



Popular deep learning algorithms for disease prediction: a review

Zengchen Yu¹ · Ke Wang² · Zhibo Wan¹ · Shuxuan Xie¹ · Zhihan Lv³

Received: 2 February 2022 / Revised: 7 July 2022 / Accepted: 3 August 2022 / Published online: 13 September 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Due to its automatic feature learning ability and high performance, deep learning has gradually become the mainstream of artificial intelligence in recent years, playing a role in many fields. Especially in the medical field, the accuracy rate of deep learning even exceeds that of doctors. This paper introduces several deep learning algorithms: Artificial Neural Network (NN), FM-Deep Learning, Convolutional NN and Recurrent NN, and expounds their theory, development history and applications in disease prediction; we analyze the defects in the current disease prediction field and give some current solutions; our paper expounds the two major trends in the future disease prediction and medical field—integrating Digital Twins and promoting precision medicine. This study can better inspire relevant researchers, so that they can use this article to understand related disease prediction algorithms and then make better related research.

Keywords Artificial neural network · Factorization machine · Convolutional neural network · Recurrent neural network

1 Introduction

In recent years, with the development of medical detection technology, a large amount of health data has been generated, which requires corresponding big data analysis methods to process these data and generate valuable information, which is helpful for disease diagnosis, personalized medicine and other medicine activities. Artificial intelligence (AI) and machine learning can be used to

identify, analyze, predict and classify medical data [1], so in the past 10 years various AI algorithms have been effectively applied to process data generated in healthcare [2, 3], such as applying logistic regression to heart disease prediction to achieve early detection of heart disease [4]. However, when the data reaches a certain level, the efficiency of traditional Machine Learning algorithms will be significantly reduced, that is, these Machine Learning algorithms lack certain big data analysis capabilities. And deep learning algorithms, namely deep neural networks (DNNs), can solve this problem. The DNN simulates the conduction of the human brain neural network (NN), and defines the input and output through complex layers composition. Each layer composition includes corresponding neurons and nonlinear functions (activation functions) [5]. Compared with traditional machine learning, the advantage of deep learning is that it can learn from the original data and has multiple hidden layers. It can learn abstract information based on input, process massive data and obtain high accuracy and performance. Therefore, it has been applied to the medical field by many scholars.

This article will divide deep learning into two types according to data types: structured data algorithms and unstructured data algorithms. Structured data algorithms include Artificial Neural Network (ANN) and Factorization Machine-Deep Learning (FM-Deep Learning), which can play a better role in processing structured medical record

✉ Zhibo Wan
wzbdata@163.com

Zengchen Yu
1580065638@qq.com

Ke Wang
438565158@qq.com

Shuxuan Xie
473755474@qq.com

Zhihan Lv
lvzhihan@gmail.com

¹ College of Computer Science and Technology, Qingdao University, Ningxia Road, Qingdao 266071, China

² Psychiatric Department, Qingdao Municipal Hospital, Zhuhai Road, Qingdao 266071, China

³ Department of Game Design, Faculty of Arts, Uppsala University, 75105 Uppsala, Sweden

data. After the combination of FM and DNN, it can solve many problems that ordinary DNN cannot solve. FM is developed from the matrix factorization algorithm. Singular Value Decomposition (SVD), non-negative matrix decomposition and probability matrix decomposition are traditional matrix decomposition methods. They can decompose high-dimensional matrices into two or more low-dimensional matrices, which is convenient to study the properties of high-dimensional data in a low-dimensional space. These matrix factorization methods are widely used in prediction, recommendation and other fields because of their high scalability and good performance. However, traditional matrix factorization methods lack the effective use of context information. In this context, the FM model was proposed and popularized. FM was proposed by Rendle [6]. It is a supervised learning model [7], which combines the advantages of matrix decomposition and Support Vector Machine (SVM). Similar to SVM, the difference is that FM models pairwise feature interaction as the inner product of hidden vectors between features through matrix decomposition, so as to better mine feature interaction information, to reduce complexity, to solve sparsity and improve performance. FM was first applied to the Click-through Rate (CTR) prediction information behind the user's click behavior. But in real-life data are often highly non-linear, so capture high-order feature interaction information can significantly improve performance. Although FM can theoretically model high-order feature interaction, it will cause parameter explosion and huge amount of calculation, resulting in significant increase in time complexity and storage space consumption. Therefore, only second-order feature interaction modeling is usually considered. If the high-order feature combination is performed manually, there are the following disadvantages: (1) experts in related fields need to spend a lot of time to study the correlation between features, which is time-consuming and laborious; (2) for large-scale prediction system, the amount of data is huge, and it is unrealistic to extract features manually; (3) it is impossible to generalize feature interactions that are not in the training set. Deep learning can automatically perform various combinations and nonlinear transformations on the input features, so as to learn high-order feature interaction information. Therefore, the combination of deep learning and FM can capture low-order to high-order features, and can better predict whether patients have diseases and disease types.

Unstructured data algorithms include Convolutional NNs (CNN) and Recurrent NNs (RNN), etc. This article will only explore the development of CNN and RNN and their applications in the medical field. CNN [8] is a DNN structure including convolutional computation, which has the ability of representation learning and can realize translation-invariant classification of input information

according to hierarchical structure. CNNs generally include convolutional layers, batch-normalization layers, pooling layers, fully connected layers, etc. The core of which is the convolutional layer. The function of the convolution layer is to perform feature extraction on the input image. The convolution layer contains multiple convolution kernels. Each element that constitutes the convolution kernel has a corresponding weight coefficient and bias value, similar to the neurons of a feed-forward NN. Convolution calculation means that the convolution kernel slides on the image, and its corresponding elements are multiplied and summed with the covered image features. This process can achieve the effect of extracting local features and reducing parameters. Because the CNN can extract local features and reduce parameters (through weight sharing), it is particularly suitable for the field of image processing. Because there are a lot of image data in the medical field, the application range of CNN in the medical field exceeds that of other models. CNN can solve the problem of spatial dimension, but cannot process data in time dimension. The RNN [9] came into being, which consists of neurons and feedback loops. RNN has unique advantages for scenarios where the previous input and the next input have dependencies. Specifically, the network will remember the previous information and apply it to the current output calculation, that is, the nodes between the hidden layers are connected, and the input of the hidden layer includes not only the output of the input layer, but also the output of the hidden layer at the previous time. RNN can process time series data well, and is widely used in natural language processing, machine translation, speech recognition, image description generation, text similarity calculation and other fields.

This paper will explore the theories, development and disease application cases of these algorithms. Specifically, the contributions and characteristics of this paper are as follows:

- (1) According to the type of main processing data, the algorithm is divided into structured data algorithm and unstructured data algorithm.
- (2) CNN and RNN papers account for a high proportion in the field of in-depth learning, and papers on structured data processing methods are rare. Therefore, readers can understand the processing algorithms of structured data in detail through this article.
- (3) Different from the summary of classification according to disease types, this paper is classified according to the characteristics of algorithms. For example, in CNN's disease application section, some paragraphs focus on transfer learning, some paragraphs focus on combinatorial algorithms, and some paragraphs focus on combining attention mechanism.

- (4) This paper probes into the problems existing in the current research of disease prediction, such as poor interpretability, unbalanced data, poor data quality and few samples in some cases, and gives the current feasible solutions.
- (5) The two major trends in future medical care, integrating Digital Twins and promoting precision medicine, are analyzed, indicating that deep learning disease prediction has a bright future.
- (6) This paper will help relevant researchers to understand the characteristics and development trends of related disease prediction algorithms, and ensure that they can purposefully select the most appropriate algorithm in the process of doing research.

Section 2 of this paper will introduce the theories, development and disease application cases of two kinds of structured data algorithms, ANN and FM-Deep Learning. Section 3 will introduce the theories, development and disease application cases of CNN and RNN. Section 4 will respectively introduce the current defects in the field of disease prediction algorithms and the coping strategies. Section 5 analyzes the two major trends of medical treatment in the future, that is, integrating Digital Twins and promoting precise medical treatment. Section 6 summarizes the full text.

2 Structured data algorithms

2.1 Artificial neural network

2.1.1 Theory

ANN consists of multiple layers, each layer has one or more artificial neurons. Each neuron receives one or more inputs. First, each input is multiplied by a network weight (network parameter), which is generally randomly initialized. Calculate the sum of all weighted inputs and deviation values of each neuron, and then input this value into the activation function (nonlinear variation function). Activation function is the core of NN. It introduces non-linearity into the network and makes it possible for the network to learn more complex functions. The output of the activation function is the output of neurons, and the output of each layer of neurons is used as the input of the next layer of neurons. In the iterative training process, the whole network will find the optimal weight distribution, and the loss function is used to measure whether the network weight is optimal. Figure 1 is a schematic diagram of a three-layer ANN. The whole network has an input layer, hidden layers (generally multiple) and an output layer. In practical application, the number of layers of the network will reach dozens or even hundreds of layers.

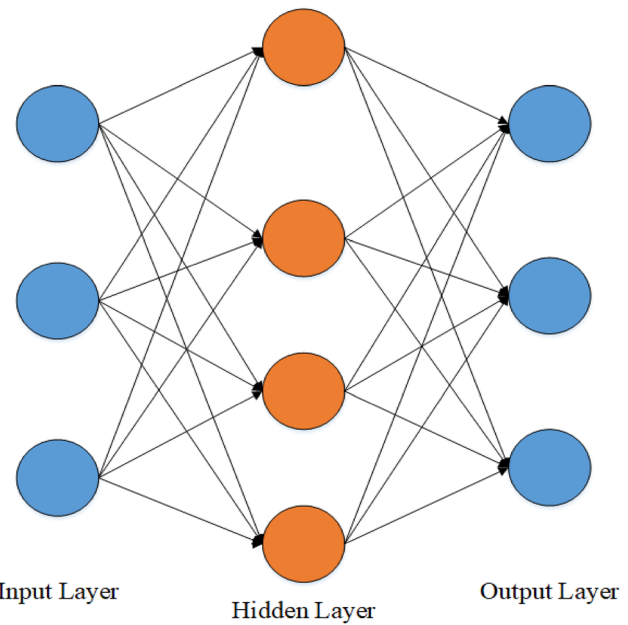


Fig. 1 Artificial neural network diagram

2.1.2 Disease application

Because the structure of ANN is relatively simple, it does not have the excellent characteristics of CNN and RNN, so there are few researches in this area [10, 11]. Khanam and Foo [12] implemented a NN model for diabetes prediction, using 1, 2, and 3 hidden layers in the NN model and changing their epochs to 200, 400, and 800, respectively. Hidden layer 2 has 400 epochs and provides 88.6% accuracy, surpassing machine learning models such as Decision Tree, K-Nearest Neighbor (KNN), Random Forest, Logistic Regression, SVM, etc. In 2021, Soundarya et al. [13] used ANN to compare with machine learning models to detect Alzheimer's Disease (AD) and found that ANN achieved the highest accuracy with sufficient data. Pasha et al. [14] used ANN to improve the prediction accuracy of cardiovascular disease. When dealing with large datasets, traditional machine learning models do not perform well, while ANN can play an advantage. These all indicate that ANN is one of the future trends, and deep learning represented by ANN will become the mainstream algorithm for disease prediction.

2.2 FM-deep learning

2.2.1 Theory

To capture second-order interactions between features, a second-order cross term is usually added to the linear regression formula:

$$y_{FM} = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n w_{ij} x_i x_j. \tag{1}$$

There are $n(n - 1)/2$ parameters in the second-order intersection part, but when finding w_{ij} , it is necessary that the features x_i and x_j are not 0 at the same time, and the sparse data (especially after one-hot code) satisfies that x_i and x_j are not 0 at the same time. There are few cases, so there are fewer samples of corresponding feature interactions in the training set, resulting in inaccurate learned w_{ij} and over-fitting. In order to solve this problem, FM decomposes w_{ij} into hidden vectors v_i and v_j , that is, $w_{ij} = \langle v_i, v_j \rangle$, where $v_i = (v_{i1}, v_{i2}, \dots, v_{ik})$ (k is a hyper-parameter, indicating the length of the hidden vector). The matrix W composed of w_{ij} can be expressed as follows:

$$W = \begin{pmatrix} v_1 \\ v_2 \\ \dots \\ v_k \end{pmatrix} (v_1 \ v_2 \ \dots \ v_k). \tag{2}$$

Now there are $n * k$ binomial parameters, far less than the original number of w_{ij} .

Why do we say that hidden vectors can solve data sparsity? Because all samples containing non-zero feature combinations of x_h can be used to learn v_h . For example, the parameters of $x_h x_i$ and $x_h x_j$ are $\langle v_h, v_i \rangle$ and $\langle v_h, v_j \rangle$, respectively. They have a common item v_h , so the value of v_h can be estimated reasonably. This can greatly reduce the impact of data sparsity.

The implicit vector mechanism can also increase the generalization of the model. According to the principle that FM can solve sparsity, when FM learns the embedded hidden vector weight of a single feature, it does not depend on whether a specific feature combination has occurred. For the feature combination $x_i x_j$ that has never appeared

before, as long as FM learns the hidden vectors corresponding to x_i and x_j , the weight of this feature combination can be calculated through the inner product, so FM has strong generalization ability. The formula of FM is as follows [15]:

$$y_{FM} = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j. \tag{3}$$

It can be seen that the complexity of FM is $O(n^2 k)$, and its complexity can be reduced to $O(n * k)$ by the following steps:

$$\begin{aligned} & \sum_{i=1}^{n-1} \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \langle v_i, v_j \rangle x_i x_j - \frac{1}{2} \sum_{i=1}^n \langle v_i, v_i \rangle x_i^2 \\ &= \frac{1}{2} \left(\sum_{i=1}^n \sum_{j=1}^n \sum_{f=1}^k v_{if} v_{jf} x_i x_j - \sum_{i=1}^n \sum_{f=1}^k v_{if}^2 x_i^2 \right) \\ &= \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{if} x_i \right) \left(\sum_{j=1}^n v_{jf} x_j \right) - \sum_{i=1}^n v_{if}^2 x_i^2 \right) \\ &= \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{if} x_i \right)^2 - \sum_{i=1}^n v_{if}^2 x_i^2 \right). \end{aligned} \tag{4}$$

The final FM equation is:

$$y_{FM} = w_0 + \sum_{i=1}^n w_i x_i + \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{if} x_i \right)^2 - \sum_{i=1}^n v_{if}^2 x_i^2 \right). \tag{5}$$

In fact, the essence of FM is embedding plus interaction, by assigning each feature x_i (discrete features will be one-hot encoded before) a implicit vector $v_i = (v_{i1}, v_{i2}, v_{i3}, v_{i4})$ (assuming here $k = 4$), change the original high-dimensional data into a low-dimensional dense vector e through the embedding layer, that is, multiply x_i by the corresponding hidden vector v_i to obtain e_i , as shown in Fig. 2.

The entire Embedding layer is shown in Fig. 3:

In summary, the overall structure of FM can be drawn, as shown in Fig. 4, where $y_{Linear} = w_0 + \sum_{i=1}^n w_i x_i$, $y_{FM2} = \frac{1}{2} \sum_{f=1}^k \left(\left(\sum_{i=1}^n v_{if} x_i \right)^2 - \sum_{i=1}^n v_{if}^2 x_i^2 \right)$.

2.2.2 Development history

In 2016, Zhang et al. [16] proposed a FM Supported NN (FNN). The model uses a DNN with embedded layers to complete the CTR prediction, which obtains the dense vector of each feature through pre training the FM model. Then all embedded vectors of the sample are spliced and

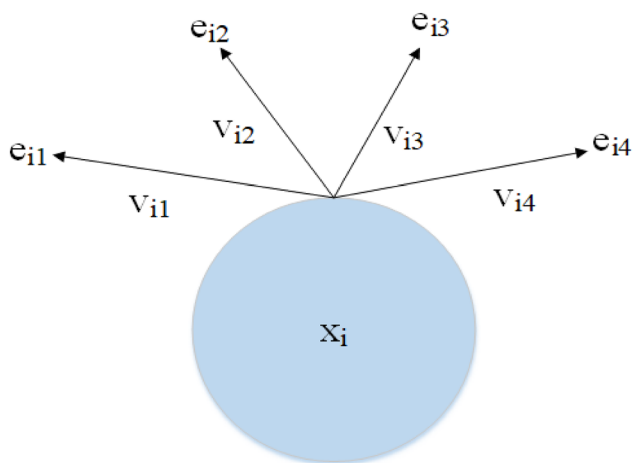


Fig. 2 Embedding of feature x_i

Fig. 3 Embedding layer of FM

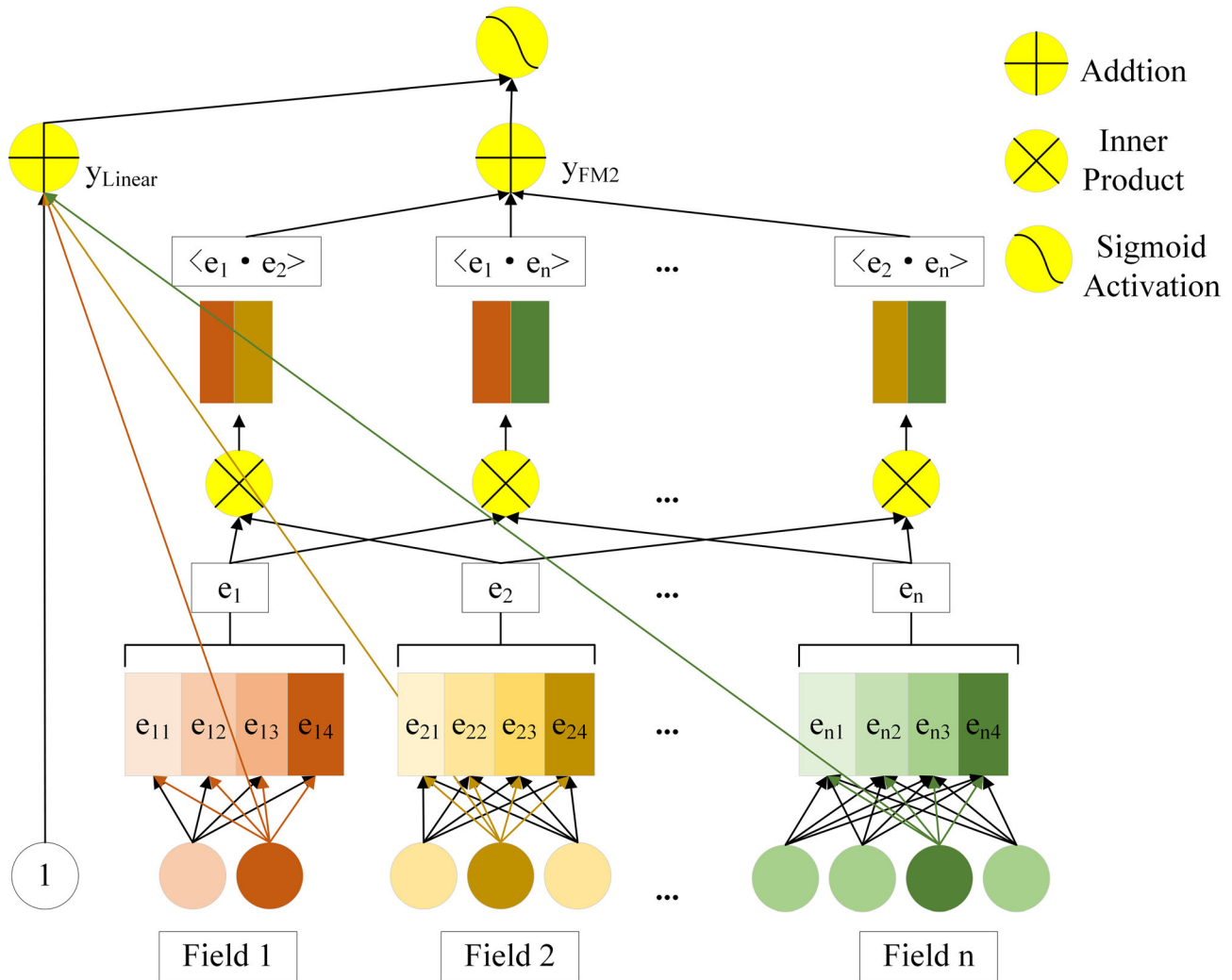
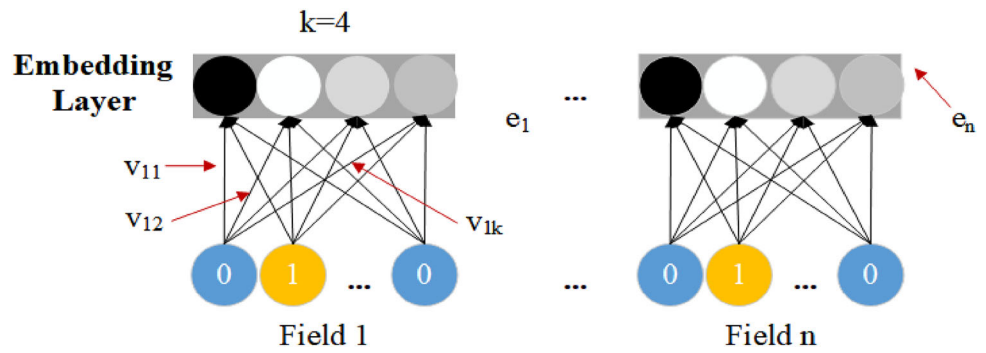


Fig. 4 Overall structure diagram of FM

input to DNN for training. The feature of FNN is that the embedding vector of each feature is trained by FM model in advance. Therefore, when training DNN model, the overhead is reduced and the model can converge faster. However, the performance of the whole network is limited by the performance of FM. In the same year, Qu et al. [17] introduced a product layer between the embedding layer

and the fully connected layer to propose Product-based Neural Network (PNN). PNN finds the relationship between features through inner product or outer product between features, but it lacks low-order feature interaction, so it may ignore the valuable information contained in the original vector. He et al. studied the recommendation problem in the case of sparse input data, and proposed

Neural FM (NFM) [18]. NFM adopts a framework similar to Wide&Deep [19], and it uses Bi-Interaction Layer (Bi-linear interaction) structure to process the second-order cross information, so that the information of the cross features can be better learned by the DNN structure, reducing the difficulty of the DNN learning higher-order cross feature information. In order to learn low-level feature interaction, Guo et al. [20] proposed DeepFM, which combined Deep and FM, used FM for low-level interaction of features, and DNN for high-level feature interaction, combining the two methods in parallel. And both parts share the same input. The final first-order features and second-order and higher-order feature interactions are simultaneously input to the output layer, and the whole process does not require pre-training and feature engineering. He et al. proposed Attention FM (AFM) [21] by extend NFM. They introduce the attention mechanism into the Bi-Linear interactive pooling operation, which further improved the representation ability and interpretability of NFM. AFM only adds an attention mechanism on the basis of FM, and the quadratic term does not enter the deeper network, so AFM does not take advantage of DNN. Zhang et al. [22] combined DeepFM and AFM, and proposed Deep AFM (DeepAFM), which combined the AFM and deep learning in a new NN structure for learning. Compared with existing deep learning models, this method can effectively learn the weighted interaction between features without feature engineering by introducing the feature domain structure. There are also many explorations of attention mechanism. Zhang et al. [23] proposed a new model FAT-DeepFFM, which dynamically captures the importance of each feature before the explicit feature interaction process by introducing CENet domain attention, thus enhancing the DeepFFM. Tao et al. [24] proposed Higher-order AFM (HoAFM), by explicitly considering the interaction of high-order sparse features, they designed a cross interaction layer, updated the representation of features by aggregating the representation of other co-occurrence features, and implemented a bit by bit attention mechanism to determine the different importance of co-occurrence features in dimensional granularity. Yu et al. [25] proposed Gated AFM (GAFM) based on dual factors of accuracy and speed, using the structure of gates to control speed and accuracy. Wen et al. [26] proposed Neural Attention Model (NAM), which deepens the FM by adding fully connected layers. Through the attention mechanism, NAM can learn the different importance of low-order feature interactions. By adding fully connected layers on top of the attention component, NAM can model higher-order feature interactions in a non-linear fashion. In 2019, Yang and colleagues [27] proposed Empirical Mode Decomposition and FM based NN (EMD2FNN). Empirical mode decomposition helps to overcome the non

stationarity of data, and the FM helps to master the non-linear interaction between inputs. Zhang et al. [28] proposed High-order Cross-Factor FM (HCFM). They designed Cross-Weight Network (CWN) to achieve high-order display interactions. The cross and compression layers of CWN are designed to effectively learn important feature combinations, and the weight pooling layer aims to learn the weights of different interaction orders to balance the different weights between high-order and low-order feature interactions. Lu et al. [29] proposed Dual-Input FMs (DIFM), which can efficiently and adaptively learn different representations of a given feature according to different input instances, and can efficiently learn input-aware factors simultaneously at the bit-wise and vector levels (using for re-weighting the original feature representation). The DIFM strategically integrates various components including multi-head self-attention, residual networks, and DNN into a unified end-to-end model. Deng et al. [30] proposed a new Deep Field-weighted FM (DeepFwFM), which itself combines FwFM components and ordinary DNN components, shows unique advantages in structure pruning, using this combination can greatly reduce inference time. Yu et al. [31] proposed Neural Pairwise Ranking FM (NPRFM), which integrates a multilayer perceptual NN into Pairwise Ranking Factorization Machine model. Specifically, to capture higher-order and nonlinear interactions between features, a multi-layer perceptual neural network is superimposed on a double-interaction layer to encode the second-order interactions between features. Pande [32] proposed Field Embedding FM (FEFM) and Deep FEFM (DeepFEFM). FEFM learns the symmetric matrix embedding of each field pair and the single vector embedding of each feature. DeepFEFM combines the FEFM interaction vector learned by FEFM components with DNN to learn high-order feature interaction. Qi and Li [33] proposed Deep Field-Aware Interaction Machine (DeepFIM) to solve the “short expression” problem and better capture multi-density feature interactions. They proposed a new feature interaction expression based on field identifier, namely “hierarchy expression”. On this basis, they designed a cross interaction layer to identify field and field interaction, and used attention mechanism to distinguish the importance of different features. A dynamic bi pool layer is introduced to enhance the acquisition of high-order features.

There is also a combination of FM and CNN. Zhang et al. [34] proposed Deep Generalized Field-aware FM (DGFFM), which uses a wide-deep framework to jointly train Generalized Field-aware FM (GFFM) and DenseNet. It aims to combine the advantages of traditional machine learning methods, including their faster learning speed for low-rank features and the ability to extract high-dimensional features, where GFFM can significantly reduce

computation time by exploiting the corresponding positional relationship between field indices and feature indices. Chanaa and El Faddouli [35] proposed Latent Graph Predictor FM (LGPFM), which utilizes CNN to capture interaction weights for each pair of features. LGPFM combines the advantages of FM and CNN, and CNN can work efficiently in the grid topology, which will significantly improve the accuracy of the results.

Metric learning can also be combined with the FM algorithm. Guo et al. [36] proposed an FM framework based on generalized metric learning technology. The metric method based on Mahalanobis distance uses semi positive definite matrix to project features into a new space, so that the features obey certain linear constraints. The distance function based on DNN is designed to capture the nonlinear feature correlation, which can benefit from the strong representation ability of metric learning method and NN. At the same time, a learnable weight is introduced for the interaction of each attribute pair, which can greatly improve the performance of the distance function.

2.2.3 Disease application

Chen and Qian [37] proposed NN and FM for the diagnosis of children's sepsis. NN can better process the test index result value of patients, and FM can better process the test index state data of patients with sparse structure. Rongge et al. [38] developed a deep FM model for AD diagnosis, which consists of three parts: an embedding layer that handles sparse categorical data, Factorization Machine that efficiently learns pairwise interactions, DNN that implicitly model higher-order interactions. The above are simple combinations of NN and FM, and the FM-Deep Learning algorithms with better performance mentioned in Section 2.2.2 are not used. While Fan et al. [39] applied DeepFM to predict the recurrence of Cushing's disease after transsphenoidal surgery, predicted the recurrence of 354 patients with initial postoperative remission in Peking Union Medical College Hospital, and obtained the highest AUC value (0.869) and the lowest logistic loss value (0.256), which exceeded other models.

3 Unstructured data algorithm

3.1 Convolutional neural network

3.1.1 Theory

CNN is particularly suitable for learning image features. Before CNN was proposed, the fully-connected network was generally used to extract image features, but the entire fully-connected network often had a particularly large number of connections, which would lead to an explosive increase in the number of parameters and training time. It can be noted

that it is not necessary for each neuron to perceive the entire image, the image has a strong 2D local structure, that is, spatially adjacent variables (or pixels) are highly correlated. So people put forward the concept of CNN, which combines three ideas: local receptive field, shared weight and down sampling. The size of convolution kernel is called receptive field. The convolution kernel slides on the image and extracts the features of its coverage area, which can achieve the purpose of forcibly extracting local features, and extract visual features such as edges and corners. Because each region of the image is scanned by a convolution kernel with the same weight, the weight sharing is realized and the number of parameters is greatly reduced. Therefore, the convolution layer of CNN can extract local features well and reduce the number of parameters.

CNN also includes batch-normalization layers, activation layers, and pooling layers. The batch-normalization layer standardizes the small batch data to make it conform to the standard normal distribution, and performs scaling and migration operations, which effectively avoids the disappearance of the gradient, speeds up the decline of the gradient and accelerates the convergence. The activation layer non-linearly processes the input through the activation function, which enables the whole NN to fit any function. The formula is as follows:

$$y = ax + b. \quad (6)$$

Here a is the activation function, x is the input, and both w and b are weight parameters.

Figure 5 is a simple schematic diagram of CNN.

3.1.2 Development history

In 1989, LeCun et al. [40] designed CNN with two convolutional layers (with convolution kernel size of 5×5), trained on the handwritten zip code dataset of the United States Post Office, and the generalization performance of the model reached best at the time. This network is actually the prototype of LeNet, but the whole network only has convolution layer and full connection layer. In 1998, LeCun et al. [41] formally put forward LeNet5, which includes convolution layer, pooling layer and full connection layer. There are seven layers in total. The convolution layer uses 5×5 convolution kernels, and the activation function uses sigmoid. LeNet-5 has a total of 340,908 connections, but the number of trainable parameters is reduced to 60,000 due to weight sharing. After LeNet-5 was proposed, the research of CNN in speech recognition, object detection, face recognition and other application fields has gradually been carried out. After 2012, the CNN entered the stage of large-scale application and in-depth research. The sign was that Krizhevsky et al. [42] proposed AlexNet-8, and its ImageNet Top5 error rate reached

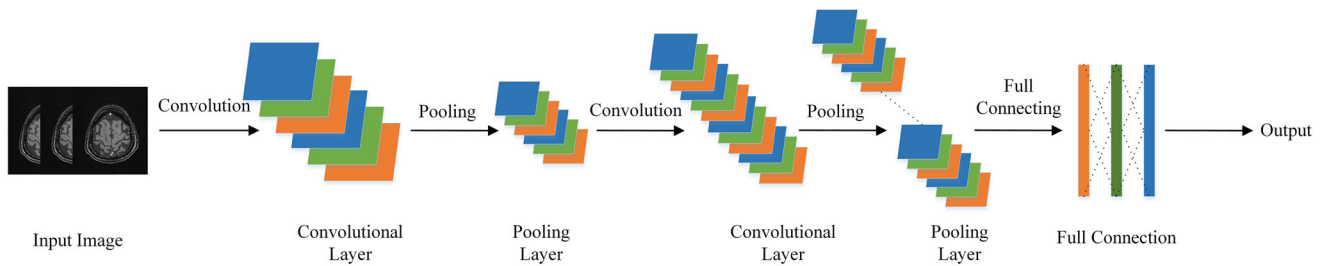


Fig. 5 Convolutional neural network diagram

15.3% in the 2012 ILSVRC competition. AlexNet-8 consists of five convolutional layers, which are filled with all zeros and use ReLU as the activation function. Some convolutional layers are followed by a maximum pooling layer, which can better extract feature textures. AlexNet also uses Dropout to prevent over-fitting. Simonyan and Zisserman [43] proposed VGGNet-16 and VGGNet-19, which used a small convolution kernel (3×3 receptive field), which improved the recognition accuracy while reducing parameters. VGGNet also adds a batch-normalization layer to speed up the training process, and its number of layers exceeds the previous network, reaching 16–19 layers, which can better learn sample features. The entire network structure is regular and suitable for parallel acceleration. In the 2014 ILSVRC competition, VGGNet reduced the ImageNet Top5 error rate to 7.3%. In the same year, InceptionNet, that is, GoogleNet, was proposed [44], with a depth of 22 layers, and using convolution kernels of different sizes in one layer to improve the perception of the model. InceptionNet uses a 1×1 convolution kernel to change 1 output features. The number of channels of the graph (can reduce network parameters). Its ImageNet Top5 error rate was reduced to 6.7%.

Although the depth increase is the development trend of CNN, the gradient will disappear as the number of layers increases to a certain extent. At this time, the accuracy of the depth learning model reaches saturation, and then the training error and test error will decrease significantly, resulting in the inability of the model to converge. So in 2015, the Kaiming He team [45] proposed Residual NN (ResNet), which is connected by residual skip connections between layers, which is mainly to add several identity mapping layers (input equal to output) after some layers. In this way, the forward information can be introduced, which can suppress the disappearance of the gradient, which enables the number of layers of the NN to exceed the previous constraints, reaching hundreds of layers and improving the accuracy. ResNet evaluated on the ImageNet dataset are 152 layers deep-8 times deeper than VGGNet, but still less complex. In addition, the model also uses a

global pooling layer to replace the fully connected layers, which can also achieve the purpose of reducing parameters.

3.1.3 Disease application

Acharya et al. [46] were the first to use CNN for Electroencephalogram (EEG) signal analysis. In this work, the authors implement a 13-layer CNN to detect normal, pre-ictal and epileptic seizure categories without separate feature extraction and feature selection steps. Muhammad et al. [47] proposed CNN-based fusion model for EEG pathology detection. Hossain et al. [48] uses Deep Learning techniques for Epilepsy Seizure Detection. Chanu and Thongam [49] proposed a computer-aided 2D cellular neural network classification technique to classify MR images into two categories: normal and tumor. This method is suitable for inclusion in clinical decision support systems for the initial diagnosis of brain tumors by clinical experts. In 2022, Seven et al. [50] used the deep learning of Endoscopic Ultrasonography (EUS) images to predict whether the malignant potential of gastrointestinal stromal tumors. First let the EUS image be resized in $28 \times 28 \times 1$ format through Lanczos interpolation. The deep learning part uses 20 CNN kernels for the first layer and 50 for the second layer. After each kernel layer, the image resolution is halved. After these convolutional processes, the feature image information is put into the ANN model to train the AI system. The results show that the AI of deep learning based on EUS images can predict the malignant potential of gastric stromal tumors with high accuracy. Yin [51] constructed two 50-layer ResNets based on different building blocks to classify skin lesion images. Although these studies have no major innovations, they have exerted the unique image feature extraction ability of CNN and achieved good results. Rahman et al. uses CNN with relevant adversarial examples (AEs) for COVID-19 diagnosis [52].

Transfer learning refers to transferring the parameters of the trained model (pre training model) to a new model to help train the new model. Because transfer learning can ensure that the model has a higher starting point (before

fine tuning, the initial performance of the model is higher), a higher slope (during the training process, the promotion rate of the model is faster) Higher asymptotic (the model converges better after training), so it often plays a role in the field of disease prediction in combination with CNN. In 2019, Amin et al. [53] proposed a new method to classify tumor/non-tumor Magnetic Resonance Images (MRI), where the segmented images are fed to a pre-trained CNN model where feature learning is performed by AlexNet and GoogleNet. Fully connected layer are used for feature mapping and score vectors are obtained from each trained model. In addition, the score vector is provided to the softmax layer and multiple classifiers. In 2020, Wang et al. [54] proposed two CNN models, which can automatically distinguish benign and malignant masses, lipomas, benign schwannomas and vascular malformations by learning image features. The author chose VGGNet-16 architecture pre-trained on ImageNet dataset to build two CNN models, so as to improve performance by using transfer learning and DNN architecture. Chelghoum et al. [55] used nine pre-trained deep networks, including AlexNet, GoogleNet, VGG-16, VGG-19, ResNet-18, ResNet-50, ResNet-101, ResNet-Inception-V2, and SENET to solve the problem of brain tumor classification by using transfer learning method. The results show that when the number of training samples is small and the number of iterations is small, the performance of the model is still good and the time consumption can be reduced. Similar to the research of Chelghoum et al., Kaur and Gandhi [56] also explored different pre trained classical CNN models to explore the transfer learning ability in pathological brain image classification. The author uses various pre trained DCNN, namely AlexNet, ResNet-50, GoogleNet, VGGNet-16, ResNet-101, VGGNet-19, Inception V3 and Inception ResNet V2. The last layers of these models are replaced to adapt to the training set. Compared with other models, AlexNet shows the best performance in a shorter time. Rehman et al. [57] also aimed at the problem of brain tumors, combined with the traditional machine learning model, adopted three classical CNNs (AlexNet, GoogleNet and VGGNet) to classify brain tumors such as meningioma, glioma and pituitary tumor. The author took these three CNNs as pre-training models and used their different freezing layers respectively. Finally, SVM is used for classification. The results show that the fine tuned VGGNet-16 architecture achieves the highest accuracy in classification and detection, reaching 98.69%. Kumar and Nandhini [58] adopted the entropy image slicing method to select the most informative MRI slices during the training phase. Transfer learning training was performed on the ADNI dataset, and the VGGNet-16 network was used to classify AD of normal individuals. By introducing the MRI slice method, the model can effectively reduce the

preprocessing complexity, and use the VGG-16 network transfer learning technique to solve the unreliability problem. Extracting the parameters of the pre-training model for processing is also one of the methods of transfer learning. Tsai and Tao [59] trained the deep Convolution NN model, and extracted the modified parameters in the network layer to identify the abundant different tissue types in the histological images of colorectal cancer. Eweje et al. [60] utilized a deep learning approach combining conventional MRI images and clinical features to develop a model to classify the malignancy of bone lesions. The method consists of three parts: (1) Imaging data model: By adopting the EfficientNet deep learning architecture, an image classification model is developed. EfficientNet models initialized with weights pre-trained on the ImageNet database can extract features from imaging data. (2) Clinical data model: logistic regression model using clinical variables. Inputs are patient age, gender, and lesion location. For 21 locations (clavicular, skull, proximal femur, distal femur, foot, proximal radius, distal radius, proximal ulna, distal ulna, hand, hip, proximal humerus, distal humerus, proximal tibia end, distal tibia, proximal fibula, distal fibula, mandible, rib/chest wall, scapula, or spine) were one-hot encoded so that the model received 23 different input variables for data quantification. (3) Ensemble model: (1) and (2) are combined using a stacking ensemble approach, where the voting ensemble receives as input the malignancy probability from the imaging and clinical feature models and creates an output based on the sum of the predicted probabilities.

Previously, Rehman et al. combined AlexNet, GoogleNet, and VGGNet with traditional machine learning models, and achieved good results, but if two different deep learning models can be combined, better results can be achieved. In 2021, Kokkalla et al. [61] proposed a deep dense initial residual network model for the three-class classification of brain tumors, which customized the output layer of inception ResNet V2 with fully connected networks and softmax layer. In the same year, Ning et al. [62] proposed an automatic Congestive Heart Failure (CHF) detection model based on a hybrid deep learning algorithm of CNN and Recursive Neural Network. Normal sinus heart rate signals and CHF signals were classified according to ECG and time spectrum. The author carries out feature extraction of ECG signal, mainly extracts RR interval sequence, calculates the time spectrum of ECG signal, and uses CNN to automatically identify the spectrum and related features crossed with time domain. Srinivasu et al. [63] introduced MobileNet V2 with LSTM components to accurately classify skin diseases from images captured from mobile devices. MobileNet V2 is used to classify skin disease types, and LSTM is used to enhance the performance of the model by maintaining state

information of features encountered in previous generation image classification.

The attention mechanism can assign different weights to the input features, so that the model can focus on more important features and information. Therefore, some scholars combine the attention mechanism with CNN for disease prediction. Toğaçar et al. [64] proposed a deep learning model BrainMRNet for brain cancer detection. BrainMRNet is a feedforward end-to-end convolution model, including super column technology, attention module and residual block. Using the super column technology, the features of the input image extracted through the convolution layer of each pixel are combined through the super vector, and the most effective features in the vector are selected and transferred to the next layer; Through the attention module, BrainMRNet attracts attention to the important areas of input data, while unnecessary areas are ignored, which can increase the verification success rate of BrainMRNet; the whole model is composed of residual blocks, which can improve the performance of the model by updating the weight parameters of back propagation. Metric learning is also called similarity learning, which is to classify by comparing the similarity between samples. Some scholars combine CNN with metric learning. Jiao et al. [65] adopted a deep distance metric to learn breast mass classification. The model contains convolutional layers and metric layers. Firstly, the model trains and fine tunes the level of CNN. The CNN structure can provide a good depth feature extraction network and a baseline for breast mass classification. Then, the large edge metric learning method with hinge loss is used to initialize the ensemble learning layer, and the ensemble learning layer is trained to make the characteristics of different breast masses more separable. The metric layers benefits from the representative characteristics of the convolutional layers, and the data flow between them is limited by one-way transmission. The relationship between the two layers is similar to the parasitic relationship in biology/ecology. Therefore, the proposed method is called parasitic metric learning network.

Shallow CNNs can reduce spatial and temporal constraints. Tripathi and Singh [66] proposed a hybrid, flexible deep learning architecture, OLConvNet, which combines the interpretability and depth of traditional object-level features by using a shallower CNN named CNN3L. Extract DL features from the original input image. Then the two sets of features are fused together to generate the final feature set. Multilayer perceptron uses the final fused feature set as input to classify the histopathological nuclei into one of four categories.

Although CNN is mainly used in the image field, some scholars also apply it to structured medical record data and speech data. In 2016, Cheng et al. [67] proposed a deep

learning method for phenotypic analysis from patients' Electronic Medical Records (EHR). Firstly, the EHR of each patient is expressed as a time matrix, with time in one dimension and events in another dimension. Then a four layer Convolution NN model is established for phenotypic extraction and prediction. The first layer consists of these EHR matrices. The second layer is a unilateral convolution layer from which the phenotype can be extracted. The third layer is the largest aggregation layer that introduces sparsity to the detected phenotypes, so as to retain only those significant phenotypes. The fourth layer is the fully connected softmax prediction layer. In order to integrate the temporal smoothness of patients' EHR, the author also studied three different temporal fusion mechanisms in the model: early fusion, late fusion and slow fusion.

In 2019, Gunduz [68] proposed two frameworks based on CNNs to classify Parkinson's Disease (PD) using sound (speech) feature sets. Although the two frameworks are used to combine various feature sets, they are different in combining feature sets. The first framework combines different feature sets and provides them as input to 9-layer CNN, while the second framework transfers the feature sets to the parallel convolution layer. The second framework can learn deep features from each feature set through parallel convolution layer. The extracted deep features can not only successfully distinguish patients with PD from healthy people, but also effectively enhance the discrimination ability of the classifier.

In 2020, Sajja and Kalluri [69] proposed a CNN to predict whether a patient has heart disease. The convolutional architecture adopted by the authors consists of two convolutional layers, two Dropout layers, and an output layer. The model predicts disease with 94.78% accuracy on the UCI-ML Cleveland dataset, outperforming logistic regression, KNN, Naive Bayes, SVMs, and NNs. This is also an application of CNN to structured data.

3.2 Recurrent neural network

3.2.1 Theory and development

RNN [70] is used for pattern recognition of streaming or sequential data such as speech, handwriting and text. There is a circular connection in the hidden layer of RNN. The RNN performs cyclic calculation in the cyclic connection of these hidden units to process the input data in sequence. Each previous input data is stored in a state vector in the hidden unit, and these state vectors are used to compute the output. In summary, RNN calculates a new output considering the current input and the previous input. Although RNN has good performance, in the back-propagation of RNN, when calculating the gradient adjustment weight matrix, due to many partial derivatives multiplied

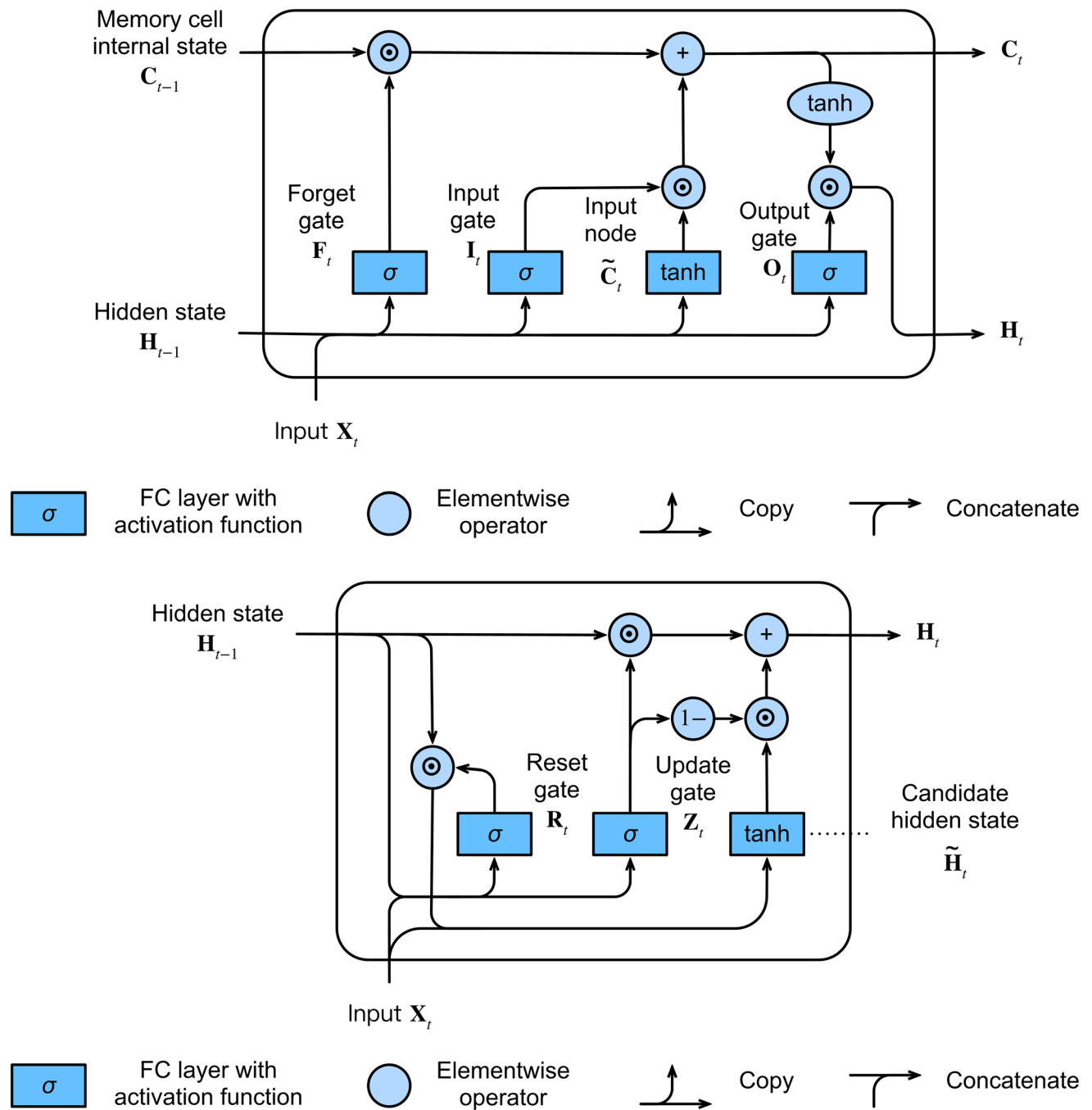


Fig. 6 LSTM and GRU structure diagram. upper: LSTM; lower: GRU

continuously, the gradient in the network will become very small and gradually disappear, or become too large, which makes it difficult for RNN to learn long-distance information. In order to solve this problem, some scholars proposed long short-term memory (LSTM) network [71], which can store sequence data for a long time and solve the problem of gradient disappearance. As shown in the upper part of Fig. 6, LSTM uses a gating mechanism and introduces an input gate, a forget gate and an output gate. When the gate is closed, it will prevent changes to the current

information, so that the previous dependency information will be learned; when the door is open, it does not completely replace the previous information, but makes a weighted average between the previous information and the current information. Therefore, no matter how deep the network is and how long the input sequence is, as long as the door is open, the network will remember these input information. The input gate controls the information of the current word to be integrated into the cell state. The current cell state integrates the information of the current word and

the cell state of the previous moment, and represents the long-term memory. The input gate determines how much information about the current word will be stored in the current cell state. The forget gate controls the information of the cell state at the previous moment to be integrated into the current cell state. When understanding a sentence, the current word may continue to describe the meaning above, or it may start to describe new content from the current word, which has nothing to do with the above, so it is necessary to do the corresponding forgetting operation. The forget gate is responsible for selectively forgetting the information of the cell state. The output gate is responsible for selectively outputting the cell state information. Gated Recurrent Unit, GRU [72] is a simplified version of LSTM. As shown in the lower part of Fig. 6, GRU changes the original three gates into two gates—update gate and reset gate. The reset gate is used to control the influence of the hidden layer state at the previous moment (representing the past information) on the current word. The update gate is a merger of the forget gate and the input gate in LSTM, and is responsible for assigning the importance of past and present information. In this way, the structure of GRU is simpler and matrix operations are less in calculation. Therefore, GRU can save more time than LSTM in the case of large training data.

3.2.2 Disease application

RNNs with LSTM hidden units, pooling, and word embeddings are used in DeepCare [73], an end-to-end deep dynamic network that infers current disease states and predicts future medical outcomes, the authors also conditioned LSTM cell with decay effect to handle irregularly timed events. In 2018, Chu et al. [74] proposed a new context-aware attention mechanism for detecting Adverse Medical Events (AME) of cardiovascular diseases to learn the local context information of words in medical texts. The attention mechanism enables the keywords related to the target AME to get more attention signals, and then drives the model to locate prominent parts of medical texts. The proposed neural attention network is combined with the standard Bi-LSTM model to detect AMEs from a large number of EHR data. The combination of global order-dependent signals of words captured by standard Bi-LSTM and local context signals of words captured by context attention mechanism can significantly improve the performance of AME detection in medical texts.

Some scholars use LSTM for Electrocardiogram (ECG) signal processing. In 2018, Tran et al. [75] proposed a feature extraction-based method to process ECG signals from Internet of Things (IoTs)-specific devices, employing an Auto-Encoder (AE) model to reduce data dimensionality, by combining LSTM extracts top ECG features.

Finally, the full connection layers were used to distinguish normal ECG from abnormal ECG.

Some medical record data with time characteristics (i.e. serialized data) can also be analyzed by LSTM. In 2018, Reddy and Delen [76] used RNN–LSTM method to predict the readmission probability of lupus patients within 30 days by extracting the time relationship from longitudinal EHR clinical data. RNN–LSTM method can make use of the relationship between patients' disease state and time, which makes the model have higher performance. In 2019, Wang et al. [77] used LSTM to predict 6-month, 1-year and 2-year mortality in dementia patients. The deep learning model proposed by the authors consists of two stacked LSTM layers and two attention layers: one between the input layer and the LSTM layer, and the other between the LSTM layer and the output layer. Stacked LSTM layers support hierarchical abstraction of the input data. Attention layers are used to improve model performance as well as keep track of the importance of temporal inputs as the model makes predictions.

There are also several examples of GRU applications. There are also several application cases of GRU. In 2017, Choi et al. [78] used GRU for heart failure diagnosis. Compared with popular methods such as logistic regression, Multi-Layer Perception (MLP), SVM and KNN, GRU performed well in heart failure diagnosis. The results show that the deep learning model suitable for using time relationship improves the performance of the model for detecting sudden heart failure in a short observation window of 12–18 months. Choi et al. [79] used RNN with GRU to develop doctor AI, an end-to-end model that uses patient history to predict subsequent diagnosis and drug treatment.

Some scholars have proposed that RNN is lighter than CNN and it can also be used for image processing. In 2020, Amin et al. [80] proposed an automatic classification method for brain tumors based on LSTM of MRI. First, N4ITK of size 595 and Gaussian filter are used to improve the quality of multi-sequence MRI. The classification is performed using the proposed four-layer deep LSTM model. In each layer, 200, 225, 200 and 225 are selected as the optimal number of hidden units, respectively. The lightweight four layer LSTM model proposed by the author has achieved better results in temporal data processing, which is conducive to the learning of multi sequence MRI.

4 Existing defects and solutions

Here we list several problems in current disease research, which will correspondingly affect the diagnosis rate of disease prediction algorithms. These problems are: Poor Interpretability, Data Imbalance, Data Quality Issues, Too

Little Data. Among them, Poor Interpretability is about deep learning algorithms. Poor interpretability leads to low reliability of deep learning disease prediction algorithms, which is not good for helping doctors analyze pathological causes. The remaining three problems are related to the data. Data Imbalance will cause the classifier to lose its classification ability. Data Quality Issues, poor quality datasets will lower the performance limit of deep learning algorithms on specific problems. Too Little Data, a small amount of data will lead to over-fitting and seriously reduce the quality of deep learning algorithms. In addition to enumerating these problems, this section also presents the current corresponding solutions.

4.1 Poor interpretability

Traditional statistical methods are usually based on manual feature engineering of medical related domain knowledge. These methods are closely combined with medical knowledge. Although the effect is not very outstanding, they give doctors reliable interpretability. Deep learning algorithms are like a black box and are driven by data which we cannot see the feature extraction and screening process. Therefore, although deep learning improves the feature extraction ability and classification ability of the model, its interpretability is very poor, which is easy to lead to the unreliability of the results and bring risks. Only by solving the interpretability problem of the model, deep learning can be more widely used in the actual disease prediction, better serve doctors and patients, and make them have confidence in the diagnosis results of the model.

The general solution is to add attention mechanism, which is suitable for both structured and unstructured data. Attention mechanism was first applied to the field of natural language processing, which can better find the relationship between words in sentences and better predict the next words. AFM and DeepAFM are the application of attention mechanism in FM algorithm; Woo et al. [81] proposed the Convolutional Block Attention Module (CBAM) in 2018. Woo et al. [81] proposed convolutional attention module (CBAM) in 2018. Given an intermediate feature map, CBAM module will infer the attention map along two independent dimensions (channel and space), and then multiply the attention map with the input feature map. CBAM is a lightweight general module, which can be seamlessly integrated into any CNN architecture for end-to-end training with basic CNN without excessive additional over-head.

The Local Interpretable Model Agnostic Explanations (LIME) can also be adopted to solve the problem of poor interpretability. LIME establishes a linear separable model locally in the model through local disturbance sampling and linear approximation, and estimates the importance of

each feature through the feature weight of the linear model [82, 83].

For images, interpretability methods based on activation mapping can be adopted, such as Class Activation Mapping (CAM) [84], Grad-CAM [85], Grad-CAM++ [86], and Score-CAM [87], etc. This method generates saliency map by linear weighted combination of activation mapping to highlight important areas in image space. The saliency map is used to highlight the features in the input considered to be related to the prediction of the learning model, which does not need training data or modify the model.

4.2 Data imbalance

There is always an imbalance in medical data because there are fewer people who are sick than those who are not. When the data is severely unbalanced, the model always classifies the samples into the majority class, for example if a model is trained to predict whether a patient has a tumor, when the number of negative samples (patients without a tumor) in the training set is much higher than When the number of positive samples is positive, when predicting whether a new patient has a tumor, the model always diagnoses the patient as not having a tumor, which is obviously not what we want.

For image data, Generative Adversarial Networks (GAN) [88] can be used. GAN can generate minority class samples that are close to real samples and solve the problem of data imbalance. For binary classification problems, the method of Synthetic Minority Oversampling Technique, SMOTE [89] can also be used. SMOTE can up-sample or down-sample the training set, so that the proportion of positive and negative samples reaches a balanced state.

Structured data can also use the SMOTE method, but up-sampling will destroy the discreteness of the data, making discrete features into continuous features, resulting in inconsistent data types in the training and test sets, which is not conducive to the learning of FM algorithms. If the number of minority class samples is too small, using down-sampling will lead to a serious shortage of training samples. These are questions to be studied in the future.

4.3 Data quality issues

Data quality remains the biggest challenge in model training. The excellent performance of deep learning models in disease prediction relies on high-quality medical data. While medical data is readily available under existing conditions, the quality of the data remains low. Moreover, there may be problems such as the mismatch between the training samples and the real samples and the existence of some abnormal features, which will affect the model effect.

There is also a lot of medical data that requires experienced medical experts to give sample labels.

For image, speech and other types of data, the quality can be improved by using GAN, up-sampling, Fourier transform and other methods. For structured data, methods such as filling in missing values, deleting duplicate values, and outliers are often used for data cleaning, and methods such as discretization, filter, wrapper, and Principal Component Analysis (PCA) are used for feature selection to obtain higher-quality samples. Since we are talking about deep learning algorithms, it is possible to build end-to-end deep learning algorithms like DeepFM, without feature engineering, and let deep learning exert automatic feature learning capabilities to overcome data quality issues. The automatic learning ability of deep learning can also be applied to sample label processing, which involves unsupervised learning and is beyond the scope of this article.

4.4 Too little data

Although a large amount of health data has been generated at present, many medical data sets involve privacy issues, which are stored in independent institutions and are not made public. Therefore, a large number of data sets can't be used for practical research, so the model can't be fully trained, and it's hard to exert its real effect. Here we only discuss how to solve the problem from the aspect of algorithms.

For images, the method of Few-shot Learning [90–92] can be used, that is, the model is trained through a large number of tasks to improve the generalization ability of the model. When faced with similar new tasks, the model can be trained after a small number of iterations. achieve better results. Few-shot Learning includes the following methods in total: (1) model fine-tuning [93, 94], obtaining a pre-trained model on a source dataset with a large number of samples, and then fine-tuning the pre-trained model on a target dataset with a small number of samples. This method is more suitable for scenarios where the source dataset and target dataset are similar, but in practical scenarios, the two datasets are usually dissimilar, which often leads to overfitting. (2) Data augmentation refers to the use of some additional datasets or information to expand the target data set or enhance the characteristics of the samples in the target data set [95, 96]. In the early stage, the data set was expanded through spatial transformation, but this could not expand the types of samples. Later, people used methods such as GANs for data augmentation. Meta learning refers to letting the model learn meta-knowledge from a large number of tasks, and use this meta-knowledge to quickly adapt to different new tasks. Meta learning includes Memory NN [97, 98], Meta Network [99], Model-Agnostic Meta-Learning (MAML) [100] and other algorithms.

Metric learning, also known as similarity learning, calculates the distance between two samples through a distance function, so as to measure the similarity between them and determine whether they belong to the same category. The metric learning algorithm consists of an embedding module and a measurement module. The embedding module converts the samples into vectors in a low-dimensional vector space, and the measurement module gives the similarity between samples. Metric learning is divided into fixed distance based metric learning [101] and learnable network based metric learning [102].

However, few-shot learning is mainly applied to images, and it is often ineffective in structured data. Because the idea of Few-shot Learning is similar to that of a child distinguishing animals, after seeing a lot of animal pictures, give him a picture of a rhino, and he can find a rhino among many animals. Images have certain similarity and have a general large data set, so they can meet the requirements of a large number of similar tasks. However, different diseases have different features, and these features have different characteristics. Therefore, there is no general large data set, which is difficult to meet the requirements of a large number of similar tasks. At present, there are traditional machine learning algorithms (low complexity), Boosting sampling algorithms, and feature selection to solve the problem of small amount of structured data. Among them, traditional machine learning algorithms and feature selection make up for the overfitting problem caused by the small amount of data by reducing the complexity (model complexity or feature complexity). There is no more effective way to solve this problem.

5 Future works and prospects

5.1 Incorporating Digital Twins

Digital Twins refers to building the same entity in the digital world through digital means to realize the understanding, analysis and optimization of the physical entity. With the development of technologies such as AI, Big Data, Virtual reality, IoT, and cloud computing [103, 104]. Digital Twins have begun to shine in industrial, medical and other fields. The application of Digital Twins in medical care is usually to create a model based on real medical data in the virtual world, and then observe and analyze the stimulus changes of the model to various conditions, such as the feedback generated by the intervention of new drugs or new treatment regimens. These real medical data come from EHRs, daily behavior databases, medical wearable devices, and more. Therefore, through Digital Twins, medical activities such as health detection, telemedicine, early disease diagnosis, and

disease treatment can be realized [105, 106], providing revolutionary solutions in the field of healthcare [107]. Health monitoring is an important means in modern medicine. The use of various wearable sensors in the Digital Twins can realize ubiquitous monitoring of the health status of patients [108], and can also reduce medical costs, reduce the number of hospitalizations, and improve the quality of life of patients [109, 110].

Digital Twins can be combined with deep learning algorithm of disease prediction to realize faster and more developed electronic medical treatment and automated medical treatment. The general realization methods are as follows: firstly, collect data, and use various sensors, especially various convenient wearable sensors to collect various health information [111, 112], and transmit these data to the cloud. It can also collect various electronic medical record data and daily behavior database data. Then, using these collected medical data, a digital model of disease prediction is established in the cloud by deep learning algorithm. Finally, the digital model is used to process and analyze the health data, so as to predict the patient's physical condition, whether he is ill or not, the probability of illness, etc. In the process of analysis, new knowledge and new information will be generated [113], which will help to adjust and upgrade the model, and help related researchers to better understand the mechanism behind the disease, so as to find a better treatment.

Many scholars have proposed a combination of Digital Twins and deep learning. For example, Chakshu et al. [114] proposed a method to achieve cardiovascular Digital Twins using reverse analysis, which uses a virtual patient database. By inputting pressure waveforms from three non-invasively accessible blood vessels (carotid, femoral, and brachial), the blood pressure waveforms in various blood vessels of the body are calculated backwards with the help of LSTM cells. The reverse analysis system established by this method is mainly used for the detection of abdominal aortic aneurysm and its severity. Quilodrán-Casas et al. [115] created two Digital Twins systems of SEIRS models and applied them to simulate the spatial and temporal propagation of COVID-19, and compared their prediction results with real data. They compared the performance of the two digital twin models [also known as Non-invasive Reduced Order Model (NIROM)]. The first method is to use PCA for dimensionality reduction and Bi-LSTM with data correction (through optimal interpolation) for prediction. The second NIROM uses PCA for dimensionality reduction again and GAN for prediction. In addition, there are many related studies.

In the future, we should realize a more intelligent processing mode through Digital Twins and deep learning model, realize a truly automatic and intelligent medical system, and greatly reduce the workload of doctors. At the

same time, more Digital Twins medical system platforms need to be developed to achieve a wider range of intelligent medical treatment. Intelligent medical treatment is one of the important links of smart city, and intelligent medical treatment is indispensable to the realization of smart city. Therefore, on the basis of ensuring the security of Digital Twins medical platform, we should further broaden the scope of application and serve the user group more comprehensively. Intelligence is one of the core elements of future medical and urban development. To truly realize comprehensive medical intelligence, we must better integrate medical Digital Twins and deep learning algorithm technology.

5.2 Promoting precision medicine

Precision medicine is the principle and practice of integrating modern medical technology and traditional medical methods, scientifically understanding human body functions and the nature of diseases, systematically optimizing the principles and practices of human disease prevention and control, and maximizing individual and social health benefits with efficient, safe and economical health care services. In clinical practice, precision medicine pursues accurate and reasonable diagnosis and treatment methods for each patient in order to minimize iatrogenic damage, minimize medical costs and maximize patient benefits. Compared with traditional medicine, it can provide patients with more effective, cheaper and more timely medical services. Since it was proposed in 2015, it has been the key to global healthcare and one of the important goals of many sustainable development plans around the world [116, 117]. The concept of precision medicine opens up new ideas for human health and healthcare [118, 119].

Like personalized medicine, precision medicine focuses on individual differences [120], exploring the impact of individual factors on disease [121]. Assessment of personal health from genomics, living environment, etc., coupled with clinical data analysis, will have higher performance. For example, Panayides et al. [122] proposed that starting from the methods of radiomics and radiogenomics, combined with precision medicine, some abnormal diseases can be found more quickly when dealing with disease problems. Precision medicine also has good performance in preventing malignant diseases, such as cancer [123, 124], tumor [125] and so on. It can be said that disease prediction and disease treatment are moving towards the era of precision medicine [126].

At present, there are many researches on precision medicine in Western countries, but the research on precision medicine in the Asia-Pacific region is still in the initial stage. On the one hand, it is necessary to ensure the diversity and high quality of gene collection. On the other

hand, it is necessary to extract the genetic characteristics consistent with the population of the Asia–Pacific region. These two are both urgent problems to be solved at present, and they are also the reasons that hinder development.

In the next era, precision medicine will be combined with multi-field applications. Realize the systematic operation of medical diagnosis and promote the development of medical care in a more intelligent direction. For example, Lu and Harrison [127] pointed out that CNN can realize large-scale medical image analysis and labeling, and can accurately obtain pathological information of different patients. Laplante and Akhloufi [128] proposed a deep NN classifier to identify the anatomical location of tumors. Using the 27 TCGA miRNA stem cell ring cohort, tumors at 20 anatomical sites were classified with 96.9% accuracy. Therefore, deep learning can be combined with precision medicine [129] to better process big data and fundamentally promote the development of precision medicine [130]. As part of precision medicine, accurate prediction of disease embodies enormous advantages and value, and can advance the development of modern medical technology. However, the current precision medicine is still in the stage of exploration and development [131–133], the research situation of different diseases is very different, and the application of deep learning technology is still in the development stage. In the future, AI-related researchers should focus more on precision medicine and build deep learning models that better meet the requirements of precision medicine in combination with the research on radiomics and genomics in the medical field. While promoting the progress of precision medicine, it also drives the multi-faceted development of deep learning, which is more in line with social needs.

6 Conclusion

This paper reviews the deep learning algorithms in the field of disease prediction. According to the type of data processed, the algorithms are divided into structured data algorithms and unstructured data algorithms. Structured data algorithms include ANN and FM-Deep Learning algorithms. Unstructured data algorithms include CNN, RNN, etc. This paper expounds the principle, development history and application of these algorithms in disease prediction. In the application part of disease prediction of each algorithm, try to analyze the literature according to the characteristics of the algorithm. Although these algorithms are the mainstream algorithms at present and in the future, there will be some problems in the current research, such as poor interpretability, sample imbalance, data quality, few samples in some cases, etc. This paper gives some temporary solutions, hoping to have better solutions in the

future. At the end of the article, we elaborate and analyze the two development trends of disease prediction in the future. The future medical technology should be combined with Digital Twins to realize real intelligent medical treatment, pay more attention to personalized medical treatment, integrate with precision medical treatment, and serve individuals more conveniently. This paper can enlighten relevant researchers, help them understand the current development, existing problems and future development trend of disease prediction algorithms, and let them focus on hot spot algorithms, combine current advanced technologies and concepts, and make more efficient, effective and reasonable research with the goal of medical development trend.

Author contributions The contributions of all authors are the same.

Funding Not applicable.

Data availability Not applicable.

Code availability Not applicable.

Consent for publication We agree to publish.

Declarations

Conflict of interest Not applicable.

Ethical approval Our research does not address ethical issues.

Informed consent We agree to participate.

References

1. Maurya, M.R., Riyaz, N.U., Reddy, M., Yalcin, H.C., Ouakad, H.M., Bahadur, I., Al-Maadeed, S., Sadasivuni, K.K.: A review of smart sensors coupled with Internet of Things and artificial intelligence approach for heart failure monitoring. *Med. Biol. Eng. Comput.* **59**(11), 2185–2203 (2021)
2. Shamshirband, S., Fathi, M., Dehzaangi, A., Chronopoulos, A.T., Alinejad-Rokny, H.: A review on deep learning approaches in healthcare systems: taxonomies, challenges, and open issues. *J. Biomed. Inform.* **113**, 103627 (2021)
3. Hossain, M.S., Muhammad, G.: Deep learning based pathology detection for smart connected healthcare. *IEEE Netw.* **34**(6), 120–125 (2020)
4. Kumar, P.M., Gandhi, U.D.: A novel three-tier Internet of Things architecture with machine learning algorithm for early detection of heart diseases. *Comput. Electr. Eng.* **65**, 222–235 (2018)
5. Bakator, M., Radosav, D.: Deep learning and medical diagnosis: a review of literature. *Multimodal Technol. Interact.* **2**(3), 47 (2018)
6. Rendle, S.: Factorization machines. In: 2010 IEEE International Conference on Data Mining, pp. 995–1000. IEEE (2010)

7. Lin, X., Zhang, W., Zhang, M., Zhu, W., Pei, J., Zhao, P., Huang, J.: Online compact convexified factorization machine. In: Proceedings of the 2018 World Wide Web Conference, 2018, pp. 1633–1642 (2018)
8. Al-Galal, S.A.Y., Alshaiikli, I.F.T., Abdulrazzaq, M.: MRI brain tumor medical images analysis using deep learning techniques: a systematic review. *Health Technol.* **11**, 1–16 (2021)
9. Leevy, J.L., Khoshgoftaar, T.M., Villanustre, F.: Survey on RNN and CRF models for de-identification of medical free text. *J. Big Data* **7**(1), 1–22 (2020)
10. Hossain, M.S., Muhammad, G., Guizani, N.: Explainable AI and mass surveillance system-based healthcare framework to combat COVID-19 like pandemics. *IEEE Netw.* **34**(4), 126–132 (2020)
11. Shorfuzzaman, M., et al.: MetaCOVID: a Siamese neural network framework with contrastive loss for n-shot diagnosis of COVID-19 patients. *Pattern Recognit.* **113**, 107700 (2021)
12. Khanam, J.J., Foo, S.Y.: A comparison of machine learning algorithms for diabetes prediction. *ICT Express* **7**(4), 432–439 (2021)
13. Soundarya, S., Sruthi, M., Bama, S.S., Kiruthika, S., Dhiyaneswaran, J.: Early detection of Alzheimer disease using gadolinium material. *Mater. Today Proc.* **45**, 1094–1101 (2021)
14. Pasha, S.N., Ramesh, D., Mohmmad, S., Harshavardhan, A., et al.: Cardiovascular disease prediction using deep learning techniques. *IOP Conf. Ser. Mater. Sci. Eng.* **981**, 022006 (2020)
15. Chen, C., Dongxing, W., Chunyan, H., Xiaojie, Y.: Exploiting social media for stock market prediction with factorization machine. In: 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), vol. 2, pp. 142–149. IEEE (2014)
16. Zhang, W., Du, T., Wang, J.: Deep learning over multi-field categorical data. In: European Conference on Information Retrieval, pp. 45–57. Springer (2016)
17. Qu, Y., Cai, H., Ren, K., Zhang, W., Yu, Y., Wen, Y., Wang, J.: Product-based neural networks for user response prediction. In: 2016 IEEE 16th International Conference on Data Mining (ICDM), 2016, pp. 1149–1154. IEEE (2016)
18. He, X., Chua, T.-S.: Neural factorization machines for sparse predictive analytics. In: Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 355–364 (2017)
19. Cheng, H.-T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., Anderson, G., Corrado, G., Chai, W., Ispir, M., et al.: Wide and deep learning for recommender systems. In: Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, 2016, pp. 7–10 (2016)
20. Guo, H., Tang, R., Ye, Y., Li, Z., He, X.: DeepFM: a factorization-machine based neural network for CTR prediction. arXiv preprint (2017). [arXiv:1703.04247](https://arxiv.org/abs/1703.04247)
21. Xiao, J., Ye, H., He, X., Zhang, H., Wu, F., Chua, T.-S.: Attentional factorization machines: learning the weight of feature interactions via attention networks. arXiv preprint (2017). [arXiv:1708.04617](https://arxiv.org/abs/1708.04617)
22. Zhang, J., Wu, Z., Li, F., Li, W., Ren, T., Li, W., Chen, J.: Deep attentional factorization machines learning approach for driving safety risk prediction. *J. Phys. Conf. Ser.* **1732**, 012007 (2021)
23. Zhang, J., Huang, T., Zhang, Z.: FAT-DeepFFM: field attentive deep field-aware factorization machine. arXiv preprint (2019). [arXiv:1905.06336](https://arxiv.org/abs/1905.06336)
24. Tao, Z., Wang, X., He, X., Huang, X., Chua, T.-S.: HoAFM: a high-order attentive factorization machine for CTR prediction. *Inf. Process. Manag.* **57**(6), 102076 (2020)
25. Yu, H., Yin, J., Li, Y.: Gate attentional factorization machines: an efficient neural network considering both accuracy and speed. *Appl. Sci.* **11**(20), 9546 (2021)
26. Wen, P., Yuan, W., Qin, Q., Sang, S., Zhang, Z.: Neural attention model for recommendation based on factorization machines. *Appl. Intell.* **51**(4), 1829–1844 (2021)
27. Zhou, F., Zhou, H.-M., Yang, Z., Yang, L.: EMD2FNN: a strategy combining empirical mode decomposition and factorization machine based neural network for stock market trend prediction. *Expert Syst. Appl.* **115**, 136–151 (2019)
28. Zhang, W., Zhang, X., Wang, H.: High-order factorization machine based on cross weights network for recommendation. *IEEE Access* **7**, 145746–145756 (2019)
29. Lu, W., Yu, Y., Chang, Y., Wang, Z., Li, C., Yuan, B.: A dual input-aware factorization machine for CTR prediction. In: Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence, 2021, pp. 3139–3145 (2021)
30. Deng, W., Pan, J., Zhou, T., Flores, A., Lin, G.: A sparse deep factorization machine for efficient CTR prediction. arXiv preprint (2020). [arXiv:2002.06987](https://arxiv.org/abs/2002.06987)
31. Yu, Y., Jiao, L., Zhou, N., Zhang, L., Yin, H.: Enhanced factorization machine via neural pairwise ranking and attention networks. *Pattern Recognit. Lett.* **140**, 348–357 (2020)
32. Pande, H.: Field-embedded factorization machines for click-through rate prediction. arXiv preprint (2020). [arXiv:2009.09931](https://arxiv.org/abs/2009.09931)
33. Qi, G., Li, P.: Deep field-aware interaction machine for click-through rate prediction. *Mob. Inf. Syst.* (2021). <https://doi.org/10.1155/2021/5575249>
34. Zhang, Q.-L., Rao, L., Yang, Y.: DGFFM: generalized field-aware factorization machine based on DenseNet. In: 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1–8. IEEE (2019)
35. Chanaa, A., El Faddouli, N.-E.: Latent graph predictor factorization machine (LGPFM) for modeling feature interactions weight. In: Proceedings of the 13th International Conference on Intelligent Systems: Theories and Applications, 2020, pp. 1–5 (2020)
36. Guo, Y., Cheng, Z., Jing, J., Lin, Y., Nie, L., Wang, M.: Enhancing factorization machines with generalized metric learning. *IEEE Trans. Knowl. Data Eng.* **34**(8), 3740–3753 (2020)
37. Chen, X., Qian, J.: An assistant diagnosis system for sepsis in children based on neural network and factorization. *Sci. Technol. Eng.* (2017)
38. Ronge, R., Nho, K., Wachinger, C., Pölsterl, S.: Alzheimer's disease diagnosis via deep factorization machine models. In: International Workshop on Machine Learning in Medical Imaging, 2021, pp. 624–633. Springer (2021)
39. Fan, Y., Li, D., Liu, Y., Feng, M., Chen, Q., Wang, R.: Toward better prediction of recurrence for Cushing's disease: a factorization-machine based neural approach. *Int. J. Mach. Learn. Cybern.* **12**(3), 625–633 (2021)
40. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1**(4), 541–551 (1989)
41. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
42. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25**, 1097–1105 (2012)
43. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint (2014). [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
44. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper

- with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9 (2015)
45. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778 (2016)
 46. Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adeli, H.: Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Comput. Biol. Med.* **100**, 270–278 (2018)
 47. Muhammad, G., et al.: EEG-based pathology detection for home health monitoring. *IEEE J. Sel. Areas Commun.* **39**(2), 603–610 (2021)
 48. Hossain, M.S., Amin, S.U., Alsulaiman, M., Muhammad, G.: Applying deep learning for epilepsy seizure detection and brain mapping visualization. *ACM Trans. Multimed. Comput. Commun. Appl.* **15**(1s), 1–17 (2019)
 49. Chanu, M.M., Thongam, K.: Computer-aided detection of brain tumor from magnetic resonance images using deep learning network. *J. Ambient Intell. Humaniz. Comput.* **12**(7), 6911–6922 (2021)
 50. Seven, G., Silahtaroglu, G., Kochan, K., Ince, A.T., Arici, D.S., Senturk, H.: Use of artificial intelligence in the prediction of malignant potential of gastric gastrointestinal stromal tumors. *Dig. Dis. Sci.* **67**(1), 273–281 (2022)
 51. Yin, X.: Pigmented skin lesions image classification based on residual network. In: 2021 6th International Conference on Machine Learning Technologies, 2021, pp. 74–81 (2021)
 52. Rahman, A., et al.: Adversarial examples—security threats to COVID-19 deep learning systems in medical IoT devices. *IEEE Internet Things J.* **8**(12), 9603–9610 (2021)
 53. Amin, J., Sharif, M., Yasmin, M., Saba, T., Anjum, M.A., Fernandes, S.L.: A new approach for brain tumor segmentation and classification based on score level fusion using transfer learning. *J. Med. Syst.* **43**(11), 1–16 (2019)
 54. Wang, B., Perronne, L., Burke, C., Adler, R.S.: Artificial intelligence for classification of soft-tissue masses at us. *Radiol. Artif. Intell.* **3**(1), 200125 (2020)
 55. Chelghoum, R., Ikhlef, A., Hameurlaine, A., Jacquir, S.: Transfer learning using convolutional neural network architectures for brain tumor classification from MRI images. In: IFIP International Conference on Artificial Intelligence Applications and Innovations, 2020, pp. 189–200. Springer (2020)
 56. Kaur, T., Gandhi, T.K.: Deep convolutional neural networks with transfer learning for automated brain image classification. *Mach. Vis. Appl.* **31**(3), 1–16 (2020)
 57. Rehman, A., Naz, S., Razzak, M.I., Akram, F., Imran, M.: A deep learning-based framework for automatic brain tumors classification using transfer learning. *Circuits Syst. Signal Process.* **39**(2), 757–775 (2020)
 58. Kumar, S.S., Nandhini, M.: Entropy slicing extraction and transfer learning classification for early diagnosis of Alzheimer diseases with sMRI. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* **17**(2), 1–22 (2021)
 59. Tsai, M.-J., Tao, Y.-H.: Deep learning techniques for colorectal cancer tissue classification. In: 2020 14th International Conference on Signal Processing and Communication Systems (ICSPCS), 2020, pp. 1–8. IEEE (2020)
 60. Eweje, F.R., Bao, B., Wu, J., Dalal, D., Liao, W.-H., He, Y., Luo, Y., Lu, S., Zhang, P., Peng, X., et al.: Deep learning for classification of bone lesions on routine MRI. *EBioMedicine* **68**, 103402 (2021)
 61. Kokkalla, S., Kakarla, J., Venkateswarlu, I.B., Singh, M.: Three-class brain tumor classification using deep dense inception residual network. *Soft Comput.* **25**(13), 8721–8729 (2021)
 62. Ning, W., Li, S., Wei, D., Guo, L.Z., Chen, H.: Automatic detection of congestive heart failure based on a hybrid deep learning algorithm in the Internet of Medical Things. *IEEE Internet Things J.* **8**(16), 12550–12558 (2020)
 63. Srinivasu, P.N., SivaSai, J.G., Ijaz, M.F., Bhoi, A.K., Kim, W., Kang, J.J.: Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. *Sensors* **21**(8), 2852 (2021)
 64. Toğaçar, M., Ergen, B., Cömert, Z.: Tumor type detection in brain MR images of the deep model developed using hyper-column technique, attention modules, and residual blocks. *Med. Biol. Eng. Comput.* **59**(1), 57–70 (2021)
 65. Jiao, Z., Gao, X., Wang, Y., Li, J.: A parasitic metric learning net for breast mass classification based on mammography. *Pattern Recognit.* **75**, 292–301 (2018)
 66. Tripathi, S., Singh, S.K.: Cell nuclei classification in histopathological images using hybrid OLConvNet. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* **16**(1s), 1–22 (2020)
 67. Cheng, Y., Wang, F., Zhang, P., Hu, J.: Risk prediction with electronic health records: a deep learning approach. In: Proceedings of the 2016 SIAM International Conference on Data Mining, 2016, pp. 432–440. SIAM (2016)
 68. Gunduz, H.: Deep learning-based Parkinson’s disease classification using vocal feature sets. *IEEE Access* **7**, 115540–115551 (2019)
 69. Sajja, T.K., Kalluri, H.K.: A deep learning method for prediction of cardiovascular disease using convolutional neural network. *Rev. d’Intell. Artif.* **34**(5), 601–606 (2020)
 70. Zaremba, W., Sutskever, I., Vinyals, O.: Recurrent neural network regularization. *arXiv preprint* (2014). [arXiv:1409.2329](https://arxiv.org/abs/1409.2329)
 71. Ma, Q., Lin, Z., Yan, J., Chen, Z., Yu, L.: Mode-LSTM: a parameter-efficient recurrent network with multi-scale for sentence classification. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020, pp. 6705–6715 (2020)
 72. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using RNN encoder–decoder for statistical machine translation. *arXiv preprint* (2014). [arXiv:1406.1078](https://arxiv.org/abs/1406.1078)
 73. Pham, T., Tran, T., Phung, D., Venkatesh, S.: DeepCare: a deep dynamic memory model for predictive medicine. In: Pacific–Asia Conference on Knowledge Discovery and Data Mining, 2016, pp. 30–41. Springer (2016)
 74. Chu, J., Dong, W., He, K., Duan, H., Huang, Z.: Using neural attention networks to detect adverse medical events from electronic health records. *J. Biomed. Inform.* **87**, 118–130 (2018)
 75. Tran, D.T., Vo, H.T., Nguyen, D.D., Nguyen, Q.M., Huynh, L.T., Le, L.T., Do, H.T., Quan, T.T.: A predictive model for ECG signals collected from specialized IoT devices using deep learning. In: 2018 5th NAFOSTED Conference on Information and Computer Science (NICS), 2018, pp. 424–429. IEEE (2018)
 76. Reddy, B.K., Delen, D.: Predicting hospital readmission for lupus patients: an RNN–LSTM-based deep-learning methodology. *Comput. Biol. Med.* **101**, 199–209 (2018)
 77. Wang, L., Sha, L., Lakin, J.R., Bynum, J., Bates, D.W., Hong, P., Zhou, L.: Development and validation of a deep learning algorithm for mortality prediction in selecting patients with dementia for earlier palliative care interventions. *JAMA Netw. Open* **2**(7), 196972–196972 (2019)
 78. Choi, E., Schuetz, A., Stewart, W.F., Sun, J.: Using recurrent neural network models for early detection of heart failure onset. *J. Am. Med. Inform. Assoc.* **24**(2), 361–370 (2017)
 79. Choi, E., Bahadori, M.T., Schuetz, A., Stewart, W.F., Sun, J.: Doctor AI: predicting clinical events via recurrent neural

- networks. In: Machine Learning for Healthcare Conference, 2016, pp. 301–318. PMLR (2016)
80. Amin, J., Sharif, M., Raza, M., Saba, T., Sial, R., Shad, S.A.: Brain tumor detection: a long short-term memory (LSTM)-based learning model. *Neural Comput. Appl.* **32**(20), 15965–15973 (2020)
 81. Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: CBAM: convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3–19 (2018)
 82. Holzinger, A., Langs, G., Denk, H., Zatloukal, K., Müller, H.: Causability and explainability of artificial intelligence in medicine. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **9**(4), 1312 (2019)
 83. Ribeiro, M.T., Singh, S., Guestrin, C.: “why should I trust you?” Explaining the predictions of any classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144 (2016)
 84. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: Learning deep features for discriminative localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2921–2929 (2016)
 85. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-CAM: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 618–626 (2017)
 86. Chattopadhyay, A., Sarkar, A., Howlader, P., Balasubramanian, V.N.: Grad-CAM++: generalized gradient-based visual explanations for deep convolutional networks. In: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 839–847. IEEE (2018)
 87. Wang, H., Wang, Z., Du, M., Yang, F., Zhang, Z., Ding, S., Mardziel, P., Hu, X.: Score-CAM: score-weighted visual explanations for convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 24–25 (2020)
 88. Mehrotra, A., Dukkupati, A.: Generative adversarial residual pairwise networks for one shot learning. *arXiv preprint (2017)*. [arXiv:1703.08033](https://arxiv.org/abs/1703.08033)
 89. Han, H., Wang, W.-Y., Mao, B.-H.: Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In: International Conference on Intelligent Computing, 2005, pp. 878–887. Springer (2005)
 90. Xinye, L., Shenpeng, L., Jing, Z.: Survey of few-shot learning based on deep neural network. *Appl. Res. Comput.* **37**(08), 2241–2247 (2020)
 91. Wang, Y., Yao, Q., Kwok, J.T., Ni, L.M.: Generalizing from a few examples: a survey on few-shot learning. *ACM Comput. Surv. (CSUR)* **53**(3), 1–34 (2020)
 92. Lee, K., Maji, S., Ravichandran, A., Soatto, S.: Meta-learning with differentiable convex optimization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 10657–10665 (2019)
 93. Gidaris, S., Komodakis, N.: Dynamic few-shot visual learning without forgetting. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 4367–4375 (2018)
 94. Nakamura, A., Harada, T.: Revisiting fine-tuning for few-shot learning. *arXiv preprint (2019)*. [arXiv:1910.00216](https://arxiv.org/abs/1910.00216)
 95. Dixit, M., Kwitt, R., Niethammer, M., Vasconcelos, N.: AGA: attribute-guided augmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 7455–7463 (2017)
 96. Shen, W., Shi, Z., Sun, J.: Learning from adversarial features for few-shot classification. *arXiv preprint (2019)*. [arXiv:1903.10225](https://arxiv.org/abs/1903.10225)
 97. Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D., Lillicrap, T.: One-shot learning with memory-augmented neural networks. *arXiv preprint (2016)*. [arXiv:1605.06065](https://arxiv.org/abs/1605.06065)
 98. Santoro, A., Bartunov, S., Botvinick, M., Wierstra, D., Lillicrap, T.: Meta-learning with memory-augmented neural networks. In: International Conference on Machine Learning, 2016, pp. 1842–1850. PMLR (2016)
 99. Munkhdalai, T., Yu, H.: Meta networks. In: International Conference on Machine Learning, 2017, pp. 2554–2563. PMLR (2017)
 100. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: International Conference on Machine Learning, 2017, pp. 1126–1135. PMLR (2017)
 101. Wu, X., Sahoo, D., Hoi, S.: Meta-RCNN: meta learning for few-shot object detection. In: Proceedings of the 28th ACM International Conference on Multimedia, 2020, pp. 1679–1687 (2020)
 102. Xiao, J., Xu, H., Zhao, W., Cheng, C., Gao, H.: A prior-mask-guided few-shot learning for skin lesion segmentation. *Computing (2021)*. <https://doi.org/10.1007/s00607-021-00907-z>
 103. El Saddik, A., Laamarti, F., Alja’Afreh, M.: The potential of digital twins. *IEEE Instrum. Meas. Mag.* **24**(3), 36–41 (2021)
 104. Hossain, M.S., Muhammad, G.: Emotion-aware connected healthcare big data towards 5G. *IEEE Internet Things J.* **5**(4), 2399–2406 (2018)
 105. Rathore, M.M., Shah, S.A., Shukla, D., Bentafat, E., Bakiras, S.: The role of AI, machine learning, and big data in digital twinning: a systematic literature review, challenges, and opportunities. *IEEE Access* **9**, 32030–32052 (2021)
 106. El Saddik, A.: Digital twins: the convergence of multimedia technologies. *IEEE MultiMed.* **25**(2), 87–92 (2018)
 107. Erol, T., Mendi, A., Dogan, D.: The digital twin revolution in healthcare, pp. 1–7 (2020). <https://doi.org/10.1109/ISM-SIT50672.2020.9255249>
 108. Pantelopoulos, A., Bourbakis, N.G.: A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Trans. Syst. Man Cybern. C* **40**(1), 1–12 (2009)
 109. Nguyen, H.H., Mirza, F., Naeem, M.A., Nguyen, M.: A review on IoT healthcare monitoring applications and a vision for transforming sensor data into real-time clinical feedback. In: 2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD), 2017, pp. 257–262. IEEE (2017)
 110. Hossain, M.S.: Cloud-supported cyber-physical localization framework for patients monitoring. *IEEE Syst. J.* **11**(1), 118–127 (2017)
 111. Vesnic-Alujevic, L., Breitegger, M., Pereira, A.G.: ‘Do-it-yourself-healthcare?’ Quality of health and healthcare through wearable sensors. *Sci. Eng. Ethics* **24**(3), 887–904 (2018)
 112. Hossain, M.S., Muhammad, G., Alamri, A.: Smart healthcare monitoring: a voice pathology detection paradigm for smart cities. *Multimed. Syst.* **25**(5), 565–575 (2019)
 113. Wickramasinghe, N., Jayaraman, P.P., Zelcer, J., Forkan, A.R.M., Ulapane, N., Kaul, R., Vaughan, S.: A vision for leveraging the concept of digital twins to support the provision of personalised cancer care. *IEEE Internet Comput.* (2021). <https://doi.org/10.1109/MIC.2021.3065381>
 114. Chakshu, N.K., Sazonov, I., Nithiarasu, P.: Towards enabling a cardiovascular digital twin for human systemic circulation using inverse analysis. *Biomech. Model. Mechanobiol.* **20**(2), 449–465 (2021)
 115. Quilodrán-Casas, C., Silva, V.L., Arcucci, R., Heaney, C.E., Guo, Y., Pain, C.C.: Digital twins based on bidirectional LSTM and GAN for modelling the COVID-19 pandemic. *Neurocomputing* **470**, 11–28 (2022)

116. Afzal, M., Islam, S.R., Hussain, M., Lee, S.: Precision medicine informatics: principles, prospects, and challenges. *IEEE Access* **8**, 13593–13612 (2020)
117. Shorfuzzaman, M., et al.: Towards the sustainable development of smart cities through mass video surveillance: a response to the COVID-19 pandemic. *Sustain. Cities Soc.* **64**, 102582 (2021)
118. Llovet, J.M., Montal, R., Sia, D., Finn, R.S.: Molecular therapies and precision medicine for hepatocellular carcinoma. *Nat. Rev. Clin. Oncol.* **15**(10), 599–616 (2018)
119. Le Tourneau, C., Borcoman, E., Kamal, M.: Molecular profiling in precision medicine oncology. *Nat. Med.* **25**(5), 711–712 (2019)
120. Fujiwara, N., Friedman, S.L., Goossens, N., Hoshida, Y.: Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine. *J. Hepatol.* **68**(3), 526–549 (2018)
121. Zhang, S., Bamakan, S.M.H., Qu, Q., Li, S.: Learning for personalized medicine: a comprehensive review from a deep learning perspective. *IEEE Rev. Biomed. Eng.* **12**, 194–208 (2018)
122. Panayides, A.S., Pattichis, M.S., Leandrou, S., Pitris, C., Constantinidou, A., Pattichis, C.S.: Radiogenomics for precision medicine with a big data analytics perspective. *IEEE J. Biomed. Health Inform.* **23**(5), 2063–2079 (2018)
123. Loomans-Kropp, H.A., Umar, A.: Cancer prevention and screening: the next step in the era of precision medicine. *NPJ Precis. Oncol.* **3**(1), 1–8 (2019)
124. Regel, I., Mayerle, J., Ujjwal Mukund, M.: Current strategies and future perspectives for precision medicine in pancreatic cancer. *Cancers* **12**(4), 1024 (2020)
125. Steuer, C.E., Ramalingam, S.S.: Tumor mutation burden: leading immunotherapy to the era of precision medicine? *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **36**(7), 631–632 (2018)
126. Hamamoto, R., Suvarna, K., Yamada, M., Kobayashi, K., Shinkai, N., Miyake, M., Takahashi, M., Jinnai, S., Shimoyama, R., Sakai, A., et al.: Application of artificial intelligence technology in oncology: towards the establishment of precision medicine. *Cancers* **12**(12), 3532 (2020)
127. Lu, L., Harrison, A.P.: Deep medical image computing in preventive and precision medicine. *IEEE MultiMed.* **25**(3), 109–113 (2018)
128. Laplante, J.-F., Akhlofi, M.A.: Predicting cancer types from miRNA stem-loops using deep learning. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2020, pp. 5312–5315. *IEEE* (2020)
129. Ahmed, Z., Mohamed, K., Zeeshan, S., Dong, X.: Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database* (2020). <https://doi.org/10.1093/database/baaa010>
130. Hulsen, T., Jamuar, S.S., Moody, A.R., Karnes, J.H., Varga, O., Hedensted, S., Spreafico, R., Hafler, D.A., McKinney, E.F.: From big data to precision medicine. *Front. Med.* **6**, 34 (2019)
131. Hey, S.P., Gerlach, C.V., Dunlap, G., Prasad, V., Kesselheim, A.S.: The evidence landscape in precision medicine. *Sci. Transl. Med.* (2020). <https://doi.org/10.1126/scitranslmed.aaw7745>
132. Dienstmann, R., Vermeulen, L., Guinney, J., Kopetz, S., Tejpar, S., Tabernero, J.: Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat. Rev. Cancer* **17**(2), 79–92 (2017)
133. Dayem Ullah, A.Z., Oscanoa, J., Wang, J., Nagano, A., Lemoine, N.R., Chelala, C.: SNPnexus: assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Res.* **46**(W1), 109–113 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



Zengchen Yu was born in Shandong, China in 1997. In 2020, he participated in School Uniform Management and Food Mall Project Development in Shandong Ciyun Technology Company. He is currently studying at the College of Computer Science and Technology of Qingdao University, pursuing a Master's Degree in Software Engineering. His main research directions are deep learning and image processing. Possess good mathematical foundation and deep learning development experience, and has excellent practical experience in software development.

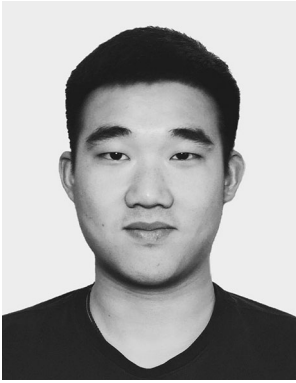


counselling for many years.

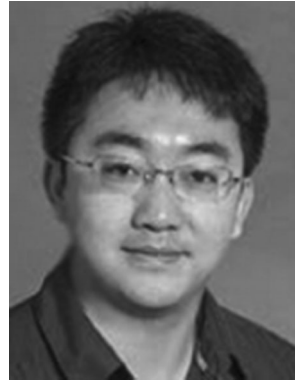
Ke Wang is a Deputy Chief Physician of the Department of Psychology, a National Second-Class Psychological Consultant, a Psychotherapist and a Marriage and Family Therapist. Vice President of Qingdao Psychological Counsellor Association, Executive Director of Qingdao Mental Health Association, Expert in Youth Posts of Qingdao Health and Family Planning Commission, and has been engaged in psychiatric clinical and psychological



Zhibo Wan was born in Jimo City, Shandong Province in 1973. He is a doctor of engineering. Currently, he is a lecturer of the College of Computer Science & Technology of Qingdao University and the director of the Ship and Offshore Engineering Equipment Intelligent Manufacturing Technology Research Institute.



Shuxuan Xie was born in Shandong, China in 1997. He is now studying as a Postgraduate in Software Engineering at the College of Computer Science and Technology of Qingdao University. The research direction is mainly Deep Learning and Transfer Learning. He has a deep interest in the direction of medical disease diagnosis and is keen on model research.



Zhihan Lv is currently an Associate Professor with the Department of Game Design, Faculty of Arts, Uppsala University, Sweden. His research interests include multimedia and virtual reality. He is the Editor-Chief of Internet of Things and Cyber-Physical Systems and an Associate Editor of 18 journals, including IEEE Transactions on Intelligent Transportation System, IEEE Transactions on Network and Service Management.