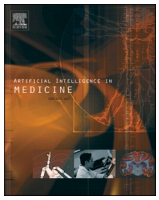
Artificial Intelligence In Medicine 140 (2023) 102548

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed



Research paper

Deep cross-modal feature learning applied to predict acutely decompensated heart failure using in-home collected electrocardiography and transthoracic bioimpedance

Xiang Pan [a,g,1], Chuangqi Wang [a,1], Yudong Yu [b], Natasa Reljin [d], David D. McManus [e], Chad E. Darling [f], Ki H. Chon [d,**], Yitzhak Mendelson [a,c,**], Kwonmoo Lee [a,g,h,*]

[a] Department of Biomedical Engineering, Worcester Polytechnic Institute, MA 01609, USA
[b] Robotics Engineering Program, Worcester Polytechnic Institute, MA 01609, USA
[c] Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, MA 01609, USA
[d] Department of Biomedical Engineering, University of Connecticut, CT 06269, USA
[e] Department of Medicine, University of Massachusetts Medical School, Worcester, MA 01655, USA
[f] Department of Emergency Medicine, University of Massachusetts Medical School, Worcester, MA 01655, USA
[g] Vascular Biology Program, Boston Children's Hospital, Boston, MA 02115, USA
[h] Department of Surgery, Harvard Medical School, Boston, MA 02115, USA

ARTICLE INFO

Keywords:
Deep learning
ECG analysis
Wearable device
Heart failure

ABSTRACT

Background: Deep learning has been successfully applied to ECG data to aid in the accurate and more rapid diagnosis of acutely decompensated heart failure (ADHF). Previous applications focused primarily on classifying known ECG patterns in well-controlled clinical settings. However, this approach does not fully capitalize on the potential of deep learning, which directly learns important features without relying on a priori knowledge. In addition, deep learning applications to ECG data obtained from wearable devices have not been well studied, especially in the field of ADHF prediction.

Methods: We used ECG and transthoracic bioimpedance data from the SENTINEL-HF study, which enrolled patients (≥21 years) who were hospitalized with a primary diagnosis of heart failure or with ADHF symptoms. To build an ECG-based prediction model of ADHF, we developed a deep cross-modal feature learning pipeline, termed ECGX-Net, that utilizes raw ECG time series and transthoracic bioimpedance data from wearable devices. To extract rich features from ECG time series data, we first adopted a transfer learning approach in which ECG time series were transformed into 2D images, followed by feature extraction using ImageNet-pretrained DenseNet121/VGG19 models. After data filtering, we applied cross-modal feature learning in which a regressor was trained with ECG and transthoracic bioimpedance. Then, we concatenated the DenseNet121/VGG19 features with the regression features and used them to train a support vector machine (SVM) without bioimpedance information.

Results: The high-precision classifier using ECGX-Net predicted ADHF with a precision of 94 %, a recall of 79 %, and an F1-score of 0.85. The high-recall classifier with only DenseNet121 had a precision of 80 %, a recall of 98 %, and an F1-score of 0.88. We found that ECGX-Net was effective for high-precision classification, while DenseNet121 was effective for high-recall classification.

Conclusion: We show the potential for predicting ADHF from single-channel ECG recordings obtained from outpatients, enabling timely warning signs of heart failure. Our cross-modal feature learning pipeline is expected to improve ECG-based heart failure prediction by handling the unique requirements of medical scenarios and resource limitations.

* Correspondence to: K. Lee, Vascular Biology Program, Boston Children's Hospital, MA 02115, USA.
** Co-corresponding authors.
E-mail addresses: ki.chon@uconn.edu (K.H. Chon), ym@wpi.edu (Y. Mendelson), kwonmoo.lee@childrens.harvard.edu (K. Lee).
[1] These authors equally contributed to this work.

https://doi.org/10.1016/j.artmed.2023.102548
Received 21 July 2022; Received in revised form 30 March 2023; Accepted 4 April 2023
Available online 11 April 2023

1. Introduction

Acute decompensated heart failure (ADHF) is a worsening of heart failure (HF)-related symptoms and signs, which include dyspnea, weight gain, edema, and fatigue. The hallmark of ADHF is a reduction in the ability of the heart to pump blood to vital organs, which causes serious symptoms, such as difficulty breathing and organ failure [1]. In the U.S., incidence increases with age, and the condition is associated with substantial morbidity, mortality, and impaired quality of life [2]. Chronic HF is the most common cause of hospitalization in elderly individuals, and the population burden of HF is increasing with the aging of the population [3]. In 2014, primary and comorbid HF caused 978,135 and 3,370,856 hospitalizations, respectively [3].

An episode of ADHF is life-threatening, but with proper and timely treatment, the risks can be minimized. At the onset of ADHF, healthcare providers often use medication to address intravascular volume, such as using oral or intravenous loop diuretics to increase urine output and reduce vascular congestion [4]. Treatment for ADHF during hospitalization greatly improves symptom burden, but a recent study demonstrated that early recurrent ADHF after hospital discharge is common (51 % and 27 % for moderate and severe congestion in the first month after discharge) [5], and the mortality rate for individuals with recurrent ADHF is approximately 20% [6,7]. To identify HF patients at risk for recurrence, it is critical to accurately predict the likelihood of recurrent ADHF.

Because ADHF is often characterized by severe dyspnea due to pulmonary edema, ADHF detection requires a method to measure fluid accumulation in the lung. Transthoracic bioimpedance (TBI) used to measure fluid accumulation in the lung has been shown to be a more accurate measure of intrathoracic volume retention than alternative contemporary clinical approaches, including monitoring total body weight as a proxy for total volume status [8–10]. TBI is measured by using electrodes to inject a small (typically <1 mA) alternating current (AC) into the tissues and measuring the response. Along with bioimpedance values, one-lead electrocardiogram (ECG) data can also be obtained with electrodes that are positioned in the thorax. Thus, the aim of this analysis is to examine both bioimpedance and ECG data, with the aid of machine learning and deep learning, to predict ADHF.

Recent progress in machine learning (ML), particularly deep learning (DL), has shown that computers (artificial intelligence algorithms) can outperform humans in fields that involve the analysis of complex high-dimensional datasets, such as image classification, audio recognition, and games [11,12]. Deep neural network (DNN) models, particularly convolutional neural network (CNN) [13,14] models, have shown high feasibility for medical applications in which the diagnosis consists of making observations from images, such as breast cancer classification and lung area detection, where they have unique advantages [15,16]. Although DNNs have been successfully applied to ECG data, previous applications have mainly focused on classifying rhythms, such as atrial fibrillation, from ambulatory patients in well-controlled clinical settings [17–19]. Although accurate and automatic detection of abnormal ECG patterns will greatly aid in making health care decisions, this approach does not fully capitalize on the potential of DL, which directly learns important features from raw input data without relying on a priori knowledge. Recent studies suggest that a patient's risk of cardiovascular death [20] or congestive HF [21] and cardiac contractile dysfunction [22] can be predicted by analyzing raw ECG signals. However, while HF is a progressive disease, detecting ADHF and related complications is attracting more attention in the field. To date, the prediction of worsening ADHF and risk for rehospitalization using single-channel ECG signals from wearable devices has not been extensively studied.

While DL is known to require a large amount of training data, clinical data are often limited and expensive to obtain. In studies with limited data sizes, transfer learning has been proven to be an effective way to reduce overfitting [23]. In particular, transfer learning has been widely adopted to build robust classifiers, mainly with image datasets. Notably, CNN models pretrained on the large numbers of natural images available in ImageNet [24] have been used as general-purpose image feature extractors to build classifiers for different domains [25–30]. However, effective pretrained models for time series data are still limited because no large-scale time series datasets are available (especially ECG datasets), making it challenging to extract rich features from time series data.

We used the ECG data from the SENTINEL-HF study by Dovancescu et al. [31] to predict worsening HF, such as recurring hospitalization. SENTINEL-HF was a prospective study of patients discharged after hospitalization for ADHF, and it employed a wearable bioimpedance vest to monitor patients after discharge to identify sudden or gradual worsening of HF symptoms [32]. Decreased bioimpedance as a result of intrathoracic fluid accumulation was correlated with decompensating HF. Data from the SENTINEL-HF study provide an opportunity for cross-modal training, where better ECG features can be learned when feature learning utilizes ECG and bioimpedance together. This allows us to extract refined ECG features that enable us to predict whether patients experience another ADHF occurrence.

In this study, we present a new ECG cross-modal feature learning pipeline, termed ECGX-Net (ECG X-modal Network), which utilizes image-based transfer learning for ECG time series for cross-modal feature learning. We leverage image-based pretrained models (DenseNet121 or VGG19) [33] for time series transfer learning by transforming 1-D ECG time series data into 2-D image data using the Gramian Angular Summation Field (GASF) [34]. After data filtering, we perform cross-modal feature learning, where a multilayer perceptron (MLP) regressor predicts transthoracic bioimpedance based on ECG. We demonstrate that our time series cross-modal feature learning pipeline can predict ADHF based on raw ECG recordings.

2. Methods

2.1. Data preparation

We used 1318 ECG recordings from 37 different volunteers in the SENTINEL-HF study by Dovancescu et al. [31] as the source of the data. The SENTINEL-HF study enrolled patients ($\geq$21 years) who were hospitalized with a primary diagnosis of heart failure or with ADHF symptoms. The participants provided informed consent, and the institutional review committee approved the study protocol.

Each ECG recording was approximately 5 min in length, contained approximately 250–350 heartbeats, and constituted 65,000–85,000 readings, ranging from 2048 mV to −2048 mV. The patients were anonymized, and the data were identified by a pseudonym of the device name. We used the MATLAB built-in function *findpeaks* with *'MinPeakHeight'* as the 5 % quantile and *'MinPeakDistance'* set to 100 to locate the S wave peaks, which were used to splice the five heartbeat ECG segments. In total, we extracted 7414 heartbeat ECG segments, 2999 of which were discarded due to artifacts during manual data selection. After feature extraction and PCA dimensional reduction, we discarded the ECG segments associated with other unrelated minor conditions, and selected the ECG segments associated with ongoing recurrent ADHF. These 1462 five heartbeat ECG time series were processed by clinical and morphological data filtering. Finally, 441 ECG samples taken after patient discharge were selected after data filtering to construct the classifiers.

2.2. Image transformation from time series

The five heartbeat ECG segments were processed through the min-max normalization shown in Eq. (1):

$$X = \frac{time\ series - min(time\ series)}{max(time\ series) - min(time\ series)}. \tag{1}$$

We transformed the normalized time series X to 2D grayscale images with the Gramian Angular Summation Field (using the *GASF()* and

fit_transform() functions in the *pyts.image 0.7.0* Python package). The GASF transformed the 1D time series to 2D images by scaling the time series to a normalized interval (Eq. (2)), and this was followed by conversion into polar coordinates (Eqs. (3)–(4)) and then calculation of the angular perspective between each pair of time points, as shown in Eq. (5) [34]:

$$\tilde{x}_i = \frac{x_i - max(X) + x_i - min(X)}{max(X) - min(X)}, \tag{2}$$

$$\phi = arccos(\tilde{x}_i),\ -1 \le \tilde{x}_i \le 1, \tilde{x}_i \in \tilde{X}, \tag{3}$$

$$r = \frac{t_i}{N}, t_i \in N, \tag{4}$$

$$GASF_{ij} = cos(\phi_i + \phi_j). \tag{5}$$

After the time series were transformed into grayscale images, we duplicated the images three times to create the RGB images that formed the input to the CNN. We also compared the feature extraction ability and predictability with those of other image transformation methods: Markov Transition Field and Recurrence Plot (using *MTF()* and *RP()* from the same Python package). For the MTF, a time series X is divided into Q quantile bins, and each x_i is assigned to the corresponding q_j. By applying a first-order Markov chain, we can calculate a Q × Q adjacency matrix W with frequency $w_{i,j}$ (where a point in q_j also falls into q_i). In this way, a time series can be converted to a 2D MTF matrix by applying Eq. (6). Using the RP, a time series is first converted into its extracted trajectories (Eq. (7)), where m is the dimension of the trajectories and τ is the time delay, and the RP is calculated as a pairwise distance matrix between trajectories (Eq. (8)) [35]. We used 10 % (the default) as the percentage threshold, ϵ:

$$mTF = \begin{bmatrix} w_{ij}|_{x_1 \in q_i, x_1 \in q_j} & \cdots & w_{ij}|_{x_1 \in q_i, x_n \in q_j} \\ \vdots & \ddots & \vdots \\ w_{ij}|_{x_n \in q_i, x_1 \in q_j} & \cdots & w_{ij}|_{x_n \in q_i, x_n \in q_j} \end{bmatrix}, \tag{6}$$

$$\vec{x} = \left(x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}\right), \forall_i \in \{1, \ldots, n-(m-1)\tau\}, \tag{7}$$

$$RP_{ij} = \theta\left(\epsilon - \left|\left|\vec{x}_i - \vec{x}_j\right|\right|\right), \forall_{i,j} \in \{1, \ldots, n-(m-1)\tau\}. \tag{8}$$

For all three methods, the resultant RGB images were transformed and resized from the ECG data with varying temporal lengths into 224 × 224-pixel images using the *resize* function from the *cv2* module. The final images were used as the input to the CNN.

2.3. Feature extraction and reduction

We used CNN-based pretrained models (DenseNet121 [36] or VGG19 [33]) with ImageNet weights as our main approach to extract features from the resized RGB images of the ECG time series data. The VGG19 network we used consists of 16 convolutional layers and 1 FC layer. The 4096 features were extracted after layer FC1. Then, using the entire dataset (all 4415 data points), PCA was performed to reduce the feature dimension to 250, which was determined by plotting the explained variance and cumulative explained variance curves to reduce the effect of noise. For the DenseNet121 network, we used the same dataset and extracted 1024 features before the last FC layer. We used PCA to reduce the feature dimension to 170.

We also used the pretrained methods VGG16 [33], ResNet50 [37], ResNet101 [37], and DenseNet169 [36] to replace DenseNet121 and VGG19. The models were imported from the *keras.applications* module, and their weights were trained on ImageNet. We also used the ImageNet-pretrained AlexNet [14] available in *MATLAB*. The features from ResNet50, ResNet101, and DenseNet169 were extracted before the last FC layer with dimensions of 2048, 2048 and 1664, respectively. The features from VGG16 and AlexNet were extracted after layer FC1, similar to VGG19, and had 4096 dimensions. PCA was then performed to reduce the feature dimensions to 500 for VGG16 and 150 for ResNet50, ResNet101, DenseNet169, and AlexNet.

2.4. MLP regressor with bioimpedance

We constructed a conventional MLP regressor to introduce transthoracic bioimpedance to the ECG-DenseNet191-PCA or ECG-VGG19-PCA features. The regressor contained three dense layers, with dimensions of 128, 64 and 1 in sequence. The regressor had two purposes. First, the outcome of the model was used as the predicted bioimpedance for the test set. Second, all DenseNet121-PCA or VGG19-PCA features were processed through the model, and new features were extracted after the second dense layer for classification purposes. The new features had 64 dimensions and preserved both the ECG and bioimpedance information.

2.5. Classifier training

In our main approach, the features of DenseNet121-PCA or VGG19-PCA were concatenated with features extracted from the MLP regressor, and the combined features (ECGX-Net features) were classified by an SVM classifier. For the DenseNet121 and VGG19 combined features, the model consisted of a radial basis function (RBF) kernel, with optimum parameter values of gamma and C selected by grid search. The grid search also included linear and polynomial kernels. The optimum kernels and parameters are shown in Supplementary Fig. 3C.

We used the data from the participants and dates that contained existing bioimpedance records as the training set and the remaining data that did not contain bioimpedance records as the test set. The ECGs associated with bioimpedance values were used for training and validation (type 1, 229 samples), and the ECGs without bioimpedance values were used for testing (type 2, 212 samples). Therefore, the testing set was completely independent from the training set. For the bioimpedance regression, we used the type 1 data as the training and validation set (random split with a 4:1 split ratio), and we used the type 2 data as the test set. For SVM training, we utilized a grid-search process to optimize the parameters for the SVM classifier using the SVC function from the sklearn.svm module and the GridSearchCV function from sklearn.model_selection. The parameter search was conducted using type 1 data and five-fold cross-validation. The optimized classifier was then applied to the type 2 data for testing. The process was repeated 50 times, with each repetition involving randomly selecting a 1:1 HF to N ratio from the type 1 data for training. The model's performance was evaluated by averaging the results of all 50 repetitions.

2.6. Hyperparameter searching

The hyperparameters of the bioimpedance MLP regressor are the numbers of dense layers and of the neurons of each layer for the MLP regressor. We tested ten different configurations of the MLP regressor. The numbers of hidden neurons in each configuration were set as 128 → 64 → 1, 128 → 32 → 1, 128 → 16 → 1, 64 → 32 → 1, 64 → 16 → 1, 32 → 16 → 1, 128 → 64 → 32 → 1, 128 → 64 → 16 → 1, 128 → 32 → 16 → 1, and 64 → 32 → 16 → 1. We used DenseNet121-PCA features as input, and the batch size was 64. With the optimum structure, we tested different optimizers such as Adam, Adadelta, Adagrad and RMSprop with the fixed learning rate 0.1, 0.01, 0.001 and 0.0001. The parameters were evaluated according to the classification performance using combined features that involved regressor features from each configuration (Supplementary Fig. 3A–C). Based on this, we selected the MLP structure of 128 → 64 → 1, Adam optimizer with the fixed learning rate of 0.001, and the batch size of 64.

The hyperparameters of the SVM classifier are the types of kernels and the kernel parameters. We carried out hyperparameter search for the SVM classifier among three kernels: linear, radial basis function

(RBF), and polynomial. We associated the SVC function from the *sklearn.svm* module with the *GridSearchCV* function from *sklearn.model_selection*. For the linear kernel, grid search was performed on the optimization parameter C (among 0.01, 0.1, 1, 10, 100, and 1000). For the RBF kernel, grid search was conducted on C (among 0.01, 0.1, 1, 10, 100, and 1000) and the curvature parameter gamma for the decision boundary (among 0.0001, 0.001, 0.01, 0.1, 1). For the polynomial kernel, grid search was performed on the degree (1, 2, and 3), C (among 0.01, 0.1, 1, 10, 100, and 1000), and gamma (0.0001, 0.001, 0.01, 0.1, and 1). We evaluated the classification performance by five-fold cross-validation using the F1-score to select the optimal parameters (Supplementary Fig. 3D).

2.7. Model performance analysis

The model's performance and prediction robustness were evaluated by calculating the accuracy, recall, precision, F1-score, and AUC-PR (area under the precision-recall curve). The calculations were based on the data using MATLAB's predicted score *perfcurve* function. During visualization, the output AUC was interpolated from 0 to 1 with a 0.008 interval. The Wilcoxon signed-rank test was used to compare the performance results from our approach and those of the baseline classifiers.

2.8. Model interpretability

To explain the features' contributions in ECGX-Net, we used SHAP (Shapley Additive exPlanations) [38], which is based on Shapley values [39]. We used KernelSHAP (the KernelExplainer class in the SHAP Python package) to identify the features' contributions in the SVN classifier. We calculated the summation of the loading values across the highest-contributing principal components to evaluate the importance of the output features of the pretrained model. Then, we used Deep SHAP (the GradientExplainer class in the SHAP Python package) for the DenseNet121 and VGG19 pretrained models. The SHAP classes were first initialized in the training set and then applied to the test set for feature interpretation. To save computational and memory resources, only the top 50 features associated with heart failure were used as the outputs of the pretrained models. We generated the SHAP values on the 7th layer of VGG19 with a size of 56 × 56 and those on the 13th layer of DenseNet121 with a size of 56 × 56. We calculated the average SHAP values across all samples in the N and HF classes.

3. Results

3.1. An overview of the data and the ML pipeline

The ECG recordings from the SENTINEL-HF study of Dovancescu et al. [31] were used as the main source of data in this study. The SENTINEL-HF study enrolled 180 patients who were previously hospitalized with ADHF. After the participants had recovered and were discharged from the hospital, ECG and other measurements were recorded regularly by wearable devices. The participants were followed to determine clinical 'events,' such as visits to the Emergency Department, rehospitalizations, and medications used to treat ADHF, such as uptitration of diuretic therapy. The charts were adjudicated by physicians to determine whether the clinical events were due to worsening HF or other etiologies [40].

In the SENTINEL-HF study, the bioimpedance vest was designed to monitor the health condition of the participants during daily activities after hospital discharge. When participants wore the vest, four textile electrodes in the vest measured changes in bioimpedance, which reflects the intrathoracic fluid accumulation level. Additionally, the vest recorded ECG signals in the form of a single-channel time series. Sample individuality, which presents as differences in ECG patterns, for example, the amplitudes and shapes of QRS intervals, was observed among different participants (Supplementary Fig. 1). The participants underwent routine/daily procedures with the instruments during the study. Each procedure produced one ECG time series approximately 5 min in length; these collected time series were used as the raw dataset in this study. The raw dataset consists of 1318 unprocessed ECG time series from 37 participants. Each five-minute ECG time series contains approximately 250–350 heartbeats. The ECG amplitude ranges from −2048 mV to +2048 mV. The large noise that exists at the beginnings of all the segments was likely caused by poor contact between the skin and ECG electrodes.

Fig. 1A shows a schematic diagram of the worsening HF prediction procedure used in this paper. A peak detection algorithm was used to record S wave locations in MATLAB (MathWorks). At each S wave location in the ECG signal, subsegments with five consecutive heartbeats were extracted (Fig. 1B). The follow-up subsegments started consecutively at one heartbeat after the start of the previous subsegment. During training set preparation, we visually inspected the ECG data and removed data containing artifacts (including large noise and missing/extra parts).

We randomly selected the ECG data from those participants who had experienced later HF or worsening HF indicators during the Dovancescu et al. study (including the follow-up period). For participants who were not rehospitalized or did not experience an HF event, we randomly selected data during the span of the Dovancescu et al. study. In this way, the labels were assigned based on the participants' health statuses at the end of the SENTINEL-HF study, and the ECG subsegments were categorized into two classes:

- Class HF (rehospitalized, worsening ADHF): The data from participants who experienced an ongoing recurrent ADHF event and/or cardiac ischemic-related heart failure during the study.
- Class N: The data from participants who were not rehospitalized or did not experience ADHF during the study and follow-up period, representing the normal group.

Initially, we verified 1462 five raw heartbeat ECG subsegments (before the data filtering step; see the Method for details). The samples were selected from 18 of the volunteers. After data filtering (see Data filtering in the Results section for details), the data from the ADHF group contained 186 ECG subsegments (from four participants), and the normal data (N) contained 255 ECG subsegments (from four participants, Fig. 1C).

Then, we transformed the ECG subsegments into 2-D gray-scaled images by applying the Gramian Angular Summation Field (GASF) algorithm (Fig. 1D; see the Methods section for details). Because the lengths of the time series vary between samples, the difference in the length of a time series and the corresponding influence was eliminated by setting a uniform size parameter of 224 × 224 pixels during the GASF transformation. To conduct feature extraction, the resulting images were copied into each RGB channel of the ImageNet-pretrained DenseNet121 or VGG19 models (Fig. 1E). This transfer learning approach has both high simplicity and proven high accuracy on the classification and segmentation of biomedical images [30,41–45]. Then, the dimension of the extracted pretrained model features was reduced by PCA. Next, we refined the ECG features by performing cross-modal feature learning by predicting the bioimpedance using the pretrained model features of ECG (Fig. 1F). The features extracted from the MLP-based regressor were then concatenated with the pretrained CNN-PCA features, followed by SVM classification using the ECG features for cross-modal learning.

3.2. VGG19 features of image-based ECG signals

We first used the VGG19 model in our pipeline to explore the dataset. The pretrained VGG19 model takes 2D images containing five heartbeats of the ECG time series as input and produces features of these images as output. These features are extracted after the first fully connected layer (FC1), which has a dimension of 4096. Then, principal

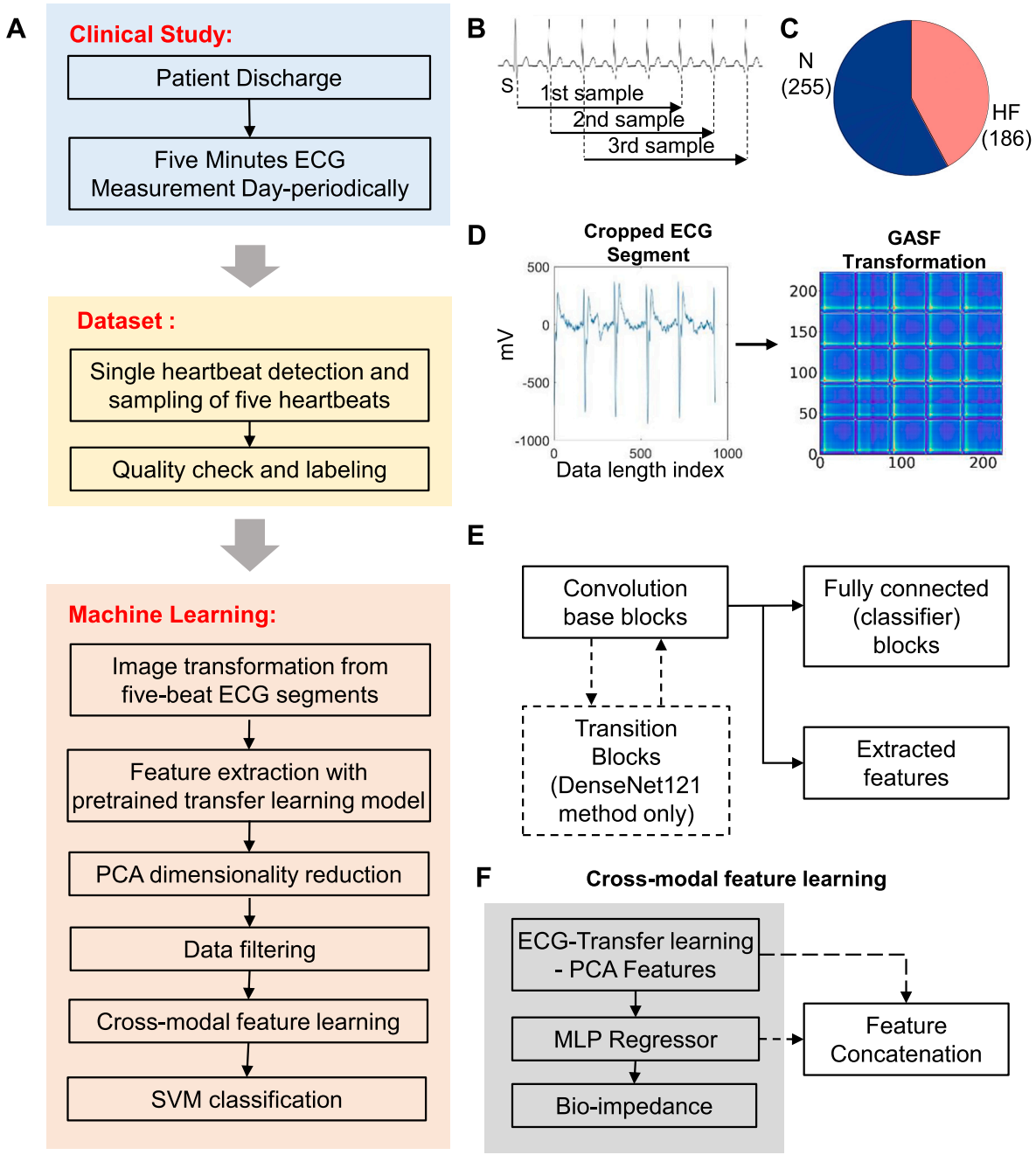


Fig. 1. The schematics of worsening heart failure detection with transfer learning: (A) flowchart of the pipeline; (B) schematic of raw ECG time series cropping; (C) distribution of data among Class HF and N after data filtering; (D) transformation of ECG time series into 2D images by GASF; (E) schematic structure of feature extraction using transfer learning models. (F) Cross-modal feature learning between ECG and bioimpedance.

component analysis (PCA) was conducted to reduce the feature dimension from 4096 to 500 (Figs. 1E, 2A–B), thus emphasizing the most important features.

We visualized the distribution of these VGG19-PCA features using t-distributed stochastic neighbor embedding (t-SNE) and uniform manifold approximation and projection (UMAP) and found that instead of large clusters, separate small clusters that belonged to either Class HF or N appeared on the t-SNE (Fig. 2C, left) and UMAP (Fig. 2D, left) plots. However, each cluster of ECG signals also corresponded to individual participants (Fig. 2C–D, right). Even though these VGG19-PCA features were highly specific to individual participants, it was not clear whether those features were related to ADHF or were merely patient bias. Therefore, a data filtering step is necessary to select the samples that can be distinguished by ADHF-related features.

3.3. *Data filtering*

Fig. 2E shows a schematic diagram of our data filtering procedure. First, we divided the entire dataset into ECG of regular and irregular heartbeats. An irregular heart rate could be due to atrial fibrillation, which should be considered separately. We identified 386 samples (from 7 participants) with irregular heart rates. In this way, we separated the data into two groups:

- Irregular heartbeat data: The data that exhibit dramatic changes in heart rate (Fig. 2G, right).

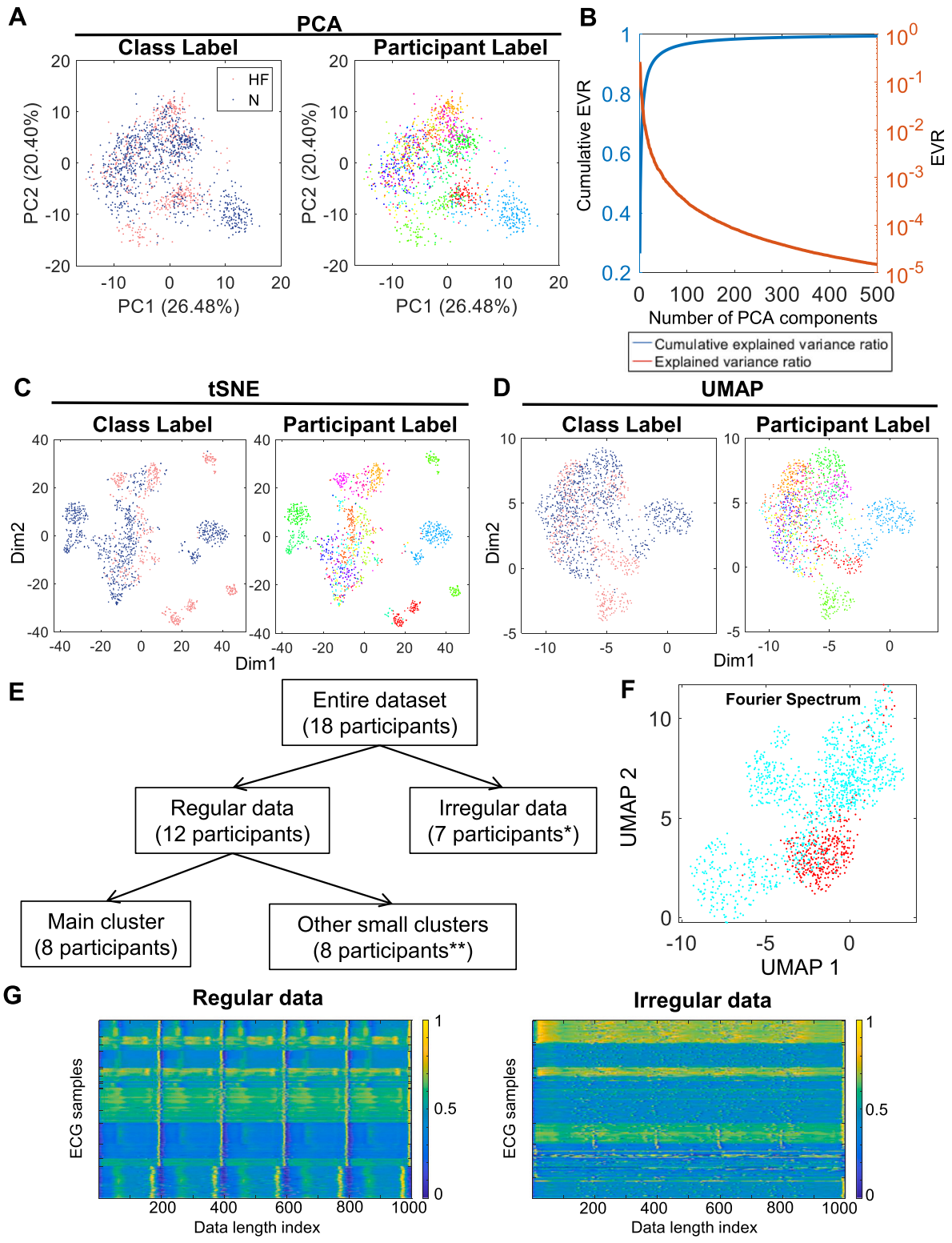


Fig. 2. Features extracted by pretrained VGG19 and data filtering: (A) first two components of PCA on VGG19 features; (B) explained variance ratio (EVR) and cumulative EVR plots with the number of PCA components; (C–D) t-SNE (C) and UMAP (D) plots of VGG19-PCA features. Right of (C), (D): Each color represents individual participants; (E) tree diagram for data filtering (*: 1 participant contributed data to both irregular and regular data groups; ** 4 participants contributed data in the main cluster and the small clusters); (F) UMAP plot of the single-sided spectrum from Fast Fourier Transform (blue: regular data; red: irregular data); (G) heatmaps of interpolated ECG of regular and irregular data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

- Regular heartbeat data: The data that show a constant heart rate (Fig. 2G, left).

To fully eliminate the effect of atrial fibrillation in the regular-data group, we separated the data in accordance with the source date and participant. If more than half of the data from a participant on a given day were labeled as irregular data, the rest of the data collected from this patient on the same day were also labeled as irregular data. Fig. 2F shows the distribution of data with regular and irregular heartbeats in a UMAP visualization of min-max normalized single-sided spectrum features from the Fourier transform.

Next, we selected the regular dataset and visualized its UMAP distribution to further filter the data (Fig. 3A–C). We performed K-means clustering to separate the data into 8 clusters (Fig. 3C), resulting in one large main cluster that contained 441 samples (from 8 participants) and 7 small clusters that contained 635 samples (from 8 participants). The small clusters contained either one or two participants, and none of them were diverse enough to have data distributed evenly among the HF and N groups for robust classification. Due to this lack of available data, we filtered out the data in the small clusters and used only the data in the main cluster to build a classifier. Fig. 3D–F shows the UMAP visualization of the features from the main cluster of the regular samples. There exists a clear boundary between Groups HF and N (Fig. 3D). Moreover, low and high bioimpedance values tend to be associated with HF and N, respectively, suggesting that the ECG features in the main clusters can reflect the pathophysiology of ADHF (Fig. 3F, where the gray dots represent the data for which a bioimpedance reading was not available).

3.4. Cross-modal feature learning of ECG with transthoracic bioimpedance

We performed cross-modal feature learning, where we trained the MLP regressor with DenseNet121-PCA or VGG19-PCA features as inputs and the bioimpedance data as outputs (Fig. 4A–B). We used the ECG features that had recorded bioimpedance readings (229 samples, 155 HF and 74 N) as the training and validation groups at a ratio of 0.8:0.2 and then predicted bioimpedance on the part of the DenseNet121-PCA or VGG19-PCA features with no recorded bioimpedance records (212 samples, 31 HF and 181 N). As shown in Fig. 4B, there was a clear correlation between the actual and predicted values of bioimpedance after training with both the DenseNet121 and VGG19 pretrained models. The results demonstrate that ECG features from the main cluster in the regular dataset can be used to recognize the bioimpedance values, thereby allowing the detection of worsening heart failure associated with low bioimpedance.

For the irregular ECG data, 89 bioimpedance recordings were available for the irregular data. However, approximately 69.7 % of the bioimpedance recordings were associated with a single participant in Group N (62 samples), and for the data in Group HF (27 samples), 14 bioimpedance recordings were associated with a single participant. Because more subjects are needed to validate the feasibility of classification for irregular data, we did not use the irregular ECG dataset for cross-modal training (Supplementary Fig. 2).

3.5. Prediction of ADHF using a support vector machine

The cross-modal features from the MLP regressor model were extracted after the second dense layer and were 64 dimensions in length. After we concatenated them with the DenseNet121-PCA or VGG19-PCA features (Fig. 1F) to obtain the features of ECGX-Net (the cross-modal pipeline that uses DenseNet121-PCA or VGG19-PCA feature extraction and MLP regression), we trained the support vector machine (SVM) classifier. We used the ECGX-Net features extracted from the data associated with bioimpedance records as the training set and the rest that did not contain bioimpedance records as the test set, thereby isolating the test set from the training set. We repeated the training 50

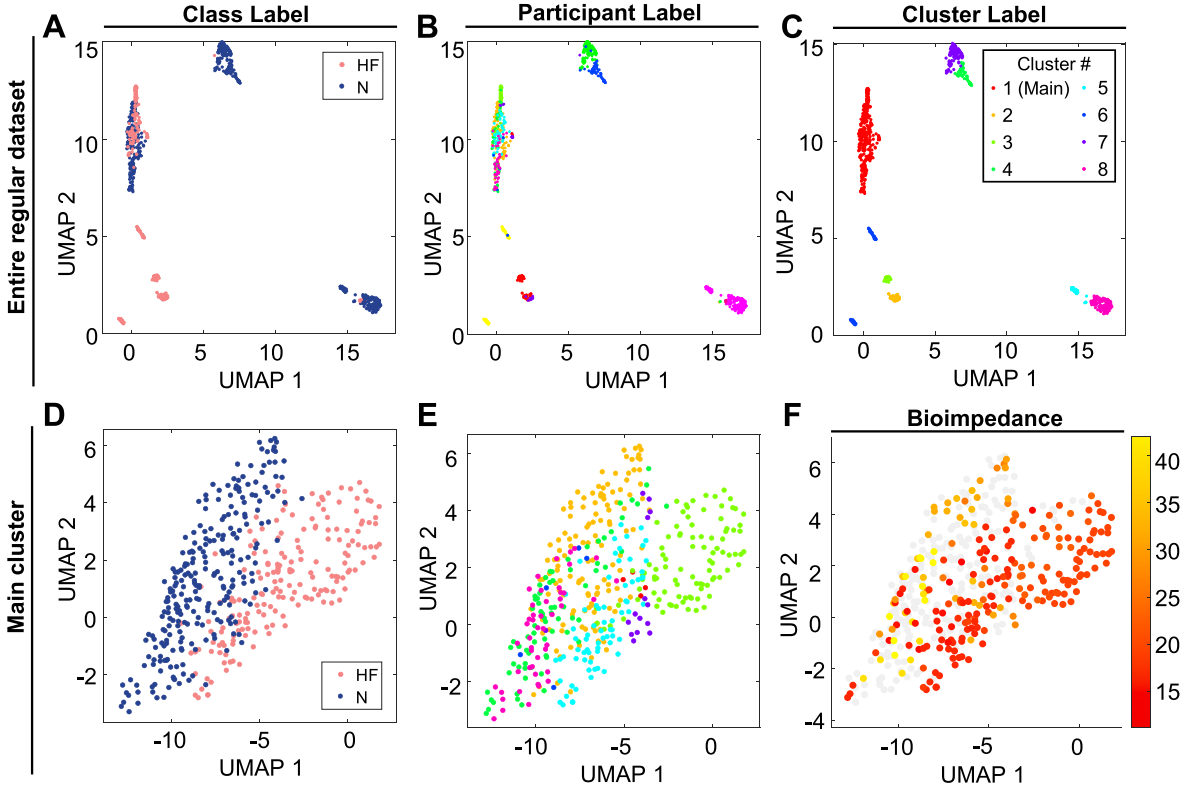


Fig. 3. UMAP visualization of VGG19-PCA features for data filtering with the regular dataset: (A–C) The entire regular dataset with the class labels (A), participant labels (B), and cluster labels (C). (D–F) The main cluster in the regular dataset with the class labels (D), participant labels (E), and bioimpedance (the gray dots correspond with the data missing bioimpedance measurement).

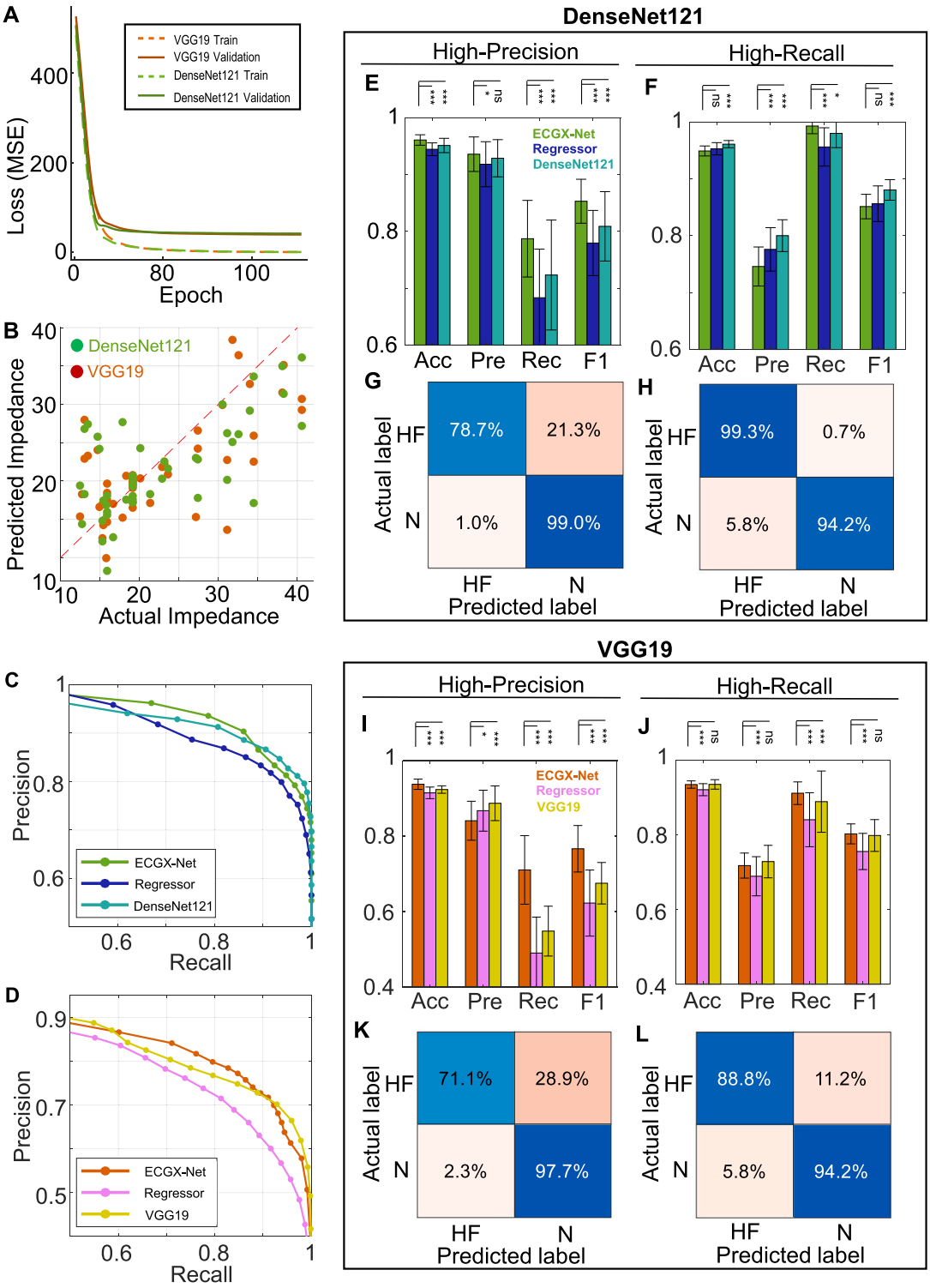


Fig. 4. Training of transthoracic bioimpedance MLP regressor: (A) Training loss curve of bioimpedance MLP regressor (green: using DenseNet121 features; orange: using VGG19 features; dotted line: training loss; solid line: validation loss); (B) Actual and predicted bioimpedance values of R0 transthoracic bioimpedance on the validation set (green: DenseNet121, correlation = 0.68, p-value = 1.95×10^{-7}; green: VGG19, correlation = 0.66, p-value = 7.16×10^{-7}). (C–D) Precision-recall curves of (C) DenseNet121 (green: DenseNet121-PCA + bioimpedance regression features; blue: bioimpedance regression features; cyan: DenseNet121-PCA features) and (D) VGG19 (orange: VGG19-PCA + bioimpedance regression features; red: bioimpedance regression features; yellow: VGG19-PCA features); (E–H) DenseNet121 SVM classification comparison. Performance (E) and confusion matrix (G) in high-precision classification (85 % HF estimate probability threshold). Performance (F) and confusion matrix (H) in high-recall classification (50 % HF estimate probability threshold). (I–L) VGG19 SVM classification comparison. Performance (I) and confusion matrix (K) in high-precision classification (85 % HF estimate probability threshold). Performance (J) and confusion matrix (L) in high-recall classification (35 % HF estimate probability threshold). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

times with random downsampling of the data to avoid class imbalance.

To characterize the role of the ECG-bioimpedance cross-modal features, we calculated the precision-recall (PR) curves of the models from DenseNet121 (Fig. 4C) and VGG19 (Fig. 4D). The mean area under the curve (AUC) of the PR curves with DenseNet121 was 0.9249 for ECGX-Net, 0.9169 for DenseNet121-PCA features, and 0.9086 for Regressor features. The AUC-PR with VGG19 was 0.8371 for ECGX-Net, 0.8323 for VGG19-PCA features, and 0.7842 for Regressor features. In the high-precision regions, where the precision is >0.9 with DenseNet121 and 0.75 with VGG19, the PR curves of ECGX-Net are above those of the classifiers using only pretrained model features. However, in the high-recall regions, where the recall is >0.9 with DenseNet121 or VGG19, the PR curves of ECGX-Net are below those of the classifiers using only pretrained model features. This suggests that the ECG-bioimpedance cross-modal features can be beneficial for high-precision classification.

Based on this finding, we built high-precision classifiers whose precision was >90 % with an 85 % threshold for HF (Fig. 4E and G, Supplementary Fig. 4A–B, E–F, I–J). ECGX-Net with DenseNet121 significantly outperformed DenseNet121 alone (see Supplemental Information 1 for the statistical testing) by achieving 93.58 % precision, 78.71 % recall, and 0.85 F1-score, whereas the sole DenseNet121 features achieved 92.87 % precision, 72.32 % recall, and 0.81 F1-score. We also built high-recall classifiers whose recall was >95 % with a 50 % threshold for HF (Fig. 4F and H). ECGX-Net with DenseNet121 achieved 74.59 % precision, 99.29 % recall, and 0.85 F1-score, whereas DenseNet121 features alone achieved 80.01 % precision, 98 % recall, and 0.88 F1-score. When we used the VGG19 pretrained model, ECGX-Net exhibited similar behaviors even though its overall performance was lower than that of DenseNet121 (Fig. 4I–L, Supplementary Fig. 4C–D, G–H, K–L, Supplemental Information 1). Therefore, ECGX-Net is preferred when high-precision classification is necessary. On the other hand, DenseNet121 features alone without cross-modal training can be used if a high-recall classifier is desired.

In heart failure detection, a high-recall classifier is typically preferred because false negative results can have severe consequences, potentially leading to a worsening of the condition and increased morbidity or mortality. However, if false positive results lead to unnecessary further testing or treatment, then a high-precision classifier may be preferred. Therefore, even though the cross-modal features may not be necessary for simple HF screening purposes, the high-precision performance of ECGX-Net can be applied to downstream diagnosis stages or low-risk patients to ensure that fewer healthy individuals are subjected to unnecessary tests or treatments.

3.6. Interpreting the model performance with SHAP values

We used SHAP (Shapley Additive exPlanations) values [38] to assess the contribution and predictiveness of the features, uncovering the relationship between GASF-transformed ECG images and heart failure. The SHAP values can be used to evaluate the contributions of image patterns and bioimpedance regressor features in the classification process. Our model consists of four key components, starting with deep feature extraction using a pretrained model such as DenseNet121 or VGG19, followed by PCA, then bioimpedance prediction using MLP, and ending with an SVM classifier (Fig. 5A). We calculated the SHAP values to reveal the importance of the input features in the prediction process for the SVM classifier, and then the SHAP values were backpropagated to the bioimpedance regressor. Then, the loading values from PCA were used to evaluate the importance of the deep features from the DenseNet121 or VGG19 pretrained models. By backpropagation through the model, we identified the ECG image regions contributing to the prediction process.

From the trained SVM classifier, we identified the top 20 features ranked by the SHAP mean absolute values. With DenseNet121, the top features included 14 Dense121-PC features obtained directly from DenseNet121 (denoted as PC in Fig. 5B) and 6 features from the bioimpedance regressor (denoted as Bio in Fig. 5B), meaning that the bioimpedance cross-modal features played a significant role in SVM classification. With VGG19, all top features were VGG19-PC features obtained directly from VGG19 (Fig. 5B), and none of the highly contributing features were from the bioimpedance regressor, meaning that the cross-modal features affected the SVM classification in a subtle way.

Then, we utilized the loading values in PCA to quantify the contributions of the features from DenseNet121 or VGG19 and visualized the SHAP values of the top 50 features (denoted as deep features in Fig. 5C). For example, VGG19 features 2552 and 551 and DenseNet121 features 863 and 532 contributed significantly to the SHAP values. Then, we backpropagated these SHAP values to the DenseNet121 or VGG19 pretrained models to quantify the contribution of the regions of the transformed ECG images and average them over the N or HF classes. Fig. 5D and Supplementary Fig. 5 demonstrate that high absolute SHAP values in our study exhibit grid patterns. This is due to the GASF-transformed ECG images (Fig. 1D) displaying grid patterns as a result of the temporal correlation between paired times of ECG signals. In these 2D SHAP plots, the off-diagonal regions correspond to temporal correlation between paired times of ECG signals, while the on-diagonal lines correspond to time-domain contributions. Notably, we observe significantly more SHAP values in the off-diagonal regions compared to the on-diagonal SHAP values. The grid patterns in the SHAP 2D plot suggest that the temporal correlation between ECG pulses at different time points plays a more important role in heart failure classification compared to time-domain signals. Moreover, in both DenseNet121 and VGG19, the average SHAP images of each feature of DenseNet121 or VGG19 exhibit marked differences between the N and HF classes, confirming the effectiveness of the classification. Intriguingly, the SHAP values of some features (F867 and F789) of DenseNet121 are localized in specific time intervals, while most of the VGG19 SHAP values are distributed over the entire image. This suggests that DenseNet121's local feature contribution could be a more effective way to predict heart failure. In summary, we demonstrate that the pretrained models on the transformed ECG images obtained the meaningful features associated with the temporal correlation of ECG to predict heart failure.

3.7. Roles of the pretrained model and image transformation in ECGX-Net

To understand the roles of the image-based DenseNet121 or VGG19 feature extraction in our pipeline, we replaced them with other pretrained models and the trained ECGX-Net with the 50 % threshold for HF. We found that DenseNet121 (Precision: 0.7459, Recall: 0.9929, F1-score: 0.85) significantly outperformed all the other considered pretrained models (Fig. 6A, Supplementary Fig. 6; see Supplementary Information 1 for statistical testing). The second-best models were VGG16 (Precision: 0.7248, Recall: 0.8961, F1-score: 0.80) and VGG19 (Precision: 0.7424, Recall: 0.8800, F1-score: 0.81). The other models, DenseNet169, ResNet50, ResNet101, and AlexNet, could not match the performance of DenseNet121. Additionally, the AUC-PR showed the same trend (Fig. 6B; DenseNet121: 0.925, VGG19: 0.837, VGG16: 0.815, DenseNet169: 0.803, ResNet101: 0.790, ResNet50: 0.610, AlexNet: 0.553). These results suggest that DenseNet121, VGG16, and VGG19 provide better features to the downstream analysis methods, resulting in a significant boost in performance compared to the other models.

To compare the computational efficiency of the different pretrained models, we measured their training times (Supplementary Fig. 7). The time required to convert the entire GASF image dataset to the pretrained model features was similar for all the methods, with the exception of AlexNet. However, the time for PCA dimensional reduction varied due to differences in the output dimensions (VGG16/VGG19/AlexNet: 4096; ResNet50/ResNet101: 2048; DenseNet121: 1024; DenseNet169: 1664). In SVM classification using the combined features (pretrained model-PCA features + regressor features), VGG16, VGG19, and DenseNet121, which performed better with radial basis function (RBF) kernels, took

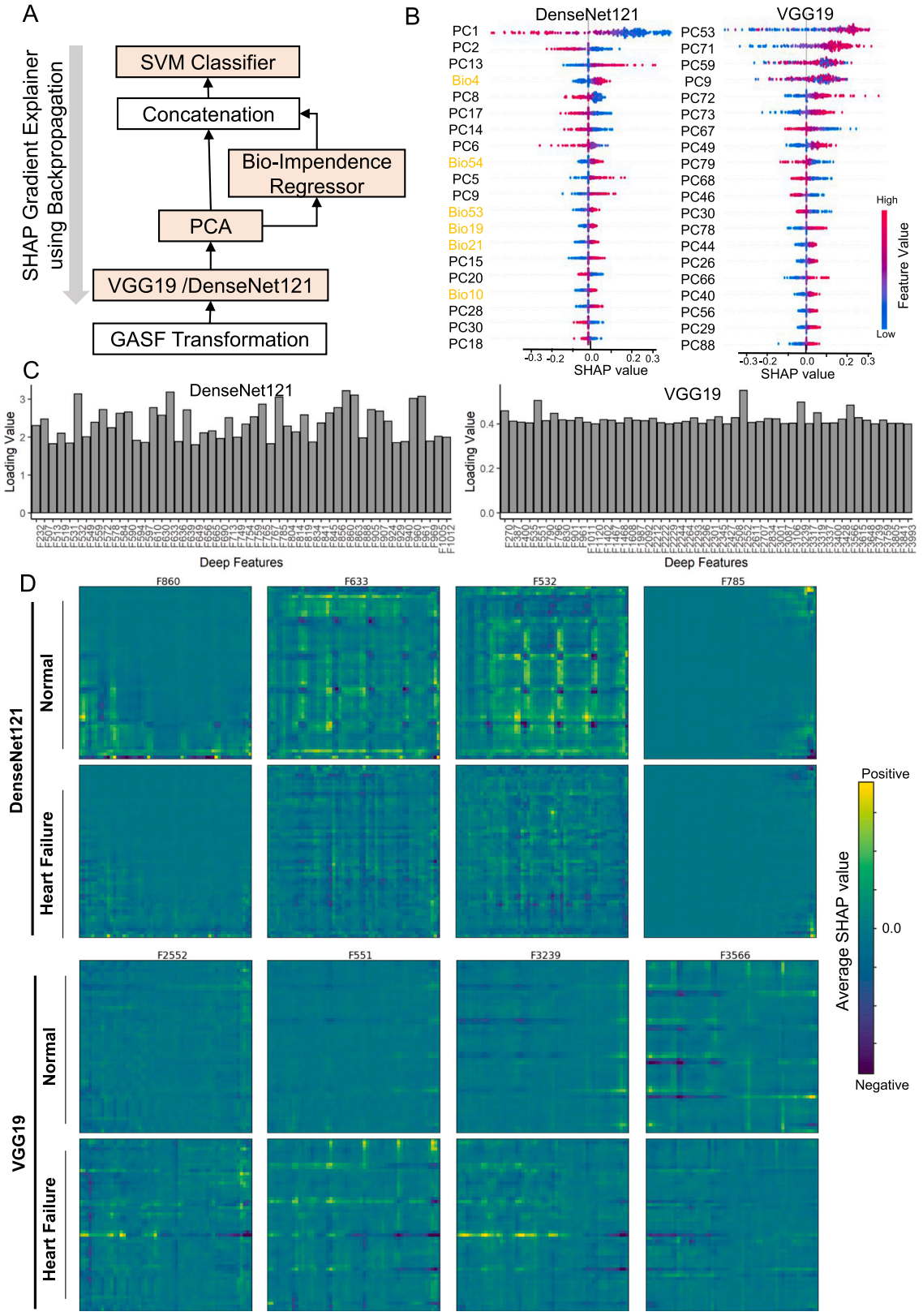


Fig. 5. Model interpretation using SHAP values. (A) Backpropagation of SHAP values through ECGX-Net. (B) SHAP values of top 20 features for the SVM classifier. (Left: DenseNet121. Right: VGG19). PC: principal component of the pretrained model features. Bio (orange): bioimpedance regressor features. (C) The top 50 deep features (pretrained model features) are based on the summation of PCA loading values from the SVM classifier and bioimpedance regressor (Left: DenseNet121. Right: VGG19). (D) Four deep features with the largest SHAP contributions on the GASF transformed ECG images in N and HF classes (Top: DenseNet121, Bottom: VGG19).

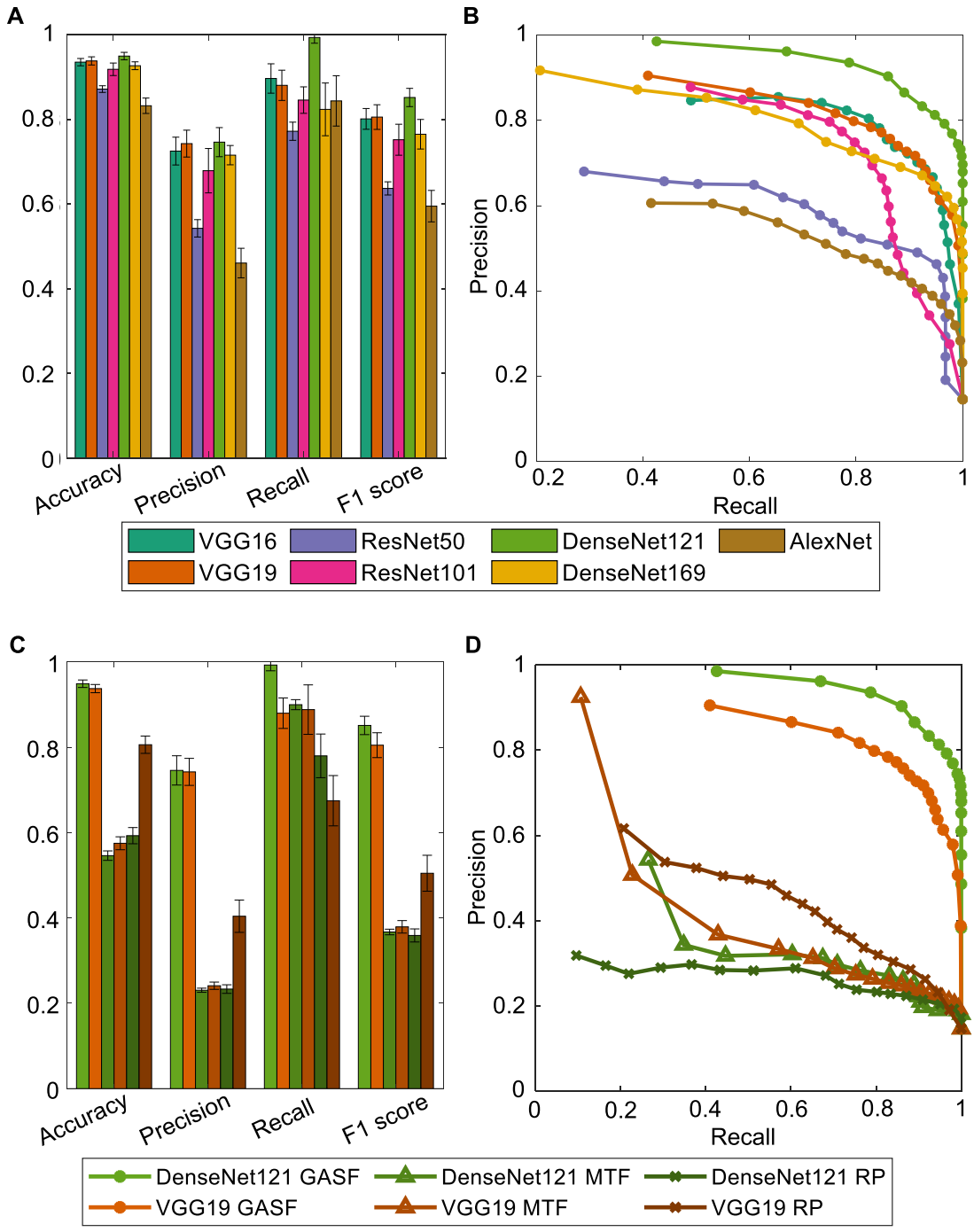


Fig. 6. Classification performance comparison among different pretrained models and image transformation methods: (A) classification performances with different pretrained models;(B) precision-recall curves of different pretrained models. (VGG16: 0.815; VGG19: 0.837; ResNet50: 0.610; ResNet101: 0.790; DenseNet121: 0.925; DenseNet169: 0.803; AlexNet: 0.553 for AUC); (C) classification performances with different image transformation methods;(D) precision-recall curves of different image transformation methods (MTF + DenseNet121: 0.417; MTF + VGG19: 0.400; RP + DenseNet121: 0.269; RP + VGG19: 0.467 for AUC).

longer than the remaining models, which showed optimal performance with a polynomial kernel. Overall, DenseNet121 had the third longest total time, and VGG19 had the longest time among the seven models.

We also tested the importance of GASF image transformation in our pipeline by replacing it with the Markov Transition Field (MTF) or Recurrence Plot (RP) (see the Methods section for details). In DenseNet121 or VGG19-based ECGX-Net, the classification performance of GASF was substantially better than that of MTF and RP (Fig. 6C–D). In particular, the precision performance of GASF was much better than that of MTF and RP, which made their F1-scores and AUC-PR significantly greater. In contrast to MTF and RP, GASF provided a more interpretable encoding of ECG time series with the temporal relationship between different time points. Because these features make important contributions to the prediction of heart failure, as demonstrated in our model interpretation, GASF plays a vital role in our ECGX-Net.

Finally, we tested a conventional 1D CNN for the prediction of heart failure. The 1D CNN model used a simple, one-layer design due to the limited data size, with a filter number of 64 and kernel size of 11. The model was trained for 128 epochs using binary cross-entropy as the loss function. As the loss curve suggests in Supplementary Fig. 8A, significant

overfitting occurred during training because of the small amount of data. The average performance of the 50 repetitions is substantially lower than that of our pipeline, as shown in Supplementary Fig. 8B. Therefore, our transfer learning approach using ImageNet-trained pre-trained models plays an essential role in our study.

4. Discussion

We demonstrated that our time-series cross-modal feature learning pipeline effectively classifies ECG data into important clinical conditions without detecting known abnormal ECG rhythms. We also implemented time series transfer learning based on transforming the ECG signals into image-based data. When it was used to screen for ADHF, our method achieved high accuracy using inputs that consist of 2D images transformed from 1D time series, followed by cross-modal feature learning with transthoracic bioimpedance. In this study, we filtered out the irregular data group and small outlier clusters due to the limited data size or the unavailability of bioimpedance measurements. With increasing data size and bioimpedance measurements, we expect that our ECGX-Net will be applicable to a variety of patient clusters.

Most deep learning studies in ECG analysis [20,21] have focused on detecting ADHF events in well-controlled clinical settings with proper multielectrode positioning, which makes it challenging to make timely predictions of ADHF. Here, we show the feasibility of predicting ADHF from a single-channel ECG recorded by patients at home, enabling timely alerts and interventions in the home or ambulatory clinic when patients experience the symptoms. From a clinical perspective, this methodology provides a potential solution and evaluation method outside of the inpatient setting using a portable health care device after discharge for an initial admission for HF. With proper, routine measurement of ECG, patient status and risk can be assessed remotely while the patient remains at home rather than having to visit hospitals in person for frequent re-examinations and treatment adjustments.

We demonstrated that ECGX-Net was particularly effective for high-precision classification, while the pretrained model features alone were sufficient for high-recall classification. Balancing precision and recall performance requires careful consideration of the costs involved in false positives and false negatives. For example, high-recall classifiers are preferred for a patient group with a high risk of heart failure to reduce patient morbidity and mortality. On the other hand, high-precision classifiers are preferred for low-risk patient groups to reduce unnecessary testing and treatment. Our pipeline allows precision and recall to be balanced to handle the unique requirements of different medical scenarios and resource limitations. The current study focuses on using our pipeline for small datasets; however, increasing the data size by recruiting more patients is necessary. The study did not include ECG data with irregular patterns due to a lack of participants (Fig. 2G). Increasing the sample size will help reduce bias and make the model more robust and sensitive in heart failure classification and prediction.

Our cross-modal feature learning method using image-based transfer learning can achieve high-performance classification. We demonstrate that ImageNet pretrained with DenseNet121 and VGG19 can efficiently extract the GASF-based features of temporal dependencies between different time points with a small amount of data. This increases the feasibility of machine learning applications in the medical field, including the analysis of electroencephalogram (EEG) data for epilepsy detection [46] and electroretinogram (ERG) data for congenital stationary night blindness diagnosis [47] and the classification/clustering of dynamic cellular phenotypes [48,49].

Supplementary data to this article can be found online at https://doi.org/10.1016/j.artmed.2023.102548.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Microsoft for providing us with Azure cloud computing resources (Microsoft Azure Research Award, K.L.), and Boston Scientific for providing us with a gift for deep learning research (K.L.). This work was supported by the NSF (Grant Number: 1522084 for Y.M. and K.L. and 1522087 for K.C.) and the NIH, United States (Grant Number: R35GM133725 for K.L.).

References

[1] Allen LA, O'Connor CM. Management of acute decompensated heart failure. Cmaj 2007;176:797–805.
[2] Mosterd A, Hoes AW. Clinical epidemiology of heart failure. Heart 2007;93:1137–46.
[3] Jackson SL, et al. National burden of heart failure events in the United States, 2006 to 2014. Circ Heart Fail 2018;11:e004873.
[4] Writing Group to Review New E. 2009 focused update: ACCF/AHA guidelines for the diagnosis and management of heart failure in adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines: developed in collaboration with the International Society for Heart and Lung Transplantation. Circulation 2009;119:1977–2016.
[5] Lala A, et al. Relief and recurrence of congestion during and after hospitalization for acute heart failure: insights from diuretic optimization strategy evaluation in acute decompensated heart failure (DOSE-AHF) and cardiorenal rescue study in acute decompensated heart failure (CARESS-HF). CirculationHeart Failure 2015;8:741–8.
[6] Ahmed A, et al. Incident heart failure hospitalization and subsequent mortality in chronic heart failure: a propensity-matched study. J Card Fail 2008;14:211–8.
[7] Solomon SD, et al. Influence of nonfatal hospitalization for heart failure on subsequent mortality in patients with chronic heart failure. Circulation 2007;116:1482–7.
[8] Baumgartner RN, Chumlea WC, Roche AF. Estimation of body composition from bioelectric impedance of body segments. Am J Clin Nutr 1989;50:221–6.
[9] Grimnes S, Martinsen OG. Sources of error in tetrapolar impedance measurements on biomaterials and other ionic conductors. J Phys D Appl Phys 2007;40:9–14.
[10] Seppa VP, Viik J, Hyttinen J. Assessment of pulmonary flow using impedance pneumography. IEEE Trans Biomed Eng 2010;57:2277–85. https://doi.org/10.1109/TBME.2010.2051668.
[11] Silver D, et al. Mastering the game of go with deep neural networks and tree search. Nature 2016;529:484–9. https://doi.org/10.1038/nature16961.
[12] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44. https://doi.org/10.1038/nature14539.
[13] LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. In: The handbook of brain theory and neural networks. 3361; 1995.
[14] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv Neural Inf Proces Syst 2012;1097–1105.
[15] Sun W, Zheng B, Qian W. Computer aided lung cancer diagnosis with deep learning algorithms. In: Medical imaging 2016: computer-aided diagnosis. 9785; 2016. 97850Z.
[16] Kirienko M, et al. Convolutional neural networks promising in lung cancer T-parameter assessment on baseline FDG-PET/CT. Contrast Media Mol Imaging 2018;2018.
[17] Hannun AY, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med 2019;25:65–9. https://doi.org/10.1038/s41591-018-0268-3.
[18] Hajeb-M S, Cascella A, Valentine M, Chon K. Deep neural network approach for continuous ECG-based automated external defibrillator shock advisory system during cardiopulmonary resuscitation. J Am Heart Assoc 2021;10:e019065.
[19] Khurshid S, et al. ECG-based deep learning and clinical risk factors to predict atrial fibrillation. Circulation 2022;145:122–33. https://doi.org/10.1161/CIRCULATIONAHA.121.057480.
[20] Shanmugam D, Blalock D, Gong JG, Guttag J. Multiple instance learning for ECG risk stratification. arXiv; 2018. preprint arXiv:1812.00475.
[21] Porumb M, Iadanza E, Massaro S, Pecchia L. A convolutional neural network approach to detect congestive heart failure. Biomed Signal Process Control 2020;55:101597.
[22] Attia ZI, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. Nat Med 2019;25:70–4. https://doi.org/10.1038/s41591-018-0240-2.
[23] Zhao, W. in AIP Conference Proceedings.1 edn 020018 (AIP Publishing LLC).
[24] Deng J, et al. Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition; 2009. p. 248–55.
[25] Kim SJ, et al. Deep transfer learning-based hologram classification for molecular diagnostics. Sci Rep 2018;8:17003. https://doi.org/10.1038/s41598-018-35274-x.
[26] Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? Adv Neural Inf Proces Syst 2014;3320–3328.
[27] Sharif Razavian A, Azizpour H, Sullivan J, Carlsson S. CNN features off-the-shelf: an astounding baseline for recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops; 2014. p. 806–13.

[28] Donahue J, et al. Decaf: a deep convolutional activation feature for generic visual recognition. Int Conf Mach Learn 2014;647–655.
[29] Oquab M, Bottou L, Laptev I, Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2014. p. 1717–24.
[30] Jang J, et al. A deep learning-based segmentation pipeline for profiling cellular morphodynamics using multiple types of live cell microscopy. Cell Reports Methods 2021;1:100105.
[31] Dovancescu S, et al. Detecting heart failure decompensation by measuring transthoracic bioimpedance in the outpatient setting: rationale and design of the SENTINEL-HF study. JMIR Res Protoc 2015;4:e121. https://doi.org/10.2196/resprot.4899.
[32] Joseph SM, Cedars AM, Ewald GA, Geltman EM, Mann DL. Acute decompensated heart failure: contemporary medical management. Tex Heart Inst J 2009;36:510–20.
[33] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv; 2014. preprint arXiv:1409.1556.
[34] Wang Z, Oates T. Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. In: Workshops at the twenty-ninth AAAI conference on artificial intelligence; 2015.
[35] Eckmann JP, Kamphorst SO, Ruelle D. Recurrence plots of dynamical systems. World Sci Ser Nonlinear Sci Ser A Monogr Treatises 1995;16:441–6.
[36] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4700–8.
[37] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–8.
[38] Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Adv Neural Inf Proces Syst 2017;30.
[39] Shapley LS. A value for n-person games. In: Contributions to the theory of games (AM-28). II; 1953. p. 307–17.
[40] Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City cardiomyopathy questionnaire: a new health status measure for heart failure. J Am Coll Cardiol 2000;35:1245–55.
[41] Li H, Pang F, Shi Y, Liu Z. Cell dynamic morphology classification using deep convolutional neural networks. Cytometry A 2018;93:628–38.
[42] Khamparia A, et al. An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning. Trans Emerg Telecommun Technol 2021;32:e3963.
[43] Alhindi TJ, Kalra S, Ng KH, Afrin A, Tizhoosh HR. In: International joint conference on neural networks (IJCNN). IEEE; 2018. p. 1–7.
[44] Habib N, Hasan MM, Reza MM, Rahman MM. Ensemble of CheXNet and VGG-19 feature extractor with random forest classifier for pediatric pneumonia detection. SN Comput Sci 2020;1:1–9.
[45] Vaidyanathan K, et al. A machine learning pipeline revealing heterogeneous responses to drug perturbations on vascular smooth muscle cell spheroid morphology and formation. Sci Rep 2021;11:23285.
[46] Shoeb AH. Application of machine learning to epileptic seizure onset detection and treatment. 2009.
[47] Utz VM, et al. Presentation of TRPM1-associated congenital stationary night blindness in children. JAMA Ophthalmol 2018;136:389–98.
[48] Wang C, et al. Deconvolution of subcellular protrusion heterogeneity and the underlying actin regulator dynamics from live cell imaging. Nat Commun 2018;9:1–17.
[49] Choi HJ, et al. Emerging machine learning approaches to phenotyping cellular motility and morphodynamics. Phys Biol 2021;18:041001.