Clinical Big Data and Deep Learning: Applications, Challenges, and Future Outlooks

Ying Yu, Min Li, Liangliang Liu, Yaohang Li, and Jianxin Wang*

Abstract: The explosion of digital healthcare data has led to a surge of data-driven medical research based on machine learning. In recent years, as a powerful technique for big data, deep learning has gained a central position in machine learning circles for its great advantages in feature representation and pattern recognition. This article presents a comprehensive overview of studies that employ deep learning methods to deal with clinical data. Firstly, based on the analysis of the characteristics of clinical data, various types of clinical data (e.g., medical images, clinical notes, lab results, vital signs, and demographic informatics) are discussed and details provided of some public clinical datasets. Secondly, a brief review of common deep learning models and their characteristics is conducted. Then, considering the wide range of clinical research and the diversity of data types, several deep learning applications for clinical data are illustrated: auxiliary diagnosis, prognosis, early warning, and other tasks. Although there are challenges involved in applying deep learning techniques to clinical data, it is still worthwhile to look forward to a promising future for deep learning applications in clinical big data in the direction of precision medicine.

Key words: deep learning; clinical data; Electronic Health Record (EHR); medical image; clinical note

1 Introduction

With the rapid development of medical informatization, diverse healthcare information systems have been employed in hospitals, including Hospital Information Systems (HIS), Clinical Information Systems (CIS), Picture Archiving and Communication Systems (PACS), Laboratory Information Systems (LIS), and Radiology Information Systems (RIS). Digital clinical data has increased rapidly over recent decades, with the massive amounts of clinical records being accumulated in medical institution. It brings tremendous opportunities for healthcare research. More specifically, reliable systems based on clinical big data and data-driven approaches are critical to study the relationship between diseases and clinical symptoms, to observe diseases development trends, and to explore risk factors related to various diseases.

Over the last few decades, most analyses of clinical data have used traditional statistical machine learning methods, such as Logistic Regression (LR)[1], Support Vector Machines (SVM)[2, 3], Decision Trees (DT)[4], Random Forests (RF)[5], Conditional Random Fields (CRF)[6], Bayesian Networks (BN)[7], Principal Component Analysis (PCA)[8, 9], and Latent Dirichlet Allocation (LDA)[8, 10]. Those methods have achieved some reliable results across many subfields of clinical

● Ying Yu is with the School of Computer Science and Engineering, Central South University, Changsha 410083, China, and the School of Computer Science and Technology, University of South China, Hengyang 421001, China. E-mail: yuying@mail.csu.edu.cn.

● Min Li, Liangliang Liu, and Jianxin Wang are with the School of Computer Science and Engineering, Central South University, Changsha 410083, China. E-mail: {limin, liuliang_double, jxwang}@mail.csu.edu.cn.

● Yaohang Li is with the Department of Computer Science, Old Dominion University, Norfolk, VA 23529, USA. E-mail: yaohang@odu.edu.

* To whom correspondence should be addressed.
  Manuscript received: 2019-03-08; accepted: 2019-03-15

studies. In recent years, deep learning[11] has been under vigorous development and has achieved great success in many competitions[12–15], due to its superior performance in learning the feature representation from raw data in an end-to-end way. Adding more hidden layers to a neural network allows a deep architecture to express more complex hypotheses, because the hidden layers can capture the nonlinear relationships[16]. In combination with the large volume of data and Graphical Processing Units (GPU), deep learning makes it possible to explore clinical data in the direction of precision medicine.

The motivation of our review is to provide a valuable overview for researchers concerned with the application of deep learning methods to clinical studies. First, the characteristics of clinical data and the diverse types of clinical data are discussed. Second, after a brief introduction to the deep learning models which have been most commonly used in clinical data analysis, we divide the applications into several categories and summarize the applications of deep learning models to different types of clinical data. Finally, we discuss the challenge of applying deep learning to clinical analysis and propose future research directions. This paper is different from other recent surveys, some of which are devoted to review deep learning applications in the context of biomedical informatics[16, 17], ranging from genomic analysis to biomedical image analysis, and some others of which are only concerned about deep learning applications for a specific data type, such as medical images[18, 19]. In general, this paper pays close attention to the variety of different types of clinical data and the application of deep learning methods to them.

2　Clinical Data

2.1　Characteristics of clinical data

An Electronic Health Record (EHR) of a patient includes many kinds of clinical data and plays an important role in medical research. It is reported that the world's electronic health data has reached 500 petabytes in 2012 and the total is expected to reach 25 exabytes by 2020[20]. In recent years, intelligent wearable devices have arisen to collect personal daily health data (such as blood sugar, blood pressure, and heart rate), which further accelerates data growth. We are witnessing clinical data growing at an unprecedented rate, accompanied by a growing variety of data sources (structured relational forms,

unstructured images or text, semi-structured records). At the same time, some real-time data is produced in the moment and needs to be retrieved and analyzed immediately for timely decision making. It is obvious that clinical data has the "4Vs" characteristics of big data: big volume, large data variety, high velocity of data generation/update, and high veracity leading to big value[21]. Moreover, privacy is an important and specific characteristic of clinical data. Medical records include Protected Health Information (PHI) of patients, e.g., real name, gender, age, location, phone number, and birth date. This makes it different from user information in social networks, which can include a nickname or fake identification. Some of these PHIs (such as gender and age) are meaningful for clinical analysis. However, it is prohibited to release data arbitrarily for research, considering the protection of patient privacy. This presents as an obstacle to organizing and sharing public clinical databases.

2.2　Different types of clinical data

Clinical data is quite heterogeneous. During a medical treatment process, a large amount of information relating to a patient's health status is generated and presented in various forms. In what follows, clinical data is discussed by dividing it into three main categories (medical images, clinical notes, and all others).

2.2.1　Medical images

With the development of medical imaging techniques, there is much medical image data being generated by X-rays, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Optical Coherence Tomography (OCT), microscopy image, and Positron Emission Tomography (PET)[22]. Due to the great successes that deep learning has achieved in computer vision, researchers have been quick to apply deep learning methods to the processing of medical image data.

Medical images are direct raw data or first-hand information reflecting a patient's status. Different clinical imaging techniques have their own advantages in detecting pathological changes in different organs. Ophthalmic imaging is specific for the eyes[23–25], MRI is good at the detection of pathologies of the brain[26–28], cardiac system[29–31], and bone and joints[32–34], CT is better on the abdominal organs[35–37] and chest[38–40], while X-ray performs well on the chest[41–43] and breast[44–46].

However, the labeled samples of medical images that can be used for training are far lower in number than social network image samples. The input of medical images can be categorized as vector format, 2D pixel or 3D voxel values, which are suitable for typical multilayer neural networks and convolutional networks, respectively. Deep learning methods have been applied to medical images for computer-aided pathology detection, image segmentation, and disease classification.

2.2.2 Clinical notes

Besides images, text is one of the most important categories of EHRs, in the form of discharge summaries[47–52], various kinds of measurement reports[53–55], and death certificates[56]. Several studies based on clinical text have been carried out on discharge summaries. Because a discharge summary usually records the whole procedure of medical treatment. It contains rich information, including descriptions of lab test results, physician diagnoses, drugs, and treatments. In addition, some further information is indirectly recorded in discharge summaries, for example, chief complaint, family history, medical history, and allergies are attained from patients. Measurement reports usually summarize the examination results, which are clues to diseases inference.

Natural Language Processing (NLP) technology and machine learning methods have been applied to various clinical notes. The free text of clinical notes is usually turned into a sequence of vectors by one-hot or distributed representation. The clinical free text has been used for some significant tasks, such as medical concept extraction, named entity recognition, disease classification, drug-allergy reactions, and phenotype extraction and de-identification. Furthermore, clinical notes are combined with medical images in some studies to improve system performance.

2.2.3 Other types

Outside of medical images and clinical notes, there are other types of clinical data that are rarely used for independent analysis. To name a few, the results of physiological measurements (lab results, vital signs), demographic information, payment and insurance information, and so forth. Most of this data is stored in two-dimensional relational tables.

The results of laboratory tests are also part of EHR. Generally, those results are presented as numeric values, which can provide objective evidence for physicians to observe the status of a patient. Nevertheless, diverse lab items have different units and different reference ranges of normal values (considering age and gender). For that reason, normalization is necessary before analyzing lab test results. Lab results are usually combined with clinical notes or demographic information to predict the risk of disease[57] or mortality[58], to model patient trajectory[59], and to recommend medication dosing[60].

A large volume of real-time data is generated by monitoring devices, i.e., electrocardiogram and multi-parameter monitors. These are essential devices for Intensive Care Unit (ICU) and postoperative monitoring[61, 62]. Smart wearable devices can also provide real-time healthcare data in the conditions of everyday life[63]. These vital signs can be collected to construct retrospective temporal datasets, which are useful for trajectory analyses of disease development[64] and can improve the effectiveness of clinical decision making[65].

In general, demographic information, such as gender, location, age, and marital status, is recorded in a structured table. Although demographic information involves issues of patient privacy, it is a good data resource for statistical analysis in order to observe distribution characteristics of diseases. For example, Cheng et al.[66] used the in-house data from the national disease surveillance for a population-based longitudinal analysis of trends in Traumatic Brain Injury (TBI) mortality from 2006 through 2013 in China that reported differences based on location (urban/rural), sex, age group, and external cause of TBI.

In addition to the aforementioned data types, there are some data indirectly related to disease, such as payment, charge, and insurance information. These data are yielded to and stored in healthcare systems each day, and usually utilized to perform statistical analysis for commercial purposes.

Although masses of clinical data are generated every second in hospitals, researchers can have very limited access to even a small amount of it. Many clinical studies that have declared their experimental data are drawing on private datasets collected from specific hospitals or research institutions. Some high-quality datasets supported by specific projects are available for purchase, such as some nationwide databases provided by the Healthcare Cost and Utilization Project (HCUP)[67]. Those datasets collect data beginning in 1988 and contain encounter-level information on basic health care[68].

Fortunately, there are still some datasets available for public access after receiving approval. Here, we list six public datasets that are used commonly in clinical studies and contain different types of clinical data (see Table 1). Some of them focus on specific diseases such as the Alzheimer's Disease Neuroimaging Initiative (ADNI)[69, 70], Osteoarthritis Initiative (OAI)[71], and Study of Osteoporotic Fractures (SOF)[72]. Some of them collectd specific data type. For example, the ADNI and the Grand-Challenges datasets[73] mainly consist of medical images, while the i2b2 challenge datasets[74–78] are made up of clinical free texts. The Medical Information Mart for Intensive Care (MIMIC)[79–81] is a large, public, and freely-available de-identified database from ICU, containing comprehensive data which is widely used in clinical studies. Overall, the data of the first four datasets in Table 1 is more systematic, while the last two datasets are released to meet some pecific challenge. Recently, several researchers are interested in pursuing studies based on the fusion of data. Multiple data sources can bring a more comprehensive understanding of patients and improve the accuracy of disease prediction[82, 83].

3 Deep Learning Models

In recent years, deep learning methods have shown high effective for a broad range of big data analyses and consequently have become a center of research attention. Deep learning involves a wide variety of approaches and has achieved promising early successes in medical research. In this paper, we briefly introduce the most common deep learning models that have been used to analyze clinical data. The details and architecture of each model are not mentioned here, since they have been discussed at length in the existing literature.

3.1 Deep neural networks

A Deep Neural Network (DNN) is an Artificial Neural Network (ANN) with multiple hidden layers between the input and output layers. Each hidden layer can be seen as a process of feature extraction. The higher layers can composite the features from lower layers, potentially modeling complex data with fewer units than a similarly performing shallow network. This enables a DNN to model complex non-linear relationships using training data. To solve different problems, multiple variants and extensions have been developed. Thus far, several extended models have been introduced in the literature, as follows.

3.2 Convolutional neural networks

Convolutional Neural Networks (CNNs) are known as shift invariant and space invariant ANNs, which are based on their shared-weights convolutional kernels and translation invariance characteristics[84]. CNNs are widely used in computer vision and NLP, and have achieved promising results on many tasks in both of these fields. It is also the first and most widely used deep learning model in clinical fields. CNNs have greatly promoted medical image research for imaging diagnosis, image segmentation, and Electrocardiograph (ECG) monitoring[85]. CNNs and their variants can capture overall structural features from feature-rich images. The most recent well-known CNN architecture is U-net, proposed by Fu et al.[86], which can derive the full context of an image by combining an equal amount of upsampling and downsampling layers[87]. ResNet[88] is another popular variant of CNN utilized for medical

Table 1   Public clinical dataset sources.

Dataset	Description	Data type	Size	Website
ADNI	Imaging data of Alzheimer's disease	MRI image, PET image, clinical data, biospecimen	1070–2000 participants (ADNI3)	http://adni.loni.usc.edu
OAI	Various data of Osteoarthritis	Clinical examination, radiological images, a biospecimen repository	4796 participants	https://oai.epi-ucsf.org/datarelease/
SOF	20 years of prospective data about osteoporosis	Self-administered questionnaire, clinic interview, clinic examination	10 366 older women, 9 visits	https://sofonline.epi-ucsf.org/interface/
MIMIC	Clinical data of ICU	Demographic information, clinical note, physiological measurements	46 520 patients (MIMIC III)	https://mimic.physionet.org/about/mimic/
Grand-Challenges	Challenge datasets of medical and biomedical images	All kinds of biomedical images	158 challenge datasets	https://grand-challenge.org/challenges/
i2b2	Challenge datasets in NLP for clinical data	Clinical notes	8 challenge datasets	https://www.i2b2.org/

images, which can ease gradient backpropagation flow and improve optimization convergence speed[89]. Recently, dense CNNs[90] have alleviated the vanishing-gradient problem, strengthened feature propagation, encouraged feature reuse, and substantially reduced the number of parameters[91]. For medical texts, CNN can be used to predict diagnosis codes from clinical notes[48].

Much research has shown that CNNs are better at extract medical terminology information than traditional methods like SVM[47,50]. In addition, there are some basic tasks of NLP on biomedical texts that have been explored with CNNs, for example, named entity recognition in biomedical texts[92] and Mesh indexing[93,94].

3.3   Recurrent neural networks

A Recurrent Neural Network (RNN) is a kind of deep feed-forward ANN. The connections between the units of an RNN form a directed graph along a sequence. RNNs have taken a central place in sequence data processing despite the gradient vanishing problem generated by long input sequences. A variant of RNN called a Long Short-Term Memory (LSTM) has been proposed to solve the problem by introducing gate mechanism and memory cell[95]. In addition, Gated Recurrent Unit (GRU) is a simplified version of LSTM. These two varieties can achieve a better performance than the plain model for long sequence processing. RNN is mainly used to solve sequence related problems, such as sentence classification, speech recognition, and machine translation. It has been used for the retrospective analysis of clinical temporal sequence data, and Choi et al.[96] applied the LSTM model to address clinical decision support problems by predicting future disease diagnosis along with corresponding timely medication interventions. The feature extraction of clinical text is another important application of RNN. In particular, it has been used to predict diagnosis codes based on medical text[49,50] and has been applied to clinical name entity recognition[97,98].

3.4   Restricted Boltzmann machines

A Restricted Boltzmann Machine (RBM) is a generative stochastic ANN that can learn the probability distribution over a set of inputs. It is an unsupervised learning model and designed for feature extraction in unlabeled training data. Feature extraction by an RBM is helpful for the original classification problem. For example, Khatami et al.[99] proposed a robust RBM-based classification model for X-ray images. Hua et al.[100] introduced a model for a Deep Belief Network (DBN), which is a hierarchical stack of RBM, to address the nodule classification for supervised learning tasks, especially with CT images.

3.5   Deep autoencoders

Deep autoencoders have the same purpose as RBMs. Proposed by Hinton and Salakhutdinov[101], they are mainly designed for feature extraction or dimensionality reduction. The purpose is to reconstruct the inputs. In order to improve the capability of mastering important information and learning richer representations from training data, a number of variants have been introduced over recent years, such as the Denoising AutoEncoder (DAE)[24], Stacked AutoEncoder (SAE)[102], and Variational AutoEncoder (VAE). In relation to clinical data, Sharma et al.[103] presented a feature extraction technique for medical images using SAEs and found the method helpful for improving the performance of medical image retrieval.

In summary, the intention of deep learning is representation, which means representing the features of input data effectively by multilayer neural network training. The deep learning methods mentioned above have made great progress in the study of clinical data. Moreover, the combination of models has brought more opportunities for effective analysis. For instance, Kong et al.[30] detected the end-diastole and end-systole frames in cine-MRI of the heart based on a combination of LSTM and CNN. The combinational model was extensively validated on thousands of cardiac sequences in which the average difference is merely 0.4 frames. Chen et al.[104] combined a bi-directional LSTM with 2D U-net-like-architectures to improve the structure segmentation of anisotropic 3D electron microscopy images. The combination model achieved promising results compared to the known deep learning 3D segmentation approaches.

4   Applications

Mass clinical data is crucial for carrying out clinical studies. It contains the complete information associated with a patient's health status, clinical features, clinical therapies, and treatment effects[105]. Recently, deep learning models have been successfully applied in the secondary use of clinical informatics to improve

knowledge discovery in healthcare. According to the clinical studies literature, multimodal clinical data (i.e., images, text, test results) have been utilized separately or in fusion in different clinical tasks. According to the target of application, we discuss four domains of application of deep learning models on all kinds of clinical data: auxiliary diagnosis, prognosis, early warning, and other tasks.

4.1   Deep learning for auxiliary diagnosis

Most studies in this domain look to capture patterns of patients from clinical data, aiming to support clinical decisions by inferring the risk of a specific disease or identifying the characteristics of a certain disease from a control cohort. This can be illustrated in three mainly auxiliary diagnosis applications: disease classification or prediction, inference of disease code, and risk factor identification.

4.1.1   Disease classification or prediction

Classification is one of the prominent issues for which deep learning has made a great contribution based on kinds of clinical data. Medical imaging classifications typically use one or multiple images as inputs for a training model with a single diagnostic variable as output (e.g., disease onset or not). Transfer learning is popular in this task, because plenty of natural image resources in computer vision can be used to pre-train deep networks. Gulshan et al.[24] applied CNN to automatically detect diabetic retinopathy and diabetic macular edema in retinal fundus photographs. The model had been pre-trained on the ImageNet dataset. An impressive work was published by Esteva et al.[12], who fine-tuned a pre-trained version of Google's Inception v3 architecture on medical data and performed classification of skin lesions using a single CNN. The model achieved near-human-expert performance by training in an end-to-end manner based directly on images (using only pixels and disease labels as inputs). Diverse deep models have been used for classification based on clinical exam images. Brosch et al.[27] and Plis et al.[28] applied DBNs to classify patients of nervous system diseases. Suk and Shen[102] used SAEs to classify Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI). However, CNN and its variants are playing leading roles in the most recent papers on disease classification based on medical images. For example, Guerrero et al.[89] proposed a CNN model which can classify white matter hyperintensities and stroke lesions. Prediction[96, 106] or

inference[107], according to EHRs or clinical notes, usually concerns the future onset of diseases. RNNs are more commonly employed for this[108]. Choi et al.[96] and Ma et al.[106] proposed deep learning architectures separately, named "doctor AI" and "dipole", both of them based on LSTM for diagnosis prediction with input of historical time-order sequence of patients' past visit records.

4.1.2   ICD code assignment

International Classification of Diseases (ICD) codes have always been used in EHRs to serve as common labels identifying symptoms and signs, diseases and abnormal findings, and operations on clinical records. ICD code assignment is a fundamental classification task for EHRs. Duarte et al.[109] addressed the assignment of ICD-10 codes using a two-level hierarchical approach based on GRU. They inferred the cause of death (ICD code) by analyzing death certificates together with the association of autopsy reports and clinical bulletins. Shi et al.[49] proposed a hierarchical RNN model with an attention mechanism that can automatically assign ICD diagnostic codes according to a given written diagnosis description. Li et al.[48] combined Doc2vec and CNN for ICD-9 automatic coding based on MIMIC datasets. Recently, Guo et al.[110] proposed a disease inference model based on bidirectional LSTM considering the symptom-disease associations. Yu et al.[111] performed automatic coding experiments on Chinese clinical notes. According to the characteristic of Chinese character, they proposed a multilayer bidirectional LSTM combining with attention mechanism for automatic ICD code assignment.

4.1.3   Risk factor identification

As defined by the World Health Organization, a risk factor refers to any attribute, characteristic, or exposure of an individual that increases the likelihood of developing a disease or injury. The analysis of risk factors can help to find the cause of disease and improve disease prevention or treatment. A Google research team, Poplin et al.[112], attempted to predict cardiovascular risk factors which were not previously thought to be present or quantifiable in retinal images using DNNs. Li et al.[113] built a framework to construct an integrated representation of features from all available risk factors in the EHR data. They used these integrated features to effectively predict osteoporosis and bone fractures. At the same time, they

developed a framework for the informative risk factor selection of bone diseases.

4.2 Deep learning for prognosis

With the accumulation of longitudinal clinical data, and by gathering monitoring data over a period of time, researchers can observe clinical data in a temporal sequential way for the purpose of prognosis.

4.2.1 Readmission or length of stay prediction

Hospital readmission is common and expensive[114]. To predict potentially avoidable readmissions is of benefit for patients and an aid to the effective and rational utilization of medical resources. Futoma et al.[115] compared different models of predicting hospital readmissions based on a large EHR database, confirming that DNNs have significantly higher prediction accuracies than conventional approaches. Nguyen et al.[116] used a simple word embedding layer as the input of a CNN architecture to predict unplanned readmission following the discharge. Pham et al.[117] constructed a framework called DeepCare that employed LSTM to predict future readmission based on past diagnoses and interventions.

4.2.2 Survival prediction and mortality prediction

Mortality and survival prediction are important clinical issues, which commonly focus on the temporal sequence data in order to predict the outcome of treatment. Grnarova et al.[118] developed a two-layer CNN architecture to represent documents for ICU mortality prediction. Carneiro et al.[119] proposed a new prognostic method based on CNNs, which can predict 5-year mortality in elderly individuals using chest CTs. van der Burgh et al.[120] attempted to combine clinical characteristics with MRI data to predict the survival of amyotrophic lateral sclerosis patients using a deep neural network. Most recently, Rajkomar et al.[13] proposed an ensemble model of three deep learning neural networks for the prediction of four outcomes (in-hospital mortality, 30-day unplanned readmission, prolonged length of stay, and final diagnosis). They represented patients' entire raw EHRs using the Fast Healthcare Interoperability Resources (FHIR) standard. In all cases, the deep learning models based on LSTM outperformed state-of-the-art traditional predictive models.

4.3 Deep learning for early warning

Early detection or diagnosis of disease can reduce healthcare costs, and save lives through early warning.

The essence of preventive treatment of diseases is "the earlier the better". Prospective healthcare is therefore one of the most valuable fields of clinical research. Some deep learning methods have been used in the early warning of disease. Ju et al.[121] used a deep autoencoder network to make early diagnosis of Alzheimer's disease based on Resting-state functional Magnetic Resonance Imaging (R-fMRI) data and relevant text information. It achieved 86.47% prediction accuracy, a 31.21% improvement over traditional approaches. On the same task, Lu et al.[122] utilized a multimodal and multiscale deep neural network to distinguish subjects on an Alzheimer's trajectory (progressive normal control, progressive mild cognitive impairment, and stable AD and those without cognitive deficits (stable normal control). It was found to deliver an 85.68% accuracy in the prediction of subjects within 3 years to conversion. Choi et al.[123] attempted to model temporal relations among events in EHRs using RNN and improved model performance in the early detection of heart failure.

4.4 Deep learning for other tasks

Extracting information from unstructured texts is very difficult, with traditional methods relying on manual engineering or rule-based systems. Several studies focus on clinical Name Entity Recognition (NER), identification of entity relations, and de-identification of clinical notes by employing deep learning methods with the support of terminology databases.

4.4.1 Clinical NER

To extract medical concepts (symptoms, disease or diagnosis, body parts, treatments, medications, clinical events, etc.) from clinical notes is a basic task for clinical text processing. Wu et al.[97] used deep neural networks to learn word embeddings and perform recognition of four types of clinical entities (problems, lab tests, procedures, and medications) based on Chinese clinical notes. By treating clinical event detection as a task of sequence labeling, Jagannatha and Yu[98] empirically evaluated the performance of LSTM and GRU on detecting medication, diagnosis, and adverse drug events.

4.4.2 Entity relation extraction

Entity relation extraction aims to structure relationships between medical concepts, including treatment-disease, disease-symptom, treatment-symptom, test-disease, test-symptom, and so on. Relationships between medical concepts can reveal the correlation of clinical

entities. Especially, it can help in analyzing the cause of a disease; for example, whether a treatment improves, worsens, or causes a disease, or whether a test reveals a disease or symptom. Luo[124] proposed a model based on LSTM for classifying three relations (relations between medical problems and treatments, between medical problems and tests, and between medical problems and other medical problems), which was the task of the 2010 i2b2/VA challenge.

4.4.3 De-identification

To remove a patient's private information from the clinical data before using it, researchers need to identify PHI (i.e., name, age, birth date, telephone number, address, and occupation) in clinical data and remove it at the very beginning; this is known as de-identification and is a critical step in clinical data utilization. The purpose is to make medical records more accessible for researchers and available for open access. However, it is a very difficult and painstaking task. Liu et al.[125] developed a hybrid system for de-identification, consisting of four individual subsystems based on LSTM, bidirectional LSTM, conditional random fields, and a rule-based subsystem. Dernoncourt et al.[126] introduced a de-identification system based on a bidirectional LSTM, which requires no handcrafted features or rules. Both of these systems outperformed state-of-the-art methods on the dataset of the 2014 i2b2 NLP challenge.

4.4.4 Image detection and segmentation

Deep learning models have achieved outstanding performance in medical image detection and segmentation which are traditional applications in computer vision. Detection involves the localization and identification of an organ, region, object, lesion, or landmark. It is not only an important preprocessing step for segmentation tasks but also one of the most labor-intensive tasks for clinicians. Segmentation is the most common topic for applying deep learning methods to medical images, involving quantitative analysis of clinical parameters related to volume and shape. It is performed by identifying a set of voxels making up either the contour or the interior of an object. CNN and its variants have been a great advancement for medical image detection and segmentation. Fakhry et al.[88] proposed a ResNet for medical image segmentation without prior knowledge. Payer et al.[127] applied CNNs to directly regress landmark locations. de Vos et al.[128] detected anatomical regions of interest in

image slices by coming three orthogonal 2D CNNs in order to localize them in 3D. Dou et al.[129] used a 3D CNN to find micro-bleeds in brain MRIs. Milletari et al.[87] proposed a U-net architecture for 3D image segmentation.

4.4.5 Content-based medical image retrieval

Large numbers of medical images make it necessary to organize them logically and retrieve them efficiently. Content-Based Image Retrieval (CBIR) helps physicians to understand rare disorders and identify similar case histories, which can ultimately improve diagnosis and treatment. The core of CBIR is to extract effective feature representations from the pixel information of a medical image. Considering the ability of deep CNN models to learn rich features at multiple levels, it has been widely used for CBIR. Liu et al.[130] composed a retrieval model using a custom CNN, which can obtain the descriptive feature vector of X-rays based on three CNN layers and two fully-connected layers. Shah et al.[131] implemented a CBIR system by combining CNN feature descriptors with hashing-forests.

The applications of deep learning to various clinical data are summarized in Tables 2, 3, and 4. They show that deep learning models have been widely applied in many clinical studies. Particularly, it is widespread in medical image processing, and is beginning to be applied to clinical notes. Different types of clinical data mixed together as the input of a model is an aspect that many studies have set out to explore via deep methods. As for the variety of deep learning models being used, CNNs are commonly used to represent the features of clinical data in most applications, especially for medical images processing, while RNNs and its variants are good at tasks involving text data or temporal sequence of EHRs. DBNs and deep autoencoders are occasionally also used for processing medical images or text.

Although deep learning models have obtained the best performance in most clinical studies, expert knowledge usually provides advantages that go far beyond adding more layers to a model. Novel forms of data preprocessing or feature representation can improve the result when applying the same architecture to a certain task. Currently, the fusion of diverse data resources, the combination of various models, and the analysis of multiple tasks are receiving more attention (see Table 4).

Table 2   Overview of deep learning methods applied on medical image.

Data type	Application	Method	Task&Reference
OCT	Classification	Deep CNN	Age-related macular degeneration detection[23]
			Diabetic retinopathy detection[24]
			Macular degeneration and diabetic retinopathy classification[25]
	Image detection	Fast CNN	Hemorrhage detection[132]
	Risk factor identification	DNN	cardiovascular risk factors[112]
	Image segmentation	CNN	Optic Disc (OD) and Optic Cup (OC) segmentation[86]
PET	Classification	SAE&CNN	Diagnosis of disease status[133]
		DNN	Diagnosis of Alzheimer's disease[134]
		CNN&RNN	Diagnosis of Alzheimer's disease[135]
Microscopy	Classification	CNN	Skin cancer classification[12]
MRI	Classification	DBN	AD /Health Control (HC) classification[27]
			Schizophrenia and Huntington's disease classification[28]
			AD/MCI/HC classification[136]
		SAE	AD and MCI classification[102]
		CNN	AD/MCI/HC classification[137]
		ANN	AD/MCI/HC classification[138]
	Image detection	CNN	Left ventricle slice detection[29]
			Automatic spine scoring, vertebral discs analysis, lesion hotspots detection[32]
			Vertebrae detection and labeling[33]
		RNN	Identification of end-diastole and end-systole frames[30]
	Image segmentation	DBN	Left ventricle segmentation[31]
		CNN	Knee cartilage segmentation[34]
			Prostate lesions segmentation[87]
	Image detection & segmentation	CNN	Localization, identification, and segmentation of vertebrae[139]
	Survival prediction	DNN	Survival prediction of amyotrophic lateral sclerosis patients[120]
CT	Image segmentation	CNN	Segmentation of liver, spleen, and kidneys[35]
			Kidney segmentation[36]
			Liver segmentation[37]
			Pancreas segmentation[140]
	Image detection	CNN	Colon polyp detection[141]
	Classification	CNN	Interstitial lung texture classification[38]
			2D interstitial pattern classification[39]
			Lung texture pattern classification[40]
		DBN&RBM	Classification of lung texture and airways[142]
	Mortality prediction	CNN	prediction of 5-year mortality in elderly individuals[119]
X-ray	Classification	CNN	Tissue classification[44]
			Detection of cardiovascular disease[143]
			Discriminate malignant masses from (benign) cysts[45]
			Pathology detection[41]
			Five common abnormalities' detection[42]
			Tuberculosis detection[43]
			Frontal/lateral classification[144]
		SAE	Breast density classification[145]
	Image detection	CNN	Mass detection[46]

5   Discussion

5.1   Challenges

Up to now, deep learning has delivered great improvements in the studies of clinical data in comparison with traditional machine learning approaches. Nevertheless, there are many obstacles in applying deep learning to clinical data.

First of all, clinical big data is hard to obtain. In order to train a reliable and effective model, large sets of

Table 3   Overview of deep learning methods applied on clinical notes.

Data type	Application	Method	Task&Reference
Discharge summary	ICD assignment	RNN	Automated ICD coding/ Diagnosis code assignment[49, 50, 110, 111]
		CNN	Automated ICD-9 coding[48]
	Classification	CNN	Phenotype extraction[52]
	Entity relation extraction	RNN	Classifying three categories' relations[124]
Pathology report	Classification	RNN&CNN	Multiple information extraction[53]
Death certificate	ICD assignment	RNN	ICD-10 codes assignment for the underlying cause of death[66]
Narrative medical record	De-identification	RNN	De-identification[125, 126]
Nonspecific	Name entity recognition	RNN	extracting medical events (medication, disease)[98]
	Mortality prediction	CNN	ICU mortality prediction[118]

Table 4   Overview of deep learning methods applied on mixed clinical data.

Data type	Application	Method	Task&Reference
Radiology report & image	Disease prediction	RNN&CNN	Radiology image identification[54]
			Retinal microaneurysm detection[55]
Admission notes and discharge summaries	Name entity recognition	DNN	Name entity recognition in Chinese clinical text[97]
Death certificates and autopsy reports	ICD assignment	RNN	Assignment of ICD-10 codes for causes of death[109]
R-Fmri & clinical text	Early warning	DAE	Early diagnosis of Alzheimer's disease[121]
Structural MR and FDG-PET images	Early warning	DNN	Identify individuals at risk of developing Alzheimer's disease[122]
EHR (medical codes)	Disease prediction	RNN	Diagnosis prediction[96, 106, 108]
EHR (medical codes & demographic information)	Disease prediction	DNN	Diagnosis prediction[83]
EHR (medical codes, demographic information, lab results)	Risk factor identification	DBN	To identify the risk factors of osteoporosis[113]
	Readmission prediction	DNN	To assess patient readmission risk[115]
		CNN	Unplanned readmission following discharge prediction[116]
		RNN	Future readmission prediction[117]
	Early warning	RNN	Early detection of heart failure[123]
	Prognosis	RNN	To predict in-hospital mortality, 30-day unplanned readmission, prolonged length of stay, and final diagnoses[13]

training data are required. Although we are witnessing an explosion of healthcare data, very little is available in open access datasets for clinical studies. One aforementioned major reason for this is the protection of patient privacy. Most of the valuable datasets are held by specific hospitals or research institutions. In addition, the real data gathered first-hand from hospital information systems has some deficiencies, such as errors, noise, and missing values. These two reasons make it difficult to gather useful big data sets.

However, the main challenge is the acquisition of relevant annotations or labels for data, which requires extensive manual labor supported by medical staff. The lack of large labeled data sets is often mentioned as an obstacle, and is particularly noticeable for clinical image data. A clinical image dataset usually consists of dozens or hundreds of labeled samples. Moreover, the scarcity of samples of some rare diseases makes it hard to use deep learning approaches.

Clinical data comes in various formats (image, text, waveform, numeric, Boolean type), and combining data from different sources is not easy. For example, medical images are seldom used together with the other types of clinical data. The representation of differently formatted data and the mechanism for fusing them are both vital issues.

On the other hand, raw data from hospitals cannot be directly used as input for deep learning models. The consistency and normalization of data is another tough problem. For instance, the same diagnosis

description may be labeled with different ICD code versions by two different institutions. Preprocessing or normalization of data is necessary before training. There are several classification schemes and ontologies that are used for recording relevant medical information and events, for example, the ICD code for diseases, signs and symptoms, abnormal findings, complaints, social circumstances, and external causes of injury; Current Procedural Terminology (CPT) for procedure; Logical Observation Identifiers Names and Codes (LOINC) for laboratory measurements; and RxNorm for medication codes. However, those schemes are inconsistent between institutions. Partial mappings are maintained in terminology databases such as the United Medical Language System (UMLS) and the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT). In order to use multi-source clinical data effectively, analyzing and harmonizing data across ontologies and among institutions is an ongoing area of research.

Moreover, deep learning theories have not yet provided complete solutions. For that reason, some researchers remain skeptical about the utilization of deep learning for clinical analysis. The main criticism of deep learning is the interpretability of models, which is essential for clinical application. Researchers always describe their models as an "end-to-end" way or a "black box". There are sets of hyperparameters in a deep learning model, such as the size and the number of filters in a CNN, the depth of a model, the learning rate, and the value of dropout. The setting of these hyperparameters that control the architecture of a model remains a blind exploration process that usually requires accurate validation.

Traditional machine learning methods, such as LR, SVM, DT, RF, CRF, and BN, have also been used effectively to analyze clinical data[71, 146]. Putting aside their performance, they have better interpretability and are less time-consuming than multilayer deep learning models in most cases. In fact, an interpretable model that can analyze data quickly enough and provide decision support efficiently is what we are seeking in clinical research.

5.2 Outlook

With the development of deep learning techniques in the big data era, there are opportunities for future deep clinical research. The fusion of heterogeneous data and the reasonable combination of different approaches have come to the foreground as promising ways

forward. The integration of clinical data, genomics data, and social behavior data may make precision medicine reality just as we expected, realizing the goal is "to provide the right treatment to the right patient at the right time"[147]. Furthermore, various combinations of deep learning models can be expected to achieve better performance. In order to model an interpretable human-like computation system, it can be helpful to join deep learning methods with medical ontologies, rule-based systems, and traditional machine learning solutions. Moreover, developing systems which can perform real-time analysis on continuous vital signs would help medical staff observe life-threatening pathological changes in a timely fashion and provide appropriate treatment as early as possible.

In conclusion, this paper provided a brief overview of deep learning applications on clinical data. It presented the categories of clinical data and their characteristics, an introduction to the common deep learning models used in clinical studies, a summary of the various applications, and a discussion of challenges and the further outlook. The paper's aim was to provide valuable insights for researchers concerning clinical data studies. The characteristics of clinical data and the variety of types of data bring both opportunities and challenges. It is encouraging that deep learning methodologies have improved predictive models in many cases. However, the interpretability of such models remains an elusive goal. With a further understanding of deep learning architecture, there is the hope of better understanding the predictions and recommendations given by deep learning models. Collaboration with other approaches can result in greater achievements in clinical analysis and provide effective assistance to clinical decision making in the foreseeable future.

Acknowledgment

References

[1]   I. Kurt, M. Ture, and A. T. Kurum, Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease, *Expert Syst. Appl.*, vol. 34, no. 1, pp. 366–374, 2008.

[2] L. V. Lita, S. P. Yu, S. Niculescu, and J. B. Bi, Large scale diagnostic code classification for medical patient records, in *Proc. 3rd Int. Joint Conf. Natural Language Processing*, Hyderabad, India, 2008, pp. 877–882.

[3] B. Koopman, G. Zuccon, A. Nguyen, A. Bergheim, and N. Grayson, Automatic ICD-10 classification of cancers from free-text death certificates, *Int. J. Med. Inform.*, vol. 84, no. 11, pp. 956–965, 2015.

[4] D. X. Wang, X. Liu, and M. D. Wang, A DT-SVM strategy for stock futures prediction with big data, in *Proc. 16th Int. Conf. Computational Science and Engineering*, Sydney, Australia, 2013, pp. 1005–1012.

[5] A. Casillas, K. Gojenola, A. Pérez, and M. Oronoz, Clinical text mining for efficient extraction of drug-allergy reactions, in *Proc. 2016 IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM)*, Shenzhen, China, 2016, pp. 946–952.

[6] G. Moharasar and T. B. Ho, A semi-supervised approach for temporal information extraction from clinical text, in *Proc. 2016 IEEE RIVF Int. Conf. Computing & Communication Technologies, Research, Innovation, and Vision for the Future (RIVF)*, Hanoi, Vietnam, 2016, pp. 7–12.

[7] K. Orphanou, A. Stassopoulou, and E. Keravnou, Temporal abstraction and temporal Bayesian networks in clinical domains: A survey, *Artif. Intell. Med.*, vol. 60, no. 3, pp. 133–149, 2014.

[8] A. Subasi and M. I. Gursoy, EEG signal classification using PCA, ICA, LDA and support vector machines, *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8659–8666, 2010.

[9] M. Pechenizkiy, A. Tsymbal, and S. Puuronen, PCA-based feature transformation for classification: Issues in medical diagnostics, in *Proc. 17th IEEE Symp. Computer-Based Medical Systems*, Bethesda, MD, USA, 2004, pp. 535–540.

[10] A. Rumshisky, M. Ghassemi, T. Naumann, P. Szolovits, V. M. Castro, T. H. McCoy, and R. H. Perlis, Predicting early psychiatric readmission with natural language processing of narrative discharge summaries, *Transl. Psychiatry*, vol. 6, no. 10, p. e921, 2016.

[11] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[12] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, Corrigendum: Dermatologist-level classification of skin cancer with deep neural networks, *Nature*, vol. 546, no. 7660, p. 686, 2017.

[13] A. Rajkomar, E. Oren, K. Chen, A. M. Dai, N. Hajaj, M. Hardt, P. J. Liu, X. B. Liu, J. Marcus, M. M. Sun, et al., Scalable and accurate deep learning for electronic health records, *npj Digit. Med.*, vol. 1, no. 1, p. 18, 2018.

[14] M. Zeng, M. Li, Z. H. Fei, F. X. Wu, Y. H. Li, Y. Pan, and J. X. Wang, A deep learning framework for identifying essential proteins by integrating multiple types of biological information, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, doi: 10.1109/TCBB.2019.2897679.

[15] L. L. Liu, S. W. Chen, F. H. Zhang, F. X. Wu, Y. Pan, and J. X. Wang, Deep convolutional neural network for automatically segmenting acute ischemic stroke lesion in multi-modality MRI, *Neural Comput. Appl.*, doi: 10.1007/s00521-019-04096-x.

[16] D. Ravì, C. Wong, F. Deligianni, M. Berthelot, J. Andreu-Perez, B. Lo, and G. Z. Yang, Deep learning for health informatics, *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 4–21, 2017.

[17] R. Fang, S. Pouyanfar, Y. M. Yang, S. C. Chen, and S. S. Iyengar, Computational health informatics in the big data age: A survey, *ACM Comput. Surv.*, vol. 49, no. 1, p. 12, 2016.

[18] D. G. Shen, G. R. Wu, and H. I. Suk, Deep learning in medical image analysis, *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–248, 2017.

[19] J. G. Lee, S. Jun, Y. W. Cho, H. Lee, G. B. Kim, J. B. Seo, and N. Kim, Deep learning in medical imaging: General overview, *Korean J. Radiol.*, vol. 18, no. 4, pp. 570–584, 2017.

[20] J. M. Sun and C. K. Reddy, Big data analytics for healthcare, in *Proc. 19th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Chicago, IL, USA, 2013, p. 1525.

[21] A. Belle, R. Thiagarajan, S. M. R. Soroushmehr, F. Navidi, D. A. Beard, and K. Najarian, Big data analytics in healthcare, *BioMed Res. Int.*, vol. 2015, p. 370194, 2015.

[22] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, A survey on deep learning in medical image analysis, *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[23] P. Burlina, D. E. Freund, N. Joshi, Y. Wolfson, and N. M. Bressler, Detection of age-related macular degeneration via deep learning, in *Proc. 13th Int. Symp. Biomedical Imaging (ISBI)*, Prague, Czech Republic, 2016, pp. 184–188.

[24] 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*, vol. 316, no. 22, pp. 2402–2410, 2016.

[25] D. S. Kermany, M. Goldbaum, W. J. Cai, C. C. S. Valentim, H. Y. Liang, S. L. Baxter, A. McKeown, G. Yang, X. K. Wu, F. B. Tan, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[26] J. Liu, Y. Pan, M. Li, Z. Y. Chen, L. Tang, C. Q. Lu, and J. X. Wang, Applications of deep learning to MRI images: A survey, *Big Data Mining and Analytics*, vol. 1, no. 1, pp. 1–18, 2018.

[27] T. Brosch, R. Tam, and Alzheimer's Disease Neuroimaging Initiative, Manifold learning of brain MRIs by deep learning, in *Proc. 16th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Nagoya, Japan, 2013, pp. 633–640.

[28] S. M. Plis, D. R. Hjelm, R. Salakhutdinov, E. A. Allen, H. J. Bockholt, J. D. Long, H. J. Johnson, J. S. Paulsen, J. A. Turner, and V. D. Calhoun, Deep learning for

neuroimaging: A validation study, *Front. Neurosci.*, vol. 8, p. 229, 2014.

[29]   O. Emad, I. A. Yassine, and A. S. Fahmy, Automatic localization of the left ventricle in cardiac MRI images using deep learning, in *Proc. 37th Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, 2015, pp. 683–686.

[30]   B. Kong, Y. Q. Zhan, M. Shin, T. Denny, and S. T. Zhang, Recognizing end-diastole and end-systole frames via deep temporal regression network, in *Proc. 19th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Athens, Greece, 2016, pp. 264–272.

[31]   T. A. Ngo, Z. Lu, and G. Carneiro, Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance, *Med. Image Anal.*, vol. 35, pp. 159–171, 2017.

[32]   A. Jamaludin, T. Kadir, and A. Zisserman, SpineNet: Automatically pinpointing classification evidence in spinal MRIs, in *Proc. 19th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Athens, Greece, 2016, pp. 949–954.

[33]   D. Forsberg, E. Sjöblom, and J. L. Sunshine, Detection and labeling of vertebrae in MR images using deep learning with clinical annotations as training data, *J. Digit. Imaging*, vol. 30, no. 4, pp. 406–412, 2017.

[34]   A. Prasoon, K. Petersen, C. Igel, F. Lauze, E. Dam, and M. Nielsen, Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network, in *Proc. 16th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Nagoya, Japan, 2013, pp. 246–253.

[35]   P. J. Hu, F. Wu, J. L. Peng, Y. Y. Bao, F. Chen, and D. X. Kong, Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets, *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 3, pp. 399–411, 2017.

[36]   W. Thong, S. Kadoury, N. Piché, and C. J. Pal, Convolutional networks for kidney segmentation in contrast-enhanced CT scans, *Comput. Methods Biomech. Biomed. Eng.: Imaging Vis.*, vol. 6, no. 3, pp. 277–282, 2018.

[37]   F. Lu, F. Wu, P. J. Hu, Z. Y. Peng, and D. X. Kong, Automatic 3D liver location and segmentation via convolutional neural network and graph cut, *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 2, pp. 171–182, 2017.

[38]   M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, and S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1207–1216, 2016.

[39]   S. Christodoulidis, M. Anthimopoulos, L. Ebner, A. Christe, and S. Mougiakakou, Multisource transfer learning with convolutional neural networks for lung pattern analysis, *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 76–84, 2017.

[40]   S. R. Tarando, C. Fetita, A. Faccinetto, and P. Y. Brillet, Increasing CAD system efficacy for lung texture analysis using a convolutional network, in *Proc. Medical Imaging 2016: Computer-Aided Diagnosis*, San Diego, CA, USA, 2016, p. 97850Q.

[41]   Y. Bar, I. Diamant, L. Wolf, and H. Greenspan, Deep learning with non-medical training used for chest pathology identification, in *Proc. Medical Imaging 2015: Computer-Aided Diagnosis*, Orlando, FL, USA, 2015, p. 94140V.

[42]   M. Cicero, A. Bilbily, E. Colak, T. Dowdell, B. Gray, K. Perampaladas, and J. Barfett, Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs, *Invest. Radiol.*, vol. 52, no. 5, pp. 281–287, 2017.

[43]   S. Hwang, H. E. Kim, J. Jeong, and H. J. Kim, A novel approach for tuberculosis screening based on deep convolutional neural networks, in *Proc. Medical Imaging 2016: Computer-Aided Diagnosis*, San Diego, CA, USA, 2016, p. 97852W.

[44]   A. Dubrovina, P. Kisilev, B. Ginsburg, S. Hashoul, and R. Kimmel, Computational mammography using deep neural networks, *Comput. Methods Biomech. Biomed. Eng.: Imaging Vis.*, vol. 6, no. 3, pp. 243–247, 2018.

[45]   T. Kooi, B. van Ginneken, N. Karssemeijer, and A. den Heeten, Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network, *Med. Phys.*, vol. 44, no. 3, pp. 1017–1027, 2017.

[46]   T. Kooi, G. Litjens, B. van Ginneken, A. Gubern-Mérida, C. I. Sánchez, R. Mann, A. den Heeten, and N. Karssemeijer, Large scale deep learning for computer aided detection of mammographic lesions, *Med. Image Anal.*, vol. 35, pp. 303–312, 2017.

[47]   A. Perotte, R. Pivovarov, K. Natarajan, N. Weiskopf, F. Wood, and N. Elhadad, Diagnosis code assignment: Models and evaluation metrics, *J. Am. Med. Inform. Assoc.*, vol. 21, no. 2, pp. 231–237, 2014.

[48]   M. Li, Z. H. Fei, M. Zeng, F. X. Wu, Y. H. Li, Y. Pan, and J. X. Wang, Automated ICD-9 coding via a deep learning approach, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, doi: 10.1109/TCBB.2018.2817488.

[49]   H. R. Shi, P. T. Xie, Z. T. Hu, M. Zhang, and E. P. Xing, Towards automated ICD coding using deep learning, arXiv preprint arXiv: 1711.04075, 2017.

[50]   T. Baumel, J. Nassour-Kassis, R. Cohen, M. Elhadad, and N. Elhadad, Multi-label classification of patient notes a case study on ICD code assignment, arXiv preprint arXiv: 1709.09587, 2017.

[51]   M. Zeng, M. Li, Z. H. Fei, Y. Yu, Y. Pan, and J. X. Wang, Automatic ICD-9 coding via deep transfer learning, *Neurocomputing*, vol. 324, pp. 43–50, 2019.

[52]   S. Gehrmann, F. Dernoncourt, Y. R. Li, E. T. Carlson, J. T. Wu, J. Welt, J. Jr. Foote, E. T. Moseley, D. W. Grant, P. D. Tyler, et al., Comparing rule-based and deep learning models for patient phenotyping, arXiv preprint arXiv: 1703.08705, 2017.

[53] S. Gao, M. T. Young, J. X. Qiu, H. J. Yoon, J. B. Christian, P. A. Fearn, G. D. Tourassi, and A. Ramanthan, Hierarchical attention networks for information extraction from cancer pathology reports, *J. Am. Med. Inform. Assoc.*, vol. 25, no. 3, pp. 321–330, 2018.

[54] H. C. Shin, L. Lu, L. Kim, A. Seff, J. H. Yao, and R. M. Summers, Interleaved text/image Deep Mining on a large-scale radiology database for automated image interpretation, *J. Mach. Learn. Res.*, vol. 17, no. 1, pp. 3729–3759, 2016.

[55] L. Dai, R. G. Fang, H. T. Li, X. H. Hou, B. Sheng, Q. Wu, and W. P. Jia, Clinical report guided retinal microaneurysm detection with multi-sieving deep learning, *IEEE Trans. Med. Imaging*, vol. 37, no. 5, pp. 1149–1161, 2018.

[56] F. Duarte, B. Martins, C. S. Pinto, and M. J. Silva, A deep learning method for ICD-10 coding of free-text death certificates, in *Proc. 18th EPIA Conf. Artificial Intelligence*, Porto, Portugal, 2017, pp. 137–149.

[57] A. Perotte, R. Ranganath, J. S. Hirsch, D. Blei, and N. Elhadad, Risk prediction for chronic kidney disease progression using heterogeneous electronic health record data and time series analysis, *J. Am. Med. Inform. Assoc.*, vol. 22, no. 4, pp. 872–880, 2015.

[58] M. Carrara, G. Baselli, and M. Ferrario, Mortality prediction in septic shock patients: Towards new personalized models in critical care, in *Proc. 37th Annu. Int. Conf. of the Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, 2015, pp. 2792–2795.

[59] M. M. Ghassemi, S. E. Richter, I. M. Eche, T. W. Chen, J. Danziger, and L. A. Celi, A data-driven approach to optimized medication dosing: A focus on heparin, *Intensive Care Med.*, vol. 40, no. 9, pp. 1332–1339, 2014.

[60] S. Nemati, M. M. Ghassemi, and G. D. Clifford, Optimal medication dosing from suboptimal clinical examples: A deep reinforcement learning approach, in *Proc. 38th Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, USA, 2016, pp. 2978–2981.

[61] M. Dunitz, G. Verghese, and T. Heldt, Predicting hyperlactatemia in the MIMIC II database, in *Proc. 37th Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, 2015, pp. 985–988.

[62] K. Gunnarsdottir, V. Sadashivaiah, M. Kerr, S. Santaniello, and S. V. Sarma, Using demographic and time series physiological features to classify sepsis in the intensive care unit, in *Proc. 38th Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, USA, 2016, pp. 778–782.

[63] A. Lanata, G. Valenza, M. Nardelli, C. Gentili, and E. P. Scilingo, Complexity index from a personalized wearable monitoring system for assessing remission in mental health, *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 132–139, 2015.

[64] L. W. Li, R. P. Adams, L. Mayaud, G. B. Moody, A. Malhotra, R. G. Mark, and S. Nemati, A physiological time series dynamics-based approach to

patient monitoring and outcome prediction, *IEEE J. Biomed. Health Inform.*, vol. 19, no. 3, pp. 1068–1076, 2015.

[65] Z. Feng, R. R. Bhat, X. Y. Yuan, D. Freeman, T. Baslanti, A. Bihorac, and X. L. Li, Intelligent perioperative system: Towards real-time big data analytics in surgery risk assessment, in *15th Intl Conf. on Dependable, Autonomic and Secure Computing, 15th Intl Conf. on Pervasive Intelligence and Computing, 3rd Intl Conf. on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC / PiCom / DataCom / CyberSciTech)*, Orlando, FL, USA, 2017, pp. 1254–1259.

[66] P. X. Cheng, P. Yin, P. S. Ning, L. J. Wang, X. J. Cheng, Y. N. Liu, D. C. Schwebel, J. M. Liu, J. L. Qi, G. Q. Hu, et al., Trends in traumatic brain injury mortality in China, 2006–2013: A population-based longitudinal study, *PLoS Med.*, vol. 14, no. 7, p. e1002332, 2017.

[67] Healthcare Cost and Utilization Project (HCUP), https://www.hcup-us.ahrq.gov, 2019.

[68] F. O. Otite, P. Khandelwal, A. M. Malik, S. Chaturvedi, R. L. Sacco, and J. G. Romano, Ten-year temporal trends in medical complications after acute intracerebral hemorrhage in the United States, *Stroke*, vol. 48, no. 3, pp. 596–603, 2017.

[69] J. Y. Zhou, J. Liu, V. A. Narayan, and J. P. Ye, Modeling disease progression via fused sparse group lasso, in *Proc. 18th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Beijing, China, 2012, pp. 1095–1103.

[70] J. Liu, J. X. Wang, Z. J. Tang, B. Hu, F. X. Wu, and Y. Pan, Improving Alzheimer's disease classification by combining multiple measures, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 5, pp. 1649–1659, 2018.

[71] L. L. Liu, J. X. Wang, M. Li, F. X. Wu, H. D. Li, Y. Yu, and Z. H. Fei, An interpretable model for predicting side effects of analgesics for osteoarthritis, in *Proc. 2017 IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM)*, Kansas City, MO, USA, 2017, pp. 861–864.

[72] D. C. Mackey, L. Y. Lui, P. M. Cawthon, K. Ensrud, K. Yaffe, and S. R. Cummings, Life-space mobility and mortality in older women: Prospective results from the study of osteoporotic fractures, *J. Am. Geriatr. Soc.*, vol. 64, no. 11, pp. 2226–2234, 2016.

[73] Grand-Challenges, Grand challenges in biomedical image analysis, http:www.grand-challenge.org, 2019.

[74] Ö. Uzuner, B. R. South, S. Y. Shen, and S. L. DuVall, 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text, *J. Am. Med. Inform. Assoc.*, vol. 18, no. 5, pp. 552–556, 2011.

[75] Ö. Uzuner, A. Bodnari, S. Y. Shen, T. Forbush, J. Pestian, and B. R. South, Evaluating the state of the art in coreference resolution for electronic medical records, *J. Am. Med. Inf. Assoc.*, vol. 19, no. 5, pp. 786–791, 2012.

[76] W. Y. Sun, A. Rumshisky, and Ö. Uzuner, Evaluating temporal relations in clinical text: 2012 i2b2 challenge, *J. Am. Med. Inform. Assoc.*, vol. 20, no. 5, pp. 806–813, 2013.

[77] A. Stubbs and Ö. Uzuner, Annotating risk factors for

heart disease in clinical narratives for diabetic patients, *J. Biomed. Inform.*, vol. 58, no. Suppl, pp. S78–S91, 2015.

[78] M. Filannino, A. Stubbs, and Ö. Uzuner, Symptom severity prediction from neuropsychiatric clinical records: Overview of 2016 CEGS N-GRID shared tasks track 2, *J. Biomed. Inform.*, vol. 75, no. Suppl, pp. S62–S70, 2017.

[79] G. B. Moody and R. G. Mark, A database to support development and evaluation of intelligent intensive care monitoring, in *Proc. 1996 Computers in Cardiology*, Indianapolis, IN, USA, 1996, pp. 657–660.

[80] M. Saeed, M. Villarroel, A. T. Reisner, G. Clifford, L. W. Lehman, G. Moody, T. Heldt, T. H. Kyaw, B. Moody, and R. G. Mark, Multiparameter intelligent monitoring in intensive care II: A public-access intensive care unit database, *Crit. Care Med.*, vol. 39, no. 5, pp. 952–960, 2011.

[81] J. Lee, D. J. Scott, M. Villarroel, G. D. Clifford, M. Saeed, and R. G. Mark, Open-access MIMIC-II database for intensive care research, in *Proc. 2011 Annu. Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Boston, MA, USA, 2011, pp. 8315–8318.

[82] R. Miotto, L. Li, B. A. Kidd, and J. T. Dudley, Deep patient: An unsupervised representation to predict the future of patients from the electronic health records, *Sci. Rep.*, vol. 6, p. 26094, 2016.

[83] E. Choi, M. T. Bahadori, E. Searles, C. Coffey, M. Thompson, J. Bost, J. Tejedor-Sojo, and J. M. Sun, Multi-layer representation learning for medical concepts, in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 1495–1504.

[84] Y. LeCun and Y. Bengio, Convolutional networks for images, speech, and time series, in *Handbook of Brain Theory and Neural Networks*, M. A. Arbib, ed. Cambridge, MA, USA: MIT Press, 1995.

[85] W. F. Yin, X. Z. Yang, L. Zhang, and E. Oki, ECG monitoring system integrated with IR-UWB radar based on CNN, *IEEE Access*, vol. 4, pp. 6344–6351, 2016.

[86] H. Z. Fu, J. Cheng, Y. W. Xu, D. W. K. Wong, J. Liu, and X. C. Cao, Joint optic disc and cup segmentation based on multi-label deep network and polar transformation, *IEEE Trans. Med. Imaging*, vol. 37, no. 7, pp. 1597–1605, 2018.

[87] F. Milletari, N. Navab, and S. A. Ahmadi, V-Net: Fully convolutional neural networks for volumetric medical image segmentation, in *Proc. 4th Int. Conf. 3D Vision (3DV)*, Stanford, CA, USA, 2016, pp. 565–571.

[88] A. Fakhry, T. Zeng, and S. W. Ji, Residual deconvolutional networks for brain electron microscopy image segmentation, *IEEE Trans. Med. Imaging*, vol. 36, no. 2, pp. 447–456, 2017.

[89] R. Guerrero, C. Qin, O. Oktay, C. Bowles, L. Chen, R. Joules, R. Wolz, M. C. Valdés-Hernández, D. A. Dickie, J. Wardlaw, et al., White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks, *NeuroImage Clin.*, vol. 17, pp. 918–934, 2018.

[90] D. M. Pelt and J. A. Sethian, A mixed-scale dense convolutional neural network for image analysis, *Proc. Natl. Acad. Sci. USA*, vol. 115, no. 2, pp. 254–259, 2018.

[91] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, Densely connected convolutional networks, in *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 2261–2269.

[92] Q. L. Zhu, X. L. Li, A. Conesa, and C. Pereira, GRAM-CNN: A deep learning approach with local context for named entity recognition in biomedical text, *Bioinformatics*, vol. 34, no. 9, pp. 1547–1554, 2018.

[93] K. Liu, S. W. Peng, J. Q. Wu, C. X. Zhai, H. Mamitsuka, and S. F. Zhu, MeSHLabeler: Improving the accuracy of large-scale MeSH indexing by integrating diverse evidence, *Bioinformatics*, vol. 31, no. 12, pp. i339–i347, 2015.

[94] S. W. Peng, R. H. You, H. N. Wang, C. X. Zhai, H. Mamitsuka, and S. F. Zhu, DeepMeSH: Deep semantic representation for improving large-scale MeSH indexing, *Bioinformatics*, vol. 32, no. 12, pp. i70–i79, 2016.

[95] S. Hochreiter and J. Schmidhuber, Long short-term memory, *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[96] E. Choi, M. T. Bahadori, A. Schuetz, W. F. Stewart, and J. M. Sun, Doctor AI: Predicting clinical events via recurrent neural networks, in *Proc. 2016 Machine Learning for Healthcare Conf.*, Los Angeles, CA, USA, 2016, pp. 301–318.

[97] Y. H. Wu, M. Jiang, J. B. Lei, and H. Xu, Named entity recognition in Chinese clinical text using deep neural network, *Stud. Health Technol. Inform.*, vol. 216, pp. 624–628, 2015.

[98] A. N. Jagannatha and H. Yu, Bidirectional RNN for medical event detection in electronic health records, in *Proc. Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, San Diego, CA, USA, 2016, pp. 473–482.

[99] A. Khatami, A. Khosravi, T. Nguyen, C. P. Lim, and S. Nahavandi, Medical image analysis using wavelet transform and deep belief networks, *Exp. Syst. Appl.*, vol. 86, pp. 190–198, 2017.

[100] K. L. Hua, C. H. Hsu, S. C. Hidayati, W. H. Cheng, and Y. J. Chen, Computer-aided classification of lung nodules on computed tomography images via deep learning technique, *Onco Targets Ther.*, vol. 8, pp. 2015–2022, 2015.

[101] G. E. Hinton and R. R. Salakhutdinov, Reducing the dimensionality of data with neural networks, *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[102] H. I. Suk and D. G. Shen, Deep learning-based feature representation for AD/MCI classification, in *Proc. 16th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Nagoya, Japan, 2013, pp. 583–590.

[103] S. Sharma, I. Umar, L. Ospina, D. Wong, and H. R. Tizhoosh, Stacked autoencoders for medical image

search, in *Proc. 12th Int. Symp. Visual Computing*, Las Vegas, NV, USA, 2016, pp. 45–54.

[104] J. X. Chen, L. Yang, Y. Z. Zhang, M. Alber, and D. Z. Chen, Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation, in *Proc. 30th Conf. Neural Information Processing Systems (NIPS)*, Barcelona, Spain, 2016, pp. 3036–3044.

[105] G. S. Birkhead, M. Klompas, and N. R. Shah, Uses of electronic health records for public health surveillance to advance public health, *Annu. Rev. Public Health*, vol. 36, pp. 345–359, 2015.

[106] F. L. Ma, R. Chitta, J. Zhou, Q. Z. You, T. Sun, and J. Gao, Dipole: Diagnosis prediction in healthcare via attention-based bidirectional recurrent neural networks, in *Proc. 23rd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Halifax, Canada, 2017, pp. 1903–1911.

[107] A. Prakash, S. Y. Zhao, S. A. Hasan, V. Datla, K. Lee, A. Qadir, J. Liu, and O. Farri, Condensed memory networks for clinical diagnostic inferencing, in *Proc. 31st AAAI Conf. Artificial Intelligence*, San Francisco, CA, USA, 2017, pp. 3274–3280.

[108] E. Choi, M. T. Bahadori, J. A. Kulas, A. Schuetz, W. F. Stewart, and J. M. Sun, RETAIN: An interpretable predictive model for healthcare using reverse time attention mechanism, in *Proc. 30th Int. Conf. Neural Information Processing Systems*, Barcelona, Spain, 2016, pp. 3504–3512.

[109] F. Duarte, B. Martins, C. S. Pinto, and M. J. Silva, Deep neural models for ICD-10 coding of death certificates and autopsy reports in free-text, *J. Biomed. Inform.*, vol. 80, pp. 64–77, 2018.

[110] D. L. Guo, M. Li, Y. Yu, Y. H. Li, G. H. Duan, F. X. Wu, and J. X. Wang, Disease inference with symptom extraction and bidirectional recurrent neural network, in *Proc. 2018 Int. Conf. Bioinformatics and Biomedicine (BIBM)*, Madrid, Spain, 2018, pp. 864–868

[111] Y. Yu, M. Li, L. L. Liu, Z. H. Fei, F. X. Wu, and J. X. Wang, Automatic ICD code assignment of Chinese clinical notes based on multilayer attention BiRNN, *J. Biomed. Inform.*, vol. 91, p. 103114, 2019.

[112] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado, L. Peng, and D. R. Webster, Predicting cardiovascular risk factors from retinal fundus photographs using deep learning, arXiv preprint arXiv: 1708.09843, 2017.

[113] H. Li, X. Y. Li, M. Ramanathan, and A. D. Zhang, Prediction and informative risk factor selection of bone diseases, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 12, no. 1, pp. 79–91, 2015.

[114] J. Donzé, D. Aujesky, D. Williams, and J. L. Schnipper, Potentially avoidable 30-day hospital readmissions in medical patients: Derivation and validation of a prediction model, *JAMA Intern. Med.*, vol. 173, no. 8, pp. 632–638, 2013.

[115] J. Futoma, J. Morris, and J. Lucas, A comparison of models for predicting early hospital readmissions, *J. Biomed. Inform.*, vol. 56, pp. 229–238, 2015.

[116] P. Nguyen, T. Tran, N. Wickramasinghe, and S. Venkatesh, Deepr: A convolutional net for medical records, *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 22–30, 2017.

[117] T. Pham, T. Tran, D. Phung, and S. Venkatesh, DeepCare: A deep dynamic memory model for predictive medicine, in *Proc. 20th Pacific-Asia Conf. Knowledge Discovery and Data Mining*, Auckland, New Zealand, 2016, pp. 30–41.

[118] P. Grnarova, F. Schmidt, S. L. Hyland, and C. Eickhoff, Neural document embeddings for intensive care patient mortality prediction, arXiv preprint arXiv: 1612.00467, 2016.

[119] G. Carneiro, L. Oakden-Rayner, A. P. Bradley, J. Nascimento, and L. Palmer, Automated 5-year mortality prediction using deep learning and radiomics features from chest computed tomography, in *Proc. 14th Int. Symp. Biomedical Imaging (ISBI)*, Melbourne, Australia, 2017, pp. 130–134.

[120] H. K. van der Burgh, R. Schmidt, H. J. Westeneng, M. A. de Reus, L. H. van den Berg, and M. P. van den Heuvel, Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis, *NeuroImage Clin.*, vol. 13, pp. 361–369, 2017.

[121] R. H. Ju, C. H. Hu, P. Zhou, and Q. Z. Li, Early diagnosis of Alzheimer's disease based on resting-state brain networks and deep learning, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 1, pp. 244–257, 2019.

[122] D. H. Lu, K. Popuri, G. W. Ding, R. Balachandar, M. F. Beg, and Alzheimer's Disease Neuroimaging Initiative, Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images, *Sci. Rep.*, vol. 8, p. 5697, 2018.

[123] E. Choi, A. Schuetz, W. F. Stewart, and J. M. Sun, Using recurrent neural network models for early detection of heart failure onset, *J. Am. Med. Inform. Assoc.*, vol. 24, no. 2, pp. 361–370, 2017.

[124] Y. Luo, Recurrent neural networks for classifying relations in clinical notes, *J. Biomed. Inform.*, vol. 72, pp. 85–95, 2017.

[125] Z. J. Liu, B. Z. Tang, X. L. Wang, and Q. C. Chen, De-identification of clinical notes via recurrent neural network and conditional random field, *J. Biomed. Inform.*, vol. 75 Suppl, pp. S34–S42, 2017.

[126] F. Dernoncourt, J. Y. Lee, Ö. Uzuner, and P. Szolovits, De- identification of patient notes with recurrent neural networks, *J. Am. Med. Inform. Assoc.*, vol. 24, no. 3, pp. 596–606, 2017.

[127] C. Payer, D. Štern, H. Bischof, and M. Urschler, Regressing heatmaps for multiple landmark localization using CNNs, in *Proc. 19th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Athens, Greece, 2016, pp. 230–238.

[128] B. D. de Vos, J. M. Wolterink, P. A. de Jong, M. A. Viergever, and I. Išgum, 2D image classification for 3D anatomy localization: Employing deep convolutional

neural networks, in *Proc. Medical Imaging 2016: Image Processing*, San Diego, CA, USA, 2016, p. 97841Y.

[129] Q. Dou, H. Chen, L. Q. Yu, L. Zhao, J. Qin, D. F. Wang, V. C. T. Mok, L. Shi, and P. A. Heng, Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks, *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1182–1195, 2016.

[130] X. R. Liu, H. R. Tizhoosh, and J. Kofman, Generating binary tags for fast medical image retrieval based on convolutional nets and Radon Transform, in *Proc. 2016 Int. Joint Conf. Neural Networks (IJCNN)*, Vancouver, Canada, 2016, pp. 2872–2878.

[131] A. Shah, S. Conjeti, N. Navab, and A. Katouzian, Deeply learnt hashing forests for content based image retrieval in prostate MR images, in *Proc. Medical Imaging 2016: Image Processing*, San Diego, CA, USA, 2016, p. 978414.

[132] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images, *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1273–1284, 2016.

[133] T. D. Vu, H. J. Yang, V. Q. Nguyen, A. R. Oh, and M. S. Kim, Multimodal learning using convolution neural network and Sparse Autoencoder, in *Proc. 2017 Int. Conf. Big Data and Smart Computing (BigComp)*, Jeju, South Korea, 2017, pp. 309–312.

[134] D. H. Lu, K. Popuri, G. W. Ding, R. Balachandar, and M. F. Beg, Multiscale deep neural network based analysis of FDG-PET images for the early diagnosis of Alzheimer's disease, *Med. Image Anal.*, vol. 46, pp. 26–34, 2018.

[135] D. N. Cheng and M. H. Liu, Combining convolutional and recurrent neural networks for Alzheimer's disease diagnosis using PET images, in *Proc. 2017 IEEE Int. Conf. Imaging Systems and Techniques (IST)*, Beijing, China, 2017, pp. 1–5.

[136] A. Ortiz, J. Munilla, J. M. Górriz, and J. Ramírez, Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease, *Int. J. Neural Syst.*, vol. 26, no. 7, p. 1650025, 2016.

[137] A. Payan and G. Montana, Predicting Alzheimer's disease: A neuroimaging study with 3D convolutional neural networks, arXiv preprint arXiv: 1502.02506, 2015.

[138] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. H. Ying, Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for

[139] diagnosis of Alzheimer's disease, *IEEE J. Biomed. Health Inform.*, vol. 22, no. 1, pp. 173–183, 2018.

[139] A. Suzani, A. Rasoulian, A. Seitel, S. Fels, R. N. Rohling, and P. Abolmaesumi, Deep learning for automatic localization, identification, and segmentation of vertebral bodies in volumetric MR images, in *Proc. Medical Imaging 2015: Image-Guided Procedures, Robotic Interventions, and Modeling*, Orlando, FL, USA, 2015, pp. 773–778.

[140] H. R. Roth, L. Lu, A. Farag, A. Sohn, and R. M. Summers, Spatial aggregation of holistically-nested networks for automated pancreas segmentation, in *Proc. 19th Int. Conf. Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Athens, Greece, 2016, pp. 451–459.

[141] J. J. Nappi, T. Hironaka, D. Regge, and H. Yoshida, Deep transfer learning of virtual endoluminal views for the detection of polyps in CT colonography, in *Proc. Medical Imaging 2016: Computer-Aided Diagnosis*, San Diego, CA, USA, 2016, p. 97852B.

[142] G. van Tulder and M. de Bruijne, Combining generative and discriminative representation learning for lung CT analysis with convolutional restricted Boltzmann machines, *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1262–1272, 2016.

[143] J. Wang, H. J. Ding, F. M. Azamian, B. Zhou, C. Iribarren, S. Molloi, and P. Baldi, Detecting cardiovascular disease from mammograms with deep learning, *IEEE Trans. Med. Imaging*, vol. 36, no. 5, pp. 1172–1181, 2017.

[144] A. Rajkomar, S. Lingam, A. G. Taylor, M. Blum, and J. Mongan, High-throughput classification of radiographs using deep convolutional neural networks, *J. Digit. Imaging*, vol. 30, no. 1, pp. 95–101, 2017.

[145] M. Kallenberg, K. Petersen, M. Nielsen, A. Y. Ng, P. F. Diao, C. Igel, C. M. Vachon, K. Holland, R. R. Winkel, N. Karssemeijer, et al., Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring, *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1322–1331, 2016.

[146] L. L. Liu, Y. Yu, Z. H. Fei, M. Li, F. X. Wu, H. D. Li, Y. Pan, and J. X. Wang, An interpretable boosting model to predict side effects of analgesics for osteoarthritis, *BMC Syst. Biol.*, vol. 12, no. Suppl 6, p. 105, 2018.

[147] The White House, Precision medicine: Health care tailored to you, https://obamawhitehouse.archives.gov/blog/2016/02/25/precision-medicine-health-care-tailored-you, 2016.
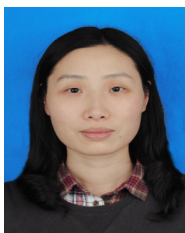
Ying Yu received the BEng and MEng degrees from University of South China in 2002 and 2009, respectively. She is studying for the PhD degree at Central South University, China. Her current research focuses on the analysis of clinical texts and electronic health records by using machine learning or deep learning methods.

Liangliang Liu received the BEng and MEng degrees from Henan University in 2006 and 2014, respectively. He is studying for the PhD degree at Central South University, China. His research focuses on machine learning, deep learning methods, medical image analysis, and electronic health records analysis.

Min Li received the PhD degree from Central South University, China, in 2008. She is currently the vice dean and a professor at the School of Computer Science and Engineering, Central South University, China. Her research interests include computational biology, systems biology, and bioinformatics. She has published more than 80 technical papers in refereed journals such as *Bioinformatics*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *Proteomics*, and conference proceedings such as BIBM, GIW, and ISBRA.

Jianxin Wang received the BEng and MEng degrees from Central South University, China, in 1992 and 1996, respectively, and the PhD degree from Central South University, China, in 2001. He is the dean and a professor in the School of Computer Science and Engineering, Central South University, China. His current research interests include algorithm analysis and optimization, parameterized algorithm, bioinformatics, and computer network. He has published more than 150 papers in various international journals and refereed conferences. He is a senior member of the IEEE.

Yaohang Li received the MS and PhD degrees from Florida State University, Tallahassee, FL, USA, in 2000 and 2003, respectively. He is an associate professor in the Department of Computer Science at Old Dominion University (ODU), Norfolk, VA, USA. His research interests are in computational biology, Monte Carlo methods, and scientific computing.