Incorporating Artificial Intelligence into Medical Cyber Physical Systems: A Survey

Omid Rajabi Shishvan, Daphney-Stavroula Zois, and Tolga Soyata

Abstract Medical Cyber Physical Systems (MCPSs) prescribe a platform in which patient health information is acquired by the emerging Internet of Things (IoT) sensors, pre-processed locally, and processed via advanced machine intelligence algorithms in the cloud. The emergence of MCPSs holds the promise to revolutionize remote patient healthcare monitoring, accelerate the development of new drugs or treatments, and improve the quality-of-life for patients who are suffering from various medical conditions among other various applications. The amount of raw medical data gathered through the IoT sensors in an MCPS provides a rich platform that artificial intelligence algorithms can use to provide decision support for either medical experts or patients. In this paper, we provide an overview of MCPSs and the data flow through these systems. This includes how raw physiological signals are converted into features and are used by machine intelligence algorithms, the types of algorithms available for the healthcare domain, how the data and the decision support output are presented to the end user, and how all of these steps are completed in a secure fashion to preserve the privacy of the users.

Keywords IoT · Cloud computing · Sensors · Physiological signals · Machine learning · Cyber-physical systems

1 Introduction

The rapid emergence of IoT devices in conjunction with advances in computational capabilities have led to a growing interest of MCPSs in both research and commercial fields [1]. With applications spanning from general areas such as fitness tracking and personal health [2] to more technical fields such as remote health monitoring and medical decision support systems [3, 4], MCPSs have emerged as an

O. R. Shishvan · D.-S. Zois · T. Soyata (✉)
University at Albany, SUNY, Albany, NY, USA
e-mail: orajabishishvan@albany.edu; dzois@albany.edu; tsoyata@albany.edu

A. El Saddik et al. (eds.), *Connected Health in Smart Cities*,
https://doi.org/10.1007/978-3-030-27844-1_8

effective technology that can not only improve general medical practice [5] but also create new business opportunities [6]. These systems consist of a network of sensors worn on the body of patients and gather physiological and environmental signals; these signals are first pre-processed at a location that is close to the acquisition source, and transmitted to either a private or public cloud. A private cloud is owned directly by a Healthcare Organization (HCO), while a public cloud is rented from cloud service providers, such as Amazon EC2. The primary purpose of the cloud is to execute a set of machine intelligence algorithms and provide decision support to healthcare professionals (e.g., doctors and nurses). In addition to the previously gathered corpus of relevant information, MCPSs use the acquired data to train machine intelligence algorithms and make inferences regarding the potential medical conditions of a patient based on his/her physiological data.

These aforementioned emerging clinical and personal healthcare applications both benefit from devices that acquire medical data and process them with varying degrees of intelligence; in this way, an MCPS is not limited to IoT devices, but any sensory device, coupled with a platform that can run the decision support algorithms.

Given the sensitive nature of the personal medical information, measures must be taken within the MCPS to remain in compliance with confidentiality laws that protect personal health information. For example, Health Information Privacy and Accountability Act (HIPAA) laws in the USA [7] have strict restrictions requiring that medical information can only be released to authorized users. This privacy issue is especially important if the HCO uses public infrastructures as a computation platform where the hardware is shared with multiple unknown users.

In this chapter, an overview of an MCPS and the artificial intelligence algorithms that are used in it are presented. Specifically, in Sect. 2, the general structure of an MCPS is described. In Sect. 3, the data acquisition component of the MCPS is described. Details of the decision support process are presented in Sect. 4, in which the algorithms are studied based on their goals. Sections 4.1 through 4.4 describe these goals, including knowledge discovery, classification, regression, and sequential decision-making, respectively. Section 5 is where visualization is elaborated on, which is an important stage of the decision support process. Issues surrounding data privacy and security are discussed in Sect. 6. Section 7 discusses the challenges and open issues in incorporating artificial intelligence in MCPSs, and summary and concluding remarks are provided in Sect. 8.

2 Medical Cyber-Physical Systems

One of the most promising applications for an MCPS is real-time, long-term health monitoring [6, 8]. The general structure of an MCPS consists of multiple components as shown in Fig. 1. The data acquisition, aggregation, and pre-processing layer gathers all relevant and necessary information through wearable sensors in a Wireless Body Area Network (WBAN), environmental sensors, and other external

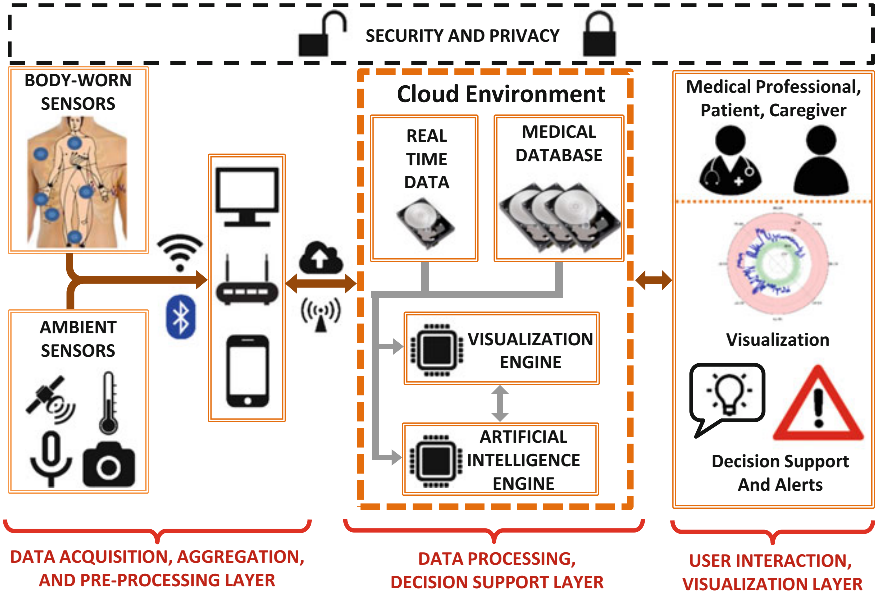

Fig. 1 An overview of different components of a MCPS. The raw data is acquired from body-worn sensors. After a pre-processing phase, the signals are aggregated and transferred to the cloud, where they are processed by machine intelligence algorithms. The outcomes of the algorithms are then presented to the users in the form of visualization, decision support suggestions, or alerts. All of these steps are subject to security measures to ensure that the data is kept secure and private

sources of data such as the Internet. This data is then aggregated, and features are extracted and transmitted to the cloud for use by machine intelligence algorithms. These algorithms yield outcomes that can be then used for decision support or visualization purposes. All of these components are subject to privacy guidelines to ensure the security and privacy of the data that is traveling within the MCPS. We discuss these layers in more detail in the sections to follow.

2.1 *Data Acquisition, Aggregation, and Pre-processing*

Data acquisition is the first component of an MCPS, which gathers physiological and ambient signals, primarily through WBAN sensors but also sensors deployed in the environment. This component also aggregates the data, performs the necessary pre-processing, and transmits the resulting data to the cloud. This transmission is usually done through conventional communication links such as 3G, 4G, or 5G networks [9, 10] or other emerging communication links such as the communication links provided by Low Power Wide Area Network (LoRaWAN) protocol [11]. Note that this data can either be in raw format (cf. Sect. 3.1) or represented by a set of features that is extracted from it (cf. Sect. 3.2). The format of the data depends on many factors, such as the application, the bandwidth of the data connection, or the system's power supply limits. We will discuss this component further in Sect. 3.

2.2 Decision Support Using Machine Intelligence

The decision support component of an MCPS is responsible for the analysis of the data to detect informative patterns. Depending on the application, the goal of this layer may be to find novel patterns in the input data (Sect. 4.1), classify the input data into a limited number of *classes* (Sect. 4.2), provide an estimate for given input data (Sect. 4.3), or provide a suggestion for a decision-making task (Sect. 4.4).

2.3 Data Visualization

It is vital that data is reported in a format comprehensible to both medical professionals and patients. This includes all acquired data, decision support suggestions, and emergency alerts. Medical data gathered through sensors is typically too voluminous for the human brain to process [12], which in turn necessitates the generation of informative summaries. It is essential that these summaries highlight sections of data that require attention without sacrificing accuracy; this can significantly reduce human errors and ensure efficient diagnosis. Section 5 discusses data visualization in depth.

2.4 Data Privacy

All personal medical data in the USA is protected under the HIPAA regulations [7]. Through its transition between the components of an MCPS, personal medical data should be treated in such a way that its integrity remains intact without compromising the patient's privacy. This requires special provisions during the design of an MCPS, such as encrypting data and enforcing restricted access to personnel. Section 6 provides detailed explanation on data privacy and security.

3 Data Acquisition, Aggregation, and Pre-processing

Machine intelligence algorithms are used by MCPSs to generate desired outcomes based on a variety of different data. In this section, the different stages of the data acquisition, aggregation, and pre-processing process are discussed along with the different types of data used.

3.1 Raw Data

WBAN sensors collect and communicate (through wireless protocols such as IEEE 802.11ah [13], WirelessHART [14], Bluetooth Low Energy (BLE) [15, 16], and ZigBee [17]) a variety of raw medical data including but not limited to data from respiration sensors, heart-rate monitors, blood pressure, glucose, and oxygen saturation monitoring sensors, and muscle activity sensors [1]. Note that other ambient sensors such as environmental temperature, location, and sound may be useful for some applications. These signals are usually gathered with relatively high sampling frequency and are prone to environmental noise, whether induced by patients' movements or by other electrical devices in the vicinity [18]. As a result, redundant information is included in the raw data that can be omitted without losing any valuable information. Thus, direct transmission of raw data to the next layer of the system can not only waste valuable bandwidth, but also expend battery resources in mobile devices that operate under severe energy constraints. To address both drawbacks, *features* are extracted from raw data to retain information that is useful for decision support and visualization.

3.2 Features

The process during which *features* are extracted from raw data is studied in Sect. 3.3. These features are presented as variables, which are used as an input to machine intelligence algorithms or the visualization system. Section 3.4 studies generic (application-independent) features (such as the average of the signal), which can be used in almost any application, although application-specific features (such as certain time intervals of an electrocardiogram (ECG)), as studied in Sect. 3.5, can have a much better representative power.

3.3 Feature Extraction

Feature extraction [19] can be accomplished via different dimensionality reduction techniques that extract the most statistically significant information from raw data. Common feature extraction methods include

- **Principal Component Analysis (PCA)** is a technique for dimensionality reduction that transforms a collection of correlated data to a collection of uncorrelated data points. PCA is widely used in the literature; for example, authors in [20] detect sensorineural hearing loss in Magnetic Resonance Imaging (MRI) images by transforming these images into features through wavelet decomposition and dimensionality reduction via PCA.

- **Kernel PCA (KPCA)** is based on PCA and uses pre-defined kernel functions, such as polynomial or Gaussian functions, to perform non-linear data transformation. For example, authors in [21] use kernel PCA to reduce 700 features to only 5 features, which are then used to assess depressive symptoms in different individuals.
- **Canonical Correlation Analysis (CCA)** is a method for analyzing the correlation between two different multivariate inputs. In [22], authors develop a hybrid brain-computer interface smart glass that is used for controlling electronic devices. They use CCA as part of their system to find the most similar Electroencephalograph (EEG) recordings and classify the user's action based on that.
- **Multidimensional Scaling (MDS)** is a dimensionality reduction method that transforms data points into a lower dimension, while maintaining the Euclidean distance between the data points. For example, authors in [23] use MDS as part of their process for daily activity recognition among subjects in which the data is presented as a matrix and MDS is used to convert them to a lower dimensional space that makes classifying the activity matrix easier.
- **Artificial Neural Networks (ANNs)** can be structured in such a way that they are able to extract features from the raw data. For example, autoencoders are a type of ANNs that take a high-dimensional signal as an input, convert them to a signal with lower dimension, and reconstruct the original signal from the converted signal. This lower dimensional data provides efficient feature reduction [24]. An example application of ANNs is presented in [25], where the authors use deep belief networks in an emotion recognition scheme in which the network extracts features from high dimensional audio and video input signals. Another application is discussed in [26], where authors build a fall detection system that utilizes frequency modulated radars. They use an autoencoder to extract features from the data and show that it improves the performance as compared to conventional methods such as PCA.

3.4 Application-Independent Features

In many applications, features extracted from raw data are very generic and do not depend on the application of interest. As a result, they can be extracted from a variety of medical data and should be interpreted based on the context of the application. The features can be categorized as: *temporal*, *spectral*, and *cepstral*. Temporal and spectral features are extracted from the time and frequency domain of a signal respectively, while cepstral features are extracted based on the changes in a signal's spectral bands.

3.4.1 Temporal Features

Statistical information of data, ranging from its mean and median to kurtosis and different percentiles, are typical examples of temporal features that may reveal useful information. Frequently, temporal features on their own are sufficient to accurately summarize a variety of data. For example, in [27], temporal features such as the mean of maxima and the mean of minima from acceleration and heart rhythm signals can adequately and accurately detect the patients' physical activity. Another application is presented in [28], where the median frequency among other temporal features is extracted from skin conductance, ECG, and electromyogram (EMG) signals and used to detect mental stress among people. More complex temporal features (e.g., Lempel-Ziv complexity, Hermite polynomial expansion (HPE) coefficients, central tendency measure) can be extracted from the data by applying advanced processing algorithms. For example, authors in [29] use both central tendency measure and Lempel-Ziv complexity of SpO^2 signals for real-time detection of sleep apnea.

3.4.2 Spectral Features

In addition to temporal features, analyzing the signals in the frequency domain can also provide useful information about their characteristics. Some of the features in this domain include power of the signal in various frequency bands, phase angle, and the spectral entropy. An example application of spectral features is discussed in [30], where EEG signals are analyzed. By using features such as dominant frequency and normalized spectral entropy, an epilepsy application is developed.

3.4.3 Cepstral Features

Cepstral features have proven to be useful for removing disturbance in the data induced by uncontrollable parameters such as sensor displacement. The cepstrum of a signal is the inverse Fourier transform of the logarithm of the spectrum of the signal. Cepstral features have been used for respiration problems detection from breath sound recordings [31], detecting heart rhythm arrhythmia from ECG recordings [32], and sound signal classification for assistive technologies for hearing impaired patients [33].

3.5 Application-Specific Features

Each biomarker signal has its own unique characteristics that can be extracted as features with high information content. For example, in an ECG recording (depicted in Fig. 2), some specific features with high information content are the RR interval and the QT interval [34]. Furthermore, it is sometimes useful to *fuse* together

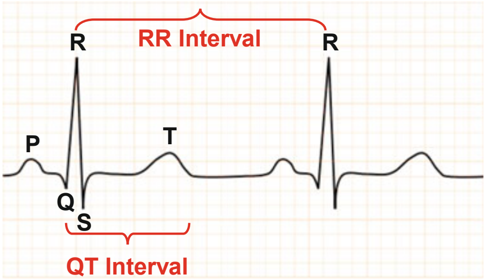

Fig. 2 A simplified example of an ECG waveform showing two consecutive heartbeats. The QT and RR intervals are indicated in the plot

simple features such as the RR and QT intervals in order to create more informative features. For instance, a corrected form of QT values is calculated by normalizing it with respect to the RR value:

$$\mathrm{QTcB} = \frac{\mathrm{QT}}{\sqrt{\mathrm{RR}}} \tag{1}$$

Another example of an application-specific feature is the power of an EEG signal in different bandwidths such as between 0.5 and 4 Hz (Delta wave) and 7.5 and 12.5 Hz (Alpha wave).

3.6 Feature Selection

Even though a large set of features can be extracted, not all of them are necessarily useful for machine intelligence algorithms. In practice, features must bear useful information and low redundancy, while at the same time, they need to be fast to process, to avoid burdening system performance [35]. Feature selection techniques can be used to select only the subset of features that contribute the most to the success of the entire system. Some feature selection methods that are commonly used are

- **Sequential Backward Selection (SBS)** starts by using all available features and sequentially eliminates the features that have minimal impact on system performance (e.g., accuracy). In [36], authors use SBS to select the most important features from all available time-domain and frequency-domain features to enable accurate sleep apnea detection using pulse oximeters.
- **Sequential Forward Selection (SFS)**, unlike SBS, starts with an empty set of features and incorporates the features that have the maximum effect on the system performance in every step. An example application of SFS is provided in [37], where the authors develop a wearable glove system to detect stress events in drivers. They use SFS to select only the features that achieve satisfactory level of accuracy in proposed system.

- **Sequential Forward Floating Selection (SFFS)** combines SBS and SFS techniques; at each step, it dynamically adds or removes a variable number of features. This approach starts with forward selection and then employs backward selection. An example application of SFFS is discussed in [38], where authors use SFFS to select the features for their proposed cognitive ability evaluation scheme. They use SFFS on a set of 33 features and show that it selects certain features more frequently while it completely ignores some other features.
- **Correlation-Based Feature Selection (CFS)** selects the subset of features that have the highest correlation with the output classes, yet they are not correlated to each other. In [39], authors detect epileptic seizures using EEG signals and use an improved version of CFS to select the best features from different domains including time and frequency. They show that CFS is able to maintain accuracy with a reduced number of features.
- **Genetic Algorithms (GA)** optimize a problem by searching among possible solutions using natural selection-based techniques. An initial set of solutions is identified, from which the best options are selected. These selections are then modified through *mutations*, converging to the optimal solution. An example application of using GA for feature selection is discussed in [40], where evolution-based algorithms select the best features from EEG signals for emotion recognition. Another example is provided in [41], where authors use GA to select ECG features for cardiac disease classification.

4 Decision Support

One of the most important functions of an MCPS is to provide decision support to aid in clinical diagnosis or personal health monitoring. Decision support systems transform the results of machine intelligence algorithms (i.e., output values) into appropriate formats that facilitate the understanding of patients and medical experts. There are various types of decision support including but not limited to providing an alert (e.g., warning for low blood sugar), estimating the likelihood of a disease (e.g., a developing arrhythmia), or displaying an intuitive visualization of the acquired signals over a long-term observation period (e.g., Holter ECG monitoring).

Decision support systems can be categorized by their respective goals as follows:

- **Knowledge discovery**: Knowledge discovery algorithms aim to identify previously unknown relations in data. Applications such as data clustering and anomaly detection fall under this category. We will elaborate on this category in Sect. 4.1.
- **Classification**: Classification algorithms work with datasets with known input-output relations and their output can be categorized into a limited number of classes. We will provide more details on this category in Sect. 4.2.

- **Regression**: Regression algorithms work with continuous outputs, as we will study in Sect. 4.3.
- **Sequential decision-making**: Sequential decision-making algorithms are used when a task requires automated decisions to be made over time in order to improve performance; we will provide more details on this category in Sect. 4.4.

4.1 Knowledge Discovery

Algorithms in this category aim to discover relations in a dataset with *unlabeled* data, i.e., data points with no previously known input/output relationships. Clustering and anomaly detection are typical and most common tasks related to knowledge discovery. Clustering involves grouping similar data points together, using techniques such as K-means, hierarchical clustering, and probabilistic clustering models also known as mixture models. On the other hand, anomaly detection focuses on identifying data points that do not conform to expected patterns when compared to other data points in a dataset. It is important to note that many anomaly detection algorithms are based on clustering algorithms that identify data points which do not belong to any major cluster.

Clustering and anomaly detection have been applied to a variety of healthcare applications including but not limited to healthcare insurance fraud, discovering unknown drug interactions, tracking epidemics, and estimating survival rates. In [42], authors employ outlier detection techniques to detect insurance frauds in health insurance claims. Specifically, they calculate the proportion of claims of fraudulent versus non-fraudulent providers, and show that fraudulent providers tend to file claims related to certain health issues more often. In [43], authors use mixture models to cluster patients into different mortality rate groups based on their physiological data gathered in Intensive Care Units (ICUs). Clustering techniques are also successful in producing trajectories of physiological data over time based on patients' individual clusters. For instance, the authors in [44] use an hierarchical clustering algorithm to detect the severity of three disease types (i.e., Crohn's disease, cystic fibrosis, down syndrome) based on lab test results. Even though patients diagnosed with Crohn's disease and cystic fibrosis can be successfully clustered based on the severity of their disease, this is not the case for down syndrome patients. The last observation is attributed to limited quantity of data. A similar application is discussed in [45], where K-means is employed on medical and mood data collected from chronic obstructive pulmonary disease (COPD) patients to track their symptoms over time and monitor the progression of their disease.

Association rule mining is another typical example of knowledge discovery, where the goal is to identify relations between variables in a dataset. For example, authors in [46] employ association rule mining on the FDA adverse event reporting system database to detect drug pairs that are associated with increased blood glucose

Table 1 Confusion matrix

	Predicted condition	
Actual condition	Positive	Negative
Positive	True positive (TP)	False negative (FN)
Negative	False positive (FP)	True negative (TN)

levels. They are able to show that a potential candidate for elevated blood glucose levels is a (previously unknown) drug combination of Paroxetine and Pravastatin, and a clinical trial verifies their findings.

4.2 Classification

In many datasets, the output values have a limited number of possibilities (i.e., *classes*), which implies that the output values are already divided into subgroups. A classification algorithm determines which subpopulation (class) each input value belongs to. The results of classification algorithms, specially the ones with binary outputs, can be presented in a confusion matrix as shown in Table 1. Based on these definitions, other metrics such as accuracy, F1-score, and the Area under receiver operating characteristic curve (AUC) are defined and used to describe the performance of classifiers. For example, accuracy is defined by Eq. (2) and F1-score is defined by Eq. (3) while AUC is defined by plotting the graph of true positive rate vs. false positive rate and calculating the area under its curve. For all of these metrics, the closer their value to "1", the better the classifier.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \tag{2}$$

$$\text{F1-score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{3}$$

An example classification application is described in [47], where the authors classify patients into one of two sub-populations, (i) patients with sleep apnea, and (ii) without apnea. Such a classification, in which there are only two possible output values is termed *binary classification*. Their classification scheme is a linear integer model that takes input features such as age, sex, smoking condition, and snoring during sleep. The authors report an AUC value of 0.785. Another classification application is presented in [48], where the authors take the ECG recordings of patients and detect whether they have long QT syndrome [34, 49] or not. They extract features such as the heart rate from the ECG data and feed them into multiple classification algorithms such as Support Vector Machines (SVMs), k-nearest neighbors, and AdaBoost. They are able to achieve accuracies higher than 70% using SVMs with radial basis function.

In [50], authors take histopathological images and classify gliomas (a type of cancer) into two classes: (i) low-grade glioma and (ii) high-grade glioma. They process the images and divide them into separate segments, where for each segment a cell-count profile is created. A decision tree algorithm is then applied to the cell-count profiles, which identifies the glioma as a low-grade or a high-grade one with 80% accuracy. Another work in [51] uses retinal images to detect Retinopathy of prematurity, a cause of blindness among children. They manually prepare a mask for the vessels in the images, fit splines into these masks, and extract feature from these splines. SVM classifiers are used to classify these vessel features into healthy and abnormal cases, where they achieve ≈95% accuracy in this task.

The work presented in [52] uses data gathered through smartphones for the purpose of remote health monitoring and physical activity classification. Authors take smartphone accelerometer data and extract features such as mean, standard deviation, and the peaks of the measurements from the signal. By using these features in a decision tree algorithm, they are able to classify multiple physical activities of the subjects, including {sitting, walking, going up or down the stairs, cycling}, with more than 80% accuracy. Authors in [27] also investigate activity recognition with an SVM classifier. They gather ECG and accelerometer data and extract time and cepstral features from these two signals. By fusing these two sets of features, they are able to distinguish nine physical activities with accuracies as high as 97.3%.

A study that uses mobile phone data in addition to wearable sensors is conducted in [28], where authors recognize stress among the participants in the study. Data for the study is gathered through a wrist sensor that has an accelerometer and measures skin conductance and mobile phone usage. They combine this information with a user survey that includes information about their mood, tiredness, and alcoholic and caffeinated beverage intake. Authors classify subjects as {stressed and not-stressed} with accuracy as high as 87.5%.

In [53], authors build a sleep apnea monitoring system that classifies the subjects based on their ECG signals. Participants undergo a sleep study, in which their ECG measurements are recorded; they are able to detect respiratory movements from these recordings in addition to extracting both time-based features and spectral features from the signals. By using an SVM classifier, they are able to achieve accuracies between 85% and 90% in sleep apnea detection. Authors in [54] build a non-intrusive mental-health tracker system. The system gathers features such as subject's head movement, heart rate, eye blinks, pupil radius, and facial expressions through a webcam and records other features that include the interactions of the subject with the computer and the content that the user views. Using this input data, they are able to classify subjects' emotion as positive, neutral, and negative with an AUC of 0.95.

Some classification applications include a large amount of data with high complexity, causing traditional feature extraction and classification techniques to fail providing acceptable results. Deep neural networks (DNNs), such as convolutional neural networks (CNNs) or recurrent neural networks (RNNs), have shown to be successful in classifying these cases. For example, in [25], authors use stacked autoencoders for feature extraction in an emotion recognition application. Since the input data includes video and audio data, authors use both conventional features and features extracted via deep belief networks to classify the emotion through an SVM classifier. They show that features extracted automatically through a DNN improve the classification accuracy of the SVM.

CNNs are one of the most commonly used DNNs for classification, due to their structure being able to capture both local and global features in multimedia inputs such as images or videos. Authors in [55] use CNNs to detect mitosis in breast cancer histology images. Their network is able to label image pixels as mitosis or non-mitosis, which results in a classification with an F1-score of 0.8. In [56], authors propose a CNN for gland segmentation in histology images. This work is able to detect benign and malignant glands; they report that under their segmentation scheme, extracted glands have less than 50 pixels of Hausdorff distance to the real ones. Another work that uses image inputs is presented in [57], where authors detect damage to retina due to diabetes in retinal fundus photographs. Their work is able to achieve an AUC of 0.99 in its classification task.

In [58], authors use a CNN with 1 dimensional input to classify different types of heart beat arrhythmia in ECG recordings. They develop a network that takes 5 min of each person's ECG in addition to multiple general heartbeat samples and train a personalized CNN to detect 5 types of arrhythmia. The network shows successful performance with accuracies as high as 99% in certain tasks. The work presented in [59] uses both CNNs and RNNs to annotate chest X-ray images with proper description. The CNN part of the paper is responsible to analyze the X-ray images and detect abnormalities in them. The output of the CNN is then fed to the RNN to produce appropriate annotations for the images to describe them, such as "normal" or "cadiomegaly/light." They show their work is able to annotate the images within acceptable range.

4.3 Regression/Estimation

In many applications, output data can take any value in a continuous range, rather than belonging to discrete groups or classes. In this case, the application is formulated as a regression or estimation problem, where the goal is to generate a continuous–valued output (contrary to a discrete output as done in classification discussed in Sect. 4.2) based on a set of input data.

An example regression application is discussed in [60], in which a robust heart rhythm estimation algorithm is proposed to combat false alarms in ICU caused by noise induced by the environment. The authors develop an estimation approach based on the Kalman filter [61, 62] that estimates the heart rate of an ICU patient from ECG and arterial blood pressure sensors and show that the proposed approach works well even when more noise is artificially added to the data. In [63], a disease trajectory prediction system is designed to predict the course of a disease in the future based on some initial patient medical data. The authors show that the proposed system can provide accurate prognosis at a personal level. The authors in [64] study the problem of prognosis of a disease in patients by focusing on the course of diabetes in diabetic patients. Their goal is to predict the possibility of a patient needing emergency care in the future. To this end, they consider lab tests, the list of the drugs that the patient uses, diagnoses, and other input features that are processed and selected by techniques such as filtering or PCA. They report the probability of a patient needing an emergency care in the future in addition to predicting their future lab test results. They show that their techniques are effective by reporting concordance indexes as high as 0.67.

A regression problem that focuses on drug discovery is discussed in [65]. The authors use DNNs with various inputs, including molecule structures of different drugs, that output on-target or off-target activities. They compare the performance of DNNs with other methods such as random forests with respect to the prediction of activities, and show that the former methods outperform the latter ones by improving the squared Pearson's correlation coefficient between the predicted activities and the observed ones from 0.42 to 0.5. In [66], the authors focus on the problem of estimating user fatigue through DNNs. They collect data from muscle and heart activity sensors, accelerometers, and a brain–computer interface that collects EEG signals. These signals are then provided as input to a DNN, which estimates the physical load of the participants and their physical fatigue.

4.4 Sequential Decision-Making

Sequential Decision-Making (SDM) models are typically used in medical applications to monitor and/or improve the medical process by estimating the task of interest as well as controlling any related variables. Example of sequential decision-making models are Markov Decision Processes (MDPs), Partially Observable Markov Decision Processes (POMDPs), and Multi Armed Bandits (MABs).

An example application of such models is discussed in [67, 68], where a WBAN [68] consisting of sensors such as accelerometers and ECGs, in addition to a mobile phone, is used for physical activity recognition. The system uses a POMDP model to select the best sensing strategy to achieve two different goals: (i) infer the physical activity of the individual accurately, and (ii) prolong mobile phone battery lifetime. The authors are able to show that the POMDP approach can lead to up 64% energy savings while losing only 10^{-4} in activity detection accuracy. Another application of sequential decision-making models is described in [69], where a video camera tracks the movements of patients with dementia to assist them with a handwashing task. To estimate the severity of dementia in patients and provide assistance in this task, a POMDP formulation is adopted, which decides when to intervene in the handwashing task. The model can do nothing and let the individual finish their task, provide cues such as task description to the individual, or call the caregiver.

In [70], a stress reduction system is introduced that uses a contextual MAB formulation to detect the relationship among different interventions to cope with stress and their outcome on different individuals for a given context. The data for the model comes from different sensors such as GPS, accelerometer, calendar, etc. in addition to other information gathered from the individual including personality traits and self-reports. The system evaluation shows that the participants in the study show lower symptoms related to depression. Another study that uses MAB approaches is presented in [71], where a personalized physical activity recommendation system is modeled as a MAB problem, which monitors the activities of individuals and provides suggestions to the users at different times for a healthier lifestyle. In [72], the authors develop a drug sensitivity prediction system that considers expert inputs to improve the efficacy of prescribed drugs for a given individual. The prediction refers to the effect of different drugs on patients with blood cancer and the features come from the genomic features of the cancer cells. To enhance the prediction, an expert provides an opinion to the prediction algorithm based on the genomic features, but the sheer number of features limits the feasibility of this input as the expert cannot provide an opinion on thousands of features. To address this issue, the authors use a MAB formulation to learn from the expert inputs and take their opinion only on the features that are considered the most important. This scheme improves the prediction accuracy by 8%.

5 Visualization

A vital part of a decision support system is presenting the important and necessary information to the users of the system in an intuitive format. A highly-effective way of achieving this goal is through data visualization that shows all relevant information in addition to machine-intelligence-based annotations for parts of data that require extra attention from the users. Visualization techniques vary based on both the application and the target users; for example, to provide feedback to medical experts, a system may need to include higher precision data with all the relevant medical information, while a lower level of technicality is sufficient for visualizing data for patients.

Despite the importance of proper data visualization and the benefits that it provides for the users, existing visualization techniques in the medical scientific disciplines are somewhat limited [73]. To date, several medical data visualization techniques have been introduced, which vary in complexity ranging from simple tables or bar graphs to advanced interactive multidimensional plotting systems. The focus of these techniques has been mostly on the data that are gathered through clinical visits. Visualization for MCPS data is more challenging due to the long duration of data acquisition and high dimensionality.

Traditional visualization techniques include lists, tables, graphs, charts, tress, pictograms, and formats to show spatial data [74] and causal relations within data elements. These formats present all information without removing any essential part

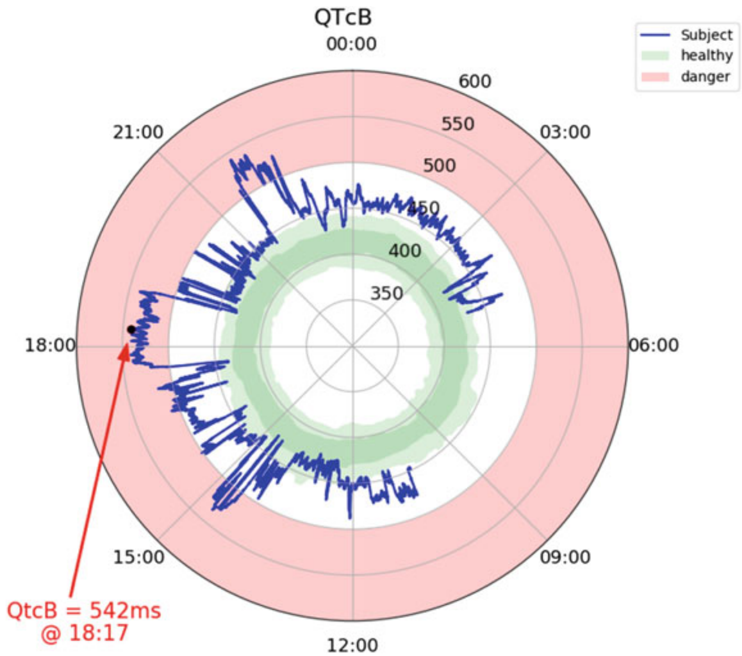

Fig. 3 A sample QTcB clock showing the calculated QTcB values of subject s30771 in [75] available in the PhysioBank database [76]. The plot shows the QTcB value for the entire recording duration and the highest value of QTcB is shown on the graph

and highlight the most important parts of the data. For example, the case shown in Fig. 3 is a 24-h visualization of a patient's ECG data [34, 49], where a "clock" shows the QTcB (Eq. (1)) value of a patient throughout the recording period. This visualization technique is designed to designate the top of the clock as midnight (00:00) and the bottom of it as noon (12:00) to visualize the entire 24-h recoding of the patient's ECG by using a single clock. The radial dimension of the clock (i.e., from the inside to the outside) represents the values of QTcB at a given time; the inner entries are colored green and represent "healthy" QTcB values (300–420 ms), whereas the values closer to the outer edges are colored red and represent "abnormal" QTcB values (500–600 ms). This visualization allows a cardiologist to view a patient's entire 24-h ECG recording period at a quick glance, which allows them to view 20–30 patient's Holter recordings within a negligible amount of time and identify the health conditions of each patient rapidly; this eliminates the need to search through traditional ECG recordings, which are printed on paper.

Multiple other techniques are proposed for plotting ECG data for different purposes. For example, authors in [77] develop an interactive ECG visualization system built on top of the research presented in [34], where multiple panels allow the medical professionals to plot different parts of the data with higher precision or show various statistical distributions related to the data. Another study that targets long-term ECG recordings is presented in [78], where heartbeats are shown in different clusters and the clusters with fewer members represent the heart beats with arrhythmia.

One of the main categories of visualization techniques involves presenting data to individuals who are self-tracking their health, physical fitness, and lifestyle. A rich body of work depends on self-logged information from users in addition to physical tracking devices to provide a visual feedback to them [79–82]. Many novel approaches are used for the feedback mechanism such as making data sculptures form the information [83], displaying them through abstract art [84], or even feedback for self-monitoring through edible chocolate [85].

Many of the proposed visualization schemes treat professional medical personnel as audience too. For example, hGraph [86] and its dependent programming libraries (like the one introduced in [87]) depict a summary of user activity, blood pressure, sleep, and in-clinic data and are dedicated to medical data visualization. Other systems such as Open mHealth [88] provide visualization as part of their overall architecture. Some other frameworks such as TimeLine [89] focus on visualizing only electronic health records and do not incorporate data from user activity in their figures. OpenICE [90] is another open platform for MCPS, which incorporates data visualization for vital signs of patients; it color-codes the vitals as being normal, not normal, and severely out of range.

PhysioEx is introduced in [91], which analyzes the streams of medical data and plots the duration, frequency, and trajectory of different events in the stream through a temporal intensity map. In the study presented in [92], a tele-rehabilitation system is designed that gathers patient activity through mobile sensors and visualizes them remotely for the care-giver for a better understanding of whether the patients adhere to their rehabilitation program or not and how their rehabilitation is progressing through time. A system that is implemented in ICU settings is proposed in [93], where all ICU data is shown and are accompanied by some notifications such as empty medication. Testing of this system shows that task completion times for nurses is significantly decreased and their situational awareness is increased.

There are many available tools that are used for creating these visualizations which are created by different companies. Plotly[1] provides different visualization tool libraries that are compatible with programming languages such as R, Python, and Java. AnyChart[2] also provides a data visualization platform that can create various forms of charts as well as real-time data streams. Tableau[3] is another tool that is widely used for creating visualizations. IBM Watson analytics[4] also provides a visualization kit for healthcare that can be used in R and Python programming languages through APIs. Some other widely used tools include software provided

[1] https://plot.ly/.

[2] https://www.anychart.com/.

[3] https://www.tableau.com/.

[4] https://www.ibm.com/watson/uk-en/health/.

by Sisense,[5] Microsoft Power BI,[6] Qlikview,[7] and SAP Lumira.[8] Note that these tools are not necessarily limited to healthcare applications and can be used to create visualizations in various industries.

6 Privacy

In the US, Health Insurance Portability and Accountability Act (HIPAA) [7] mandates the assurance of medical data privacy at every component of an MCPS. Security of an MCPS, whether concerning the general security of the system [8] or security of its specific components, has been addressed in the literature extensively and is an ongoing developing research topic.

Attacks on an MCPS can be categorized as either active or passive. *Active* attacks aim at accessing secret information by deviating from the security protocols, while a *passive* adversaries follow the security protocols, yet are able to access restricted information. The security layer of an MCPS should ensure that both of these types of threats are made ineffective in all layers of the system [8]. A general modeling of threats is presented in [94], where authors categorize the stakeholders in an MCPS and build trust and threat models based on them throughout the system. They categorize different potential threats such as confidentiality, integrity, and availability of data in different sections of the system such as the communication links and the software/hardware platform and list the possible remedies that may prevent these threats from inflicting damage on the system.

The study in [95] investigates the idea of integrating forensic principles into the design of an MCPS, which gives the HCO a means to investigate the intruders in case its MCPS is compromised. Their idea does not necessarily stop adversaries from intruding the system, but provides a means for detecting them after their attack. The authors discuss a forensic-by-design framework for an MCPS by breaking it down into different components such as risk assessment, forensic readiness principles, security and privacy requirements, relevant legislations and regulations, medical and safety requirements, and software and hardware requirements in addition to providing forensic-readiness testing criteria.

In addition to conventional privacy measures, there has been a growing interest in using machine intelligence algorithms to ensure the security of an MCSP. For example, authors in [96] develop an ANN-based intrusion detection system for an MCPS. The idea behind their system is to find anomalies in data access patterns and potentially deny access to a user request with an anomalous request for data. Based on this idea, the authors develop an evolving ANN that decides if an incoming

[5] https://www.sisense.com/.

[6] https://powerbi.microsoft.com/en-us/.

[7] https://www.qlik.com/us/.

[8] https://saplumira.com/.

request is normal or an attack; if it is classified as an attack, the request is sent to another ANN to classify the specific type of the attack. Another study presented in [97] uses machine learning algorithms for unusual behavior detection within a healthcare organization network. Their system monitors data access patterns in a computer network and detects anomalous behavior and is able to enhance system performance through feedback given to it by security analysts. Their work includes a visualization phase that helps identify the most valuable nodes for a potential attacker.

7 Future Directions, Open Issues, and Challenges

Incorporating Artificial Intelligence (AI) in MCPSs is still in its infancy with numerous possibilities for further advances in this area. As the amount of accumulated medical data increases, concurrently with the increasing computing power and storage capability of cloud platforms, AI-powered MCPSs will undoubtedly influence the medical field increasingly. While the rich datasets will help improve the accuracy of AI-based algorithms, it will facilitate the collection of much larger quantities of data. This positive feedback cycle will eventually allow the testing of more sophisticated—and data-hungry—algorithms that were not feasible to test previously.

Personalization of the algorithms is also another topic of interest. As the deployment of MCPSs becomes more mainstream, each individual will have a more detailed personal medical history. Designing AI algorithms that are adaptable to a given individual and providing their analysis based on specifics of one's medical history is an open issue which has to be studied further.

Security of an MCPS and keeping medical records private during the processing of medical data is one of the most important challenges that should be considered in all layers of an MCPS. This may lead into the emergence of AI algorithms that can be coupled with advanced encryption schemes, such as homomorphic encryption to keep the medical data secure at all times, even if the algorithm is executed on a public cloud server.

Another challenge associated with MCPSs is the power consumption of the sensors in the data acquisition layer. Although power consumption of the layers that are connected to the grid are minimally affected from this constraint, layers that operate on batteries impose severe limits to the design of an MCPS. This power consumption constraint manifests itself both in the first layer, where the battery-operated sensors acquire the data, and the battery-operated pre-processing layer, where the nodes process data at the local nodes and all of the local and even long-range communication links are powered by batteries. Designing a system that maximizes the battery life is crucial in an MCPS.

Adaptability of algorithms with newer types of data is another aspect that can be studied. As newer sensors are developed and novel methods of sensing are introduced that can be used in everyday situations, the acquired signal from these new sensors may be different than the traditional signals. Making the algorithms adapt to newer—and more advanced—sensors is another topic of interest.

8 Summary and Concluding Remarks

In this paper, we review different aspects of MCPSs and elaborate on incorporating artificial intelligence into them. We outline the general structure of an MCPS, which consists of multiple components. These components are (i) data acquisition, aggregation, and preprocessing, (ii) data processing and decision support, and (iii) visualization and user interaction.

Component (i) is responsible for acquiring patient data, extracting features from this data, aggregating it, and preparing it for transmission into the cloud. We provide a set of algorithms that enable the extraction of features from raw data. Table 2 provides a list of this set of algorithms.

Component (ii) includes the machine intelligence algorithms that process the summarized data from the previous component to prepare it for presentation to the end user. We discuss a rich set of machine intelligence algorithms that reside in this component and categorize them based on their goal. These goals are categorized into knowledge discovery, classification, regression, and sequential decision-making. Examples of algorithms falling into each of these categories are also presented. A summarized list of these algorithms is shown in Table 3.

The final component (iii) is the interface between the machine intelligence and healthcare professionals. We discuss different techniques on providing feedback to the users through data visualization with example applications. We also study the issues that relate to the privacy and security of the personal medical data that is being processed by the MCPS; we provide information about system-level and crypto-level mechanisms that ensure data security and privacy.

Table 2 Algorithms used in data acquisition, aggregation, and pre-processing components

Component stage	Algorithms, methods, and mechanisms
Feature extraction	Principal Component Analysis (PCA) Kernel PCA (KPCA) Canonical Correlation Analysis (CCA) Multidimensional Scaling (MDS) Artificial Neural Networks (ANNs)
Feature selection	Sequential Backward Selection (SBS) Sequential Forward Selection (SFS) Sequential Forward Floating Selection (SFFS) Correlation-based Feature Selection (CFS) Genetic Algorithms (GA)

Table 3 Algorithms used in the decision support component, broken down by goal

Algorithmic goal	Algorithms, methods, and mechanisms
Knowledge discovery	K-Means Hierarchical Clustering Probabilistic Clustering
Classification	Support Vector Machines (SVM) k-Nearest Neighbor Decision Tree AdaBoost Convolutional Neural Networks (CNNs) Recurrent Neural Networks (RNNs)
Regression/estimation	Kalman Filters Linear/Nonlinear Regression Deep Neural Networks (DNNs)
Sequential decision-making	Markov Decision Processes (MDPs) Partially Observable MDPs (POMDPs) Multi Armed Bandits (MABs)

References

1. M. Hassanalieragh, A. Page, T. Soyata, G. Sharma, M.K. Aktas, G. Mateos, B. Kantarci, S. Andreescu, Health monitoring and management using Internet-of-Things (IoT) sensing with cloud-based processing: opportunities and challenges, in *2015 IEEE International Conference on Services Computing (SCC)*, New York (June 2015), pp. 285–292
2. X. Chen, Z. Zhu, M. Chen, Y. Li, Large-scale mobile fitness app usage analysis for smart health. IEEE Commun. Mag. **56**(4), 46–52 (2018)
3. P. Wu, M.Y. Nam, J. Choi, A. Kirlik, L. Sha, R.B. Berlin, Supporting emergency medical care teams with an integrated status display providing real-time access to medical best practices, workflow tracking, and patient data. J. Med. Syst. **41**(12), 186 (2017)
4. J. Jezewski, A. Pawlak, K. Horoba, J. Wrobel, R. Czabanski, M. Jezewski, Selected design issues of the medical cyber-physical system for telemonitoring pregnancy at home. Microprocess. Microsyst. **46**, 35–43 (2016)
5. G. Honan, A. Page, O. Kocabas, T. Soyata, B. Kantarci, Internet-of-everything oriented implementation of secure Digital Health (D-Health) systems, in *Proceedings of the 2016 IEEE Symposium on Computers and Communications (ISCC)*, Messina (Jun 2016), pp. 718–725
6. A. Page, S. Hijazi, D. Askan, B. Kantarci, T. Soyata, Research directions in cloud-based decision support systems for health monitoring using Internet-of-Things driven data acquisition. Int. J. Serv. Comput. **4**(4), 18–34 (2016)
7. 104th Congress Public Law 191, Health Insurance Portability and Accountability Act of 1996 (1996). https://www.gpo.gov/fdsys/pkg/PLAW-104publ191/html/PLAW-104publ191.htm. Accessed 28 July 2017
8. O. Kocabas, T. Soyata, M.K. Aktas, Emerging security mechanisms for medical cyber physical systems. IEEE/ACM Trans. Comput. Biol. Bioinform. **13**(3), 401–416 (2016)

9. G. Yang, L. Xie, M. Mäntysalo, X. Zhou, Z. Pang, L. Da Xu, S. Kao-Walter, Q. Chen, L.R. Zheng, A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box. IEEE Trans. Ind. Inf. **10**(4), 2180–2191 (2014)
10. D.M. West, How 5G technology enables the health internet of things. Brookings Center for Technology Innovation **3**, 1–20 (2016)
11. A. Mdhaffar, T. Chaari, K. Larbi, M. Jmaiel, B. Freisleben, IoT-based health monitoring via lorawan, in *IEEE EUROCON 2017-17th International Conference on Smart Technologies* (IEEE, Piscataway, 2017), pp. 519–524
12. A. Page, T. Soyata, J. Couderc, M. Aktas, B. Kantarci, S. Andreescu, Visualization of health monitoring data acquired from distributed sensors for multiple patients, in *IEEE Global Telecommunications Conference (GLOBECOM)*, San Diego (Dec 2015), pp. 1–7
13. S. Aust, R.V. Prasad, I.G. Niemegeers, IEEE 802.11ah: advantages in standards and further challenges for sub 1 GHz Wi-Fi, in *2012 IEEE International Conference on Communications (ICC)* (IEEE, Piscataway, 2012), pp. 6885–6889
14. S. Han, Y.H. Wei, A.K. Mok, D. Chen, M. Nixon, E. Rotvold, Building wireless embedded internet for industrial automation, in *IECON 2013-39th Annual Conference of the IEEE Industrial Electronics Society* (IEEE, Piscataway, 2013), pp. 5582–5587
15. G. Mokhtari, Q. Zhang, G. Nourbakhsh, S. Ball, M. Karunanithi, Bluesound: a new resident identification sensor—using ultrasound array and BLE technology for smart home platform. IEEE Sens. J. **17**(5), 1503–1512 (2017)
16. W.L. Chen, L.B. Chen, W.J. Chang, J.J. Tang, An IoT-based elderly behavioral difference warning system, in *2018 IEEE International Conference on Applied System Invention (ICASI)* (IEEE, Piscataway, 2018), pp. 308–309
17. Y. Li, Z. Chi, X. Liu, T. Zhu, Passive-ZigBee: enabling ZigBee communication in IoT networks with 1000x+ less power consumption, in *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems* (ACM, New York, 2018), pp. 159–171
18. A.M. Rahmani, T.N. Gia, B. Negash, A. Anzanpour, I. Azimi, M. Jiang, P. Liljeberg, Exploiting smart e-health gateways at the edge of healthcare Internet-of-Things: a fog computing approach. Futur. Gener. Comput. Syst. **78**, 641–658 (2018)
19. M.L. Raymer, W.F. Punch, E.D. Goodman, L.A. Kuhn, A.K. Jain, Dimensionality reduction using genetic algorithms. IEEE Trans. Evol. Comput. **4**(2), 164–171 (2000)
20. Y. Chen, M. Yang, X. Chen, B. Liu, H. Wang, S. Wang, Sensorineural hearing loss detection via discrete wavelet transform and principal component analysis combined with generalized eigenvalue proximal support vector machine and Tikhonov regularization. Multimed. Tools Appl. **77**(3), 3775–3793 (2018)
21. A. Ghandeharioun, S. Fedor, L. Sangermano, D. Ionescu, J. Alpert, C. Dale, D. Sontag, R. Picard, Objective assessment of depressive symptoms with machine learning and wearable sensors data, in *Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII)*, San Antonio (2017)
22. Y. Kim, N. Kaongoen, S. Jo, Hybrid-BCI smart glasses for controlling electrical devices, in *2015 54th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)* (IEEE, Piscataway, 2015), pp. 1162–1166
23. C. Li, W.K. Cheung, J. Liu, J.K. Ng, Bayesian nominal matrix factorization for mining daily activity patterns, in *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)* (IEEE, Piscataway, 2016), pp. 335–342
24. G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006)
25. Y. Kim, H. Lee, E.M. Provost, Deep learning for robust feature generation in audiovisual emotion recognition, in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, Piscataway, 2013), pp. 3687–3691
26. B., Jokanović, M. Amin, Fall detection using deep learning in range-Doppler radars. IEEE Trans. Aerosp. Electron. Syst. **54**(1), 180–189 (2018)

27. M. Li, V. Rozgic, G. Thatte, S. Lee, A. Emken, M. Annavaram, U. Mitra, D. Spruijt-Metz, S. Narayanan, Multimodal physical activity recognition by fusing temporal and cepstral information. IEEE Trans. Neural Syst. Rehabil. Eng. **18**(4), 369–380 (Aug 2010)
28. A. Sano, R.W. Picard, Stress recognition using wearable sensors and mobile phones, in *IEEE Humane Association Conference on Affective Computing and Intelligent Interaction (ACII)* (2013), pp. 671–676
29. B. Xie, H. Minn, Real-time sleep apnea detection by classifier combination. IEEE Trans. Inf. Technol. Biomed. **16**(3), 469–477 (2012)
30. V. Srinivasan, C. Eswaran, N. Sriraam, Artificial neural network based epileptic detection using time-domain and frequency-domain features. J. Med. Syst. **29**(6), 647–660 (2005)
31. B. Lei, S.A. Rahman, I. Song, Content-based classification of breath sound with enhanced features. Neurocomputing **141**, 139–147 (2014)
32. D. Sow, A. Biem, M. Blount, M. Ebling, O. Verscheure, Body sensor data processing using stream computing, in *Proceedings of the International Conference on Multimedia Information Retrieval* (ACM, New York, 2010), pp. 449–458
33. S. Souli, Z. Lachiri, Audio sounds classification using scattering features and support vectors machines for medical surveillance. Appl. Acoust. **130**, 270–282 (2018)
34. A. Page, T. Soyata, J. Couderc, M.K. Aktas, An open source ECG clock generator for visualization of long-term cardiac monitoring data. IEEE Access **3**, 2704–2714 (2015)
35. C.M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)* (Springer, New York, 2006)
36. D. Sánchez-Morillo, M. López-Gordo, A. León, Novel multiclass classification for home-based diagnosis of sleep apnea hypopnea syndrome. Expert Syst. Appl. **41**(4), 1654–1662 (2014)
37. D.S. Lee, T.W. Chong, B.G. Lee, Stress events detection of driver by wearable glove system. IEEE Sens. J. **17**(1), 194–204 (2017)
38. W.H. Wang, Y.L. Hsu, P.C. Chung, M.C. Pai, Predictive models for evaluating cognitive ability in dementia diagnosis applications based on inertia-and gait-related parameters. IEEE Sens. J. **18**(8), 3338–3350 (2018)
39. M. Mursalin, Y. Zhang, Y. Chen, N.V. Chawla, Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier. Neurocomputing **241**, 204–214 (2017)
40. B. Nakisa, M.N. Rastgoo, D. Tjondronegoro, V. Chandran, Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors. Expert Syst. Appl. **93**, 143–155 (2017)
41. H. Li, D. Yuan, X. Ma, D. Cui, L. Cao, Genetic algorithm for the optimization of features and neural networks in ECG signals classification. Sci. Rep. **7**, 41011 (2017)
42. V. Chandola, S.R. Sukumar, J.C. Schryver, Knowledge discovery from massive healthcare claims data, in *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, New York, 2013), pp. 1312–1320
43. B.M. Marlin, D.C. Kale, R.G. Khemani, R.C. Wetzel, Unsupervised pattern discovery in electronic health care data using probabilistic clustering models, in *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium* (ACM, New York, 2012), pp. 389–398
44. D.P. Chen, S.C. Weber, P.S. Constantinou, T.A. Ferris, H.J. Lowe, A.J. Butte, Clinical arrays of laboratory measures, or “clinarrays”, built from an electronic health record enable disease subtyping by severity, in *AMIA* (2007)
45. D. Sanchez-Morillo, M.A. Fernandez-Granero, A.L. Jiménez, Detecting COPD exacerbations early using daily telemonitoring of symptoms and k-means clustering: a pilot study. Med. Biol. Eng. Comput. **53**(5), 441–451 (2015)
46. N.P. Tatonetti, J.C. Denny, S.N. Murphy, G.H. Fernald, G. Krishnan, V. Castro, P. Yue, P.S. Tsau, I. Kohane, D.M. Roden, et al., Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin. Pharmacol. Ther. **90**(1), 133 (2011)

47. B. Ustun, M.B. Westover, C. Rudin, M.T. Bianchi, Clinical prediction models for sleep apnea: the importance of medical history over symptoms. J. Clin. Sleep Med. Off. Publ. Am. Acad. Sleep Med. **12**(2), 161–168 (2016)
48. S. Hijazi, A. Page, B. Kantarci, T. Soyata, Machine learning in cardiac health monitoring and decision support. IEEE Comput. Mag. **49**(11), 38–48 (2016)
49. A. Page, M.K. Aktas, T. Soyata, W. Zareba, J. Couderc, "QT Clock" to improve detection of QT prolongation in long QT syndrome patients. Heart Rhythm **13**(1), 190–198 (2016)
50. H.S. Mousavi, V. Monga, G. Rao, A.U.K. Rao, et al., Automated discrimination of lower and higher grade gliomas based on histopathological image analysis. J. Pathol. Inform. **6**(1), 15 (2015)
51. E. Ataer-Cansizoglu, V. Bolon-Canedo, J.P. Campbell, A. Bozkurt, D. Erdogmus, J. Kalpathy-Cramer, S. Patel, K. Jonas, R.V.P. Chan, S. Ostmo, et al., Computer-based image analysis for plus disease diagnosis in retinopathy of prematurity: performance of the "i-ROP" system and image features associated with expert diagnosis. Transl. Vis. Sci. Technol. **4**(6), 5–5 (2015)
52. I. Bisio, F. Lavagetto, M. Marchese, A. Sciarrone, A smartphone-centric platform for remote health monitoring of heart failure. Int. J. Commun. Syst. **28**(11), 1753–1771 (2015)
53. M. Bsoul, H. Minn, L. Tamil, Apnea MedAssist: real-time sleep apnea monitor using single-lead ECG. IEEE Trans. Inf. Technol. Biomed. **15**(3), 416–427 (2011)
54. D. Zhou, J. Luo, V.M.B. Silenzio, Y. Zhou, J. Hu, G. Currier, H.A. Kautz, Tackling mental health by integrating unobtrusive multimodal sensing, in *AAAI*, 1401–1409 (2015)
55. D.C. Cireşan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Mitosis detection in breast cancer histology images with deep neural networks, in *International Conference on Medical Image Computing and Computer-assisted Intervention* (Springer, Berlin, 2013), pp. 411–418
56. H. Chen, X. Qi, L. Yu, P.A. Heng, DCAN: deep contour-aware networks for accurate gland segmentation (2016). Preprint arXiv:1604.02677
57. V. Gulshan, L. Peng, M. Coram, M.C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, et al., Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA **316**(22), 2402–2410 (2016)
58. S. Kiranyaz, T. Ince, M. Gabbouj, Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Trans. Biomed. Eng. **63**(3), 664–675 (2016)
59. H.C. Shin, K. Roberts, L. Lu, D. Demner-Fushman, J. Yao, R.M. Summers, Learning to read chest X-rays: recurrent neural cascade model for automated image annotation (2016). Preprint arXiv:1603.08486
60. Q. Li, R.G. Mark, G.D. Clifford, Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. Physiol. Meas. **29**(1), 15 (2007)
61. R.E. Kalman, A new approach to linear filtering and prediction problems. J. Basic Eng. **82**(1), 35–45 (1960)
62. R.E. Kalman, R.S. Bucy, New results in linear filtering and prediction theory. J. Basic Eng. **83**(1), 95–108 (1961)
63. P. Schulam, S. Saria, A framework for individualizing predictions of disease trajectories by exploiting multi-resolution structure, in *Advances in Neural Information Processing Systems* (2015), pp. 748–756
64. H. Neuvirth, M. Ozery-Flato, J. Hu, J. Laserson, M.S. Kohn, S. Ebadollahi, M. Rosen-Zvi, Toward personalized care management of patients at risk: the diabetes case study, in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, New York, 2011), pp. 395–403
65. J. Ma, R.P. Sheridan, A. Liaw, G.E. Dahl, V. Svetnik, Deep neural nets as a method for quantitative structure–activity relationships. J. Chem. Inf. Model. **55**(2), 263–274 (2015)
66. Y. Gordienko, S. Stirenko, Y. Kochura, O. Alienin, M. Novotarskiy, N. Gordienko, Deep learning for fatigue estimation on the basis of multimodal human-machine interactions (2017). Preprint arXiv:1801.06048

67. D.S. Zois, M. Levorato, U. Mitra, Energy-efficient, heterogeneous sensor selection for physical activity detection in wireless body area networks. IEEE Trans. Signal Process. **61**(7), 1581–1594 (2013)
68. U. Mitra, B.A. Emken, S. Lee, M. Li, V. Rozgic, G. Thatte, H. Vathsangam, D.S. Zois, M. Annavaram, S. Narayanan, M. Levorato, D. Spruijt-Metz, G. Sukhatme, KNOWME: a case study in wireless body area sensor network design. IEEE Commun. Mag. **50**(5), 116–125 (2012)
69. J. Hoey, C. Boutilier, P. Poupart, P. Olivier, A. Monk, A. Mihailidis, People, sensors, decisions: customizable and adaptive technologies for assistance in healthcare. ACM Trans. Interactive Intell. Syst. **2**(4), 1–36 (2012)
70. P. Paredes, R. Gilad-Bachrach, M. Czerwinski, A. Roseway, K. Rowan, J. Hernandez, PopTherapy: coping with stress through pop-culture, in *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)* (2014), pp. 109–117
71. M. Rabbi, M.H. Aung, T. Choudhury, Towards health recommendation systems: an approach for providing automated personalized health feedback from mobile data, in *Mobile Health* (Springer, Berlin, 2017), pp. 519–542
72. I. Sundin, T. Peltola, M.M. Majumder, P. Daee, M. Soare, H. Afrabandpey, C. Heckman, S. Kaski, P. Marttinen, Improving drug sensitivity predictions in precision medicine through active expert knowledge elicitation (2017). Preprint arXiv:1705.03290
73. D. Chou, Health it and patient safety: building safer systems for better care. JAMA **308**(21), 2282–2282 (2012)
74. A.A. Bui, W. Hsu, Medical data visualization: toward integrated clinical workstations, in *Medical Imaging Informatics* (Springer, Berlin, 2010), pp. 139–193
75. F. Jager, A. Taddei, G.B. Moody, M. Emdin, G. Antolič, R. Dorn, A. Smrdel, C. Marchesi, R.G. Mark, Long-term ST database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. Med. Biol. Eng. Comput. **41**(2), 172–182 (2003)
76. A. Golberger, L. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R. Mark, J. Mietus, G. Moody, P. Chung-Kan, H. Stenley, Physiobank, physiotoolkit, and physionet: component of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2000)
77. K. Xu, S. Guo, N. Cao, D. Gotz, A. Xu, H. Qu, Z. Yao, Y. Chen, ECGLens: interactive visual exploration of large scale ECG data for arrhythmia detection, in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)* (ACM, New York, 2018), Paper 663, 12 pp. https://doi.org/10.1145/3173574.3174237
78. C.A. Christmann, G. Zolynski, A. Hoffmann, G. Bleser, Effective visualization of long term health data to support behavior change, in *Digital Human Modeling. Applications in Health, Safety, Ergonomics, and Risk Management: Health and Safety. DHM 2017*, ed. by V. Duffy. Lecture Notes in Computer Science, vol. 10287 (Springer, Cham, 2017)
79. C.A. Christmann, G. Zolynski, A. Hoffmann, G. Bleser, Effective visualization of long term health data to support behavior change, in *International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management* (Springer, Berlin, 2017), pp. 237–247
80. A. Cuttone, M.K. Petersen, J.E. Larsen, Four data visualization heuristics to facilitate reflection in personal informatics, in *International Conference on Universal Access in Human-Computer Interaction* (Springer, Berlin, 2014), pp. 541–552
81. S. Theis, P. Rasche, C. Bröhl, M. Wille, A. Mertens, User-driven semantic classification for the analysis of abstract health and visualization tasks, in *International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management* (Springer, Berlin, 2017), pp. 297–305
82. K. Tollmar, F. Bentley, C. Viedma, Mobile health mashups: making sense of multiple streams of wellbeing and contextual data for presentation on a mobile device, in *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)* (IEEE, Piscataway, 2012), pp. 65–72

83. S. Stusak, A. Tabard, F. Sauka, R.A. Khot, A. Butz, Activity sculptures: exploring the impact of physical visualizations on running activity. IEEE Trans. Vis. Comput. Graph. **20**(12), 2201–2210 (2014)
84. C. Fan, J. Forlizzi, A.K. Dey, A spark of activity: exploring informative art as visualization for physical activity, in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (ACM, New York, 2012), pp. 81–84
85. R.A. Khot, D. Aggarwal, R. Pennings, L. Hjorth, F. Mueller, Edipulse: investigating a playful approach to self-monitoring through 3D printed chocolate treats, in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (ACM, New York, 2017), pp. 6593–6607
86. F. Jonathan, J. Sonin, hGraph: an open system for visualizing personal health metrics. Involution Studios, Arlington, Tech. Rep. (April 2012)
87. A. Ledesma, M. Al-Musawi, H. Nieminen, Health figures: an open source javascript library for health data visualization. BMC Med. Inform. Decis. Mak. **16**(1), 38 (2016)
88. D. Estrin, I. Sim, Open mHealth architecture: an engine for health care innovation. Science **330**(6005), 759–760 (2010)
89. A.A. Bui, D.R. Aberle, H. Kangarloo, Timeline: visualizing integrated patient records. IEEE Trans. Inf. Technol. Biomed. **11**(4), 462–473 (2007)
90. J. Plourde, D. Arney, J.M. Goldman, OpenICE: an open, interoperable platform for medical cyber-physical systems, in *2014 ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)* (IEEE, Piscataway, 2014), pp. 221–221
91. R. Kamaleswaran, C. Collins, A. James, C. McGregor, PhysioEx: visual analysis of physiological event streams, in *Computer Graphics Forum*, vol. 35 (Wiley Online Library, 2016), pp. 331–340
92. B. Maradani, H. Levkowitz, The role of visualization in tele-rehabilitation: a case study, in *2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence* (IEEE, Piscataway, 2017), pp. 643–648
93. S.H. Koch, C. Weir, D. Westenskow, M. Gondan, J. Agutter, M. Haar, D. Liu, M. Görges, N. Staggers, Evaluation of the effect of information integration in displays for ICU nurses on situation awareness and task completion time: a prospective randomized controlled study. Int. J. Med. Inform. **82**(8), 665–675 (2013)
94. H. Almohri, L. Cheng, D. Yao, H. Alemzadeh, On threat modeling and mitigation of medical cyber-physical systems, in *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)* (IEEE, Piscataway, 2017), pp. 114–119
95. G. Grispos, W.B. Glisson, K.K.R. Choo, Medical cyber-physical systems development: a forensics-driven approach, in *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)* (IEEE, Piscataway, 2017), pp. 108–113
96. N. Mowla, I. Doh, K. Chae, Evolving neural network intrusion detection system for MCPS, in *2017 19th International Conference on Advanced Communication Technology (ICACT)* (IEEE, Piscataway, 2017), pp. 183–187
97. A. Boddy, W. Hurst, M. Mackay, A. El Rhalibi, A study into data analysis and visualisation to increase the cyber-resilience of healthcare infrastructures, in *Proceedings of the 1st International Conference on Internet of Things and Machine Learning (IML '17)* (ACM, New York, 2017), Article 32, 7 pp. https://doi.org/10.1145/3109761.3109793