Review of Neural Network Algorithms

Meiying Liu
Yibo Jiang

Follow this and additional works at: https://aisel.aisnet.org/iceb2021

This material is brought to you by the International Conference on Electronic Business (ICEB) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICEB 2021 Proceedings (Nanjing, China) by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
The artificial neural network is the core tool of machine learning to realize intelligence. It has shown its advantages in the fields of sound, image, sound, picture, and so on. Since entering the 21st century, the progress of science and technology and people's pursuit of artificial intelligence have introduced the research of artificial neural networks into an upsurge. Firstly, this paper introduces the application background and development process of the artificial neural network in order to clarify the research context of neural networks. Five branches and related applications of single-layer perceptron, linear neural network, BP neural network, Hopfield neural network, and depth neural network are analyzed in detail. The analysis shows that the development trend of the artificial neural network is developing towards a more general, flexible, and intelligent direction. Finally, the future development of the artificial neural network in training mode, learning mode, function expansion, and technology combination has prospected.

Keywords: Neural network, machine learning, review.

INTRODUCTION
Modern computers work on the basis of Von Neumann’s principle through program access. It can be used to solve programmable problems with clear features, reasoning, and operation rules efficiently. Its precision and speed in numerical and logical operations greatly expand the capacity of the human brain. Since the 21st century, computer networks have been rapidly developed and popularized. While enjoying the convenience brought by computer networks, people have put forward new requirements for the development of computer networks, and the voice of replacing the brain with a machine is getting higher and higher. However, because the traditional von Neumann’s machine is limited to logical operation rules. The solution of unstructured problems is far from reaching the level of human intelligence. With the rapid development of modern information science and technology, this problem is becoming increasingly acute, which urges scientific and technical experts to find a new way to solve the problem. When people's thinking turns to the study of the human brain structure model and information processing mechanism, it promotes the in-depth development of brain science and the research of artificial neural networks and brain models. With an in-depth understanding of the biological brain, the artificial neural network has made great progress.

The artificial neural network is a mathematical model or computational model composed of a large number of artificial neurons, which imitates the corresponding structure and related functions of the biological brain. It is based on the basic principle of neural networks in biology. After understanding and abstracting the structure of the human brain and the response mechanism of external stimuli, it simulates the neural system of the human brain to process complex information based on the knowledge of network topology. The model is characterized by high fault tolerance, intelligence, adaptability, and self-learning. As an important algorithm and model to realize machine learning tasks, it has been successfully applied in the technical fields of handwriting recognition, speech recognition, image recognition, and natural language processing.

The research of artificial neural networks provides a way for machine learning subjective and informal knowledge and brings unprecedented expectations to the research of artificial intelligence. The research chain of "neural network — machine-learning — artificial intelligence" has become the core of a new round of sci-tech revolution and industrial reform and is having an extremely profound impact on the world economy, social progress, and people's life. In terms of the world economy, artificial intelligence is a strategic technology leading the future. The development of artificial intelligence has become a major strategy for major countries in the world to enhance national competitiveness and promote national economic growth. In terms of social progress, intelligent technology provides new technologies and ideas for social governance. The application of artificial intelligence in social governance is the most direct and effective way to reduce governance costs, improve governance efficiency and reduce governance interference. In terms of daily life, deep learning, image recognition, speech recognition, and other technologies have been widely used in intelligent terminals, smart homes, mobile payments, and other fields. In the future, artificial intelligence technology will also play a more significant role in education, medical treatment, travel, and other fields closely related to people's lives and provide life services with wider coverage, a better experience for ordinary people.

DEVELOPMENT PROCESS OF NEURAL NETWORK
The starting point of neural network development is the MP model. In 1943, psychologist W.S. McCulloch and mathematical logician W. Pitts established a neural network and mathematical model, which is called the MP model. They put forward the formal mathematical description and network structure method of neurons through the MP model and proved that a single neuron could perform logical functions, thus creating an era of artificial neural network research.
In the 1960s, the concept of perceptron and adaptive linear elements further promoted the development of the artificial neural network. The perceptron is a structure composed of two layers of neurons. The function of the input layer is to receive external input signals. The output layer is also called the function layer, which is MP neurons. Because it has only one layer of functional neurons, it can not deal with nonlinear separable problems, and its learning ability is limited. After carefully analyzing the functions and limitations of the single-layer neural network system represented by perceptron, M. Minsky et al. published perceptron in 1969, pointing out that perceptron can not solve the high-order predicate problem. The development of neural networks has fallen into a trough.

In 1982, J. J. Hopfield, a physicist at the California Institute of Technology, proposed the Hopfield neural network model, breaking the downturn in the field of neural network research. Hopfield neural network introduces the concept of "computational energy," expresses the problem to be solved with the state of network neurons, and makes the judgment of network stability. In 1984, he put forward the Continuous Hopfield neural network model (CHNN), which did pioneering work for the research of neural computers, created a new way for the neural network to be used in associative memory and optimization calculation and effectively promoted the research of neural network.

In 1985, Geoffrey Hinton and other scholars proposed the Boltzmann Machine (BM) and the improved Restricted Boltzmann Machine (RBM) in 1986. RBM is actually a two-layer neural network. On this basis, Depth Boltzmann Machine (DBM, stacked restricted Boltzmann machine) and Depth Belief Network (DBN, stacked restricted Boltzmann machine plus BP algorithm) are developed.

In 1986, Geoffrey Hinton, David Rumelhart, and Ronald Williams published an important paper, "Learning representations by back-propagating errors," in the journal Nature, and proposed a BP algorithm (error backpropagation algorithm). BP neural network algorithm greatly reduces the time required to train the neural network and makes the artificial neural network have the ability of self-study and generalization, but it also has obvious defects, that is, when the number of network layers is large, it is easy to fall into optimal local solution and overfitting. With the continuous development of technology, the advantages of the BP algorithm are prominent, which makes it not only a classic artificial neural network algorithm but also one of the most commonly used neural network models.

In 1986, in the book Parallel Distributed Processing: Exploration in the Microstructure of Cognition edited by Rumelhart and McClelland, they established the parallel distributed processing theory, mainly devoted to the micro research of cognition, and made a detailed analysis of BP algorithm, which solved the problem that there was no effective algorithm for weight adjustment for a long time. It can solve the problems that cannot be solved by the perceptron, answers the problem about the limitations of neural networks in perceptrons, and proves that the artificial neural network has strong computing ability in practice.

In 1989, Yann Lecun proposed a Convolutional Neural Network LeNet updated by reverse conduction. In 1998, Yann Lecun and his collaborators constructed a complete Convolutional Neural Network LeNet-5 and achieved success in handwritten numeral recognition. The success of LeNet-5 has attracted attention to the application of a convolutional neural network. Microsoft developed Optical Character Recognition (OCR) system using a convolutional neural network in 2003. Other application research based on Convolutional Neural networks has also been carried out.

In 2006, Geoffrey Hinton's research team used the Restricted Boltzmann Machine to model the continuous layer of a neural network, used the layer by layer pre-training method to extract the high-dimensional features in the model data, and later proposed the Deep Belief Network. This technique was then extended to many different neural network architectures by many scholars, which greatly improved the generalization effect of the model on the test set. At the same time, with the improvement of hardware computing power and algorithm, the hidden layer of the network can continue to increase, so the researchers represented by Hinton redefined the artificial neural network as deep learning and popularized it. Since 2011, the neural network has gained an overwhelming advantage in the benchmark test of speech recognition and image recognition. Moreover, because the structure of the Convolutional Neural Network is very suitable for image recognition, combined with those research results, it has also attracted people's attention again.

At present, there are many kinds of neural networks with different uses. The knowledge map of the current status of artificial neural network studies by CiteSpace software is shown in Figure 1, and the study focused on the following knowledge domains: BP neural network, genetic algorithm, support vector machine, forecast, machine learning, etc. Through the keyword analysis of relevant literature on the artificial neural network by CiteSpace software, the ten keywords with the highest citation frequency are obtained, as shown in Figure 2: backpropagation, knowledge base, genetic algorithm, support vector machine, classification, forecasting, machine learning, etc. The research concentration shown in the Figure 2 timeline shows that the research of artificial neural networks began to increase after 2007, entered prosperity after 2010, and continues to this day. The research trend benefited from 2006 when Hinton et al. proposed an effective strategy for training deep neural networks that not only improves the model accuracy but also greatly promotes the development of unsupervised learning. After 2010, the deep neural network has been widely used, especially in 2011, Microsoft's speech recognition system broke through the traditional speech recognition technology, and the advent of AlexNet in 2012 has brought a great breakthrough to the whole image recognition field, but also made the deep neural network sweep the whole field of artificial intelligence foreshadowing.
RESEARCH STATUS OF NEURAL NETWORK

With the deepening of the research on neural networks, the types of neural networks are also enriched. By increasing the network level, changing the network structure, and improving the activation function, the branches of the neural network are gradually growing. Due to the different functions and limitations of each branch, its application fields are also different. This section studies the branches of neural networks and summarizes the structural characteristics and applications of different branches.

Single-Layer Perceptron

The starting point of the neural network is Single-Layer Perceptron. As the simplest neural network, it includes two layers: the input layer and the output layer. The input layer is directly connected to the output layer. The model of Single-Layer Perceptron is shown in Figure 3, which can be expressed as

\[ y = \sigma(W^T X) \]  

(1)

where \( X \) represents vector \( [X_1, X_2, \ldots, X_n, 1] \), \( W \) represents vector \( [W_1, W_2, \ldots, W_n, b] \), and \( \sigma \) represents activation function.

Figure 1: Knowledge map of the current status of artificial neural network research.

Figure 2: Keyword analysis diagram of the literature related to artificial neural networks.

<table>
<thead>
<tr>
<th>Keywords</th>
<th>Year</th>
<th>Strength</th>
<th>Begin</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>eeg</td>
<td>2000</td>
<td>1.7725</td>
<td>2005</td>
<td>2008</td>
</tr>
<tr>
<td>back propagation</td>
<td>2000</td>
<td>1.3844</td>
<td>2008</td>
<td>2010</td>
</tr>
<tr>
<td>knowledge base</td>
<td>2000</td>
<td>1.3617</td>
<td>2010</td>
<td>2013</td>
</tr>
<tr>
<td>backpropagation</td>
<td>2000</td>
<td>1.3714</td>
<td>2013</td>
<td>2015</td>
</tr>
<tr>
<td>genetic algorithm</td>
<td>2000</td>
<td>1.8716</td>
<td>2015</td>
<td>2017</td>
</tr>
<tr>
<td>artificial neural network (ann)</td>
<td>2000</td>
<td>3.0454</td>
<td>2015</td>
<td>2017</td>
</tr>
<tr>
<td>support vector machine</td>
<td>2000</td>
<td>1.8556</td>
<td>2016</td>
<td>2018</td>
</tr>
<tr>
<td>classification</td>
<td>2000</td>
<td>2.2992</td>
<td>2016</td>
<td>2018</td>
</tr>
<tr>
<td>forecasting</td>
<td>2000</td>
<td>2.1931</td>
<td>2017</td>
<td>2018</td>
</tr>
<tr>
<td>machine learning</td>
<td>2000</td>
<td>2.0446</td>
<td>2018</td>
<td>2021</td>
</tr>
</tbody>
</table>

The 21st International Conference on Electronic Business, Nanjing, China, December 3-7, 2021
The Single-Layer Perceptron has the advantages of simple structure, a clear model, and a small amount of calculation. Since the activation function is a symbolic function, the output result $y$ can only be two (0/1 or -1/1). Therefore, Single-Layer Perceptron is mainly used to solve the linearly separable binary classification problem.

Due to the limitation of the structure of the perceptron, it has limitations in practical application. For example, the original Single-Layer Perceptron can only deal with linear separable problems. For linear non-separable problems, the algorithm will oscillate in the calculation process. That is, the algorithm can not stop. For the limitations of Single-Layer Perceptron, many scholars have proposed improved algorithms to improve the performance of Single-Layer Perceptron. One of the more famous is the Pocket Algorithm proposed by Gallant, that is, a pocket weight vector is introduced in the iterative process of PLA to store the perceptron weight vector with the correct runs, and its goal is to find a solution (optimal solution) with the least misclassification samples. Pocket PLA is a greedy approximation algorithm, which can deal with linear inseparable problems and better solve the noisy data and contradictory data in the samples.

The classical Single-Layer Perceptron has gradually withdrawn from the research field. At present, the research on Single-Layer Perceptron is mostly the research on the improved algorithm of Single-Layer Perceptron. However, as the simplest learning machine familiar to pattern recognition and neural network researchers, the classical perceptron algorithm is the basis of many complex algorithms developed later.

**Linear Neural Network**

Linear Neural Networks are very similar to perceptron. The main difference is that the activation function can output only two possible values (-1 or 1), and the output of linear neural networks can take any value. Its activation function is a linear function. The adaptive linear neuron is the earliest typical representative of a linear network. It uses Widrow Hoff learning rule, that is, LMS (least mean square) algorithm, to adjust the weight and bias of the network. The criterion of the LMS algorithm is to minimize the expected value of the square $E(e^2(n))$ of the difference between the desired signal and the actual output of the filter, and modify the weight coefficient vector $W(n)$ according to this criterion.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Classification</th>
<th>Perceptron</th>
<th>Linear Neural Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network transfer function</td>
<td>The perceptron transfer function is a binary threshold element, which determines that the perceptron can only do simple classification.</td>
<td>The transmission function of the linear neural network is linear, and it can also achieve fitting or approximation on a perceptron basis.</td>
<td></td>
</tr>
<tr>
<td>Learning algorithm</td>
<td>The perceptron learning algorithm stops just where it can correctly classify, making the classification boundaries too close to some modes and making the system sensitive to error.</td>
<td>The classification boundaries obtained by the LMS algorithm are often directly in the middle of the two classes of models.</td>
<td></td>
</tr>
</tbody>
</table>
of physical layer security output.

**BP Neural Network**

Due to the simplicity of structure, the application of perceptron and linear neural networks is also limited to linear separable problems. With the in-depth study of neural networks, researchers improve the structure of the neural network and put forward the concept of a multilayer neural network. BP neural network is a multi-layer perceptron, which increases the number of network layers on the basis of the perceptron. The BP neural network model is shown in Figure 4. The network is composed of multiple layers, all layers are connected, and neurons in the same layer are not connected. In function, because it contains multiple hidden layers, BP neural network can mine more information from the input, has stronger network classification and recognition ability, and has the function of dealing with linear inseparable problems. In addition, BP adopts an error backpropagation algorithm for learning. Its main characteristic is that the signal propagates forward, and the error propagates back. Specifically, for the neural network model with only one hidden layer: the process of the BP neural network is mainly divided into two stages. The first stage is the forward propagation of the signal, passing through the hidden layer from the input layer to the output layer. The second stage is the backpropagation of error. From the output layer to the hidden layer, and finally, to the input layer, the weight and bias from the hidden layer to the output layer and from the input layer to the hidden layer are adjusted in turn. With continuous learning, the final error becomes smaller and smaller.

![Figure 4: The BP neural network model.](image)

In the BP neural network model, setting the input data from the input layer as $X$, the parameters from the input layer to the hidden layer as $w$, $b_1$, the parameters from the hidden layer to the output layer as $v$, $b_2$, and the activation function as $f_1$, $f_2$. So the model is set as:

$$\hat{y} = f_2(v^Tf_1(w^Tx + b_1) + b_2)$$

(2)

where the loss function is

$$E(\theta) = \frac{1}{n}\sum_{k=1}^{n}(y_k - \hat{y}_k)^2$$

(3)

In the process of weight adjustment, there are

$$y^{(k)} = y^{(k-1)} + \eta \nabla_k v, \quad w^{(k)} = w^{(k-1)} + \eta \nabla_k w$$

(4)

BP neural network has the characteristics of multi-layer in structure and feedback in the algorithm, which makes it have strong nonlinear mapping ability and flexible network structure. The number of middle layers and the number of neurons in each layer of the network can be set arbitrarily according to the specific situation, and its performance varies with the difference of structure. However, in the process of specific application, BP neural network has no corresponding theoretical guidance in the selection of network layers and the number of neurons. Due to the large number of layers of the network, the error surface may contain multiple different local minima, and the gradient descent may lead to falling into the local minima. It is easy to overfit when the training times are too many and the spatial dimension is too high. For the problems of traditional BP neural networks, many scholars have studied the improved algorithm. Silaban et al. Applied BFGS Quasi-Newton to BP neural network and proposed that BP neural network with BFGS improved the convergence of the learning process, with an average improved convergence rate of 98.34%, and the accuracy will be improved when BFGS is used together with BP. Zhu et al. combined with Rumelhart's adding inertia pulse to dynamically adjust the learning rate can adjust the learning rate to a larger value so as to improve the learning speed of the model, affect the connection threshold and weight parameters of nodes through the learning rate, eliminate invalid iterations in the learning process of BP network, and combine the two to propose an improved algorithm to optimize BP network. It improves the local minimum value and the slow convergence speed of the traditional BP network. Ye et al. used a genetic algorithm to have excellent heuristic searchability in the multi-parameter optimization process and used a genetic algorithm to optimize the initialization of BPNN weight and threshold to make it approach the global minimum as much as possible.

At present, the network idea and algorithm model of the BP neural network are increasingly mature. The proposal and application of the improved algorithm improve the defects of the classical BP network. The variability of the network level enhances the flexibility of the network, which makes the application range of the BP neural network wider. Specific application fields include prediction and evaluation, pattern recognition and classification, data compression, etc.
The application of the BP neural network in prediction and evaluation is based on its characteristic of error backpropagation. The network weight and threshold can be adjusted according to the prediction error so that the prediction output of the BP neural network is close to the expected output. In terms of biological research, many scholars extract the surface characteristics of fruit, define the fruit defect and fruit decay, use BP neural network to fit the data, establish a defect-rot prediction model, and improve the average accuracy of fruit decay prediction. In terms of physics research, some scholars put forward a short-term wind electric power prediction algorithm (Improved Beetle Antennae Search, IBAS) optimization model of neural network parameters, introduced the Metropolis criterion into day cattle whiskers algorithm to reduce the probability of optimal local solution in the optimization process, so as to obtain the optimal initial weight and threshold of BP neural network, and improve the accuracy of short-term wind electric power prediction. In terms of engineering construction, the application of BP neural network prediction is conducive to solving the prediction problem under building regulation.

BP neural networks have strong pattern recognition capabilities through self-organized or learned-trained networks. Yin et al. used the characteristic parameters of underwater acoustic communication signals to build a modulation recognition system of underwater acoustic communication signals based on the BP neural network to realize the modulation recognition of underwater acoustic communication signals. Zhao et al. applied the trained BP network model to the actual cable trench safety monitoring system and used BP neural network to recognize the four common disturbance signals of optical fiber sensors. You et al. studied the hand motion pattern recognition of the BP neural network, introduced the cross-entropy cost function to improve the neural network so as to improve its learning rate, generalization ability, and the accuracy of motion classification.

The nonlinearity, fault tolerance, self-organization, and adaptability of BP neural network technology are widely used in data compression technology and greatly simplify the complexity of data compression. Zheng et al. used an artificial neural network to compress image data, and the network training time decreased significantly. BP neural network is applied to image compression of space fluid experiment, which achieves high compression ratio and good reconstructed image quality, and the trained network has high robustness. Chen et al. designed an improved BP neural network PID controller to adjust the parameters of the revolving door algorithm online to realize the adjustable precision of lossy compression. Through the simulation experiment, the best automatic precision lossy compression algorithm controller and its corresponding algorithm are determined.

Hopfield Neural Network

Hopfield neural network is different from the three neural networks mentioned above. As a representative of the feedback neural network, Hopfield neural network has undergone fundamental changes in structure. The structure diagram of the Hopfield neural network is shown in Figure 5. All neurons have the same status without a hierarchical difference. They can be connected with each other or feedback signals to themselves, which have the characteristics of circulation and recursion.

![Figure 5: The Hopfield neural network model.](image)

In the Hopfield network model, a Hopfield neural network composed of multiple units is set. The input of the $i^{th}$ unit at time $t$ is recorded as $x_i(t)$, the output is recorded as $y_i(t)$, the connection weight is $w_{ij}$, and the threshold is $b_i(t)$, then the output of the $i^{th}$ unit at time $t+1$ can be expressed as $y_i(t + 1)$:

$$
y_i(t) = \begin{cases}
  1 & x_i(t) > 0 \\
  y_i(t) & x_i(t) = 0 \\
  0 & x_i(t) < 0
\end{cases} \tag{5}
$$

$$
x_i(t) = \sum_{j=1}^{n} w_{ij} y_j(t) - b_i(t) \tag{6}
$$

Hopfield's learning rules are based on indoctrination learning. That is, the weight of the network is not trained but calculated according to certain rules. Once the weight is determined, it will not change, and the state of each neuron in the network is constantly updated during operation. When the network evolves to a stable state, the solution to the problem is the state of each neuron. The characteristics of the Hopfield network are related to the introduced energy function $E$.

$$
E = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} y_i y_j + \sum_{i=1}^{n} b_i x_i \tag{7}
$$

The 21st International Conference on Electronic Business, Nanjing, China, December 3-7, 2021
By observing the equation, it can be seen that the energy function $E$ is always non-increasing. That is, it decreases gradually with the increase of time until the network reaches a stable state. When the input mode is consistent with the output mode, the result of the energy function $E$ is 0.

Hopfield network can be divided into discrete and continuous network models according to different activation functions. The discrete model uses the hard limit function as the activation function, while the continuous model uses the sigmoid function as the activation function. The discrete network can converge to a stable equilibrium state and take it as the memory information of the sample. It has the ability to recall the complete memory information from a piece of incomplete information, so it is suitable for associative memory. However, the capacity of this associative memory is limited, and it is prone to a pseudo-stable point. When the memory samples are similar, the network can not recall the correct memory. By taking the energy function as the objective function, the continuous network corresponds to the optimization problem and is suitable for dealing with related problems.

For the application research of Discrete Hopfield Neural Network, karbasi et al. discussed that the internal noise of associative memory could improve the ability of associative memory when the memory mode does not reduce to a fixed threshold with the noise. Li Ya et al. combined memristor with traditional Hopfield neural network designed a memristor Hopfield neural network circuit. Danesh et al. studied emotion detection using a recursive auto-associative memory network by collecting speech, tags, and other data from social networking sites. For the research on the optimization of Discrete Hopfield Neural Network, Mofrad et al. embedded the coding technology into neural associative memory to improve the network recovery ability. Tanaka et al. proposed a Hopfield associative memory network with sparse modularization, introduced an iterative learning algorithm that depends on the network topology to change the connection weight and concluded that the network could have better associative memory ability only when the number of connections of each neuron is relatively consistent. Zhu et al. proposed a new learning method of network weight, replacing SGN with a linear function, and proved that the improved method has a great improvement in associative memory.

Hopfield (1984) successfully applied the continuous neural network model to the traveling salesman problem for the first time, which opened the door of a continuous Hopfield network in dealing with optimization problems. GARCiA L et al. studied the Hopfield neural network algorithm for optimizing topological mechanisms. Chen et al. designed a logistics path planning scheme based on a continuous Hopfield network, mapped the objective function of path optimization into the energy function of the network, designed the dynamic equation of the objective function, and obtained the optimal value of path planning by finding the minimum value of the equation. Wang et al. used the continuous Hopfield neural network algorithm for route optimization design, comprehensively considered the constraints of route freight rate and distance, mapped the optimal solution of the problem to the stable state of the neural network, and finally obtained the route optimization scheme with the lowest freight rate per unit distance.

Hopfield neural network makes machine learning tend to the function of the human brain, widens the research road in the field of machine learning, and puts forward higher requirements for a deep neural network.

**Deep Neural Network**

The branch of the neural network mentioned above belongs to the traditional artificial neural network. Compared with a biological neural network, it is a shallow structure, which is far from human brain intelligence. With the improvement of computer processing speed and storage capacity, the design and implementation of a Deep Neural Network have gradually become possible. In the research field of the deep neural network, Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) are more studied.

The recurrent neural network is mainly used to process sequence data. Because the hidden layer of RNN is cyclic, as shown in Figure 6, its biggest feature is that the output of neurons at a certain time can be input to neurons again. This series network structure is very suitable for time series data and can maintain the dependency in the data. The expanded RNN is a repeated structure, and the parameters in the network structure are shared, which greatly reduces the neural network parameters to be trained. On the other hand, shared parameters also enable the model to be extended to data of different lengths, so the input of RNN can be an indefinite length sequence. Although the purpose of RNN at the beginning of design is to learn long-term dependence, due to the iterative nature of standard RNN, it is often difficult to save information for a long time.
The LSTM proposed by Hochreiter et al. is improved on the basis of RNN, which can effectively overcome the gradient disapp
earance problem in RNN, especially in long-
distance dependent tasks. Due to its excellent properties, LSTM has been used in a large number of tasks related to sequence le
arning. Wei et al. applied MLP, RNN, LSTM, and GRU to Pore Water Pressure (PWP) respectively, and concluded that the model with
RNN structure is more accurate than MLP for time series data, especially LSTM and GRU can describe the time-delay
effect between input and response, which is more accurate and reliable than standard RNN. Bai et al. combined RNN with
CRFID for the intelligent library, predicted readers' perceived needs at different stages through readers' borrowing records and
borrowing behavior, and provided readers with book purchase and personalized services. The actual data recorded by the library
confirmed that the model was feasible to sense readers' needs at different stages. Palangi et al. use LSTM to learn the sentence
vector with semantics and use the feature vector for the document retrieval task in the network. The hidden layer of the network
provides the semantic representation of the whole sentence and can detect the keywords in the sentence. Miyamoto et al.
introduced a language model based on BLSTM, which can adaptively mix word and character level word vectors to obtain the
final word vector representation. Mikolov et al. proposed a new RNN based language model (RNN LM), which can be applied
in speech recognition. Compared with the back-off language model, the confusion can be reduced by about 50% by using mixed
models of several RNN LMS. Google uses LSTM for speech recognition on its smartphones and Google translation.

In terms of application, it can be seen that the Recurrent Neural Network has made remarkable achievements in processing
sequence data applications such as text, audio, and video, which is of great significance to promoting the development of machine
learning and artificial intelligence.

Convolutional Neural Network is a common network in machine learning. It is one of the most widely used deep neural networks.
A typical Convolutional Neural Network consists of five parts: input layer, convolution layer, pooling layer, full connection layer,
and output layer. The convolution layer is responsible for extracting local features in the image. The pooling layer is used to
greatly reduce the parameter order (dimension reduction). The full connection layer is similar to the part of a traditional neural
network to output the desired results. Because the pooling layer can reduce the data dimension, it can not only greatly reduce the
amount of computation in the process of image recognition but also effectively avoid overfitting. With the continuous
development of a convolutional neural network, the number of hidden layers increases gradually, including activation layer,
normalization layer, tangent layer, fusion layer, and so on. In the activation layer, the ReLU activation function is used to introduce
nonlinearity into the network, and the output result of the convolution layer is nonlinearly mapped. The image is segmented by
cutting layers, and a part of the region is studied independently. The fusion layer can fuse the tangent layers or the features
learned by convolution kernels of different sizes. Based on the processing function of the hidden layer, the convolutional neural
network can realize powerful applications in the field of image processing.

Convolutional neural networks are widely used in the fields of image recognition, semantic segmentation, and machine
translation. Lou et al. applied VGG16 and CNN to face recognition, collected the discarded image information, and applied it to
the original CNN. The performance of the improved model is significantly improved compared with the ICA algorithm and
traditional convolutional neural network. Hu et al. used CNN for image diagnosis. After the CT image was manually marked by
experts, the mask R-CNN (Regions with CNN) was used to automatically cut the lung in the CT image. In the experiment, mask
R-CNN and K-means are combined to obtain the best segmentation effect compared with other methods. The segmentation
accuracy is 97.68%±3.42%, and the average running time is 11.2 s.

As a kind of artificial neural network, Convolutional Neural Network contains many networks with different structures. The
early convolutional neural network structure is relatively simple, such as the Le Net-5 model. It is mainly used in handwritten
character recognition, image classification, and other relatively single computer vision applications. With the deepening of
research, the structure of the Convolutional Neural Network is continuously optimized, and its application field is gradually
extended. As an unsupervised generation model, Convolutional Deep Belief Network (CDBN), which is generated by the
combination of Convolutional Neural Network and Deep Belief Network (DBN), has been successfully applied to face feature
extraction. Alex Net has made a breakthrough in the field of massive image classification. R-CNN has achieved success in the

![Recurrent neural network structure](image)

**Figure 6: Recurrent neural network structure.**
field of target detection. A fully Convolutional Network (FCN) realizes end-to-end image semantic segmentation and greatly surpasses the traditional semantic segmentation algorithm inaccuracy. In recent years, the research on the structure of the Convolutional Neural Network is still very hot. With the continuous improvement of network structure and performance, Convolutional Neural Network has emerged in various fields.

The research shows that the development of the neural network is in the spiral stage. The newly proposed or improved neural network can improve or expand the application of the previous neural network, but it can not completely cover the previous neural network. Networks with different structures have their own advantages and limitations in dealing with different problems. For this reason, analyzing neural networks from different perspectives will promote their improvement in structure and function.

CONCLUSIONS AND OUTLOOK
As one of the most popular research directions in the field of machine learning, the neural network has attracted people from all walks of life to study and develop it continuously. This paper summarizes the development process of the neural network and introduces the development and research status of each branch of the neural network. It can be found that the development of the neural network is becoming more general, flexible, and intelligent, which is mainly reflected in the following four aspects. (1) The change of model structure and the optimization of the algorithm. (2) The exploration efficiency of neural networks in complex scenes is gradually improving and has achieved good performance in some difficult tasks. (3) Deep neural network is trying to simulate the working model of a human brain-assisted learning system in order to construct a model of autonomous memory, learning, and decision-making. (4) The progress of technology and equipment provides an objective guarantee for the research and development of the network.

In the future, neural networks will develop in the following directions. (1) They tend to train networks through incremental and combined learning methods, and hybrid neural networks will be more widely used. (2) Unsupervised learning neural network is closer to the processing mechanism of the human brain and will play a more important role in the development of the neural network. (3) Further strengthen the inspiration of cognitive neuroscience on a neural network to gradually master the functions of memory, focusing, planning, learning, and understanding possessed by the human brain. (4) Transfer learning will be more applied to network models to alleviate the lack of training data in real task scenarios. (5) With the help of cloud server-side multi-task collaborative learning will become a new trend.

It can be predicted that, with the development of hardware technology and the continuous deepening of the research on neural network theory and methods, more opportunities and challenges will usher in the development of the artificial neural network. In iterative updates, the application of artificial neural networks will be more and more extensive and more mature. In the field of artificial intelligence and others, the development of neural networks will continue to surprise people. Humans will achieve the ideal goal of "solving intelligence and solving everything with intelligence" in the near future.

ACKNOWLEDGEMENTS
Supported by the Jiangsu water conservancy science and technology project (2019054).

REFERENCES