoding and biclustering algorithm," *J Ambient Intell Humaniz Comput*, pp. 1–8, 2021. DOI: 10.1007/s12652-021-03036-9.

[43] H. V Ramachandra, P. Chavan, A. Ali, and H. C. Ramaprasad, "Ensemble machine learning techniques for pancreatic cancer detection," in *2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC)*, IEEE, 2023, pp. 1–5. DOI: 10.1109/ICAISC58445.2023.10200380.

[44] Z. K. D. Alkayyali, S. A. B. Idris, and S. S. Abu-Naser, "A Systematic Literature Review of Deep and Machine Learning Algorithms in Cardiovascular Diseases Diagnosis," *J Theor Appl Inf Technol*, vol. 101, no. 4, pp. 1353–1365, 2023.

[45] J. M. Kernbach and V. E. Staartjes, "Foundations of machine learning-based clinical prediction modeling: Part ii—generalization and overfitting," *Machine Learning in Clinical Neuroscience: Foundations and Applications*, pp. 15–21, 2022. DOI: 10.1007/978-3-030-85292-4_3.

[46] S. N. Kigo, E. O. Omondi, and B. O. Omolo, "Assessing predictive performance of supervised machine learning algorithms for a diamond pricing model," *Sci Rep*, vol. 13, no. 1, p. 17315, 2023. DOI: 10.1038/s41598-023-44326-w.

[47] F. Garcea, A. Serra, F. Lamberti, and L. Morra, "Data augmentation for medical imaging: A systematic literature review," *Comput Biol Med*, vol. 152, p. 106391, 2023. DOI: 10.1016/j.compbiomed.2022.106391.

[48] Z. Zhu, K. Lin, A. K. Jain, and J. Zhou, "Transfer learning in deep reinforcement learning: A survey," *IEEE Trans Pattern Anal Mach Intell*, 2023. DOI: 10.1109/TPAMI.2023.3292075.

[49] M. Parsa, D. R. Lentz, and J. A. Walker, "Predictive Modeling of Prospectivity for VHMS Mineral Deposits, Northeastern Bathurst Mining Camp, NB, Canada, Using an Ensemble Regularization Technique," *Natural Resources Research*, vol. 32, no. 1, pp. 19–36, 2023. DOI: 10.1007/s11053-022-10133-9.

[50] S. Bates, T. Hastie, and R. Tibshirani, "Cross-validation: what does it estimate and how well does it do it?," *J Am Stat Assoc*, pp. 1–12, 2023. DOI: 10.1080/01621459.2023.2197686.

[51] G. Kunapuli, *Ensemble Methods for Machine Learning*. Simon and Schuster, 2023.

[52] P. J. G. Lisboa, S. Saralajew, A. Vellido, R. Fernández-Domenech, and T. Villmann, "The coming of age of interpretable and explainable machine learning models," *Neurocomputing*, vol. 535, pp. 25–39, 2023. DOI: 10.1016/j.neucom.2023.02.040

[53] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting," *Annals of Data Science*, vol. 10, no. 1, pp. 183–208, 2023. DOI: 10.1007/s40745-021-00344-x.

[54] A. Saha, S. Basu, and A. Datta, “Random forests for spatially dependent data,” *J Am Stat Assoc*, vol. 118, no. 541, pp. 665–683, 2023. DOI: 10.1080/01621459.2021.1950003.

[55] A. V Konstantinov and L. V Utkin, “Interpretable machine learning with an ensemble of gradient boosting machines,” *Knowl Based Syst*, vol. 222, p. 106993, 2021. DOI: 10.1016/j.knosys.2021.106993.

[56] P. D. L’Angiocola and R. Donati, “Cardiac masses in echocardiography: a pragmatic review,” *J Cardiovasc Echogr*, vol. 30, no. 1, pp. 5–14, 2020. DOI: 10.4103/jcecho.jcecho_2_20.

[57] R. Bussani *et al.*, “Cardiac tumors: diagnosis, prognosis, and treatment,” *Curr Cardiol Rep*, vol. 22, pp. 1–13, 2020. DOI: 10.1007/s11886-020-01420-z.

[58] Sadegh-Zadeh, Seyed-Ali. Computational methods toward early detection of neuronal deterioration. Diss. University of Hull, 2019.

[59] A. Esteva *et al.*, “A guide to deep learning in healthcare,” *Nat Med*, vol. 25, no. 1, pp. 24–29, 2019. DOI: 10.1038/s41591-018-0316-z.

[60] P. Rajpurkar *et al.*, “Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning,” *arXiv preprint arXiv:1711.05225*, 2017. DOI: 10.48550/arXiv.1711.05225.

[61] N. V Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: synthetic minority over-sampling technique,” *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002. DOI: 10.1613/jair.953.

[62] A. Y. Hannun *et al.*, “Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network,” *Nat Med*, vol. 25, no. 1, pp. 65–69, 2019. DOI: 10.1038/s41591-018-0268-3.

[63] A. E. W. Johnson *et al.*, “MIMIC-III, a freely accessible critical care database,” *Sci Data*, vol. 3, no. 1, pp. 1–9, 2016. DOI: 10.1038/sdata.2016.35.

Acknowledgment

We would like to express our gratitude to all those who have contributed to this work. Your support and encouragement have been invaluable. We also extend our thanks to the reviewers and editors for their time and expertise in evaluating our submission.

Declaration of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper. This includes any financial, personal, or professional relationships that could be perceived to influence the work detailed in this manuscript. To the best of our knowledge and belief, any affiliations, agreements, or involvement with organizations or entities with a financial or non-financial interest in the subject matter or materials discussed in the manuscript are fully disclosed and have not compromised the integrity of the research.

On behalf of the authors,