ECGLens: Interactive Visual Exploration of Large Scale ECG Data for Arrhythmia Detection

Ke Xu
Hong Kong University of Science and Technology
kxuak@connect.ust.hk

Shunan Guo
IDV^X Lab, East China Normal University
g.shunan@gmail.com

Nan Cao
IDV^X Lab, Tongji University
nan.cao@gmail.com

David Gotz
University of North Carolina at Chapel Hill
gotz@unc.edu

Aiwen Xu
IDV^X Lab, Tongji University and NYUSH
ax266@nyu.edu

Huamin Qu
Hong Kong University of Science and Technology
huamin@cse.ust.hk

Zhenjie Yao
Beijing University of Technology
yaozhenjie@gmail.com

Yixin Chen
Washington University in St. Louis
yixin.chen@gmail.com

ABSTRACT
The Electrocardiogram (ECG) is commonly used to detect arrhythmias. Traditionally, a single ECG observation is used for diagnosis, making it difficult to detect irregular arrhythmias. Recent technology developments, however, have made it cost-effective to collect large amounts of raw ECG data over time. This promises to improve diagnosis accuracy, but the large data volume presents new challenges for cardiologists. This paper introduces ECGLens, an interactive system for arrhythmia detection and analysis using large-scale ECG data. Our system integrates an automatic heartbeat classification algorithm based on convolutional neural network, an outlier detection algorithm, and a set of rich interaction techniques. We also introduce A-glyph, a novel glyph designed to improve the readability and comparison of ECG signals. We report results from a comprehensive user study showing that A-glyph improves the efficiency in arrhythmia detection, and demonstrate the effectiveness of ECGLens in arrhythmia detection through two expert interviews.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces; Graphical user interfaces (GUI)

Author Keywords
Visualization; Health - Clinical; Interaction Design; Visual Design; Artifact or System

INTRODUCTION
The Electrocardiogram (ECG) is a graph of voltage versus time that shows the electrical activity of the heart over a period of time. ECGs is a standard diagnostic tool used by cardiologists to identify both the presence and type of arrhythmia in a patient [11] through a highly observational analysis on the ECG records [20]. The manual ECG analysis process performed by a cardiologist can be time-consuming when facing a large amount of data. Consider that an ECG monitor (e.g., a Holter monitor) will typically collect 60-100 beats per minute for an adult with a normal resting heart rate [5]. This means over 86,000 heartbeats can be recorded for each patient for each 24 hour period. At this scale, it takes great effort to quickly scan the captured data for irregularities, not to mention the careful inspection of wavelength and amplitude that a cardiologist must perform to make a confident diagnostic decision. Thus, very often cardiologists will focus on relatively short-term ECG measurements captured over just a few minutes. However, anomalous heartbeats can be highly irregular and intermittent, meaning many arrhythmia can be missed. To support longer-term observations which can help uncover patterns related to complex heart problems, cardiologists need a new tool to help them conduct large-scale ECG data analysis and exploration.

Recent studies have addressed the problem of analyzing large-scale ECG data using various automatic analysis algorithms. These algorithms are generally based on signal processing techniques for feature extraction such as time-domain analysis [39, 40], wavelet transform [4, 38], and machine learning methods for heartbeat classification such as Support Vector Machines [29], Hidden Markov Models [3] and Deep Neural Networks [33, 46]. Such methods can be very helpful, and our work builds upon a recently published algorithm for automated heartbeat classification model based on Convolutional Neural Networks (CNN) [37]. However, even the best fully automated classification algorithms produce significant errors. As importantly, automated classification methods do not address cardiologists' need to explore the classified ECG data.

For these reasons, it is necessary to develop interactive tools which make ECG data and the results of ECG analytic algorithms more accessible to cardiologists. This has motivated the development of a variety of desktop applications [2, 13, 31] and web-based tools [6, 26, 30]. However, these focus primarily on representing 10-second ECG records [7]. Meanwhile, the

CHI 2018, April 21–26, 2018, Montréal, QC, Canada.

http://dx.doi.org/10.1145/3173574.3174237

few methods [1,35] designed for large-scale ECG data present either highly-summarized views or focused only on certain features which prevent detailed analysis.

This paper introduces ECGLens, a visual analytics system developed to help cardiologists efficiently and flexibly explore large-scale ECG data. Major research contributions include:

- **System Design.** We introduce a comprehensive visual analytics system for the inspection and exploration of large-scale ECG data. We identify key requirements through a pilot study with target users, and propose a system design motivated by those requirements.
- **Interactive Analysis.** We propose an interactive framework which supports large-scale exploration of ECG data through three major steps: initial inspection, anomaly analysis, and diagnosis. We utilize a classification model based on CNN which classifies heartbeats within a user-selected ECG interval into four clinical categories. An outlier detection algorithm then helps identify heartbeats that are most likely to be misclassified or correlative with certain diseases. Interactive visualization tools using a novel glyph design then facilitate visual heartbeat comparison and diagnosis.
- **Evaluation.** We report results from a user study comparing users' abilities to detect arrhythmia using our glyph design compared to other representations. The results show that A-glyph outperforms baseline glyphs in arrhythmia detection with a comparable performance in heartbeat classification. We also assessed the effectiveness of ECGLens through interviews with two domain experts who used our system on two real-world datasets.

RELATED WORK

In this section, we provide an overview of prior research that is most related to our work. This includes: (1) analysis methods for ECG data, and (2) techniques for ECG data visualization and interactive exploration.

ECG Data Analysis

The automatic analysis of ECG data is essential for arrhythmia diagnosis. This is due to the widespread use of portable ECG devices, such as the Holter monitor, which produce a very large amount of data to be analyzed. Generally, the automated analysis of ECG data is composed of two crucial steps: feature extraction, and beat classification [36]. The results of these steps are then used by arrhythmia identification algorithms [11].

ECG feature extraction is of chief importance in precise heartbeat classification and diagnosis of cardiac diseases, especially in the examination of long-term recordings [19]. Various extraction algorithms and signal transformation techniques for ECG data have been proposed, and these approaches can be broadly categorized as (1) time-domain, (2) frequency-domain, or (3) time-frequency domain techniques. For example, one relatively efficient time-domain arrhythmia classification algorithm [40] utilizes RR-intervals extracted through a sliding window. Another approach called SAX (symbolic aggregate approximation) classifies heartbeats based on a symbolic representation of ECG data [23]. In a frequency-domain approach, Gothwal et al. [16] introduced an arrhythmia detection algorithm that extracting frequency-domain features using the Fast Fourier Transform. As a hybrid approach, Zhao et al. [45] designed a heart rhythm recognition algorithm using Wavelet Transforms which preserves both time-domain and frequency-domain features.

Given a set of features, a variety of machine learning and data mining algorithms have been employed in the task of heartbeat classification. Among these methods, four algorithms are most popular [24] including: Support Vector Machines [32,43], Artificial Neural Network [18,33], Linear Discriminant Analysis [11,44], and Reservoir Computing With Logistic Regression [14,15]. A recent study by Rajpurkar et al. [37], representing the state-of-art method for arrhythmia detection, leverages Convolutional Neural Networks (CNN). The authors of this study reported that the method is able to classify 15 types of heart rhythms using a single-lead ECG signal more effectively than human cardiologists. In this paper, we leverage this related work by incorporating an automated CNN-based heartbeat classifier within our system.

ECG Visualization

ECG data is most commonly depicted as a temporal chart of a heart's electrical activity over time as measured from various leads. The resulting waves depict heart beat and rhythm for a patient over a given period of time. It is these charts that provide preliminary evidence for cardiologists in making diagnosis and treatment decisions. However, the morphology changes in ECG waves that are of interest to clinicians are often too subtle to be detected using these traditional representations [17]. This has motivated various studies that have focused on developing alternative graphical representations. For example, Madias et al. [25] introduced a 13th multi-use ECG lead to help cardiologists further acquire more information of the patient's cardiac state. A mirror image 24-lead ECG is proposed in [21] to increase the sensitivity in detecting acute myocardial infarction (AMI). Chiang et al. [9] integrated waves of multiple leads into two images, which represent the precordial leads and limb leads respectively, providing integrated views over a large number of ECG signals. As these examples highlight, much of the focus in this area of research has aimed to improving standard representations of multiple leads. However, they have not addressed the challenge of facilitating the observation single-lead ECG data *over a long periods of time*, which is the challenge posed by the widespread use of portable ECG devices. An example of one exception is from Alfredo et al. [17], who proposed a single-lead ECG visualization method which remaps the ECG signal to a spatial curve with the aim of detecting important parameters such as ECG derivatives, maximal and minimal values of waves, and wave slope. However, the representation is dramatically different from the standard representation for ECG signal, resulting in reduction in readability.

Other research efforts have explored interactive methods which go beyond the limitations of static representations to facilitate exploration of larger amounts of ECG datasets. Desktop applications, such as those developed by ECGSoft [13], AMPS [2], and OFFIS [31] have been used widely in recent years. Similarly, an interactive web application named WebECG [26] was proposed for remote patient monitoring, and a similar web-based tool [30] for visualizing ECG data was developed on the basis of open source technologies. Whilst the majority of these techniques look to help users retrieve historical data of 10 seconds from database by incorporating simple interactions, Rajendra et al. [1] summarized the occurrence of 13 types of arrhythmia in single-lead ECG recordings for 24 hours by integrating an arrhythmia detection algorithm. The analysis result is visually displayed using a circle of 24 sectors representing the record of 24 hours. Each sector is color coded to mark the occurrence of different types of arrhythmia. This system depicts ECG data at a far larger scale than previous methods, making it relatively easy to localize moments of arrhythmia over a long time period. However, its utility is highly dependent on the accuracy of the underlying algorithm, and this information is not made transparent.

In contrast to these existing visualization techniques, our system enables users to explore large amounts of ECG data both comprehensively and interactively. The system incorporates an automatic classification algorithm based on CNN, as well as a local outlier detection algorithm to help users inspect and interpret the analysis results. Our approach also highlights heartbeats with high outlier scores, suggesting that they might indicate certain diseases or be mislabeled. ECG records are visually displayed in various forms and in multiple linked views, offering a set of rich interaction techniques designed for effective large-scale ECG data inspection and exploration.

SYSTEM REQUIREMENTS AND DESIGN

To inform the design and development of our proposed system, we conducted a preliminary design study with two domain experts. This study provided key insights into user needs and the process of diagnosing cardiac diseases with ECG data. A list of requirements was derived from the users' comments and used to drive the development of our prototype system.

Pilot Study

We conducted a pilot study during which we asked two medical experts to examine the initial classifications assigned to heartbeats by a preliminary heartbeat analysis algorithm. We trained an offline CNN model to classify heartbeats into four categories as recommended by the AAMI [12] (Fig. 1): (1) Sinus beat, (2) Ventricular Ectopic beat (VEB), (3) Supraventricular Ectopic beat (SVEB), and (4) a fusion of Sinus and VEB. We integrated the well-trained model for this classification task into an initial system with a basic set of features to display the classification result for a heartbeat along with the corresponding raw ECG signals. Participants were asked to inspect each heartbeat and evaluate the correctness of the classifiers labels. This study helped us find (1) the key parameters that were cardiologists relied upon most to assess the classification results, and (2) the limitations of the pilot study system and opportunities for innovation.

This study confirmed the use of four time-domain components of a heartbeat during manual classification of an ECG heartbeat signal: the P wave, the QRS complex, the ST segment, and the T wave [42]. To efficiently determine if a waveform is anomalous, clinicians must gain expertise through practice in estimating the wavelength and amplitude of each wave. This is traditionally facilitated by the inclusion of background grid in the ECG rendering. Adding our automated classification algorithm to the traditional view allowed clinicians to determine beat categories and detect anomalies more efficiently. However, the automated algorithm also did not perform perfectly.

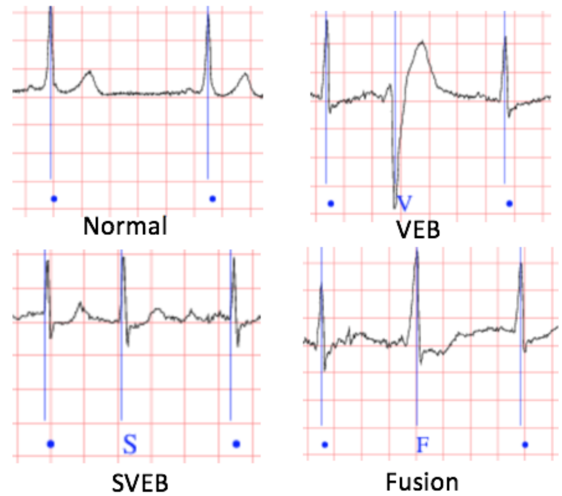

Figure 1. Four different types of heartbeats as recommended by the AAMI [12].

The expert users detected several mislabelled heartbeats during the pilot study and suggested that the system should allow corrections to the recommended labels. In addition, they suggested that the system could guide cardiologists through the analysis process, allowing them to avoid manual inspection of every single heartbeat.

Design Requirements

Based on the feedback gathered during the above pilot study, we identified the following key design requirements.

- **R1 Support for large-scale data.** The system should be designed to allow effective analysis and exploration of large-scale ECG data, enabling users to examine ECG results captured over several hours.
- **R2 Precise heartbeat classification.** Heartbeat classification algorithms should be integrated into the system to support the efficient arrhythmia detection of a massive amount of data. The system should provide methods that help users quickly identify anomalous heartbeats which may require clinical investigation and eventual diagnosis, as well as potential classification errors.
- **R3 Enhanced beat recognition and comparison.** The system should highlight the key parameters of each heartbeat and facilitate the comparison of heartbeats to help cardiologists more rapidly read the ECG waves and find patterns which may correspond to disease.
- **R4 Interactive data exploration.** To support efficient exploration and comparison across large amounts of ECG data, it is necessary to incorporate flexible interactions that help users quickly navigate a very large number of heartbeats and associated features.

Figure 2. The ECGLens system contains nine interactively coordinated views, including (1) a view illustrating the entire series of raw ECG waves in a circular timeline, (2) a view showing the distribution of heartbeats in a feature space, (3) conventional & (4) glyph representations of a segment of the ECG waves selected from view (1), (5) an inspection view depicting the details of the selected heartbeats, (6) a cluster view showing a list of clustered anomalous heartbeats with the details of each cluster shown in (7), (8) a view showing the statistics of the heartbeats under investigation, and (9) a focal view showing the details of one single selected heartbeat.

INTERACTIVE ARRHYTHMIA EXPLORATION

We designed ECGLens, an interactive arrhythmia exploration framework of large-scale ECG data (**R1, R4**), to meet the above requirements. As illustrated in Fig. 3, ECGLens consists of three basic steps which allow users to: (1) inspect and select ECG segments with anomalous QTc values; (2) detect heartbeat irregularities based on automatic classification and local outlier detection algorithms (**R2**); and (3) examine detailed heartbeat information to arrive at an accurate diagnosis (**R3**). This section provides a detailed overview of the system's functionality, as well as the corresponding visual and interactive designs which support these three steps.

Initial Inspection

We employ a circular timeline (Fig. 2(1)) introduced in [34] to display the QTc value (a well-known risk indicator for ventricular arrhythmia) of a sequence of heartbeats. The background color illustrates the risky range: green indicates healthy values while the red indicates the dangerous QTc value range. Users can select a segment with irregular QTc values to retrieve more details including a 2D MDS projection of the selected heartbeats (Fig. 2(2)) and a detailed temporal view of the corresponding ECG waves (Fig. 2(3,4)).

Anomaly Detection

The anomaly analysis procedure used within ECGLens consists of two parts: beat classification (Fig. 3(**2a**)) and outlier detection (Fig. 3(**2b, 2c**)).

Heartbeat Classification. Heartbeat classification is an important step towards arrhythmia detection [11]. However, it is a time-consuming task when facing a large amount of ECG data. ECGLens implements an automatic time-series classification model based on Multi-Scale Convolutional Neural Network [10]. The model is trained offline using a widely-used public dataset, MIT-BIH [28], which contains a large number of validated cardiologist annotations.

The model contains three stages. (1) First is the transformation stage, which includes three independent branches: (a) an identity mapping transformation that preserves the original signal; (b) a spectral transformation that generates a multi-scale branch in the time domain; and (c) a down-sampling transformation that reduces noise and enhances the signal's larger-scale features. (2) Second is the connecting stage, in which the transformation results are fed into a local convolution in which 3 convolutional layers are followed by ReLU and maxpooling layers. (3) Third is the final full convolution stage, which concatenates the data and sends it through 2 convolutional layers (each followed by a ReLU and maxpooling layer), a fully connected layer, and finally a softmax layer. The final output of the model is a probability distribution over the four possible labels (i.e., Sinus, VEB, SVEB and Fusion), and a given heartbeat is classified using the label with the highest probability (for more details, see the supplemental material[1]).

[1] http://idvx.lab.tongji.edu.cn/suppl/chi2018_ecglens.pdf

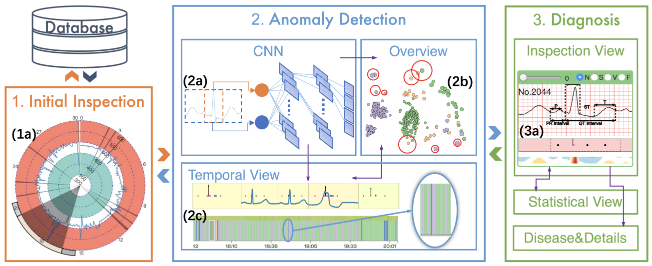

Figure 3. The process of interactive arrhythmia exploration

Outlier Detection. The classification model shows promising results and reaches a classification accuracy of 91.59% on average. However, nearly 10% of the heartbeats remain incorrectly labelled. An outlier detection algorithm is used to detect heartbeats that are most likely to be mislabeled, or which might represent a specific medical anomaly. The system uses local outlier factor (LOF) [8] to find isolated data points, and a rare category detection algorithm [22] to detect other similarly anomalous instances. The heartbeats identified by the algorithm will be highlighted for a further inspection.

Interactive Exploration

The system integrates multiple views to visually display the analysis results from various perspectives. Heartbeats of each category are represented with a specific color, with the same colormap used consistently across different views. The spatial distribution of different types of heartbeats is revealed through the overview (Fig. 2(2)). Meanwhile, a Bar View (Fig. 2(3a)) reveals the temporal distribution of heartbeat categories. These views use the output of the rare category detection algorithm to highlight anomalous heartbeats for further inspection. In the overview, users can investigate these rare categories by either: inspecting individual beats in sequential order of their LOF score, or inspecting anomalies category-by-category. In addition, users can inspect the distribution of anomalous heartbeats over time. Users can also manually brush a group of data points in the overview or a time interval in the temporal view to show more details about the corresponding heartbeats through other linked panels.

Diagnosis

To help clinicians make more accurate diagnostic decisions, we designed several additional views that link to selections made in the previously described panels. After a selection is made, detailed information about the selected heartbeats are displayed in the Inspection View (Fig. 2(5)). In this view, users can sort the order in which heartbeats are displayed by either time or LOF score. The color of each inspection box represents the category of the heartbeat as determined by the classification algorithm. If the user determines that a heartbeat has been incorrectly labeled based on manual inspection, he/she can correct the label from this interface. When a correction is made, all linked views are simultaneously updated to reflect the updated label. Each inspection box also incorporates a slider which allows users to underlay neighboring heartbeats within the visualization to allow contextualized comparisons. If a user identifies an abnormal waveform, he/she can temporarily save an average of the currently selected heartbeats in the Inspection view to Cluster View (Fig. 2(6)). In addition, a disease label can be added to allow fast retrieval at a later time via the Record View (Fig. 2(7)).

To improve the readability of ECG signals and facilitate the comparison of neighboring heartbeats during the diagnosis process, we incorporate a novel glyph, A-glyph, under each heartbeat in the Inspection View (Fig. 2(5b)). The glyph represents a heartbeat using the four components we found clinicians using in our pilot study: the P wave, QRS complex, ST segment and T wave. Each component is encoded using an independent visual element to depict its corresponding wavelength and amplitude, as illustrated in Fig. 4(e). The design uses solitary circles to represent wavelengths that fall within the expected range, and rectangles for wavelengths that are outside of the normal range. The rectangles are color-coded by wavelength, with blue representing exceptionally low values and red representing exceptionally high values. A rectangle's length indicates the magnitude of the deviation from the normal range. Rectangles are connected via vertical lines to offset circles, with the length of the line representing wave amplitude. Lines extending upwards represent higher-than-normal amplitudes, while lines extending downward show lower-than-normal amplitudes. The length of a line shows the amplitude's magnitude of deviation from normal. The absence of a vertical line suggests that the amplitude is within a normal range. Finally, the gylph's background color varies linearly from green-to-yellow-to-red (as shown in the legend), encoding the average feature deviation from normal range for the corresponding heartbeat. A series of A-glyphs are provided in the Temporal view (Fig. 2(4)), and each glyph can be clicked to reveal the corresponding waveform.

The system also incorporates a Statistics View (Fig. 2(8)) to summarize ECG data for the selected heartbeats using a scatterplot and histogram. The data points are aggregated via a parameter chosen by the user (see the parameter selection box in the figure), such as heart rate or various wave component features (e.g., characteristics of the P wave, which relate to atrial depolarization). To obtain the detailed information about a specific heartbeat, users can click on the visual elements corresponding to the heartbeat in any view. In response, the system will show a comprehensive view of the heartbeat in the Focal View (Fig. 2(9)). In additional, all views have linked highlighting, making it easy to spot visual representations for the same heartbeat across different views.

EVALUATION

This section reports results from evaluations of both the A-glyph design and the overall ECGLens system.

Evaluation of the Glyph Design

A controlled user study was conducted to compare the effectiveness of alternative glyph designs.

Baselines. The study compared our new A-glyph design against three alternative ECG glyphs: (a) the conventional ECG waveform (G_c), (b) the A-glyph (G_a) and (c) the heatmap-based glyph (G_h) which uses background color to indicate wavelength and circle radius to indicate amplitude. A fourth design based on horizon charts (Fig. 4(d)) was also considered

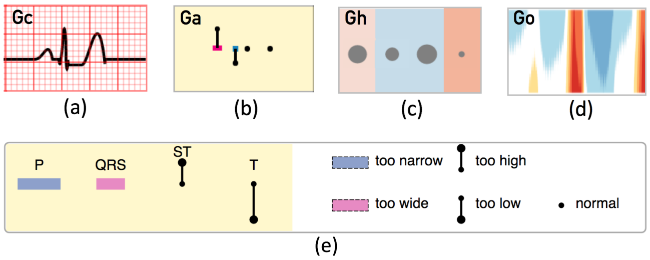

Figure 4. ECG visualization designs: (a) conventional ECG (G_c), (b) A-glyph (G_a), (c) heatmap-based glyph (G_h) and (d) horizon chart (G_o). The encoding principles of A-glyph are illustrated in (e).

	18	Users
×	3	Glyph designs
×	2	Successive Tasks (**T1** & **T2**)
×	2	Data sizes (small(2), large(35))
×	4	Number of Abnormal components (1, 2, 3, 4)
×	3	Anomaly Types (wavelength, amplitude, both)
×	3	Repetitions
	7776	Trials

Table 1. Testing Conditions.

but omitted from the study as our users found it is too confuse to interpret.

Tasks. To evaluate how well each glyph design helped users identify arrhythmia and classify heartbeats, we designed two specific tasks for our study. The tasks approximate how clinicians use ECG data to make a diagnosis. First, cardiologists were asked to read a set of ECG records which contained mostly normal heartbeats, but also a small number of potentially arrhythmic heartbeats. Whenever a cardiologist observed an arrhythmia, they were asked to identify the type of the unusual heartbeat. This process reflects an approximation of two actual medical scenarios in arrhythmia detection: (1) comparing adjacent heartbeats, and (2) inspecting a standard length ECG examination result, which leads to two formal tasks:

T1 Detect arrhythmic heartbeats from a sequence of ECG records.

T2 Classify the heartbeats detected in **T1** into one of four categories (i.e., Sinus, SVEB, VEB and Fusion) [12].

The primary variable tested for both tasks was the choice of glyph. However, the study also tested for three other factors (summarized in Table 1): (a) the number of heartbeats shown to users; (b) the number of abnormal components in a given heartbeat, including P wave, QRS complex, ST segment and T wave; and (c) the type of abnormality (i.e., deviation for wavelength, amplitude, or both) for the aforementioned four components. We determine the range for these other factors based on expert interviews and the results of our pilot study. The numbers of heartbeats displayed to users in each trial of the experiment was either 2 (small) or 35 (large). The numbers of abnormal components are in the range 1-4, reflecting that a single arrhythmic heartbeat can be abnormal in any combination of one or more of its four characteristic components (P, QRS, ST, T). Components were chosen to be abnormal in the either amplitude, wavelength, or both (as is typical in abnormal ECGs [27]).

A random testing dataset containing all of the above conditions was generated. The generated irregular heartbeats were shown to three cardiologists independently for labeling, and all discrepancies were reviewed and resolved by consensus. The determined labels served as the ground truth for our second task. In the study, $N-1$ normal heartbeats and 1 abnormal heartbeat were randomly selected for each testing trial.

Hypotheses. Based on the tasks and conditions outlined above, we designed a study to test four hypotheses:

H1 A-glyph costs less time than other glyphs (conventional ECG and heatmap-based glyph) in detecting arrhythmia.

H2 A-glyph has higher accuracy than other glyphs in detecting arrhythmia.

H3 A-glyph costs less time than other glyphs in determining the heartbeat type.

H4 A-glyph has higher accuracy than other glyphs in determining the heartbeat type.

These hypotheses were motivated by the design principle of maximizing the data-ink ratio [41] to facilitate a clean data representation, which is helpful for data comparison and interpretation of different heartbeats.

Task Performance Measures. Performance is measured using both task completion time and accuracy. For **T1**, users are required to select one single target to ensure a reasonable completion time. The completion time for each task was automatically recorded by our study system.

Participants and Apparatus. A total of 18 volunteers (16 females, 2 males) were recruited to participate in the study. All participants are board-certified cardiologists, or medical students training to be cardiologists with ages ranging from 23 to 40 (M = 25.56, SD = 1.35). The experiment was conducted in a 970×490 pixel window with the white background on a laptop computer. Glyphs were aligned at the center of the window in the experiment, with a cell size of 100×60 pixels.

Procedure. We first introduced the study, explained the A-glyph and heatmap-based glyph designs, and showed users how to perform tasks with our user study system. Next, users were asked to practice based on a small sample dataset. Users were encouraged to ask questions during the practice session. After these preparations, a within-subject study was performed based on a single experimental dataset for each user (the order of the anomalies are randomized to avoid learning effects) to ensure a fair comparison. The testing order of different designs is also counterbalanced. The task completion time and accuracy were automatically recorded by the system. At the end, the users completed a post-study questionnaire. In total, each study session lasted approximately 1.5-2 hours.

Results. This section presents both quantitative and qualitative results from the above study. We first evaluate the effect of three variables (data size, the number of abnormal components, and the type of abnormality) on overall task performance. We then compare completion time and accuracy of the three glyph designs (G_c, G_a, G_h). Finally, we present the results from the post-study questionnaire.

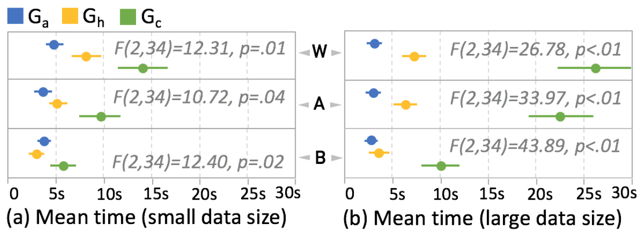

Figure 5. The mean T1 time for (a) small and (b) large data size for abnormal wavelength(W), amplitude(A), or both(B). Error bars show 95% confidence intervals (CIs).

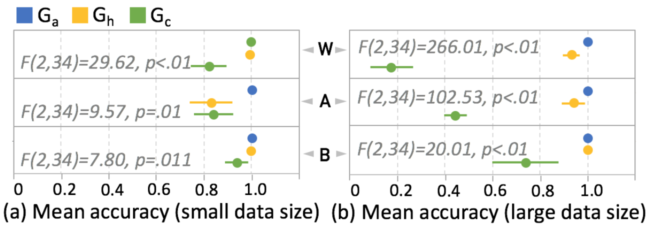

Figure 6. Accuracy of T1 for (a) small and (b) large data scales for abnormal wavelength (W), amplitude (A), or both (B). Error bars indicates 95% confidence intervals.

<u>Validation of Variables.</u> To examine the effect of data size, number of abnormal components, and type of abnormalities, we analyzed users' performance of the two tasks with each glyph under different conditions. We employed a Paired-Samples T Test to compare the difference in mean accuracy and task completion time between small and large data size for each glyph, and Repeated Measures ANOVA (RM-ANOVA) was used for comparison among different levels of other variables (number of abnormal components, type of abnormality).

The analysis results show that: (1) The mean time of **T1** was significantly affected by the data size across all glyphs except the heatmap-based glyph ($G_a : t(17) = 3.133, p < .01$; $G_c : t(17) = 6.117, p < .01$; $G_h : t(17) = -.498, p = .625$). **T1** mean time was also sensitive to the change of type of abnormality for all glyph designs ($G_c : F(2,34) = 20.04, p < .01$; $G_a : F(2,34) = 3.54, p = .04$; $G_h : F(2,34) = 26.28, p < .01$). (2) In terms of mean accuracy of **T1**, the A-glyph and

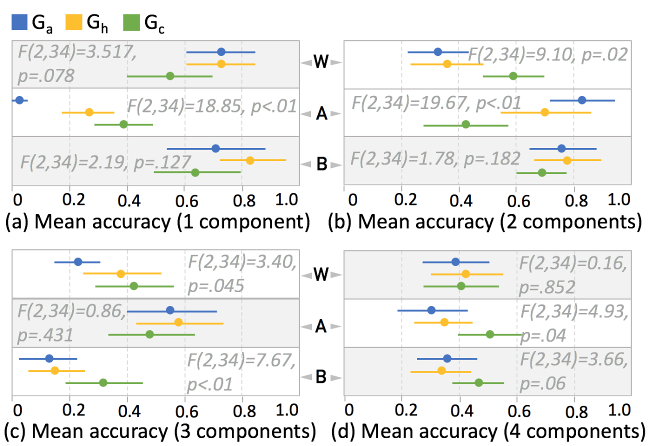

Figure 7. The T2 accuracy of (a) one, (b) two, (c) three and (4) four abnormal components, with either G_c, G_a or G_h glyphs, when abnormal parameter number was wavelength (W) or amplitude (A) or both (B). Error bars are 95% CIs.

Heatmap-based designs showed no significant difference between two data sizes, while the conventional ECG suffered a significant drop in performance when the size of data became large ($G_c : t(17) = 9.381, p < .01$). In addition, the type of abnormality had a large influence on accuracy when using the baseline glyphs ($G_c : F(2,34) = 69.39, p < .01$; $G_h : F(2,34) = 15.62, p < .01$). The A-glyph, meanwhile, showed 100% accuracy across all abnormality types.

We note that in task **T1**, we did not take the number of abnormal components into consideration. This is because during arrhythmia detection, cardiologists are focused on detecting any one anomaly to confirm an abnormal heartbeat. In task **T2**, meanwhile, we did not take data size into consideration because users were only asked to classify the individual heartbeats they selected as abnormal during **T1**.

The results also show that in terms **T2** completion time, all three glyphs performed well (with relatively small differences between them) in response to both the type of abnormality and the number of abnormal components. However, the difference in mean accuracy of **T2** was statistically significantly for different numbers of abnormal components for all glyphs ($G_c : F(3,51) = 4.56, p < .01$; $G_a : F(3,51) = 18.09, p < .01$, $G_h : F(3,51) = 21.95, p < .01$).

<u>Comparison of Glyphs.</u> We conducted a RM-ANOVA analysis to compare the mean task completion time and accuracy of different tasks across three glyph designs. We also analyzed the pairwise comparisons via Bonferroni correction. The normality of data was tested via the Shapiro-Wilk test and unsatisfied data were transformed by Normal Inverse Cumulative Distribution Function. The degrees of freedom was tested by Mauchly's test and corrected by using Greenhouse-Geisser estimate of sphericity when the assumption was violated.

The results of **T1** for each glyph are summarized in Fig. 5-6. The results are reported by data size (2, 35) and type of abnormality (wavelength, amplitude, or both).

<u>*Completion time of **T1**.*</u> When data size was either small or large (Fig. 5), significant differences were observed among the three glyphs in all abnormal parameter types. Moreover, post-hoc analysis showed that the A-glyph had a lower mean task completion time than the other two glyphs (except in the configuration of small data size and where both amplitude and wavelength were abnormal, in which the G_h is the fastest). This result supports **H1**.

<u>*Accuracy of **T1**.*</u> As shown in Fig. 6, in both small and large data size conditions, the accuracy was significantly different among different glyphs when the abnormal parameter was wavelength or/and amplitude. Post-hoc analysis showed that the A-glyph was significantly more accurate than the other two glyphs in most conditions when comparing mean accuracy. This result supports **H2**.

<u>*Completion time of **T2**.*</u> When focusing on the glyph design as the only variable, post-hoc tests using Bonferroni correction revealed that the A-glyph and heatmap-based glyphs were not statistically significant different in completion time ($T(G_a) = 7.43 \pm 0.70$s vs $T(G_h) = 6.14 \pm 0.45$s, $p = 0.168$). Both the A-

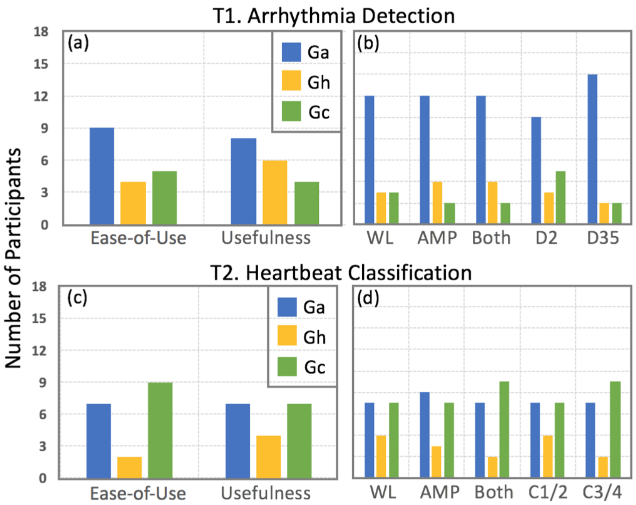

Figure 8. Users' ratings of different glyphs with respect to their usability in (a) T1 and (c) T2, and their efficiency under different conditions in (b) T1 and (d) T2. A user was asked to rate one glyph at the same time. In the figure, y-axis indicates the number of ratings, WL means Wavelength, AMP means Amplitude, D2 & D35 are different data sizes and C1/2/3/4 means the number of abnormal components.

glyph and heatmap-based designs were, however, faster than the conventional ECG ($T(G_c) = 9.33 \pm 0.85$s). This result rejects **H3**.

*Accuracy of **T2**.* As shown in Fig. 7, half of the results showed a significant difference among the three glyphs. Those showing a large difference were associated with amplitude abnormalties as well as those with two abnormal components. In some of these cases, the A-glyph was significantly better than other two glyphs, while in others the conventional ECG had better performance. This result rejects **H4**.

Overall, A-glyph was more effective than the two other glyphs in arrhythmia detection for both task completion times and accuracy rates, especially when the data size was large. In terms of heartbeat classification, however, the three glyphs have no significant difference in completion time, and the conventional ECG slightly outperformed the other two glyphs in accuracy.

Questionnaire. The post-study questionnaire contained questions designed to capture qualitative feedback about the three glyph designs. The first seven questions were focused on **T1**. Q1-Q2 in the survey asked users to choose which glyph type was most useful and easy-to-use for abnormal heartbeat detection. Q3-Q7 asked users to choose the glyph type they felt most effective for anomaly detection under various conditions (large vs small datasets, and three parameter conditions, i.e., wavelength or/and amplitude). The results are shown in Fig. 8(a&b). Overall, the A-glyph was most popular in detecting abnormal heartbeats, especially when the data size was large. The conventional ECG was least popular in this task, which aligns with the quantitative findings in this study.

The other nine questions targeted **T2**. Q8-Q9 focused on the utility and ease-of-use of classifying heartbeats. Q10-Q16 asked users to choose the glyph type they found most effective for heartbeat classification under various conditions (three abnormality types, and different numbers of abnormal components). The results are summarized in Fig. 8(c&d), which shows that the conventional ECG design was thought to be slightly more efficient in classifying heartbeat. The result is consistent with our quantitative findings which found that users had a higher accuracy in classifying heartbeat types using the conventional design.

The final three questions were free response questions asking for feedback as to the advantages and disadvantages of the A-glyph design. The most valuable feedback from these questions was as follows: (1) A-glyph was very intuitive and easy-to-understand; and (2) it could not represent the full detailed description of a complicated heartbeat.

Evaluation via Expert Interviews

To compliment the controlled study, we evaluated the usefulness and usability of our system through in-depth expert interviews with two domain experts. We first interviewed one expert and collected a variety of feedback on potential improvements to the system's interactions. After making refinements accordingly, we interviewed a second expert with our final system. This section describes the procedure for both interviews, as well as three major themes identified through an analysis of the interview transcripts.

Procedure. We conducted both interviews in the form of a brief case study with a real-world dataset from MIT-BIH. In the case study, experts were asked to diagnose any arrhythmia diseases or detect anomalous heartbeat classification results by navigating through a long-term ECG trace. We began each interview with a brief introduction, during which we clarified the goals of our study, explained the glyph designs, and provided a brief tutorial regarding the use of our system. Then, the experts were allowed to use the system and explore the data on their own. After the experts had made their diagnostic decisions, we conducted a semi-structured interview which incorporated several questions about glyph designs, overall usefulness, ease of use, general pros and cons of the prototype system and insights obtained from using the system. Each interview lasted approximately 1.5 hours, during which notes were taken and the entire procedure was recorded.

The first expert was a female graduate student (**E1**) studying cardiology with one year of clinical experience. We used modified limb lead II (MLII) data from Record 213 (male, age 61) in the MIT-BIH Arrhythmia Database. The record is slightly over 30 minutes long and contains 3251 beats in total. The second interview was conducted with a female board-certified cardiologist (**E2**) using our improved version of system. We used MLII data from Record 208 (female, age 23) in the MIT-BIH Arrhythmia Database, which contains 2955 heartbeats recorded over approximately 30 minutes.

Theme 1: Trust in Automation. A major theme observed during the expert interviews was a transition in attitude towards the CNN model. The anomaly analysis capabilities of the system are based upon the classification results produced by the CNN model. At the beginning of the interviews, the experts preferred to manually investigate the original ECG signal using the Temporal View (Fig. 2(3,4)) even though the

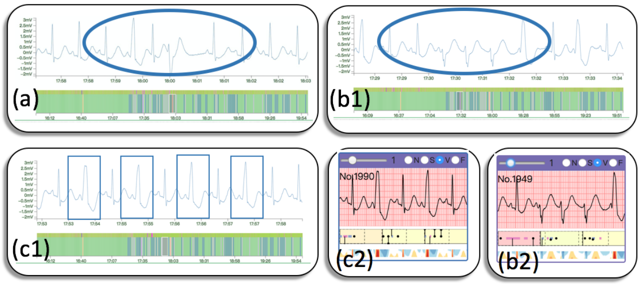

Figure 9. Diagnosis results of Interview I.

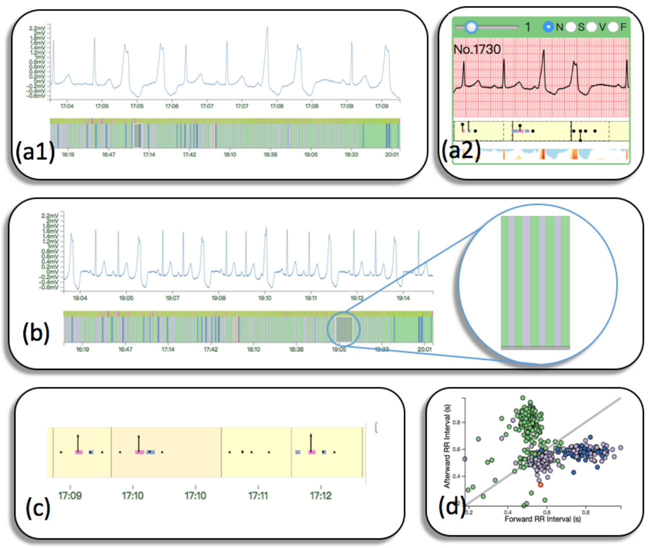

Figure 10. Diagnosis results of Interview II.

CNN model has an average accuracy of 91.59%. As a result, it took a significant amount of time for the users to identify the arrhythmia. However, after repeated use of the CNN's classification results to locate the arrhythmia, users placed increasing trust in the algorithm's results. For example, **E1** said that "the colors [encoding the results of CNN classification] hint that they are VEB and SVEB beats" when she immediately identified a number of PVCs and APC in the record. **E2** also noted a potential ventricular trigeminy in the Bar View (Fig. 2(3a)), where VEB (purple bar) and sinus (green bar) appeared alternately (Fig 10 (b)). Their trust in the CNN model increased again when they tried to detect misclassified heartbeats. It was difficult for users to identify any mislabeled heartbeats when manually investigating using the Overview (Fig. 2(2)), where the vast majority of classification results were found to be correct. This led users to shift focus to a task that depended on the model's results: analyzing the recommendations from the local outlier detection algorithm. The cardiologists found it useful to trust the CNN model, taking advantage of its results to speed their work.

Theme 2: Usability. Users discussed the usability of the system from three perspectives according to the thematically coded interview transcripts.

First, the experts regarded the system as "intelligent" and "time-saving." This was due to both (1) the automatic recommendations based on the CNN classification results, and (2) the outlier detection algorithm. For example, **E1** detected a group of similarly mislabelled heartbeats while using the rare category detection feature of the Overview (Fig. 2(2)). Referencing the automatic recommendation, she stated: "It helps us skip the normal beats and saves a lot of time.". In addition, with the automated recommendation of anomalous beats, **E2** discovered a group of ventricular couplets (16:09, 17:05, 17:30) and ventricular pre-excitation (17:07) (Fig. 10(a1)&(a2)).

Second, the system was considered "powerful" because it provides various views of the data, in some cases using novel visualization designs. Different views served different purposes for different experts. For example, referencing the Overview panel, **E1** said that "the overview of all heartbeats can prevent us from missing some abnormal heartbeats. This may also reveal the severity of the disease to some extent." Moreover, she endorsed the design of the Inspection View (Fig. 2(5)), stating that "the background grids can help us understand the ECG signals quantitatively". **E2** found the Statistic View (Fig. 2(8)) very helpful: "I like the Statistic View because the parameters displayed are important for diagnosing cardiac diseases." She made this comment as she discovered that sinus heartbeats tended to display different RR intervals than ventricular or fusion heartbeats (Fig. 10(d)).

Finally, the experts generally enjoyed the "user-friendly" interaction designs but also made suggestions for improvement. In particular, **E1** made two very valuable suggestions. First, she suggested that the Temporal View (Fig. 2(3,4)) should allow users to shift through a fixed number of heartbeats using a fixed-scale zoom, making the control of the view more agile and simple. Second, she suggested that cardiologists may often want to quickly compare a heartbeat of interest to its context (neighboring beats) when inspecting the Temporal View. In response this feedback, we added those features to our system design and made them available to **E2** in her experiment. The features were well received.

Theme 3: Adaption of Workflows. While the ECGLens system was designed to meet the needs of its target users, it was also observed that users adapted their own workflows to the capabilities of the system. For example, the ability to initially inspect the data using the Database View (Fig. 2(1)) allowed the experts to quickly navigate from a general overview to a focused analysis of areas of interest. For instance, **E1** used Database View to locate an occurrence of tachycardia around 18:00. She then brushed this 3-minute segment to make further inspections. **E2** similarly started by brushing several interesting segments in the Database View and inspecting the Temporal View simultaneously.

For anomaly analysis, both the CNN classification and outlier detection algorithms proved influential over user behaviors. **E1** used the computed results to identify some PVCs and APC around the 18:00 mark as shown in Fig. 9(a). The expert also followed the recommendation of the system to explore anomalous beats in the Temporal View (Fig. 2(3,4)). She stated "here is an obvious ventricular tachycardia" after examining the first recommendation at approximately 17:30 (Fig. 9(b1)). Based on this finding, she explored nearby waves and quickly located additional abnormal heartbeats suggesting ventricular pre-excitations (Fig. 9(c1)). In addition, by observing the A-glyphs, **E2** found many heartbeats with a prolonged QRS-complex (Fig. 10(c)) and thus inferred that this patient may suffer from PVC.

Finally, they made their diagnoses according to their personal needs. **E1** noted "Many of [the inspection boxes shown in Inspection View] verify my findings in the Temporal View". For example, the findings shown in Fig. 9(b2) and Fig. 9(c2) in the Inspection View corresponded to Fig. 9(b1) and Fig. 9(c1).

DISCUSSION

In clinical practice, the analysis of long term ECG data streams can help doctors observe many irregular and intermittent heart problems, which could not be reliably observed during a short clinical appointment with a doctor. However, exploring large scale ECG data is time-consuming while the automatic analysis algorithms can neither meet the exploratory requirements nor provide sufficient contextual information to guide the exploration process. Thus, ECGLens is aimed to facilitate the exploration of large scale ECG data through an interactive visualization system integrated with automated analysis algorithms. In this section, we discuss the limitations of our system and implications of our system designs.

Limitations

Based on the results of our user study and expert interviews, we have identified certain limitations of of our system.

Capacity of A-glyph design. Although the user study results provide evidence of the effectiveness of the A-glyph in arrhythmia detection, there is information beyond the four key components of the heartbeat that is omitted from the design. The study results for T2, as well as the questionnaire results, showed that the conventional ECG representation slightly outperformed the A-glyph in heartbeat classification. We believe that this is in part because the standard ECG representation is more familiar to cardiologists. However, it is also a factor that the A-glyph shows a simplified view. As shown in Fig. 7, when the variables (types of abnormality or the number of components) got more complicated, the difference in classification performance was decreasing in general. Therefore, we believe that providing additional detail to cardiologists will help them perform their tasks more effectively. Adding more visual elements to the A-glyph design or dividing heartbeats into smaller segments are two potential solutions.

Comparison of multiple patients. Our system is only able to display the data of one patient at a time. One expert we interviewed mentioned that our system could incorporate data from multiple patients, so that doctors can make comparisons and gain more insights regarding higher-level patterns over populations, such as the effect of a specific treatment, or typical disease progression patterns.

A demand for additional features. Although ECGLens supports a large set of features, our evaluation identified a number of additional features that our participants suggested as valuable topics for future work. For example, one expert suggested that the Database View could integrate more choices of parameters other than QTc value so as to provide cardiologists with a more comprehensive overview of the data.

Implications of design

Our evaluation findings suggest several implications for future interaction design.

Extension to other cardiac diseases. Although this work focuses on arrhythmia detection and identification, the interactive framework we proposed can be widely applicable. It provides a method for integrating automatic analysis algorithms with visual representations and facilitate the exploration of large-scale ECG data. Moreover, introducing rare category detection helps users in detecting anomalies automatically. This framework, including the analysis and visualization methods can be easily extended to diagnose other cardiac diseases.

Consistency with conventional ECG diagnosis. The evaluation results highlight the importance of conventional ECG representations in heartbeat classification. We have adhered to this principle when we design ECGLens. For example, we have added standard ECG grids to the system's visual design. These grids are both familiar to cardiologists and proven through day-to-day use to aid in diagnoses. For this reason, we have incorporated the grids into many views in our system. This design choice proved to be very helpful for users during their diagnosis. We believe that system's ECG exploration must balance novel designs with the strengths of conventional and familiar representations to reduce the learning curve and leverage trained expertise.

CONCLUSION AND FUTURE WORK

We have presented an interactive data exploration system, ECGLens, that enables cardiologists to visually identify and analyze arrhythmia in large-scale ECG records. The system design supports the exploration process by integrating a heartbeat classification algorithm, an outlier detection algorithm to our system, as well as an interactive workflow for arrhythmia detection. The design includes the novel A-glyph representation for ECG data which we show outperforms baseline designs in identifying heartbeats that exhibit arrhythmia. Moreover, our evaluation shows that the overall system successfully supported arrhythmia identification within large-scale ECG datasets. In the future, we plan on addressing the limitations of our current implementation and deploying our system to local hospitals so as to improve our system through more users' feedback. Moreover, we intend to improve our classification algorithm by feeding clinician-provided corrections back to a model re-training process in real-time.

ACKNOWLEDGEMENTS

Nan Cao is the corresponding author. We thank all the study participants and reviewers for their comments. This work is part of the research supported by NFSC Grants-61602306, the National Basic Research Program of China (973 Program)-2014CB340304, the DBI-1356669, SCH-1343896, III-1526012, and SCH-1622678 grants from the National Science Foundation and grant 1R21HS024581 from the National Institute of Health.

REFERENCES

1. U Rajendra Acharya, P Subbanna Bhat, and UC Niranjan. 2002. Comprehensive visualization of cardiac health using electrocardiograms. *Computers in Biology and Medicine* 32, 1 (2002), 49–54.
2. AMPS. 2008. A. M. P. S. llc. www.amps-llc.com/Home.php. (NOV 2008).
3. Rodrigo Varejao Andreao, Bernadette Dorizzi, and Jérôme Boudy. 2006. ECG signal analysis through hidden Markov models. *IEEE Transactions on Biomedical Engineering* 53, 8 (2006), 1541–1549.
4. Muhammad Arif and others. 2008. Robust electrocardiogram (ECG) beat classification using discrete wavelet transform. *Physiological Measurement* 29, 5 (2008), 555.
5. American Heart Association. 2014. Target Heart Rates. http://www.heart.org/HEARTORG/HealthyLiving/PhysicalActivity/Target-Heart-Rates_UCM_434341_Article.jsp#.WbI8stMjFBw. (MAY 2014).
6. Raymond Robert Bond, Dewar Darren Finlay, Chris Desmond Nugent, and George Moore. 2010. A web-based visualization tool for transforming the 12-lead ECG into a body surface potential map. In *Computing in Cardiology, 2010*. IEEE, 285–288.
7. Raymond R Bond, Dewar D Finlay, Chris D Nugent, George Moore, and Daniel Guldenring. 2013. Methods for presenting and visualising electrocardiographic data: from temporal signals to spatial imaging. *Journal of Electrocardiology* 46, 3 (2013), 182–196.
8. Markus M Breunig, Hans-Peter Kriegel, Raymond T Ng, and Jörg Sander. 2000. LOF: identifying density-based local outliers. In *ACM Sigmod Record*, Vol. 29. ACM, 93–104.
9. Huihua Kenny Chiang, Chao-Wei Chu, Gau-Yang Chen, and Cheng-Deng Kuo. 2001. A new 3-D display method for 12-lead ECG. *IEEE Transactions on Biomedical Engineering* 48, 10 (2001), 1195–1202.
10. Zhicheng Cui, Wenlin Chen, and Yixin Chen. 2016. Multi-scale convolutional neural networks for time series classification. *arXiv preprint arXiv:1603.06995* (2016).
11. Philip De Chazal, Maria O'Dwyer, and Richard B Reilly. 2004. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering* 51, 7 (2004), 1196–1206.
12. AAMI ECAR. 1987. Recommended practice for testing and reporting performance results of ventricular arrhythmia detection algorithms. *Association for the Advancement of Medical Instrumentation* (1987).
13. EcgSoft. 2008. EcgSoft home page. http://www.ecg-soft.com/. (NOV 2008).
14. M Escalona-Morán, MC Soriano, J García-Prieto, I Fischer, and CR Mirasso. 2014. Multivariate nonlinear time-series estimation using delay-based reservoir computing. *The European Physical Journal Special Topics* 223, 13 (2014), 2903–2912.
15. Miguel Angel Escalona-Morán, Miguel C Soriano, Ingo Fischer, and Claudio R Mirasso. 2015. Electrocardiogram classification using reservoir computing with logistic regression. *IEEE Journal of Biomedical and health Informatics* 19, 3 (2015), 892–898.
16. Himanshu Gothwal, Silky Kedawat, and Rajesh Kumar. 2011. Cardiac arrhythmias detection in an ECG beat signal using fast fourier transform and artificial neural network. *Journal of Biomedical Science and Engineering* 4, 04 (2011), 289.
17. Alfredo Illanes-Manriquez, Raúl Jiménez, Gustavo Dinamarca, Claudia Jiménez, and Eduardo Lecannelier. 2010. Visualizing the electrocardiogram through orbital transform. In *Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBC)*. IEEE, 5290–5293.
18. Wei Jiang and Seong G Kong. 2007. Block-based neural networks for personalized ECG signal classification. *IEEE Transactions on Neural Networks* 18, 6 (2007), 1750–1761.
19. S Karpagachelvi, M Arthanari, and M Sivakumar. 2010. ECG feature extraction techniques-a survey approach. *arXiv preprint arXiv:1005.0957* (2010).
20. John A Kastor. 2000. *Arrhythmias*. WB Saunders Company.
21. Jan A Kors and Gerard van Herpen. 2008. Mirror image electrocardiograms and additional electrocardiographic leads: new wine in old wineskins? *Journal of Electrocardiology* 41, 3 (2008), 245–250.
22. Hanfei Lin, Siyuan Gao, David Gotz, Fan Du, Jingrui He, and Nan Cao. 2017. RCLens: Interactive Rare Category Exploration and Identification. *IEEE Transactions on Visualization and Computer Graphics* (2017).
23. Jessica Lin, Eamonn Keogh, Li Wei, and Stefano Lonardi. 2007. Experiencing SAX: a novel symbolic representation of time series. *Data Mining and knowledge discovery* 15, 2 (2007), 107–144.
24. Eduardo José da S Luz, William Robson Schwartz, Guillermo Cámara-Chávez, and David Menotti. 2016. ECG-based heartbeat classification for arrhythmia detection: A survey. *Computer Methods and Programs in Biomedicine* 127 (2016), 144–164.
25. John E Madias. 2004. The 13th multiuse ECG lead: Shouldn't we use it more often, and on the same hard copy or computer screen, as the other 12 leads? *Journal of Electrocardiology* 37, 4 (2004), 285–287.
26. Farah Magrabi, Nigel H Lovell, and Branko G Celler. 1999. A web-based approach for electrocardiogram monitoring in the home. *International Journal of Medical Informatics* 54, 2 (1999), 145–153.

27. Juan Pablo Martínez, Rute Almeida, Salvador Olmos, Ana Paula Rocha, and Pablo Laguna. 2004. A wavelet-based ECG delineator: evaluation on standard databases. *IEEE Transactions on Biomedical Engineering* 51, 4 (2004), 570–581.

28. George B Moody and Roger G Mark. 2001. The impact of the MIT-BIH arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine* 20, 3 (2001), 45–50.

29. Jalal A Nasiri, Mahmoud Naghibzadeh, H Sadoghi Yazdi, and Bahram Naghibzadeh. 2009. ECG arrhythmia classification with support vector machines and genetic algorithm. In *Third UKSim European Symposium on Computer Modeling and Simulation*. IEEE, 187–192.

30. M Oefinger, W Zong, M Krieger, and RG Mark. 2004. An interactive web-based tool for multiscale physiological data visualization. In *Computers in Cardiology, 2004*. IEEE, 569–571.

31. OFFIS. 2011. Discussion forum for OFFIS DICOM tools. `http://forum.dcmtk.org/2011`. (JAN 2011).

32. Stanislaw Osowski, Linh Tran Hoai, and Tomasz Markiewicz. 2004. Support vector machine-based expert system for reliable heartbeat recognition. *IEEE Transactions on Biomedical Engineering* 51, 4 (2004), 582–589.

33. Yüksel Özbay, Rahime Ceylan, and Bekir Karlik. 2006. A fuzzy clustering neural network architecture for classification of ECG arrhythmias. *Computers in Biology and Medicine* 36, 4 (2006), 376–388.

34. Alex Page, Tolga Soyata, Jean-Philippe Couderc, and Mehmet Aktas. 2015b. An open source ECG clock generator for visualization of long-term cardiac monitoring data. *IEEE Access* 3 (2015), 2704–2714.

35. Alex Page, Tolga Soyata, Jean-Philippe Couderc, Mehmet Aktas, Burak Kantarci, and Silvana Andreescu. 2015a. Visualization of health monitoring data acquired from distributed sensors for multiple patients. In *Global Communications Conference (GLOBECOM)*. IEEE, 1–7.

36. Saurabh Pal and Swanirbhar Majumder. 2010. ECG Data Analysis. In *Intelligent Medical Technologies and Biomedical Engineering: Tools and Applications*. IGI Global, 122–144.

37. Pranav Rajpurkar, Awni Y Hannun, Masoumeh Haghpanahi, Codie Bourn, and Andrew Y Ng. 2017. Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. *arXiv preprint arXiv:1707.01836* (2017).

38. EM Tamil, NH Kamarudin, R Salleh, M Yamani Idna Idris, MN Noorzaily, and AM Tamil. 2008. Heartbeat electrocardiogram (ECG) signal feature extraction using discrete wavelet transforms (DWT). *Proceedings of CSPA* (2008), 1112–1117.

39. A Teeramongkonrasmee, C Tangwongsan, and S Sitthisook. 2009. Development of a real-time cardiac arrhythmia analyzer. In *Proceedings of 32nd Electrical Engineering Conference (EECON-32)*, Vol. 2. 1367–1370.

40. Markos G Tsipouras, Dimitrios I Fotiadis, and D Sideris. 2005. An arrhythmia classification system based on the RR-interval signal. *Artificial Intelligence in Medicine* 33, 3 (2005), 237–250.

41. Edward Tufte and P Graves-Morris. 2014. The visual display of quantitative information.; 1983. (2014).

42. Galen S Wagner. 2001. *Marriott's practical electrocardiography*. Lippincott Williams & Wilkins.

43. Can Ye, BVK Vijaya Kumar, and Miguel Tavares Coimbra. 2012. Combining general multi-class and specific two-class classifiers for improved customized ECG heartbeat classification. In *21st International Conference on Pattern Recognition (ICPR)*. IEEE, 2428–2431.

44. Yun-Chi Yeh, Wen-June Wang, and Che Wun Chiou. 2009. Cardiac arrhythmia diagnosis method using linear discriminant analysis on ECG signals. *Measurement* 42, 5 (2009), 778–789.

45. Qibin Zhao and Liqing Zhang. 2005. ECG feature extraction and classification using wavelet transform and support vector machines. In *International Conference on Neural Networks and Brain*, Vol. 2. IEEE, 1089–1092.

46. Muhammad Zubair, Jinsul Kim, and Changwoo Yoon. 2016. An automated ECG beat classification system using convolutional neural networks. In *6th International Conference on IT Convergence and Security (ICITCS)*. IEEE, 1–5.