Research
Artificial Intelligence—Review

Artificial Intelligence in Healthcare: Review and Prediction Case Studies

Guoguang Ronga,b, Arnaldo Mendezc, Elie Bou Assic, Bo Zhao d, Mohamad Sawana,b,c,e

Abstract
Artificial intelligence (AI) has been developing rapidly in recent years in terms of software algorithms, hardware implementation, and applications in a vast number of areas. In this review, we summarize the latest developments of applications of AI in biomedicine, including disease diagnostics, living assistance, biomedical information processing, and biomedical research. The aim of this review is to keep track of new scientific accomplishments, to understand the availability of technologies, to appreciate the tremendous potential of AI in biomedicine, and to provide researchers in related fields with inspiration. It can be asserted that, just like AI itself, the application of AI in biomedicine is still in its early stage. New progress and breakthroughs will continue to push the frontier and widen the scope of AI application, and fast developments are envisioned in the near future. Two case studies are provided to illustrate the prediction of epileptic seizure occurrences and the filling of a dysfunctional urinary bladder.

© 2020 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction
Artificial intelligence (AI) is defined as the intelligence of machines, as opposed to the intelligence of humans or other living species [1,2]. AI can also be defined as the study of “intelligent agents”—that is, any agent or device that can perceive and understand its surroundings and accordingly take appropriate action to maximize its chances of achieving its objectives [3]. AI also refers to situations wherein machines can simulate human minds in learning and analysis, and thus can work in problem solving. This kind of intelligence is also referred to as machine learning (ML) [4].

Typically, AI involves a system that consists of both software and hardware. From a software perspective, AI is particularly concerned with algorithms. An artificial neural network (ANN) is a conceptual framework for executing AI algorithms [5]. It is a mimic of the human brain—an interconnected network of neurons, in which there are weighted communication channels between neurons [6]. One neuron can react to multiple stimuli from neighboring neurons and the whole network can change its state according to different inputs from the environment [7]. As a result, the neural network (NN) can generate outputs as its responses to environmental stimuli—just as the human brain reacts to different environmental changes. NNs are typically layered structures of various configurations. Researchers have devised NNs that can do (1) supervised learning, where the task is to infer a function that maps an input to an output based on example pairs of inputs and outputs; (2) unsupervised learning, where the task is to learn from test data that has not been labeled, classified, or categorized, in order to identify common features in the data and, rather than responding to system feedback, to react based on the existence or inexistence of identified common features in new data; and (3) reinforced learning, where the task is to act within the given surroundings in order to maximize rewards and minimize penalties, both according to some form of accumulative nature [8]. With the advancement of computation power, NNs have become “deeper,” meaning that more layers of neurons are involved in the network to mimic a human brain and carry out learning. In addition, more functions can be incorporated into the NN, such as merging feature extraction and classification functions into a single deep network—hence the technical term “deep learning” [9,10].

From a hardware perspective, AI is mainly concerned with the implementation of NN algorithms on a physical computation platform. The most straightforward approach is to implement

https://doi.org/10.1016/j.eng.2019.08.015
2095-8099/© 2020 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
NN algorithm on a general-purpose central processing unit (CPU), in a multithread or multicore configuration [7]. Furthermore, graphical processing units (GPUs), which are good at convolutional computations, have been found to be advantageous over CPUs for large-scale NNs [11]. CPU and GPU co-processing has turned out to be more efficient than CPU alone, especially for spiking NNs [12,13]. Moreover, some programmable or customizable accelerator hardware platforms, such as field-programmable gate arrays (FPGAs) and application-specific integrated circuits (ASICs), can implement NNs toward a customized application in a more efficient way, in terms of computation capability, power efficiency, and form factor [14]. Compared with GPU and CPU, these platforms can be customized for a specific application and thus can be more power efficient and compact than GPU and CPU platforms. To deploy AI in edge devices, such as mobile phones in wireless networks or sensor nodes in the Internet of Things (IoT), further improvements in power efficiency and form factor are needed. Researchers have tried to implement AI algorithms using analog integrated circuits [15,16], spintronics [17], and memristors [18–20]. Some of these new platforms, such as memristor crossbar circuits [21], can merge computation with memory and thus avoid the problem of access to the “memory wall” of traditional von Neumann architectures. This access is mandatory in order to update needed parameters. Recently, researchers have tried to improve the efficiency of AI implementation by reducing the number of bits used for data representation. It turns out that the computation accuracy can be maintained when the data representation goes from 32 or 16 bits down to 8 bits. The advantage is faster computation, less power, and smaller form factor [22]. However, the “memory wall” limits remain. On the other hand, the adoption of appropriate training methods (e.g., deep training instead of surface-level training [23] or using pre-training techniques [24]) and the use of balanced datasets [25], sufficient amounts of data [26], and constant availability [27] of datasets are important factors to consider in order to achieve satisfactory performance of ANNs.

Due to the rapid development of AI software and hardware technologies, AI has been applied in various technical fields, such as the IoT [28], machine vision [29], autonomous driving [30,31], natural language processing [32,33], and robotics [34]. Most interestingly, researchers in the biomedical fields have been actively trying to apply AI to help improve analysis and treatment outcomes and, consequently, increase the efficacy of the overall healthcare industry [35–37]. Fig. 1 shows the number of publications in this area in the last 20 years, from 1999 to 2018. The growth of interest is obvious, especially in the last five years, and continued growth in future can be forecast. The benefits that AI can offer to biomedicine were envisioned a couple of decades earlier [38]. In fact, reviews have been published on the role of AI in biomedical engineering [36,37]. More recently, new progress has been made in AI and its applications in biomedicine.

This paper reviews recent breakthroughs in the application of AI in biomedicine, covering the main areas in biomedical engineering and healthcare.

The goal for healthcare is to become more personal, predictive, preventative, and participatory, and AI can make major contributions in these directions. From an overview of the progress made, we estimate that AI will continue its momentum to develop and mature as a powerful tool for biomedicine. The remainder of this paper is oriented toward the main AI applications. Section 2 includes a description of information processing and algorithm implementation, while Section 3 focuses on disease diagnostics and prediction. Case studies of the prediction of two medical diseases are reported in Section 4. Finally, conclusions are summarized in Section 5.

2. Information processing and algorithm implementation

The main applications of AI in biomedicine can be divided into four categories. The first three categories described in this section are intended to efficiently treat big data and provide quick access to data in order to solve issues related to healthcare. These applications deal with living assistance for elderly and disabled people, natural language processing techniques, and fundamental research activities. The last category of AI applications concerns the diagnosis and prediction of diseases and will be analyzed in Section 3.

2.1. AI for living assistance

In the area of assisted living for elderly and disabled people, AI applications using corresponding smart robotic systems are paving the way for improvements in life quality. An overview of smart home functions and tools offered for people with loss of autonomy (PLA), and intelligent solution models based on wireless sensor networks, data mining, and AI was published recently [39]. NNs can be trained with specific image-processing steps to recognize human facial expressions as commands. Furthermore, human–machine interfaces (HMs) based on facial expression analysis allow people with disabilities to control wheelchairs and robot assistance vehicles without a joystick or sensors attached to the body [40].

An “ambient intelligent system” called RUDO can help people who are blind to live together with sighted people and work in specialized fields such as informatics and electronics [41]. People who are blind can make use of multiple functions of this intelligent assistant through a single user interface. A “smart assistant” based on AI can help pregnant women with dietary and other necessary measures [42]. An advanced level through its own intelligence, combined with “cloud-based communication media between all people concerned” [43].

A fall-detection system based on radar Doppler time–frequency signatures and a sparse Bayesian classifier can reduce fall risks and complications for seniors [43]. In fact, “smart communication architecture” systems for “ambient assisted living” (AAL) have been developed to allow AI processing information to be gathered from different communication channels or technologies, and thus to determine the occurrences of events in the network environment and the assistance needs of elderly people [44]. The “ambient intelligence” of smart homes can provide activity awareness and sensor data assistance to elderly people such that AAL environments allow “aging in place”—that is, aging at home. For example, the activity-aware screening of activity limitation and safety awareness (SALSA) intelligent agent can help elders with daily medication activities [45]. ML in motion analysis and gait study can raise an alarm at hazardous actions and activate preventative measures [46,47]. Fig. 2 [39] shows the model for AAL.

In this scenario, sensors collect data about the ambient environment and human behavior, which is then analyzed by cloud computing or edge intelligence. A decision is then made regarding what actions are necessary, and this decision is used to activate alarms or preventative measures. An expert system based on AI, in conjunction with mobile devices and personal digital assistants (PDAs), can help persons with lasting memory damage by enhancing their memory capabilities so as to achieve independent daily living [48]. This is a significant extension of the expert system for memory rehabilitation (ES-MR) for “non-expert” users.

2.2. AI in biomedical information processing

Breakthroughs have been made in natural language processing for biomedical applications. In the area of biomedical question
mechanisms, protein–protein interaction information extraction in biomedical research and innovation activities [58, 59]. In this regard, AI can accelerate the screening and indexing of academic literature [51]. On a global scale, AI is recognized as a powerful tool in biomedical research [57].

On a global scale, AI is recognized as a powerful tool in biomedical research [57].

For biomedical information collected from different sources over an elongated period of time, many important tasks can dominate; these are clinical information merging, comparison, and conflict resolution [53]. These have long been time-consuming, labor-intensive, and unsatisfying tasks performed by humans. To improve efficiency and accuracy, AI has been demonstrated to be capable of performing these tasks with results as accurate as professional evaluators can do [54]. Also, natural language processing of medical narrative data is needed to free humans from the challenging task of keeping track of temporal events while simultaneously maintaining structures and reasons [55]. ML can be used to process high-complexity clinical information, which originates from word sentiment analysis—can work effectively toward information extraction from binary answers [52].

For biomedical information collected from different sources over an elongated period of time, many important tasks can dominate; these are clinical information merging, comparison, and conflict resolution [53]. These have long been time-consuming, labor-intensive, and unsatisfying tasks performed by humans. To improve efficiency and accuracy, AI has been demonstrated to be capable of performing these tasks with results as accurate as professional evaluators can do [54]. Also, natural language processing of medical narrative data is needed to free humans from the challenging task of keeping track of temporal events while simultaneously maintaining structures and reasons [55]. ML can be used to process high-complexity clinical information, which originates from word sentiment analysis—can work effectively toward information extraction from binary answers [52].

For biomedical information collected from different sources over an elongated period of time, many important tasks can dominate; these are clinical information merging, comparison, and conflict resolution [53]. These have long been time-consuming, labor-intensive, and unsatisfying tasks performed by humans. To improve efficiency and accuracy, AI has been demonstrated to be capable of performing these tasks with results as accurate as professional evaluators can do [54]. Also, natural language processing of medical narrative data is needed to free humans from the challenging task of keeping track of temporal events while simultaneously maintaining structures and reasons [55]. ML can be used to process high-complexity clinical information, which originates from word sentiment analysis—can work effectively toward information extraction from binary answers [52].

For biomedical information collected from different sources over an elongated period of time, many important tasks can dominate; these are clinical information merging, comparison, and conflict resolution [53]. These have long been time-consuming, labor-intensive, and unsatisfying tasks performed by humans. To improve efficiency and accuracy, AI has been demonstrated to be capable of performing these tasks with results as accurate as professional evaluators can do [54]. Also, natural language processing of medical narrative data is needed to free humans from the challenging task of keeping track of temporal events while simultaneously maintaining structures and reasons [55]. ML can be used to process high-complexity clinical information, which originates from word sentiment analysis—can work effectively toward information extraction from binary answers [52].

In addition to being able to act as an “eDoctor” for disease diagnosis, management, and prognosis, AI has unexplored usage as a powerful tool in biomedical research [57]. On a global scale, AI can accelerate the screening and indexing of academic literature in biomedical research and innovation activities [58, 59]. In this direction, the latest research topics include tumor-suppressor mechanisms [60], protein–protein interaction information extraction [61], the generation of genetic association of the human genome to assist in transferring genome discoveries to healthcare practices [62], and so forth. Furthermore, biomedical researchers can efficiently accomplish the demanding task of summarizing the literature on a given topic of interest with the help of a semantic graph-based AI approach [63]. Moreover, AI can help biomedical researchers to not only search but also rank the literature of interest when the number of research papers is beyond readability. This allows researchers to formulate and test to-the-point scientific hypotheses, which are a very important part of biomedical research. For example, researchers can screen and rank figures of interest in the increasing volume of literature [64] with the help of an AI to formulate and test hypotheses.

Some intelligent medical devices are becoming increasingly “conscious” [65, 66], and this consciousness can be explored in biomedical research. An intelligent agent called the computational modeling assistant (CMA) can help biomedical researchers to construct “executable” simulation models from the conceptual models they have in mind [67]. Fig. 3 [67] shows a general view of the process flow and interactions between a CMA and human researchers. The CMA is provided with various knowledge, methods, and databases. The researcher hypothesis is expressed in the form of biological models, which are supplied as input to the CMA. The intelligence of the CMA allows it to integrate all this knowledge and these models, and it transforms the hypothesis of the researchers into concrete simulation models. The researcher then reviews and selects the best models and the CMA generates simulation codes for the selected models. In this way, the CMA enables a significantly accelerated research process and enhanced productivity. In addition, some intuitive machines could guide scientific research in fields such as biomedical imaging, oral surgery, and plastic surgery [68, 69]. Human and machine consciousness and its relevance to biomedical engineering have been discussed in order to better understand the impact of this development [70].

3. Disease diagnostics and prediction

The most urgent need for AI in biomedicine is in the diagnostics of diseases. A number of interesting breakthroughs have been made in this area. AI allows health professionals to give earlier and more accurate diagnostics for many kinds of diseases [71]. One major class of diagnosis is based on in vitro diagnostics using biosensors or biochips. For example, gene expression, which is a very important diagnostic tool, can be analyzed by ML, in which AI interprets microarray data to classify and detect abnormalities
One new application is to classify cancer microarray data for cancer diagnosis [74]. With integrated AI, biosensors and related point-of-care testing (POCT) systems can diagnose cardiovascular diseases in the early stage [75]. In addition to diagnosis, AI can help to predict the survival rates of cancer patients, such as colon cancer patients [76]. Researchers have also identified some limitations of ML in biomedical diagnosis and have suggested steps to minimize the effects of these drawbacks [77]. Thus, there is still much potential for AI in diagnostics and prognostics.

Another important class of disease diagnostic is based on medical imaging (two-dimensional) and signal (one-dimensional) processing. Such techniques have been employed in the diagnosis, management, and prediction of illnesses. For one-dimensional signal processing, AI has been applied to biomedical signal feature extraction [78], such as electroencephalography (EEG) [79], electromyography (EMG) [80], and electrocardiography (ECG) [81]. An important application of EEG is epileptic seizure prediction. It is very important to predict seizures so as to minimize their impact on patients [82]. In recent years, AI has been recognized as one of the key elements of an accurate and reliable prediction system [83,84]. It is now possible to predict by means of deep learning [85], and the prediction platform can be deployed in a mobile system [86]. AI can also play an important role in diagnosis based on biomedical image processing [87]. AI has been utilized in image segmentation [88], multidimensional imaging [89], and thermal imaging [90] to improve image quality and analysis efficiency. AI can also be deployed in portable ultrasonic devices, so that untrained personal can use ultrasound as a powerful tool to diagnose many kinds of illnesses in undeveloped regions [91].

In addition to the above applications, AI can assist standard decision support systems (DSSs) [92,93] to improve diagnostic accuracy and facilitate disease management in order to reduce the burden on personnel. For example, AI has been used in the integrated management of cancer [94], to support the diagnosis and management of tropical diseases [95] and cardiovascular diseases [96], and to support the decision-making process of diagnostics [92]. These applications demonstrate that AI can be a powerful tool for early and accurate diagnostics, management, and even prediction of diseases and patient conditions. Two case studies are illustrated below.

4. Healthcare

AI is now covering a wide range of healthcare applications [97]. In particular, it has been used for signal and image processing, and for predictions of function changes such as in urinary bladder control [98], epileptic seizures [99], and stroke predictions [100]. Below, we describe two typical case studies: bladder volume prediction and epileptic seizure prediction.

4.1. Bladder volume prediction

When the storage and urination functions of the bladder fail as a result of spinal cord injury or because of other neurological diseases, health status, or aging, various complications occur in the patient’s health conditions. Nowadays, partial restoration of bladder function in drug-refractory patients can be achieved using implantable neural stimulators. To improve the efficiency and safety of neuroprostheses through conditional neurostimulation [101], a bladder sensor that detects stored urine is required as a
feedback device that applies electrical stimulation only when needed. The sensor can also be used to notify patients with impaired sensations in a timely manner when the bladder needs to be emptied or when an abnormally high residual post-micturition volume remains after an incomplete voiding.

We have proposed new methods [102] and developed a dedicated digital signal processor (DSP) [103] for sensing both the pressure and its fullness in urine by using afferent neural activities from the regular neural roots of the bladder (i.e., mechanoreceptors), which depicts the changes during filling.

Smart neuroprostheses are typically composed of two units: an internal unit implanted in the patient and an external unit, which is usually carried as a wearable device. Both units are typically connected by a wireless link that conveys data and provides power to the implant. The internal unit performs several functions: neural signal recording; on-chip processing (to a variable extent, depending on the application) of the signals conveying sensory information; neurostimulation of appropriated nerves using functional electrical stimulation (FES) techniques; logical control of the implanted unit functions; and communication with the external unit. When additional signal processing is needed to execute more complex algorithms that require extra computing capabilities, which are not suitable for implantation due to the size, power consumption, temperature rise, electromagnetic emission, and so forth, the internal unit sends the recorded signal to the external unit, which executes more complex algorithms and sends back appropriated neurostimulation commands. The external base station integrates the implant–user interface and the implant–computer interface for greater flexibility.

The work presented in this section is aimed at the design of an effective implantable volume and pressure sensor for the urinary bladder that is capable of providing the necessary feedback to the neuroprosthesis. This sensor can eventually be used for either implementing a conditional neurostimulation approach to restore bladder functions or for bladder-fullness sensing in patients with impaired sensations due to the several diseases and conditions mentioned above. To better meet patients' needs regarding bladder neuroprosthetic devices, we chose to implement within the implantable unit an optimized DSP capable of decoding in real time the bladder pressure and volume. This approach greatly influenced the choice of the most suitable prediction methods, which are described below.

Qualitative and quantitative monitoring methods based on ML algorithms were proposed to decode the afferent neural activity produced by the bladder natural sensors (i.e., mechanoreceptors responding to bladder wall stretching). To implement these methods, the neural activity recorded by the implantable unit must be detected, discriminated (i.e., classified), and then decoded in real time. The proposed qualitative method decodes the bladder fullness into just three levels (i.e., low, medium, and high). This method significantly reduces the number of operations, using fewer hardware resources, so that the power consumption is minimized. The quantitative method needs more complex algorithms to compute the bladder volume or pressure to feedback the closed-loop system for the neurostimulation, but it is optimized to run in the implanted unit with minimum power consumption, thanks to customized hardware.

The monitoring algorithms used for the qualitative and quantitative methods first perform an offline learning phase (also known as a training phase) in real time. During this phase, the sensor learns or identifies the parameters intended for real-time monitoring. For the learning phase, we can choose the best algorithms, regardless of their complexity and execution time, because they are executed offline in a computer connected to the implant through the external unit. Thanks to the learning phase, we can shift the complexity and hardware burden to offline processing, allowing the real-time monitoring phase to be implemented with lower complexity yet effective prediction algorithms and optimized power consumption. The learning phase includes:

---

Please cite this article as: G. Rong, A. Mendez, E. Bou Assi et al., Artificial Intelligence in Healthcare: Review and Prediction Case Studies, Engineering, https://doi.org/10.1016/j.eng.2019.08.015
(1) Digital data conditioning, which involves band-pass filtering with a non-causal linear-phase finite impulse response (FIR) filter;
(2) Identification of the afferent neural activity with the best correlation with bladder volume and/or pressure (Unit 1 in the example shown in Fig. 4). The Spearman’s rank correlation coefficient (\( \rho \)) was used instead of the Pearson (linear) correlation coefficient because the former assesses a monotonic dependence, which is not necessarily linear, which improves the robustness of our estimation method. Eq. (1) was used to compute the Spearman’s rank correlation coefficient:

\[
\rho_k = \frac{\sum_{i=1}^{n} (FR_{i,k} - \bar{FR}_k) (V_{i,k} - \bar{V}_k)}{\sqrt{\sum_{i=1}^{n} (FR_{i,k} - \bar{FR}_k)^2 \sum_{i=1}^{n} (V_{i,k} - \bar{V}_k)^2}}
\]

where \( \rho_k \) is the Spearman’s coefficient of the class \( k \) unit; \( k \) is the number of classes detected; \( i \) is the counter of timeframes; \( n \) is the number of the timeframes used along the recorded signals, henceforward referred to as bins; \( FR_{i,k} \) is the used values of the units’ firing rate (FR) per second; \( V_{i,k} \) is the mean volume for the same bin; \( \bar{FR}_k \) and \( \bar{V}_k \) are the means of all FRs and volume bins relative to class \( k \), respectively;

(3) Computing the background neural activity, where the background baseline is computed by the averaging of the neural FR for 60 s before the bladder is filled with saline;

(4) Volume curve quantization using bins, which results in a finite number of bins with a properly chosen width, henceforward defined as the bin width (BW);

(5) FR integration, which involves computing of the FR within the bins (i.e., the number of spikes detected within the bin for the selected unit), henceforward called the bin-integrated rate (BIR);

(6) Volume or pressure prediction, for which, as mentioned above, we used two algorithms for bladder volume or pressure prediction: a qualitative one to predict three levels of bladder fullness and another to quantify it. A short description of both algorithms is presented below.

The qualitative volume or pressure prediction algorithm predicts three levels representing bladder fullness (low, medium, and high): low volume (a comfortable level), the need-to-void medium volume level (within some predefined time), and the urge-to-void high volume (where there is a risk of urine leaking). The programmed thresholds correspond to 0.25, 0.5, and 1.0 times the full capacity of the bladder volume. By using linear-regression-based algorithms in the learning phase, we computed the BIR for these three degrees of fullness: BIR0.25, BIR0.5, and BIR1.0. Finally, the qualitative volume or pressure prediction was realized by finding the minimum distance within the real-time computed BIR and the stored reference values (BIR0.25, BIR0.5, and BIR1.0). Based on this method, each bin was assigned to one of the three fullness levels. To find the optimal bin length, the BW was swept using different intervals and the length yielding the best (i.e., lowest) qualitative estimation error \((E_{\text{qual}})\) was adopted using Eq. (2). To compute \( E_{\text{qual}} \) in Eq. (2), we calculated the overall success rate (OSR) \([104]\), which is the ratio of all good classifications of the states over all performed classifications. The OSR is evaluated by adding \( B_i \), the number of bins for which the estimated state matches the actual one, and dividing by the total number of bins \((n)\).

\[
E_{\text{qual}} = 1 - \text{OSR} = 1 - \frac{\sum_{i=1}^{n} B_i}{n}
\]

To implement the quantitative volume and pressure estimation using a frugal amount of hardware resources, a regression method model was used, as shown in Eq. (3):

\[
\hat{V} = \sum_{i=0}^{N} (c_i \times \text{BIR}_i)
\]

where \( \hat{V} \) is the estimated volume, \( c_i \) represents the regression model coefficients. We used the bisquare robust fitting method to compute \( c_i \) and minimize the impact of the outlier values \([105]\). The optimal order \((N)\) for the regression model, shown in Eq. (3), was found by running simulations (trials) with real neural recordings from animal models. By running several simulation trials, we selected the minimal \( N \) and the shortest BW, which yielded the lowest estimation error.

Fig. 4. Bladder afferent neural activity recordings (ENG) during a slow filling. The spike’s raster of three units identified is shown. Unit 1 activity exhibited the best correlation with the bladder volume in this example. ENG: electroneurogram; \( P_{\text{ves}} \): vesical pressure; U1: Unit 1; U2: Unit 2; U3: Unit 3; 1 CmH2o = 98.0655pa.

Please cite this article as: G. Rong, A. Mendez, E. Bou Assi et al., Artificial Intelligence in Healthcare: Review and Prediction Case Studies, Engineering, https://doi.org/10.1016/j.eng.2019.08.015
Eq. (3) can also be used to estimate the pressure by substituting \( \bar{V} \) by \( \bar{P} \). The parameters needed for pressure estimation can also be computed during the learning phase.

We used the root mean square error (RMSE), which is computed using Eq. (4), for the validation of the algorithms in real-time-like test runs. In Eq. (4), \( V_i \) is the present volume (or pressure), \( V_{ib} \) of the bin \( i \), \( \bar{V}_i \) (or \( \bar{P}_i \)) is the estimated bin of the volume (or pressure) value computed using Eq. (3) for the same bin \( i \), and \( n \) is the total number of bins.

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (V_i - \bar{V}_i)^2}{n}}
\]

The real-time monitoring phase runs the following steps: ① digital non-causal filtering; ② on-the-fly spike classification; ③ BIR computing using the optimal BW; ④ comparing the BIR to the baseline and setting the volume to 0 for lower values; and, if the BIR is higher, ⑤ proceeding to compute the bladder volume or pressure using Eq. (3). Several test runs were executed to test and validate our algorithms, as shown in the example in Fig. 5 [106].

Finally, all the proposed algorithms were validated using data recorded from anesthetized rats during acute experiments. More details about these experiments and prediction algorithms are given in Ref. [102].

We succeeded in qualitatively predicting the three states of bladder fullness in 100% of the trials when the recorded afferent neural activities exhibited a Spearman’s correlation coefficient of 0.6 or better. We also succeeded in quantitatively predicting the bladder volume and pressure employing time windows of appropriately chosen durations.

We implemented a dedicated DSP, shown in Fig. 6 [106], to monitor the bladder volume or pressure, running the algorithms described above. The DSP runs the on-the-fly spike sorting and sensory decoding of the afferent neural activity of the bladder in real time to predict the volume or pressure value, as described in this section. The dedicated DSP was validated using synthetic data (with known a priori ground truth), and with signals recorded from acute experiments from the bladder afferent nerves.

The on-the-fly spike-sorting algorithm runs on the dedicated DSP and compares advantageously with other works reported in the literature [103]. We achieved a higher accuracy (92%) even when using difficult synthetic signals composed of highly correlated spike waveforms and low signal-to-noise ratios. The volume or pressure prediction module showed an accuracy of 94% and 97% for quantitative and qualitative estimations, respectively.

### 4.2. Epileptic seizure prediction

Epilepsy, a neurodegenerative disease, is one of the most common neurological conditions and is characterized by spontaneous, unpredictable, and recurrent seizures [107,108]. While first lines of treatment consist of long-term medications-based therapy, more than one third of patients are refractory.

On the other hand, recourse to epilepsy surgery is still relatively low due to very modest success rates and fear of complications. An interesting research direction is to explore the possibility of predicting seizures, which, if made possible, could result in the development of alternative interventional strategies [83]. Although early seizure-forecasting investigations date back to the 1970s [109], the limited number of seizure events, the paucity of intracranial electroencephalography (iEEG) recordings, and the limited extent of interictal epochs have been major hurdles toward an adequate evaluation of seizure prediction performances.

Interestingly, iEEG signals acquired from naturally epileptic canines implanted with the ambulatory monitoring device (Neuro-Vista) have been made accessible through the ieeeg.org online portal [110]. However, the seizure onset zone was not disclosed available. Our group investigated the possibility of forecasting seizures using the aforementioned canine data. Subsequently, we performed a directed transfer function (DTF)-based, quantitative identification of electrodes located within the epileptic network [111]. A genetic algorithm was employed to select the features most discriminative of the preictal state. We proposed a new fitness function that is insensitive to skewed data distributions. An average sensitivity of 84.82% at a time-in-warning of 10% was reported on the held-out dataset, improving previous seizure prediction performances [111].

Trying to find new opportunities for seizure prediction, we also explored novel features to track the preictal state based on higher order spectral analysis. Extracted features were then used as inputs to a multilayer perceptron for classification. Our preliminary findings revealed significant differences between interictal and preictal states using each of three bispectrum-extracted characteristics \( p < 0.05 \). Test accuracies of 73.26%, 72.64%, and 78.11% were achieved for the mean of magnitudes, normalized bispectral entropy, and normalized squared entropy, respectively. In addition, we demonstrated the existence of consistent differences between the epileptic preictal and interictal states in mean phase-amplitude coupling on the same bilateral canine iEEG recordings [112].

In contrast, we also explored the possibility of using quantitative effective connectivity measures to determine the network of

![Fig. 5. Quantitative volume estimation in simulated real-time data-processing experiments [106]. The first measurement cycle was used as a training phase, followed by two consecutive cycles where the volume was computed from the training phase. The optimal BW for this fiber was 47 s, using a regression model of order six. The fitting error of the first cycle and the estimation errors of the following two cycles were very low, as shown by the RMSE_all of each cycle: 2%, 3.9%, and 4.1%, respectively.](https://doi.org/10.1016/j.eng.2019.08.015)
seizure activity in high-density recordings. The ability of the DTF to quantify causal relations between iEEG recordings has been previously validated. However, quasi-stationarity of the analyzed signals remains a must to avoid spurious connections between iEEG contacts [113]. Although the identification of stationary epochs is possible when dealing with a relatively small number of contacts, it becomes more challenging when analyzing high-density iEEG signals. Recently, a time-varying version of the DTF was proposed: the spectrum-weighted adaptive directed transfer function (swADTF). The swADTF is able to cope with non-stationarity issues and automatically identify frequency ranges of interest [113]. Subsequently, we validated the possibility of finding seizure activity generators and sinks by employing the swADTF on high-density recordings [114]. The database consisted of patients with refractory epilepsy admitted for pre-surgical evaluation at the University of Montreal Hospital Center. Interestingly, the identified seizure activity sources were within the epileptic focus and resected volume for patients who went seizure-free after surgical resection. In contrast, additional or different generators were identified in non-seizure-free patients. Our findings highlighted the feasibility of accurately identifying seizure generators and sinks using the swADTF. Electrode selection methods based on effective connectivity measures are thus recommended in future seizure-forecasting investigations.

Recent findings highlight the feasibility of predicting seizures using iEEG recordings; the transition from interictal into ictal states consists of a “buildup” that can be tracked using advanced feature extraction and AI techniques. Nevertheless, before current approaches can be translated into clinical devices, further research is needed on feature extraction, electrode selection, hardware implementation, and deep learning algorithm.

5. Conclusion

In this paper, we reviewed the latest developments in the application of AI in biomedicine, including disease diagnostics and prediction, living assistance, biomedical information processing, and biomedical research. AI has interesting applications in many other biomedical areas as well. It can be seen that AI plays an increasingly important role in biomedicine, not only because of the continuous progress of AI itself, but also because of the innate complex nature of biomedical problems and the suitability of AI to solve such problems. New AI capabilities provide novel solutions for biomedicine, and the development of biomedicine demands new levels of capability from AI. This match of supply and demand and coupled developments will enable both fields to advance significantly in the foreseeable future, which will ultimately benefit the quality of life of people in need.

Nomenclature

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAL</td>
<td>ambient assisted living</td>
</tr>
<tr>
<td>AI</td>
<td>artificial intelligence</td>
</tr>
<tr>
<td>AN</td>
<td>artificial neural network</td>
</tr>
<tr>
<td>ASIC</td>
<td>application-specific integrated circuits</td>
</tr>
<tr>
<td>BioQA</td>
<td>biomedical question answering</td>
</tr>
<tr>
<td>BIR</td>
<td>bin-integrated rate</td>
</tr>
<tr>
<td>CMA</td>
<td>computational modeling assistant</td>
</tr>
<tr>
<td>DSP</td>
<td>digital signal processor</td>
</tr>
<tr>
<td>DSS</td>
<td>decision support system</td>
</tr>
<tr>
<td>DTF</td>
<td>directed transfer function</td>
</tr>
<tr>
<td>ECG</td>
<td>electrocardiography</td>
</tr>
<tr>
<td>EEG</td>
<td>electroencephalography</td>
</tr>
<tr>
<td>EMG</td>
<td>electromyography</td>
</tr>
<tr>
<td>ES-MR</td>
<td>expert system for memory rehabilitation</td>
</tr>
<tr>
<td>CPU</td>
<td>central processing unit</td>
</tr>
<tr>
<td>GPU</td>
<td>graphical processing unit</td>
</tr>
<tr>
<td>FES</td>
<td>functional electrical stimulation</td>
</tr>
<tr>
<td>FIR</td>
<td>finite impulse response</td>
</tr>
<tr>
<td>FPGA</td>
<td>field-programmable gate array</td>
</tr>
<tr>
<td>HMI</td>
<td>human–machine interface</td>
</tr>
<tr>
<td>iEEG</td>
<td>intracranial electroencephalography</td>
</tr>
</tbody>
</table>
IoT Internet of Things
ML machine learning
NN neural network
PDA personal digital assistant
PLA people with loss of autonomy
POCT point-of-care testing
RMSE root mean square error
SALSA screening of activity limitation and safety awareness
swADTF spectrum-weighted adaptive directed transfer function
FR firing rate
BW bin width
OSR overall success rate

Acknowledgement
This work was supported by the Startup Research Fund of Westlake University (041030080118), the Research Fund of Westlake University, and the Bright Dream Joint Institute for Intelligent Robotics (10318H991901).

Compliance with ethics guidelines
Guoguang Rong, Arnaldo Mendez, Elie Bou Assi, Bo Zhao, and Mohamad S. Saman declare that they have no conflict of interest or financial conflicts to disclose.

References


