Circulation: Cardiovascular Imaging

ORIGINAL ARTICLE

Artificial Intelligence-Enhanced Analysis of Echocardiography-Based Radiomic Features for Myocardial Hypertrophy Detection and Etiology Differentiation

Inki Moon, MD*; Jina Lee, BS*; Seung-Ah Lee, MD, PhD; Dawun Jeong, BS; Jaeik Jeon, MSc; Yeonggul Jang, PhD; Sihyeon Jeong, BS; Jiyeon Kim, BS; Hong-Mi Choi, MD; In-Chang Hwang, MD; Youngtaek Hong, PhD; Goo-Yeong Cho, MD, PhD; Yeonyee E. Yoon, MD, PhD; Hyuk-Jae Chang, MD, PhD

BACKGROUND: While echocardiography is pivotal for detecting left ventricular hypertrophy (LVH), it struggles with etiology differentiation. To enhance LVH assessment, we aimed to develop an artificial intelligence algorithm using echocardiography-based radiomics. This algorithm is designed to detect LVH and differentiate its common etiologies, such as hypertrophic cardiomyopathy (HCM), cardiac amyloidosis (CA), and hypertensive heart disease (HHD), based on echocardiographic images.

METHODS: The developmental data sets from multiple medical centers included 867 subjects, with an independent external test set from a single tertiary medical center containing 619 subjects. Radiomic feature analysis was conducted on 4 echocardiographic views, extracting both conventional and harmonization-driven myocardial textures along with myocardial geographic features. Then, we developed classification models for each condition. Variable contributions were evaluated using Shapley Additive Explanations analysis.

RESULTS: The radiomics-based LightGBM model, selected from internal validation, maintained strong performance in the external test set (area under the curve of 0.96 for HCM, 0.89 for CA, and 0.86 for HHD). Compared with the logistic regression model using conventional echocardiographic parameters (left ventricular ejection fraction, left ventricular mass index, left atrial volume index, and E/e′), the final model demonstrated superior sensitivity (0.89 versus 0.80 for HCM, 0.80 versus 0.80 for CA, and 0.75 versus 0.33 for HHD) and F1-score (0.87 versus 0.57 for HCM, 0.84 versus 0.72 for CA, and 0.82 versus 0.50 for HHD). Feature analysis highlighted that harmonization-driven textures played a key role in differentiating HCM, while conventional textures and myocardial thickness were influential in differentiating CA and HHD.

CONCLUSIONS: This study confirms that artificial intelligence-enhanced echocardiography-based radiomics effectively differentiate the etiology of LVH, highlighting the potential of artificial intelligence-driven texture and geographic analysis in LVH evaluation.

GRAPHICAL ABSTRACT: A graphical abstract is available for this article.

Key Words: area under curve ■ artificial intelligence ■ heart diseases ■ radiomics ■ stroke volume

See Editorial by Author

Correspondence to: Yeonyee E. Yoon, MD, PhD, Department of Cardiology, Cardiovascular Center Seoul National University Bundang Hospital82, Gumi-ro 173 Beon-gil, Seongnam, Gyeonggi-do 13620, Republic of Korea. Email yeonyeeyoon@snubh.org
*I. Moon and J. Lee contributed equally.
Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/CIRCIMAGING.124.017436.
For Sources of Funding and Disclosures, see page XXX.

Circulation: Cardiovascular Imaging is available at www.ahajournals.org/journal/circimaging

CLINICAL PERSPECTIVE

An artificial intelligence model using echocardiography-based radiomics was developed to enhance the evaluation of left ventricular hypertrophy. It integrates advanced myocardial texture and geographic feature analysis for high accuracy in differentiating left ventricular hypertrophy etiology, including hypertrophic cardiomyopathy, cardiac amyloidosis, and hypertensive heart disease. Variables from radiomic and geographic features were evaluated using Shapley Additive Explanations, providing key diagnostic insights. Despite promising results in both internal validation and external test, further studies are necessary to confirm its generalizability across diverse populations and clinical settings.

Nonstandard Abbreviations and Acronyms

AI	artificial intelligence
CA	cardiac amyloidosis
HCM	hypertrophic cardiomyopathy
HHD	hypertensive heart disease
LV	left ventricular
LVH	left ventricular hypertrophy
SHAP	Shapley Additive Explanations

Left ventricular hypertrophy (LVH) is frequently observed in clinical settings and often indicates various cardiovascular diseases, significantly impacting morbidity and mortality.[1–3] Although echocardiography is a primary noninvasive tool for diagnosing LVH, its prevalence may lead to underestimation and oversight in clinical assessment. Furthermore, while echocardiography provides left ventricular (LV) mass, wall thickness, and cardiac function, it alone may not precisely pinpoint the etiological factor of LVH, which is crucial for specific treatment plans and outcome prediction.[4,5] The subjective nature and variability in echocardiographic interpretation necessitate advanced techniques, such as magnetic resonance imaging or endomyocardial biopsy, for deeper etiological insights.[6]

In response to these challenges, there has been a significant shift toward utilizing artificial intelligence (AI) to assist in the echocardiographic detection and categorization of LVH etiology.[7] However, existing studies predominantly rely on deep learning (DL)-based end-to-end models that, while seemingly accurate in differentiating LVH etiologies, fail to disclose which specific echocardiographic features are used for this differentiation.[7–9] This lack of interpretability hinders clinical adoption and limits their utility. In addition, many of these studies lacked independent data set evaluation or showed significantly reduced performance when assessed on separate test data sets.

The current study introduces an approach using echocardiography-based radiomic features to develop an algorithm that not only detects LVH but also differentiates its various etiologies, such as hypertrophic cardiomyopathy (HCM), cardiac amyloidosis (CA), and hypertensive heart disease (HHD). We extracted texture and geographic features using an AI-based automated segmentation algorithm. The texture features included both conventional and harmonization-driven features, the latter obtained through a novel harmonization technique to ensure vendor-independent analysis of myocardial textures. The geographic features comprised myocardial shape and thickness. Using these radiomic features, we constructed machine learning-based classification models. Unlike DL-based end-to-end models, our approach identified specific features contributing to the differentiation of LVH etiologies through Shapley Additive Explanations (SHAP) analysis, offering interpretability and actionable insights for clinical application. In addition, we utilized a multicenter data set for training and internal validation and further tested the model's performance on an independent external test set. This design aimed to achieve superior generalizability compared with prior studies.

METHODS

Our AI-based system automates LVH evaluation through a comprehensive pipeline (USfeat.ai; Ontact Health, Korea). This pipeline selects the necessary view, segments the LV myocardium, detects cardiac phases, and extracts relevant features, which are then used in the classification model. The data supporting the findings of this study are available from the corresponding author upon reasonable request. In addition, a subset of the data can be accessed at AI-Hub.[10]

Study Data set

This study aggregated echocardiography data from multiple Korean sources, including the Open AI Dataset Project,[10] Korea National Standard Cardiovascular Database,[11] Seoul National University Bundang Hospital LVH registry,[8] and the ACREDIT registry (REGISTRATION: URL: https://www.clinicaltrials.gov; Unique identifier: NCT05108168). Normal subjects had no history of cardiovascular disease, hypertension, or diabetes, with normal echocardiographic findings. HHD subjects had hypertension, LVH confirmed by echocardiography, and documented regression of LVH after blood pressure management. HCM was defined by cardiac magnetic resonance imaging and genetic test confirmation, while CA was diagnosed via endomyocardial biopsy and cardiac magnetic resonance, pyrophosphate scans. Cases with potential alternative causes for LVH were excluded from the HHD, HCM, and CA groups. Descriptions of each cohort and the inclusion criteria for each condition are detailed in Method S1.

To streamline our study's data set management, we organized the data sets according to hospital origin and intentionally excluded Seoul National University Bundang Hospital when compiling developmental data set from various hospitals, assessing 2285 subjects for eligibility, with 1171 meeting the

clinical inclusion criteria (Figure 1A). For the external test, we utilized an independent data set from Seoul National University Bundang Hospital, reviewing 1410 patients, with 713 qualifying for clinical inclusion criteria (Figure 1B).

This study complied with the ethical guidelines of the Declaration of Helsinki and was approved by each institution's institutional review board. Given its retrospective nature and the use of de-identified images, informed consent was waived. The research also followed the Proposed Requirements for Machine Learning Evaluation in Cardiovascular Imaging.[12]

Data Preparation

In this study, our algorithm was designed to extract and analyze myocardial features from images of the end-diastolic and end-systolic phases within the 4 fundamental views of echocardiography: parasternal long-axis, parasternal short-axis at mid-level, apical 4-chamber, and apical 2-chamber. To achieve this, we used our previously developed AI-based echocardiographic automatic analysis system (Sonix Health; Ontact Health, Korea) to select the necessary views[13] and automatically segment the LV myocardium.[14,15] The subsequent step involved automatic detection of the end-diastolic and end-systolic phases, followed by myocardial feature extraction from the segmented myocardium. Detailed descriptions of the automatic segmentation algorithm and representative cases are provided in Method S2, along with Video S1.

After the initial processing, which excluded cases missing any of the 4 fundamental views, poor image quality leading to failed segmentation, or those with inaccurate end-diastolic/end-systolic phase detection were excluded (Figure 1). The developmental data set was reduced from 1171 to 867 cases: 276 with HCM, 168 with CA, 135 with HHD, and 288 normal subjects. We ensured homogeneity by maintaining disease ratios when splitting the data 8:2 for training and internal validation (704:163). All steps, including feature selection, model building, and hyperparameter tuning, were performed exclusively on the training set. The internal validation set was used to evaluate the performance of different classification models and to select the final model. The independent external test set was refined from 716 to 619 patients: 46 with HCM, 66 with CA, 93 with HHD, and 414 normal subjects. This data set was used exclusively for final model performance assessment, with no modifications made to the model following training. The adjudication process, including view selection, phase detection, and segmentation, was rigorously validated by experienced cardiologists (I.M. and S.-A.L), who reached a consensus on any discrepancies.

Myocardial Feature Extraction

We delved into quantitative features by extracting texture and geographic features from each segmented LV myocardium, employing the well-established open-source Python package Pyradiomics.[16] For a clearer understanding and effective management of the diverse features, we organized them into the following categories: we extracted conventional texture features (Feat1) and additionally applied a novel harmonization technique developed by our team to derive harmonization-driven texture features (Feat2). This technique was designed to reduce vendor-specific speckle pattern variability in LV myocardium.[17] We also included geographic features, myocardial shape (Feat3), and thickness (Feat4; Figure 2). Features were extracted during the end-diastolic and end-systolic phases, and percent changes between these phases (%Δ) were calculated. In total, 19 839 features were utilized for LVH detection and etiology differentiation. For more technical details, refer to Method S3.

Feature Selection and Classification Modeling

The study utilized feature selection techniques, including the Boruta and XGBoost algorithms.[18] The Boruta algorithm was used to select the top-ranked features, and XGBoost was used to select the essential features. The 2 algorithms sequentially identified significant features for classification targets: (1) 70

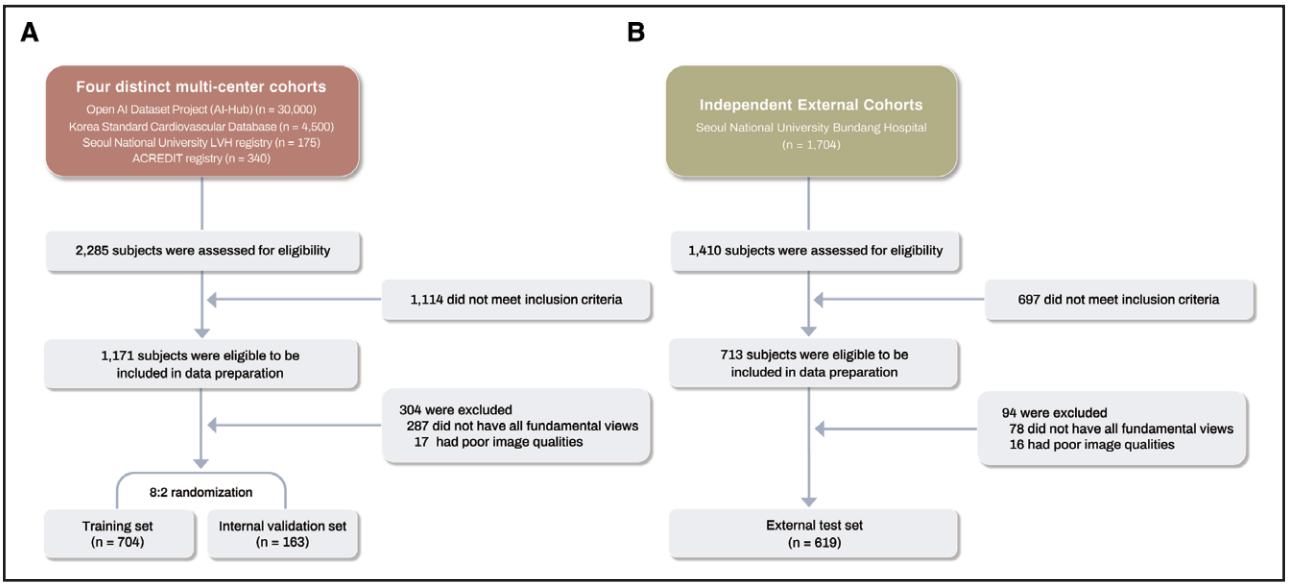

Figure 1. Flowchart of data preparation process.
The flowchart illustrates the preparation process for 2 data sets: (**A**) the multicenter developmental data set and (**B**) the independent external data set sourced from Seoul National University Bundang Hospital.

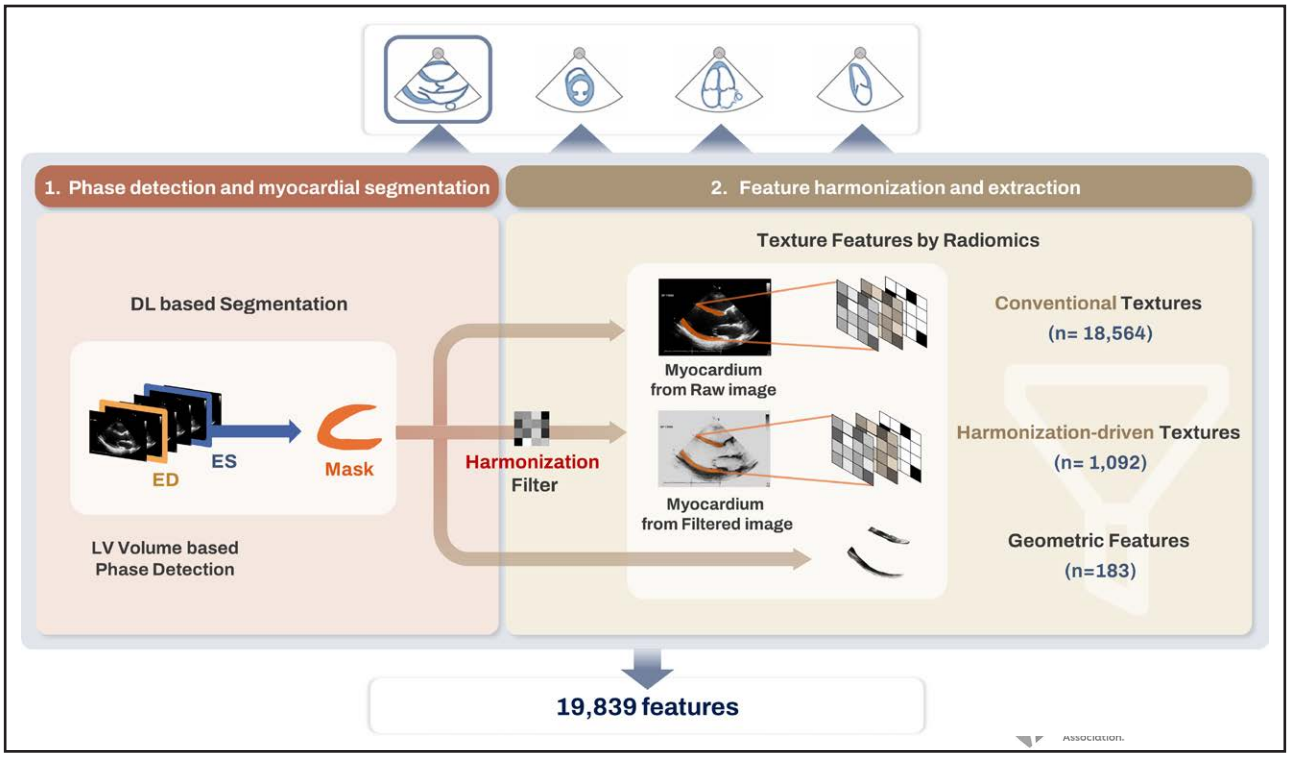

Figure 2. Automated process of echocardiographic feature extraction for developing the classification model.
The automated software segmented left ventricular (LV) myocardium from 4 echocardiographic views. It performed segmentation across the full cardiac cycle, automatically detecting end-diastole (ED) and end-systole (ES) phases and extracting radiomic features. Both conventional and newly developed harmonization-driven texture features were obtained using advanced filters. Geographic features encapsulating myocardial shape and thickness were also extracted. Upon combining data across all views and phases, a total of 19 839 features were extracted and utilized to develop the classification model. DL indicates deep learning.

features for detecting LVH from normal, (2) 82 for differentiating HCM from CA and HHD, (3) 83 for distinguishing CA from HCM and HHD, and (4) 6 for differentiating HHD from HCM and CA. A detailed description of the feature selection process, as well as the specific features used for each classification, including their feature importance, can be found in Method S4.1. After feature selection, we evaluated several models to determine the most effective for diagnosing LVH and its etiologies. The LightGBM algorithm was selected as the classification model due to its superior performance,[19] as outlined in Method S4.2. To optimize its efficiency, hyperparameters were fine-tuned using grid search, as outlined in Method S4.3. In addition, to tackle the issue of class imbalance, the synthetic minority oversampling technique was used.[20] If there is more than 1 cardiac cycle, the probability for each cardiac cycle is calculated and then converted into a patient-level probability (Method S4.4).

Upon model detection of LVH from given subjects, our study imprints dedicated models to further classify and differentiate among 3 etiological conditions: HCM, CA, or HHD (Figure 3A). If an LVH subject exhibits one of these 3 conditions exclusively, that condition is definitively classified as the etiology of LVH. For LVH subjects exhibiting 2 or more conditions, the model prioritizes the diagnosis associated with the highest probability, thereby assigning the most likely etiological classification. When LVH is detected without a positive indication for any of the specified 3 etiologies, the subject is categorized into an Others group. This designation suggests the presence of an alternative or unspecified LVH etiology not covered by the primary classifications of HCM, CA, or HHD (Method S4.5).

Statistical Analysis

Categorical variables are presented as absolute counts and percentages, and continuous variables are expressed as means±SD or medians [interquartile range], as appropriate. Categorical differences were analyzed using the χ^2 test or Fisher exact test, and continuous variables using the Student *t* test. Multiple comparisons were assessed via 1-way ANOVA or the Kruskal-Wallis method, with Bonferroni correction applied for type I errors. Diagnostic performance was evaluated with the area under the curve, sensitivity, specificity, and F1-score. The reproducibility of radiomic features was confirmed through manual segmentation of 30 randomly selected internal validation cases and intraclass correlation coefficient calculation. SHAP was used to interpret machine learning predictions (Method S5).[21] A 2-tailed $P<0.05$ was considered statistically significant. All analyses were performed using SAS, version 9.4 (SAS Institute, Inc, Cary, NC) and R 3.3.0 (R Development Core Team, 2016).

RESULTS

Study Population

The developmental data set included 867 patients (mean age, 52.0±17.9 years; 60.3% male) and the external test

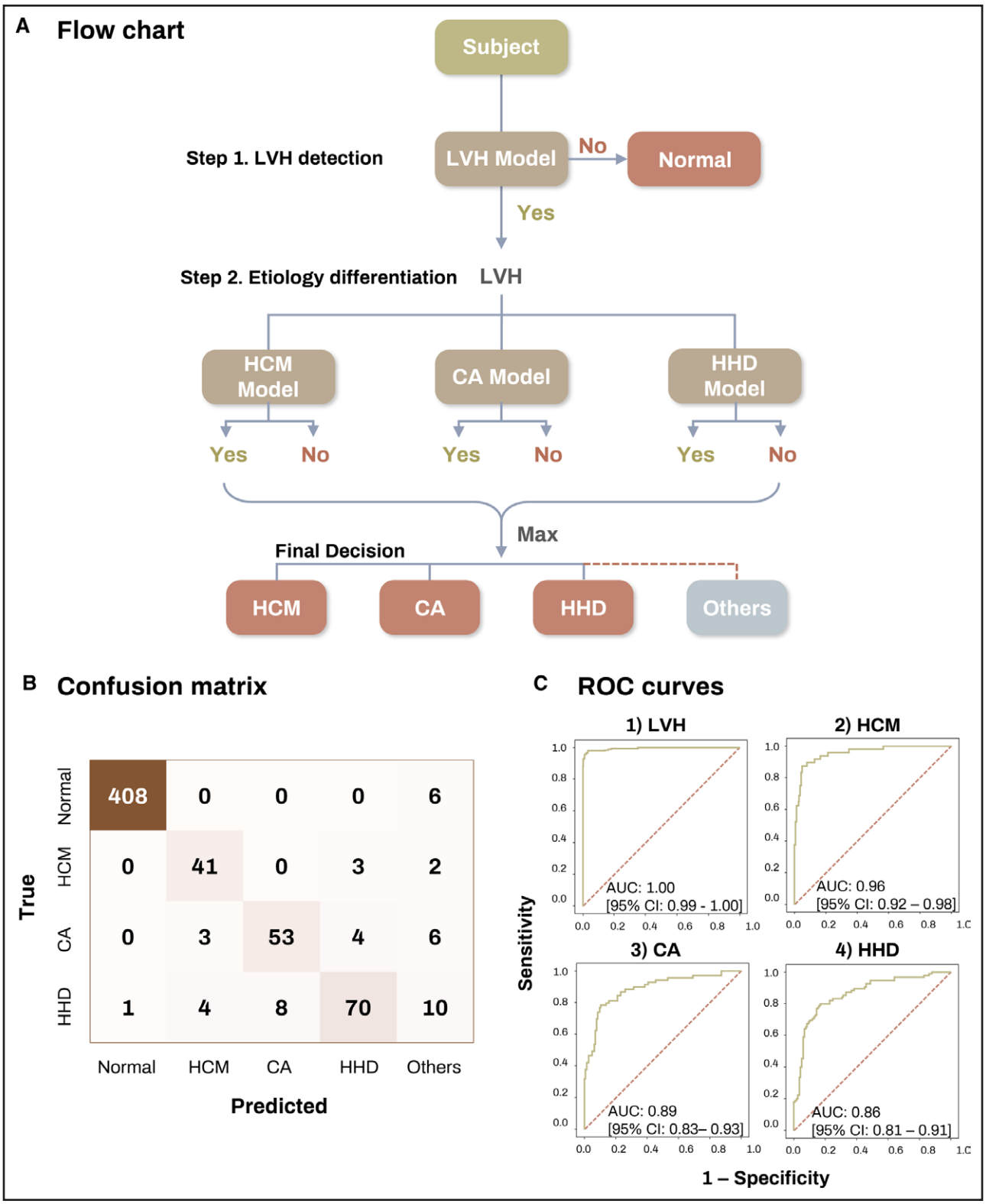

Figure 3. Hierarchical classification process and model performance in left ventricular hypertrophy (LVH) assessment. Panel (**A**) demonstrates how the classification model first differentiates LVH from normal subjects and then categorizes it into 3 distinct etiologies. The diagnostic accuracy of the models is presented through (**B**) confusion matrices and (**C**) receiver operating curve analysis in the external test cohort. CA indicates cardiac amyloidosis; HCM, hypertrophic cardiomyopathy; and HHD, hypertensive heart disease.

set comprised 619 patients (mean age, 51.3±14.6 years; 48.6% male; Table 1). LVH prevalence differed significantly between the developmental (66.8%) and external test set (33.1%, $P<0.001$). HCM was most common in the developmental data set (47.7%), while HHD was more frequent in the external test set (45.4%). Among

Table 1. Baseline Characteristics of the Study Cohorts

	Developmental set					
	Total (n=867)	Training (n=704)	Internal validation (n=163)	External test (n=619)	*P* value*	*P* value†
Age, y	52.0±17.9	52.1±18.0	51.2±17.3	51.3±14.6	0.543	0.438
Men, n (%)	523 (60.3)	424 (60.2)	99 (60.7)	301 (48.6)	0.975	<0.001
BMI, kg/m^2	23.8±4.7	23.8±4.9	23.5±3.8	23.9±3.4	0.479	0.394
LVH, n (%)	579 (66.8)	476 (67.6)	103 (63.2)	205 (33.1)	0.323	<0.001
HCM, n (%)	276 (47.7)	223 (46.8)	53 (51.5)	46 (22.4)	0.572	<0.001
CA, n (%)	168 (29.0)	139 (29.2)	29 (28.2)	66 (32.2)		
HHD, n (%)	135 (23.3)	114 (23.9)	21 (20.4)	93 (45.4)		

Values are presented as mean±SD or number (percentage). BMI indicates body mass index; CA, cardiac amyloidosis; HCM, hypertrophic cardiomyopathy; HHD, hypertensive heart disease; and LVH, left ventricular hypertrophy.

*P values were obtained using Student *t* test or χ^2 analysis, comparing across the training and internal validation sets.

†P values were obtained using Student *t* test or χ^2 analysis, comparing developmental and external test sets.

HCM, the septal type was prevalent in the developmental set (52.5%), whereas mixed or diffuse types predominated in the external test set (71.7%). For CA, AL amyloidosis was the predominant subtype across both data sets (84.5% and 89.4%, respectively). Comprehensive demographics and echocardiographic characteristics are in Result S1.

Classification Model Performance

The LightGBM model with radiomic features demonstrated the best performance in the internal validation set and was selected as the final model (detailed diagnostic performance available in Result S2). This model maintained strong performance in the external test set, achieving an overall accuracy of 92.4% in the external test set (Figure 3B). Consistent area under the curve values were observed for LVH (1.00 [95% CI, 0.99–1.00]), HCM (0.96; 0.92–0.98), CA (0.89; 0.83–0.93), and HHD (0.86; 0.81–0.91; Figure 3C). When compared with conventional echocardiographic parameters (LV ejection fraction, LV mass index, left atrial volume index, and E/e′; Table 2), the final model demonstrated superior sensitivity (0.89 versus 0.80 for HCM, 0.80 versus 0.80 for CA, and 0.75 versus 0.33 for HHD) and F1-score (0.87 versus 0.57 for HCM, 0.84 versus 0.72 for CA, and 0.82 versus 0.50 for HHD). Additional experimental results assessing the utility of Feat2 are presented in Result S3.

Reproducibility of Radiomic Features Extraction

Interobserver variability of radiomic feature extraction showed good reproducibility (intraclass correlation coefficient ≥0.6) in 79% (Feat1), 98% (Feat2), 85% (Feat3), and 79% (Feat4) of features (Figure 4A). When features with poor or moderate reproducibility were excluded, the classification model's performance in the external test set remained stable, with no significant decline in diagnostic accuracy (Figure 4B). A detailed comparison is available in Result S4.

Characteristics According to the Prediction

Patients classified as LVH exhibited smaller LV dimensions, thicker walls, lower e′ velocities, and higher E/e′ ratios (Result S5). Seven normal subjects misclassified

Table 2. Performance Comparison of Conventional and Radiomics-Based Machine Learning Model in the External Test Data sets

	AUC	Sensitivity	Specificity	F1-score	PPV	NPV
Logistic regression model with LVMI, LVEF, LAVI, E/e′						
LVH	**1.00**	**0.99**	**1.00**	**0.99**	**1.00**	**0.98**
HCM	0.78	0.80	0.92	0.57	0.45	0.98
CA	0.85	**0.80**	0.95	0.72	0.65	**0.98**
HHD	**0.83**	0.33	**1.00**	0.50	**0.97**	0.89
LightGBM model with radiomics features						
LVH	**1.00**	**0.98**	**1.00**	**0.99**	**1.00**	0.97
HCM	**0.96**	**0.89**	**0.99**	**0.87**	**0.85**	**0.99**
CA	**0.89**	**0.80**	**0.97**	**0.84**	**0.87**	**0.98**
HHD	**0.83**	**0.75**	0.99	**0.82**	0.91	**0.96**

AUC indicates area under the curve; CA, cardiac amyloidosis; HCM, hypertrophic cardiomyopathy; HHD, hypertensive heart disease; LAVI, left atrial volume index; LVEF, left ventricular ejection fraction; LVH, left ventricular hypertrophy; LVMI, left ventricular mass index; NPV, negative predictive value; and PPV, positive predictive value.

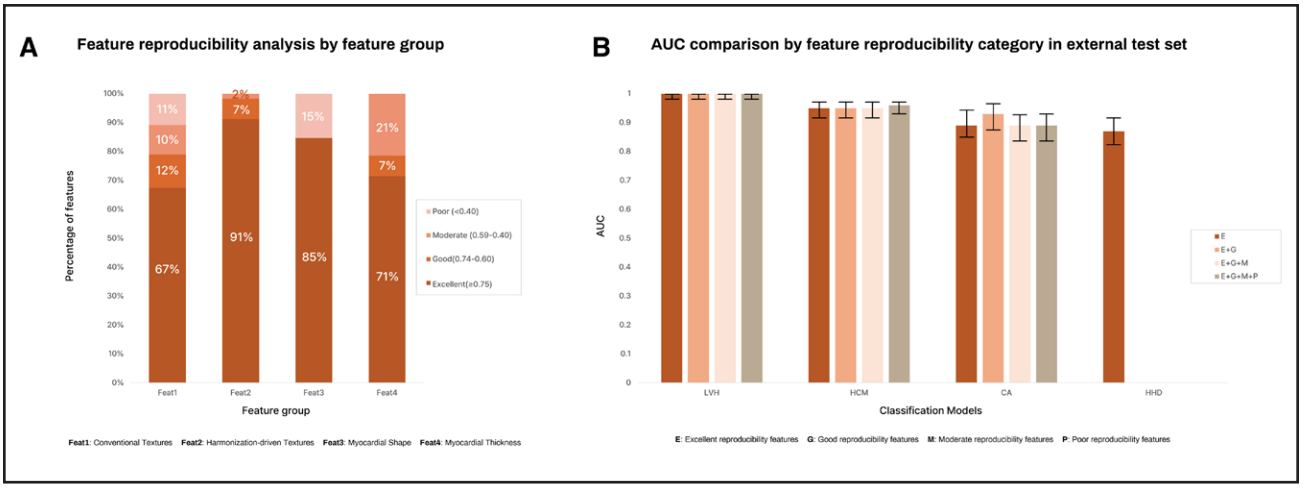

Figure 4. Reproducibility of radiomic feature extraction and its impact on model performance.
A, Interobserver variability analysis of radiomic feature extraction. **B**, Effect of excluding features with poor or moderate reproducibility on classification model performance in the external test set. The hypertensive heart disease (HHD) classification model exclusively utilized features with excellent reproducibility, eliminating the need for further validation. AUC indicates area under the curve.

as LVH (1 in the internal and 6 in the external test set) had higher LV mass indices and lower e′ velocities (Result S6); these were subsequently categorized as Others (Figure 3).

Among LVH patients, no significant differences in LV mass index were observed across HCM, CA, or HHD classifications (Result S5). Across both data sets, 28 LVH patients were misclassified into Others (6 HCM, 9 CA, and 13 HHD). Misclassified HCM exhibited higher diastolic blood pressure and lower LV ejection fraction compared with those correctly classified (Result S7); CA misclassified patients had more diabetes and higher LV ejection fraction (Result S8); HHD misclassified patients showed smaller LV cavities and higher LV ejection fraction (Result S9).

Model Interpretation by SHAP Analysis

SHAP analysis identified key features for each model (Figure 5). For HCM, significant factors were harmonization-driven textures (Feat2), myocardial thickness (Feat4), and conventional textures (Feat1). The most crucial texture feature for HCM was Run entropy from a gray-level run-length matrix after harmonization, highlighting texture heterogeneity. In CA, conventional textures (Feat1), myocardial thickness (Feat4), and percent changes in myocardial thickness (%ΔFeat3) and harmonization-driven textures (%ΔFeat2) were crucial. Run-Length Non-Uniformity with a gray-level run-length matrix captured myocardial homogeneity. In HHD detection, myocardial thickness (Feat4) and shape (Feat3) were important. The inverse difference normalized with the Gray Level Co-occurrence Matrix was an important texture feature indicating local myocardial homogeneity.

By mapping the sum of SHAP values for texture features (Feat1 and Feat2) onto the myocardium, we found that properly classified cases exhibited considerably high SHAP values (Figure 5), in contrast to the notably low SHAP values in misclassified Others cases (Result S10). This variance in myocardial texture features could have influenced the model's decision, indicating that these cases were divergent from their actual condition.

DISCUSSION

This study explored the effectiveness of utilizing echocardiography-based radiomic features analyzed through machine learning models to enhance LVH etiology differentiation. Our approach, which integrates advanced myocardial texture analysis and geographic attributes, was thoroughly validated internally and tested externally, showcasing robust performance. Furthermore, enhanced by SHAP analyses, our approach also offered clinical interpretability. These findings underscore the potential clinical relevance and adaptability of our methods for practical clinical application in real-world settings.

LVH, a common cardiovascular condition, has various etiologies. HHD results from increased afterload,[3,22] HCM has genetic roots and may cause severe complications,[23,24] and CA involves amyloid deposition, demanding specific therapies.[25,26] Despite different etiologies, morphological similarities and the lack of pathognomonic findings often challenge accurate diagnosis by visual assessments and conventional echocardiographic evaluation. Indeed, the accuracy of human experts in differentiating the etiologies of LVH ranges from 50% to 80%.[8,9] This necessitates further diagnostics such as cardiac magnetic resonance, nuclear scintigraphy, genetic testing, and myocardial biopsy.[6]

To improve the diagnostic process and minimize unnecessary tests, researchers have used DL in echocardiography to detect subtle differences in LVH etiologies

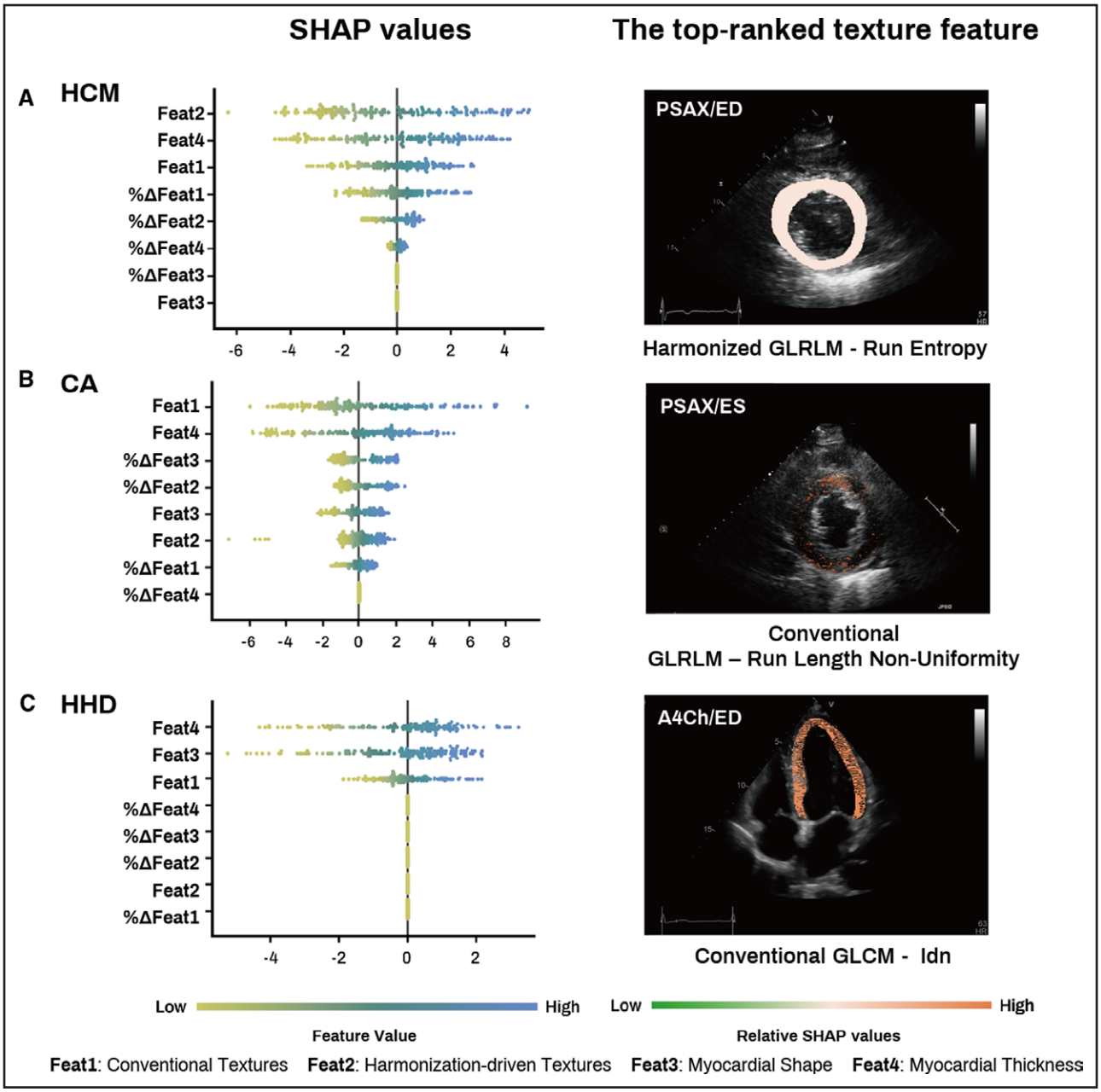

Figure 5. Insights from Shapley Additive Explanations (SHAP) values and feature contribution in classification models. We used SHAP values to the influence of features on the predictions made by classification models. Higher absolute SHAP values indicate a greater impact of belonging to the positive or negative class. We performed the analysis through the differential process (**A** through **C**). The **left** depict feature contributions across the various integrated classification models. The **right** display representative mappings of the top-ranked features from both conventional (Feat1) and harmonization-driven (Feat2) texture analyses for each classified group. GLCM indicates Gray Level Co-occurrence Matrix; and GLRLM, run entropy derived from a gray-level run-length matrix.

that are not apparent through conventional methods.[7–9] For instance, Duffy et al[7] demonstrated the reliability of using a DL algorithm trained on echocardiographic data to quantify LV wall thickness and determine the cause of LVH utilizing only the apical 4-chamber view. Although their data set was extensive, our study demonstrated superior discriminative performance for CA and HCM (area under the curve 0.89 and 0.96, respectively, compared with their 0.79 and 0.89). Moreover, unlike their end-to-end DL model, ours identifies specific features contributing to the diagnosis. Notably, Hwang et al[8] reported that DL-based differentiation of LVH etiology was superior to human expertise, achieving an overall accuracy of 92.3% compared with 80%, underscoring the potential benefits of DL in improving diagnostic accuracy. However, previous efforts in applying DL encountered remarkable limitations, including the lack of multicenter cohorts and evaluation on independent data sets.[8,9] In addition, the critical need for interpretability, essential for securing clinical trust and facilitating wider adoption, remains largely unaddressed.

In this study, we aimed to enhance diagnostic accuracy and provide significant insight by employing echocardiography-based radiomic features. Instead of directly inputting echocardiographic images into a DL model, we applied AI technology for automated segmentation of the LV myocardium, extracting and analyzing a broad array of features. This approach allowed for an in-depth examination of myocardial texture and geographic features. Specifically, myocardial shape (Feat3) and thickness (Feat4) were identified as important factors in diagnosing and differentiating LVH. Myocardial thickness (Feat4) proved crucial across all the processes of etiologic diagnosis, while the percent change in myocardial shape (%ΔFeat3) played a significant role in differentiating CA (Figure 5).

In the evaluation of LVH, assessing myocardial texture is an aspect that echocardiography has long attempted to characterize because of its potential value.[27,28] Particularly, texture analysis has shown promise in detecting LV remodeling and differentiating transthyretic CA from other cardiomyopathies.[29,30] Despite the historical challenges in quantifying myocardial texture changes, which often reduce diagnostic reproducibility, we address this issue by employing an AI-driven approach to automatically segment the LV myocardium, extract, and analyze texture features. Importantly, our method incorporated both conventional (Feat1) and novel harmonization-driven (Feat2) texture features, the latter designed to reduce variability across different imaging settings and vendors.[17] In our classification model, both texture features were pivotal: the harmonization-driven texture features (Feat2) were crucial for detecting LVH and differentiating HCM, while conventional texture features (Feat1) significantly helped in differentiating CA. Although geographic features played an essential role in identifying HHD, conventional texture features (Feat1) also extensively aided in its differentiation (Figure 5).

This approach effectively quantifies and evaluates the impact of specific myocardial features on the detection and differentiation of LVH. By identifying which echocardiography-based radiomic features are pivotal in differentiating various cardiac conditions, our classification model gains credibility and provides deeper insights into the diagnostic process. Furthermore, the model's scalability allows it to classify unknown conditions beyond the current diagnostic scope by placing atypical cases into the Others category. This method enhances diagnostic safety by preventing disease misclassification. Intriguingly, as SHAP values have highlighted, divergent texture features in misclassified cases suggest the possibility of uncovering unrecognized pathophysiologies.

LIMITATIONS

This study has some limitations. First, although the training data set was multicenter-driven and the model was externally validated, it was predominantly based on data sourced from Korean patient cohorts. This could introduce ethnic biases and suggests a need for further validation across diverse demographic populations. Extending the model to include a broader range of conditions and testing it in various settings with different equipment standards will be essential to enhance its applicability and reliability. Second, our LVH cohort was not comprehensive of all LVH cases but was confined to HHD, HCM, and CA. Notably, we did not include physiological hypertrophy, such as athlete's heart, due to challenges in gathering sufficient training data. This highlights the need for further research into physiological LVH and echocardiography-based radiomic features. In addition, despite our efforts, we were unable to include the rare causes of LVH, such as Fabry disease, Danon syndrome, and sarcoidosis due to their low prevalence. While we incorporated an Others category for atypical features of known disease groups, this approach has inherent limitations. Future work will refine the model to differentiate physiological hypertrophy and incorporate a broader spectrum of disease categories, enhancing diagnostic accuracy and applicability in diverse clinical settings. Third, while we have demonstrated the robustness of our model by applying it to an independent external test set with different compositions, we acknowledge the significant challenges of lacking prospective testing and evaluation of model efficacy during actual clinical practice.[31] Plans are underway to conduct prospective studies to address this gap.

CONCLUSIONS

In this study, we developed an AI-based classification model that utilizes echocardiography-based radiomics to enhance the diagnosis and differentiation of LVH. The model, which was developed using data from multi-institutional cohorts and validated externally, demonstrates promising results in applying radiomic analysis to real-world clinical settings. However, its application has been primarily within a specific demographic, highlighting the need for broader testing across diverse populations to ensure its efficacy and generalizability in varied clinical settings.

ARTICLE INFORMATION

Received August 7, 2024; accepted February 16, 2025.

Affiliations

Division of Cardiology, Department of Internal Medicine, Soonchunhyang University Bucheon Hospital, Republic of Korea (I.M.). Department of Internal Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea (I.M.). Department of Internal Medicine, Graduate School of Medical Science, Brain Korea 21 Project, Yonsei University College of Medicine, Seoul, Republic of Korea (J.L., D.J., S.J., J.K.). CONNECT-AI Research Center, Yonsei University College of Medicine, Seoul, Republic of Korea (J.L., S.-A.L., D.J., S.J., J.K., H.-J.C.). Ontact Health, Seoul, Republic of Korea (S.-A.L., J.J., Y.J., Y.H., H.-J.C.). Cardiovascular Center and Division of Cardiology, Department of Internal Medicine, Seoul National University

Bundang Hospital, Seongnam, Gyeonggi, Republic of Korea (H.-M.C., I.-C.H., G.-Y.C., Y.E.Y.). Division of Cardiology, Severance Cardiovascular Hospital, Yonsei University College of Medicine, Yonsei University Health System, Seoul, Republic of Korea (H.-J.C.).

Acknowledgments

Drs Lee, Hong, Yoon, and Chang contributed to the conceptualization of the study. Data curation was conducted by Dr Moon, J. Lee, D. Jeong, and Dr Hong, while formal analysis was performed by Dr Moon, J. Lee, D. Jeong, and Dr Hong. Dr Chang was responsible for funding acquisition. The investigation was carried out by Dr Moon, J. Lee, D. Jeong, and Dr Hong. J. Jeon, Drs Jang, and Hong developed the methodology. Project administration was managed by Dr Moon, S. Jeong, D. Jeong, and J. Kim. Resources were provided by Drs Moon, Choi, Hwang, Yoon, Cho, and Chang. Software development was handled by J. Jeon, Drs Jang, and Hong. Supervision was conducted by Drs Lee, Cho, and Chang. Validation was performed by Drs Moon, Lee, Hong, Choi, Hwang, Yoon, and Cho. Visualization efforts were undertaken by J. Lee, S. Jeong, D. Jeong, and J. Kim. The original draft was written by Dr Moon, J. Lee, Drs Lee, Jang, Hong, Hwang, and Yoon, while the review and editing process was carried out by Dr Moon, J. Lee, and Dr Yoon.

Sources of Funding

This study was supported by Seoul R&BD Program (BT230080) through the Seoul Business Agency funded by the Seoul Metropolitan Government and also supported by Pfizer Gran 65650777. Medical AI Clinic Program through the NIPA, funded by the MSIT (grant number: H0904-24-1002).

Disclosures

Dr Chang holds stock in Ontact Health, Inc. Dr Yoon is a Medical Consultant of Ontact Health, Inc. Dr Hong is an inventor on a patent related to this work filed by Ontact Health (Method: Method and Apparatus for Determining Whether Organs or Muscles of Body Is Abnormal). J. Jeon, Drs Jang, Hong, and Lee are currently affiliated with Ontact Health, Inc. The other authors report no conflicts.

Supplemental Material

Methods S1–S5
Results S1–S10
Video S1
References 32–40

REFERENCES

1. Schirmer H, Lunde P, Rasmussen K. Prevalence of left ventricular hypertrophy in a general population; the Tromso Study. *Eur Heart J*. 1999;20:429–438. doi: 10.1053/euhj.1998.1314
2. Cuspidi C, Facchetti R, Bombelli M, Sala C, Tadic M, Grassi G, Mancia G. Risk of mortality in relation to an updated classification of left ventricular geometric abnormalities in a general population: the Pamela study. *J Hypertens*. 2015;33:2133–2140. doi: 10.1097/HJH.0000000000000658
3. Ruilope LM, Schmieder RE. Left ventricular hypertrophy and clinical outcomes in hypertensive patients. *Am J Hypertens*. 2008;21:500–508. doi: 10.1038/ajh.2008.16
4. Weidemann F, Niemann M, Ertl G, Stork S. The different faces of echocardiographic left ventricular hypertrophy: clues to the etiology. *J Am Soc Echocardiogr*. 2010;23:793–801. doi: 10.1016/j.echo.2010.05.020
5. Angeli F, Reboldi G, Verdecchia P. Echocardiographic left ventricular hypertrophy: implications for clinicians. *J Hypertens*. 2012;30:2279–2284. doi: 10.1097/HJH.0b013e32835ac71b
6. Yilmaz A, Sechtem U. Diagnostic approach and differential diagnosis in patients with hypertrophied left ventricles. *Heart*. 2014;100:662–671. doi: 10.1136/heartjnl-2011-301528
7. Duffy G, Cheng PP, Yuan N, He B, Kwan AC, Shun-Shin MJ, Alexander KM, Ebinger J, Lungren MP, Rader F, et al. High-Throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning. *JAMA Cardiol*. 2022;7:386–395. doi: 10.1001/jamacardio.2021.6059
8. Hwang IC, Choi D, Choi YJ, Ju L, Kim M, Hong JE, Lee HJ, Yoon YE, Park JB, Lee SP, et al. Differential diagnosis of common etiologies of left ventricular hypertrophy using a hybrid CNN-LSTM model. *Sci Rep*. 2022;12:20998. doi: 10.1038/s41598-022-25467-w
9. Yu X, Yao X, Wu B, Zhou H, Xia S, Su W, Wu Y, Zheng X. Using deep learning method to identify left ventricular hypertrophy on echocardiography. *Int J Cardiovasc Imaging*. 2022;38:759–769. doi: 10.1007/s10554-021-02461-3
10. Open AI Dataset Project (AI-Hub). National Information Society. https://aihub.or.kr/ Accessed XXX
11. Korea National Standard Cardiovascular Database and Reference. Korea Research Institute of Standards and Science. https://www.srd.re.kr/css/index.do Accessed XXX
12. Sengupta PP, Shrestha S, Berthon B, Messas E, Donal E, Tison GH, Min JK, D'Hooge J, Voigt JU, Dudley J, et al. Proposed requirements for cardiovascular imaging-related machine learning evaluation (PRIME): a checklist: reviewed by the American College of Cardiology Healthcare Innovation Council. *JACC Cardiovasc Imaging*. 2020;13:2017–2035. doi: 10.1016/j.jcmg.2020.07.015
13. Jeon J, H S, Yoon YE, Kim J, Jeong H, Jeong D, Jang Y, Hong Y, Chang HJ. Improving out-of-distribution detection in echocardiographic view classification through enhancing semantic features. *arXiv preprint arXiv:230816483*. 2023; doi: 10.48550/arXiv.2308.16483
14. Jang Y, Choi H, Yoon YE, Jeon J, Kim H, Kim J, Jeong D, Ha S, Hong Y, Lee SA, et al. An artificial intelligence-based automated echocardiographic analysis: enhancing efficiency and prognostic evaluation in patients with revascularized STEMI. *Korean Circ J*. 2024;54:743. doi: 10.4070/kcj.2024.0060
15. Park J, Jeon J, Yoon YE, Jang Y, Kim J, Jeong D, Lee J, Hong Y, Ha S, Reza A, et al. Artificial intelligence-enhanced automation of left ventricular diastolic assessment: a pilot study for feasibility, diagnostic validation, and outcome prediction. *Cardiovasc Diagn Ther*. 2024;14:352–366. doi: 10.21037/cdt-24-25
16. van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, Beets-Tan RGH, Fillion-Robin JC, Pieper S, Aerts H. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. 2017;77:e104–e107. doi: 10.1158/0008-5472.CAN-17-0339
17. Lee J, Jeong D, Jang Y, Jeong S, Jung T, Yoon YE, Moon I, Lee SA, Chang HJ. Self supervised convolutional kernel based handcrafted feature harmonization: enhanced left ventricle hypertension disease phenotyping on echocardiography. *arXiv preprint arXiv:231008897*. 2023; doi: 10.48550/arXiv.2310.08897
18. Chen Q, Pan T, Wang YN, Schoepf UJ, Bidwell SL, Qiao H, Feng Y, Xu C, Xu H, Xie G, et al. A coronary CT angiography radiomics model to identify vulnerable plaque and predict cardiovascular events. *Radiology*. 2023;307:e221693. doi: 10.1148/radiol.221693
19. Ke G, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu TY. LightGBM: a highly efficient gradient boosting decision tree. *Adv Neural Inf Process Syst*. 2017;30:3149–3157.
20. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res*. 2002;16:321–357. doi: 10.1613/jair.953
21. Scott M, Lundberg SIL. A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017;4768–4777. doi: 10.48550/arXiv.1705.07874
22. Drazner MH. The progression of hypertensive heart disease. *Circulation*. 2011;123:327–334. doi: 10.1161/CIRCULATIONAHA.108.845792
23. Maron MS, Hellawell JL, Lucove JC, Farzaneh-Far R, Olivotto I. Occurrence of clinically diagnosed hypertrophic cardiomyopathy in the United States. *Am J Cardiol*. 2016;117:1651–1654. doi: 10.1016/j.amjcard.2016.02.044
24. Moon I, Lee SY, Kim HK, Han KD, Kwak S, Kim M, Lee HJ, Hwang IC, Lee H, Park JB, et al. Trends of the prevalence and incidence of hypertrophic cardiomyopathy in Korea: a nationwide population-based cohort study. *PLoS One*. 2020;15:e0227012. doi: 10.1371/journal.pone.0227012
25. Kittleson MM, Maurer MS, Ambardekar AV, Bullock-Palmer RP, Chang PP, Eisen HJ, Nair AP, Nativi-Nicolau J, Ruberg FL; American Heart Association Heart Failure and Transplantation Committee of the Council on Clinical CardiologyTransplantation Committee of the Council on Clinical C. Cardiac amyloidosis: evolving diagnosis and management: a scientific statement from the American Heart Association. *Circulation*. 2020;142:e7–e22. doi: 10.1161/CIR.0000000000000792
26. Arbelo E, Protonotarios A, Gimeno JR, Arbustini E, Barriales-Villa R, Basso C, Bezzina CR, Biagini E, Blom NA, de Boer RA, et al; ESC Scientific Document Group. 2023 ESC Guidelines for the management of cardiomyopathies. *Eur Heart J*. 2023;44:3503–3626. doi: 10.1093/eurheartj/ehad194
27. Bhandari AK, Nanda NC. Myocardial texture characterization by two-dimensional echocardiography. *Am J Cardiol*. 1983;51:817–825. doi: 10.1016/s0002-9149(83)80139-8
28. Pinamonti B, Picano E, Ferdeghini EM, Lattanzi F, Slavich G, Landini L, Camerini F, Benassi A, Distante A, L'Abbate A. Quantitative texture analysis in two-dimensional echocardiography: application to the diagnosis of myocardial amyloidosis. *J Am Coll Cardiol*. 1989;14:666–671. doi: 10.1016/0735-1097(89)90108-3
29. Hathaway QA, Yanamala N, Siva NK, Adjeroh DA, Hollander JM, Sengupta PP. Ultrasonic texture features for assessing cardiac remodeling and dysfunction. *J Am Coll Cardiol*. 2022;80:2187–2201. doi: 10.1016/j.jacc.2022.09.036

30. Datar Y, Cuddy SAM, Ovsak G, Giblin GT, Maurer MS, Ruberg FL, Arnaout R, Dorbala S. Myocardial texture analysis of echocardiograms in cardiac transthyretin amyloidosis. *J Am Soc Echocardiogr*. 2024;37:570–573. doi: 10.1016/j.echo.2024.02.005
31. Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med*. 2021;27:582–584. doi: 10.1038/s41591-021-01312-x
32. Traverso A, Wee L, Dekker A, Gillies R. Repeatability and reproducibility of radiomic features: a systematic review. *Int J Radiat Oncol Biol Phys*. 2018;102:1143–1158. doi: 10.1016/j.ijrobp.2018.05.053
33. Fortin JP, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, Adams P, Cooper C, Fava M, McGrath PJ, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*. 2018;167:104–120. doi: 10.1016/j.neuroimage.2017.11.024
34. Woo S, Debnath S, Hu R, Chen X, Liu Z, Kweon IS, Xie S. ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders. *arXiv:230100808*. 2023; doi: https://doi.org/10.48550/arXiv.2301.00808
35. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. *J Stat Softw*. 2010;36:1–13. doi: 10.18637/jss.v036.i11
36. Al'Aref SJ, Singh G, Choi JW, Xu Z, Maliakal G, van Rosendael AR, Lee BC, Fatima Z, Andreini D, Bax JJ, et al. A boosted ensemble algorithm for determination of plaque stability in high-risk patients on coronary CTA. *JACC Cardiovasc Imaging*. 2020;13:2162–2173. doi: 10.1016/j.jcmg.2020.03.025
37. Han H, Wang WY, Mao BH. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. Paper/Poster presented at: International conference on intelligent computing; 2005.
38. He H, Bai Y, Garcia EA, Li S. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. Paper/Poster presented at: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence); 2008.
39. Luo X, Piao S, Li H, Li Y, Xia W, Bao Y, Liu X, Geng D, Wu H, Yang L. Multilesion radiomics model for discrimination of relapsing-remitting multiple sclerosis and neuropsychiatric systemic lupus erythematosus. *Eur Radiol*. 2022;32:5700–5710. doi: 10.1007/s00330-022-08653-2
40. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32–35. doi: 10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3