Digital Twins for Food Nutrition and Health Based on Cloud Communication

Zhihan Lv and Liang Qiao

1 Introduction

Today in the twenty-first century, people's living standards are constantly improving. Health has become a topic of concern. Health involves physical health and healthy diets. To provide proper guidance on people's diet, the relevant departments of nutrition and diet in China have formulated dietary guidelines according to the actual situation and publicized them to the masses via various publicity methods. For example, the even more popular short videos and Applets often advertise some foods' nutritional content, calorie content, and consumption methods. Still, it is unknown whether they can provide people with personalized guidance [1, 2]. With the popularization of Artificial Intelligence technology, there is also diverse nutritional diet-related software on the market recommending suitable food combinations through the user's eating habits. However, most computer-based healthy catering systems aim to cater to particular groups, such as pregnant women, infants, and diabetic patients [3]. The catering process of ordinary people is cumbersome and professional, so it is difficult to judge the rationality of the food collocation recommended by these systems. Therefore, the application of AI, Cloud Computing, and other technologies to the general public's nutrition, health, and balanced diet has become the concern of researchers in related fields.

A long-term, scientific, reasonable, and balanced diet can keep people healthy and enable them to create more value for society and improve work efficiency, promoting national economic development. With the progress of mobile communication, intelligent terminals are increasing rapidly, and the technologies related to

---

Z. Lv (✉)
Department of Game Design, Faculty of Arts, Uppsala University, Uppsala, Sweden

L. Qiao
College of Computer Science & Technology, Qingdao University, Qingdao, China

37

Cloud Computing and Communication are even more widely used [4–6]. In the dietary nutrition field, people's increasing attention to dietary nutrition balance is in sharp contrast to the current situation of dietary nutrition imbalance. Therefore, it is urgent to apply computer technologies to food nutrition identification and health analysis. For example, apply Big Data to food nutrition management, or use Cloud Communication technology to build a cloud platform for food nutrition and diet management. In the food cloud platform, Big Data is of a large volume. It can effectively extract the relevant data in food, such as the picture and weight of food, and analyze the relevant data information with the help of the Cloud Communication platform [7]. Computer technologies face the diversification of nutrition, food types, and nutrients required by users, and the constraints of users' families, cooking menus, and regions. The convolution calculation unique to Convolution Neural Networks (CNNs) in Deep Learning (DL) can extract and classify food nutrition data features. CNN can also understand the personalized preference of users' nutritional needs and food nutrition matching through the input of many orders and autonomic learning to realize the intelligent recommendation of food nutrition matching [8, 9]. Digital Twins (DTs) technology can map the food pictures and other data collected in the physical space to the virtual space. Then, it analyzes their nutrition in the cloud platform of the virtual space to achieve the balanced food nutrition intake of users [10].

To sum up, in today's expected progress of science and technology and social economy, it is of significant practical value to make people's intake of food nutritionally balanced and healthy to promote social development. The innovation of this study lies in four aspects. Firstly, Cloud Communication technology is adopted to analyze the needs of human nutrition and health. Secondly, the improved AlexNet network extracts and classifies food images. Thirdly, DTs technology is used to map the obtained image information and human nutrition and health data. Fourth, a food DTs nutrition evaluation model is constructed based on Cloud Computing combined with AlexNet. Besides, the objective function of the Sum Throughput Maximization (STM) algorithm is built to optimize the objective function. Finally, the model algorithm is verified by experiments to provide ideas for the follow-up balanced intake and intellectual development in the dietary nutrition field.

2   Recent Related Work

2.1   Development Status of Intelligent Detection of Dietary Nutrition Intake

Regular monitoring of people's nutritional intake is essential to reduce the risk of disease-related malnutrition, especially for special groups such as pregnant women, infants, and people with diabetes. Although methods for estimating nutrient intake have been gradually developed, there is still a clear need for a more reliable, fully automated technique. Many scholars have researched the intelligent detection of

nutrition and intake in food. Jia et al. (2019) employed a self-centered wearable camera to acquire pictures and developed an automatic food detection algorithm via AI from images for dietary estimation. Finally, the authors found that this algorithm could obtain images in the real world and automatically detect food with an accuracy of more than 85% [11]. Liu et al. (2020) proposed a food ingredient joint learning module based on the Attention Fusion Network for fine-grained food and ingredient recognition. The simulation proved that the module could effectively recognize fine-grained images of food [12]. Lu et al. (2020) put forward an AI-based new system to accurately estimate food nutrient exposure by simply processing Red Green Blue Depth image pairs collected before and after meals. The experimental results showed that the estimated nutrient intakes were highly correlated with the ground truth (>0.91), with a tiny average relative error (<20%), significantly better than existing nutrient intake assessment techniques [13]. Rachakonda et al. (2020) proposed a DL model for Edge Computing platforms capable of automatically detecting, classifying, and quantifying food on a user's plate. Compared with other paradigms, the model achieved an overall accuracy of 98% and an average accuracy of 85.8% [14]. Arslan et al. (2021) investigated food classification using common DL algorithms. The authors took the public food database as the object and found that the classification accuracy reached 90.02%, meeting the practical classification needs [15].

2.2 Application Status and Development Trend of DTs

The DTs technology has been extensively applied in various industries in recent years. It can map the actual data obtained in the industrial field into the virtual space, further increasing its potential. At present, DTs technology also extends to the medical and health areas. Many scholars have studied it. Laamarti et al. (2020) put forward an ISO/IEEE11073-standardized DTs framework, including acquiring information from individual health equipment, analyzing the data, and sending feedback to users in a cyclical fashion. Finally, the authors proved that this framework could be used as the foundation for promoting intelligent medical DTs [16]. Wan et al. (2021) used DTs technology to map physical brain pictures to virtual space for many un data in brain images. On this basis, the authors established a brain image integration DTs diagnosis and forecasting model based on Semi-Supervised Support Vector Machine and improved AlexNet. The experimental results showed that the constructed model had brilliant accuracy, acceleration efficiency, and segmentation and recognition performance under the premise of ensuring low error. The research supported detecting a brain image characteristics and digital diagnosis [17]. Francisco et al. (2021) used intelligent meters to adjust building energy in DTs of smart cities, making smart energy management within the geographical scope of medium and large buildings in the smart city a critical step in smart city construction [18]. Zhang et al. (2021) proposed a social-awareness-based vehicle edge caching mechanism to cope with the rapid development of intelligent

transportation. This mechanism dynamically coordinated the capability of roadside units according to user preference similarity and service availability, and the caching capabilities of smart vehicles. At the same time, the edge cache system was mapped into the virtual space using DTs technology to facilitate the construction of a social relationship model. Through simulation, the authors found that the constructed edge cache scheme had significant advantages in optimizing cache utility [19].

2.3 Analysis and Summary of Existing Research

To sum up, through the analysis of the researches of scholars in the above-mentioned related fields, many scholars have carried out intelligent detection on the nutritional components of food. However, there are few studies on dietary intake and nutritional matching. Besides, with the continuous improvement of DTs technology, its application fields have also developed from the initial industrial production to medical health and many other areas. Therefore, this paper also introduced DTs technology to evaluate dietary nutrition using intelligent algorithms like DL to balance food intake.

3   Analysis of the Nutrition Evaluation Model of Food DTs Based on Cloud Computing and DL

3.1 Analysis of Demand for Intelligent Assessment of Food Nutrition

The nutritional needs of the human body all come from food. To ensure the healthy balance required by the human body and achieve reasonable nutrition, people should eat diversified types of food and supplement nutrition according to the standard level. The *Dietary Guidelines for Chinese Residents* (2019) [20] issued by the Chinese nutrition society points out that eating more green leafy vegetables and fruits, taking cereals as the staple food, and adequately supplementing dairy products, beans, and a reasonable diet with meat and vegetables every day can make the human body carry out normal metabolism. The average level of nutrition required by the human body (average demand), the amount of food recommended by the nutrition standard (recommended daily intake), the proportion of food suitable for intake (adequate intake), and the macronutrients like protein are within the acceptable range. Based on the nutritional needs ratio, macronutrients, namely protein, fat, and sugar, primary elements, trace elements, namely minerals and various vitamins, fat-soluble vitamins, and water-soluble vitamins, have been included in the *Dietary Reference Intakes of Chinese Residents* by the Ministry of Health. Productive
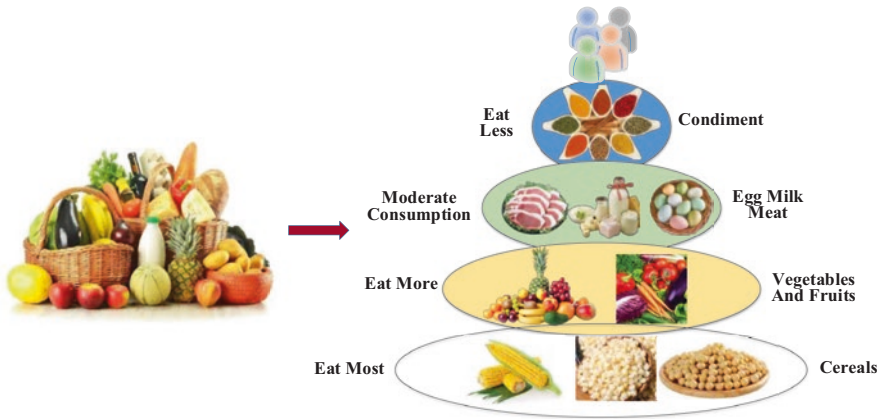
Fig. 1 Examples of food nutrition collocation

(thermal) nutrients release energy in the human body, including proteins, lipids, and carbohydrates [21]. Figure 1 displays the food nutrition collocation.

As shown in Fig. 1, human intake primarily includes five categories: spices, eggs, milk, meat, vegetables, fruits, and cereals. Due to the different nutritional content of various foods, the intake is from less to more when matching, but the demand for nutrients of different groups is different. Food recommendation is faced with the diversification of user characteristics, nutrition required by users, food types, food and nutritional components of agricultural products, and the constraints of multi-user families, cooking menus, historical menus, seasons, regions, etc. Therefore, selecting the right food from multiple agricultural products under much agricultural product classification and making intelligent recommendations to users is a practical demand. Based on the food DTs nutrition evaluation model of Cloud Computing combined with DL, this paper collects the information of relevant food data in the human body using Cloud Computing technology and transmits it to the cloud center for processing. Besides, DTs technology is used to map the data information to virtual space to design a multi-objective optimization algorithm for human nutrition demand. The methods designed here can achieve the optimal matching relationship between the nutrition required by the human body and the necessary nutrition in various foods.

3.2 Application of Cloud Computing and Communication Technology to Human Nutrition and Health Analysis

As one of the essential branches of the communication network, the Wireless Body Area Network (WBAN) takes the human body as the network information environment. It is a communication network composed of light, ultra-thin, low-power, and intelligent sensors located on the body surface or in the body. The node transmits

the critical physiological parameters obtained by the WBAN in real-time to the health monitoring center and the Cloud Computing center through the wireless communication network to realize data sharing [22, 23]. Here, the WBAN is applied to human health analysis.

A WBAN is a wireless communication network constructed by multiple sensor nodes located on the surface, inside and outside the human body. The whole system adopts the operating mode of first distributed perception information, then aggregated processing and remote analysis. The specific WBAN communication architecture is divided into three-layer: communication within the WBAN, communication between the WBAN, and communication outside the WBAN, as shown in Fig. 2.
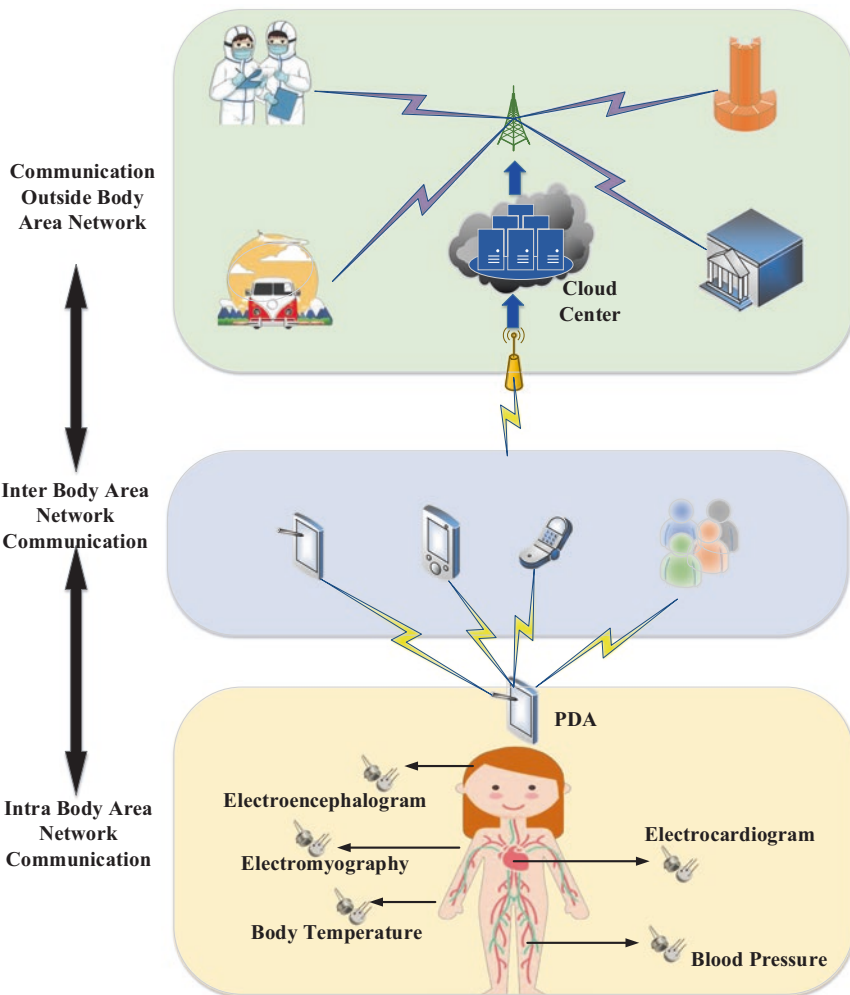

Fig. 2 WBAN communication architecture under cloud communication

As shown in Fig. 2, in the WBAN communication, various sensors (such as EEG, ECG, blood pressure sensors, etc.) distributed in the human head, limbs, and trunk communicate with the sink node in real-time. By designing a network structure suitable for the scene, they choose the corresponding communication protocol according to the receiver structure. In interbody area network communication, a Personal Digital Assistant obtains human physiological information shares information and preliminarily integrates and analyzes data with other wireless body area networks (including home data terminals) through Bluetooth or Wi-Fi communication. The communication outside the Body Area Network shares the human physiological data to hospitals, intelligent homes, smart cars, etc., through the straightforward base station and the Internet to provide people with various services. It even analyzes human data through Big Data and Cloud Computing to provide high-quality dietary balance and human health decisions.

Due to the dynamic characteristics of the human body (such as body posture, muscle movement), shadow effect, and the different environment of WBAN, the typical path loss model P_L can be described as Eq. (1).

$$P_L = PL(d) + X \tag{1}$$

In Eq. (1), according to the Fries Transfer Formula [24] in the free space, $PL(d)$ refers to the path loss corresponding to the reference distance d to the node, and X stands for the shadow factor of a normal distribution with zero mean and standard deviation. The shadow factor X is determined according to different channel environments.

In body-surface communication, the channel loss function $P_{dB}(d)$ between body-surface nodes is constructed as Eq. (2).

$$P_{dB}(d) = P_{0,dB} + 10n \log(d / d_0) \tag{2}$$

In Eq. (2), n represents the channel loss index, d_0 denotes the reference distance, and $P_{0,dB}$ corresponds to the initial path loss of the reference distance d_0.

However, the non-uniform and lossy communication transmission medium composed of human tissue layers is not uniform. Each tissue layer will absorb the electromagnetic energy of the sensor, significantly reducing the signal transmission in the body. Therefore, the channel loss model P_d between receivers and transmitters in vivo communication can be expressed:

$$P_d = \frac{P_t}{P_r} = P_0 (d / d_0)^n \chi_r \tag{3}$$

$$P_d(dB) = P_0(dB) + 10n \log(d / d_0) + \chi_r(dB) \tag{4}$$

where P_t and P_r represent the transmitted signal power and received signal power in the communication link, respectively; P_0 refers to the initial path loss corresponding

to the reference distance d_0; $\chi_r \sim N\left(0,\sigma_\chi^2\right)$ refers to the usual random variables with mean 0 and variance σ_χ^2, that is, the deviation caused by different tissue layers of the human body (such as skin, muscle, fat, etc.) and antenna gain in different directions.

When monitoring human health, WBAN can realize real-time monitoring of critical physiological parameters of the human body through the transmission and collection of radio frequency energy and the acquisition of information data. Among them, the energy and information transmission models of WBAN principally include a point-to-point transmission model and a cooperative transmission model with relay nodes, as presented in Fig. 3.

In Fig. 3, h and g refer to the channel coefficients of the downlink and uplink, respectively. The point-to-point energy and information transmission model is the most basic and simple transmission model in the WBAN. It is suitable for scenes with close distance and little shadow effect interference. However, in the scenario of long-distance, the point-to-point direct link will affect the system's overall performance. The relay node needs to forward energy and information to solve the problems of human dynamics and shadow effect to ensure the efficiency and reliability of system information transmission.

In the point-to-point transmission model, the SN P simultaneously transmits the energy flow carrying the command information to the sensor S. After passing through the power splitter at the receiving end, a part of the wireless signal is used for energy collection. The signal y_s received by the sensor node is expressed as Eq. (5).

$$y_s = \sqrt{P_a}\,hx + n_s \qquad (5)$$

In Eq. (5), P_a and n_a refer to the transmit power and antenna noise of sensor S, respectively; x_a refers to the signal sent by the sensor node when the wireless signal is used for energy harvesting. The energy E collected by the sensor can be written as Eq. (6).
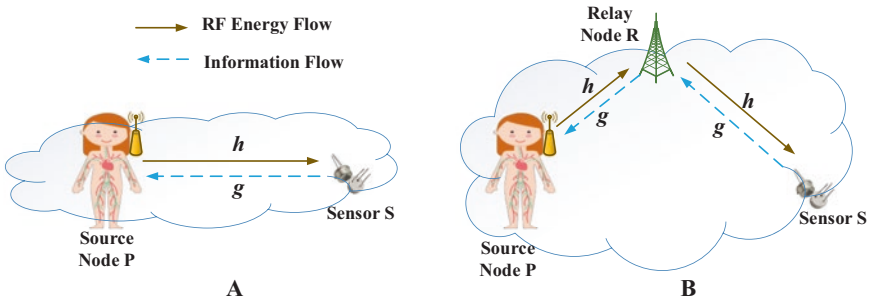
Fig. 3 Energy and information transmission model of WBAN under cloud communication (**a**. point-to-point transmission model; **b**. cooperative transmission model with relay nodes)

$$E = \eta \rho P_a |h|^2 (T / 2) \tag{6}$$

In Eq. (6), $\rho(0 < \rho < 1)$ and $\eta(0 < \eta < 1)$ refer to the power split ratio and energy conversion efficiency, respectively; T stands for the time period of information transmission. Another part of the wireless signal is used for information decoding. The signal y_s' received by sensors and the signal x_s sent by sensors can be written as:

$$y_s' = \sqrt{1-\rho} \left(\sqrt{P_a} h x_a + n_s \right) \tag{7}$$

$$x_s = \sqrt{\mu} \left(\sqrt{1-\rho} \left(\sqrt{P_a} h x_a + n_s \right) + n_s' \right) \tag{8}$$

where $n_s' \sim N(0, \sigma_n^2)$ stands for the additional processing noise on the sensor node. μ refers to the amplification and forwarding coefficient, as shown in Eq. (9).

$$\mu = \eta \rho P_a |h|^2 / (1-\rho) P_s |g|^2 + \sigma_n^2 \tag{9}$$

In Eq. (9), P_s refers to the received power of the sensor S.

Moreover, after the sensor S decodes the information and receives the energy, it sends the sensed information back to the SN P to complete the information transmission for some time T. After the amplification and forwarding coefficient μ is energy balanced, the signal received by the SN P is y_a expressed as Eq. (10).

$$y_a = g \sqrt{\mu} \left(\sqrt{1-\rho} \left(\sqrt{P_a} h x_a + n_s' \right) + n_a \right) \tag{10}$$

In Eq. (10), n_a refers to the antenna noise of SN P following a normal distribution with mean 0 and variance σ_a^2. Finally, the throughput from sensor S to SN P is expressed as Eq. (11).

$$R = \frac{1}{2} \log_2 (1 + \gamma) \tag{11}$$

In Eq. (11), γ refers to the Signal to Noise Ratio (SNR) from the sensor S to the SN P in the point-to-point transmission model.

In the relay cooperative transmission model, the relay node R is used to amplify and forward the wireless energy and information from the SN P and the sensor S. Usually, there are two power splitting and time switching protocols for the transmission of energy and information.

In the power-sharing protocol, the SN P and the sensor S transmit wireless energy and information to the relay node in the $T/2$ time slot. Then, the signal received by the relay node y_r can be described as Eq. (12).

$$y_r = \sqrt{P_a} h x_a + \sqrt{P_s} g x_s + n_r \tag{12}$$

In Eq. (12), n_r refers to the noise generated by the relay node. Another part of the signal y_r' received through power splitting is expressed as Eq. (13).

$$y_r' = \sqrt{1-\rho}\left(\sqrt{P_a}\,hx_a + \sqrt{P_s}\,gx_s + n_r\right) \tag{13}$$

The relay node forwards the information and energy to the SN P and the sensor S in the remaining time slots, respectively. Equation (14) indicates the signal $y_{a,ps}$ received by the SN P.

$$y_{a,ps} = h\sqrt{\beta_{ps}}\left(\sqrt{1-\rho}\,\sqrt{P_s}\,gx_s + n_r'\right) + n_a \tag{14}$$

In Eq. (14), $n_r' \sim N\left(0,\sigma_r^2\right)$ refers to the additional processing noise of the relay, and β_{ps} represents the balance coefficient to ensure energy collection and consumption, as shown in Eq. (15).

$$\beta_{ps} = \eta\rho P_a |h|^2 / \left((1-\rho)P_s|g|^2 + \sigma_r^2\right) \tag{15}$$

Then, the SNR generated on the SN P can be written as Eq. (16).

$$\gamma_{a,ps} = \frac{(1-\rho)P_s|h|^2|g|^2}{|h|^2\sigma_r^2 + \sigma_a^2\left(\left((1-\rho)P_s|g|^2 + \sigma_r^2\right)/\left(\rho P_a|h|^2\right)\right)} \tag{16}$$

In the power-sharing protocol, in the first time slot τT, the SN P transmits wireless energy to the relay node. $0 \le \tau \le 1$ means the time switching ratio. Equation (17) demonstrates the total energy $E_{r,ts}$ collected by the relay node.

$$E_{r,ts} = \eta P_a |h|^2 \tau T \tag{17}$$

Furthermore, in the second time slot $(1-\tau)T/2$, the information $y_{r,ts}$ received by the relay node from the sensor is expressed as:

$$y_{r,ts} = \sqrt{P_s}\,gx_s + n_r' \tag{18}$$

where $n_r' \sim N\left(0,\sigma_r^2\right)$ refers to the noise on the relay node. The relay node forwards the amplified signal to the SN P; meanwhile, the sensor decodes information and energy collection. Equation (19) describes the balance coefficient β_{ts} satisfying the energy collection and consumption.

$$\beta_{ts} = 2\tau P_a |h|^2 / \left((1-\tau)\left(P_s|g|^2 + \sigma_r^2\right)\right) \tag{19}$$

Similar to the analysis of the power shunt protocol, the SNR $\gamma_{a,ts}$ generated by the AP SN is presented as Eq. (20).

$$\gamma_{a,ts} = \frac{P_s |h|^2 |g|^2}{|h|^2 \sigma_r^2 + \sigma_a^2 \left(\left((1-\tau) \left(P_s |g|^2 + \sigma_r^2 \right) \right) / \left(2\tau\eta P_a |h|^2 \right) \right)} \tag{20}$$

The final throughput R_{ts} from the sensor to the SN P is:

$$R_{ts} = \frac{1-\tau}{2} \log_2 \left(1 + \gamma_{a,ts} \right) \tag{21}$$

3.3 Application of DL in Food Nutrition Evaluation and Analysis

DL is a typical feature extraction algorithm. CNN is a booming Feedforward Neural Network with the optimal effect. Its most significant superiority lies in local connectivity and sharing weights. A large number of neurons in the model are organized in a particular pattern and have responded to the overlapping areas in the visual field [25]. The first parameter of the CNN's operation refers to the input, the second parameter (function *w*) refers to the kernel function, and the output refers to the characteristic mapping [26]. Generally, CNNs conduct convolutions in multiple dimensions. Denote the input as a two-dimensional matrix *I*. Then, a two-dimensional kernel K is used, as shown in Eq. (22).

$$S(i,j) = (I \cdot K)(i,j)$$
$$= \sum_m \sum_n I(m,n) K(i-m,j-n) \tag{22}$$

In Eq. (22), *i*, *j*, *m*, and *n* are all fixed parameters, referring to the dimension and order of the matrix. When CNN is applied to food image feature extraction and classification, the food image information in CNN intersects several convolutional and pooling layers and gathers together by one or more fully connected layers. All neurons in the layers are connected to the neurons in the previous layer. Usually, the establishment of this layer relies on two one-dimensional network layers. The local information of the convolutional layer or the pooling layer is grouped. The ReLU function is usually used as the excitation function by all neurons to improve the performance of the network.

AlexNet [27], a deep CNN model with more network layers and more vital learning ability is selected in this paper to simply the computation and enhance the generalization ability of CNN. Figure 4 reveals the process of extracting and classifying DTs data features of food images based on AlexNet.
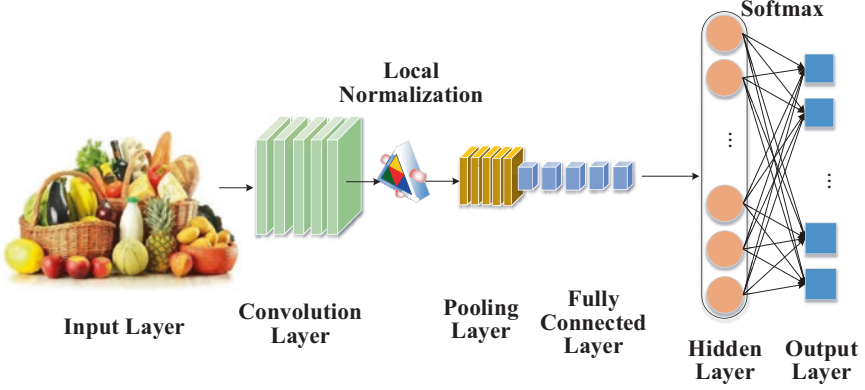
Fig. 4 Extraction and classification process of DTs data features of food images via AlexNet

In Fig. 4, the convolution operation is performed on the food image first when extracting and classifying the DTs data of food images. Then, the local normalization, pooling, and full connection operations are performed.

Moreover, the functional layer of the AlexNet model's convolutional layer is improved. Specifically, the local normalization and pooling operation operations in order are reversed. It can enhance the generalization ability of AlexNet and weaken the overfitting phenomenon, significantly reducing the training time. In addition, local normalization after Overlapping Maximum Pooling (OMP) can retain more useful information and drain redundant information in the pooling process. It also speeds up the convergence of the model training and highlights the advantages of OMP over the existing MP approaches. The calculation of each layer is as follows.

The overlapping pooling method samples the *t*-th characteristic mapping $y_t^l(i,j)$ of the *l-th* convolutional layer.

$$a_t^l(i,j) = \max\{y_t^l(i,j), i_s \leq i \leq i_s + w_c - 1, \\ j_s \leq j \leq j_s + w_c - 1\} \tag{23}$$

In Eq. (23), *s* stands for the pooling movement step size, w_c represents the pooling area's width, and $w_c > s$.

The AlexNet model performs the first and second pooling operations. Then, a local normalization layer is introduced to normalize the feature map according to Eq. (24).

$$c_t^l(i,j) = a_t^l(i,j) / \left(k + \alpha \sum_{\max(0, t - m/2)}^{\min(N-1, t+m/2)} \left(a_t^l(i,j) \right)^2 \right)^\beta \tag{24}$$

In Eq. (24), *k*, α, β, and *m* are hyperparameters, the values of which are 2, 0.78, 10^{-4}, 7, respectively. *N* denotes the amount of convolution kernels of the first

convolutional layer. The ReLU function is used to activate the convolutional product $S_t^l(i,j)$ to avoid gradient dispersion in the network model. The activation function $y_t^l(i,j)$ can be expressed as Eq. (25).

$$y_t^l(i,j) = f\left(S_t^l(i,j)\right) = \max\left\{0, S_t^l(i,j)\right\}$$ (25)

The Dropout Operation's parameters are set to 0.5 to prevent overfitting in the fully connected layer. All food feature maps are integrated into a high-dimensional single-layer neuron architecture C^5. Equation (26) describes the input Z_i^6 of the i-th neuron in the sixth fully connected layer.

$$Z_i^6 = W_i^6 C^5 + b_i^6$$ (26)

In Eq. (26), b_i^6 and W_i^6 represent the sixth fully connected layer's bias and weight of the i-th neuron, respectively.

The neurons C^l in the 6th and 7th fully-connected layers are abstained and output in $_l$ the process of generalization capability improvement; $r_j^l \sim bernoulli(dp)$; $\tilde{C}{}^l = r^l C^l$. Then, the input Z_i^{l+1} of the i-th neuron in the seventh and eighth fully-connected layers is $W_i^{l+1} \tilde{C}{}^l + b_i^{l+1}$. The output C_i^l of the i-th neuron of the 6th and 7th fully-connected layers is $f\left(Z_i^l\right)$, i.e., $\max\left\{0, Z_i^l\right\}$. Finally, the input q^i of the first neuron of the eighth fully connected layer can be obtained according to Eq. (27).

$$q^i = soft\max\left(Z_i^8\right) = \frac{e^{Z_i^8}}{\sum_{j=1}^{12} e^{Z_i^8}}$$ (27)

Meanwhile, the Cross-Entropy Loss Function (CELF) for the food image classification problem is used as the model's error function, which can be written as Eq. (28).

$$Loss = \sum_{i=1}^{K} y_i \cdot \log\left(p_i\right)$$ (28)

$$p_i = \frac{\exp\left(\tilde{y}_i\right)}{\sum_{i=1}^{K} \exp\left(\tilde{y}_j\right)}$$ (29)

In Eq. (28), K stands for the quantity of types; y_i denotes the accurate class distribution of the samples; y_i refers to the CNN's output; p_i signifies the sort results after the Softmax categorizer. The Softmax function's input is an N-dimensional real vector, denoted as x, as presented in Eq. (30).

$$\xi\left(x\right)_i = \frac{e^{x_i}}{\sum_{n=1}^{N} e^{x_i}}, i = 1, 2, \ldots, N$$ (30)

Furthermore, the essence of the Softmax function is analyzed. This function maps an N-dimensional random real vector into an N-dimensional vector with the values of each element within (0, 1) to realize the normalization of the vector. The μ companding transformation reduces the output data volume to 2^8 to simplify the neural network's computation, i.e., $\mu = 255$, to improve the model's forecasting ability.

$$f\left(x_t\right) = sign\left(x_t\right)\frac{\ln\left(1+\mu\left|x_t\right|\right)}{\ln\left(1+\mu\right)}, \left|x_t\right| < 1 \tag{31}$$

3.4 Construction and Analysis of the DTs Evaluation Model of Food Nutrition and Health Based on Cloud Computing and AlexNet

Firstly, human nutrition and health needs are analyzed through cloud communication technology to realize human nutrition and health needs and multi-targeted personalized intelligent food recommendation. The improved AlexNet is used to extract and classify the features of food images. Then, DTs technology maps the obtained image information and human nutrition and health data to the virtual space to intelligently recommend methods suitable for human nutrition and health. Figure 5 illustrates the food nutrition and health DTs evaluation model based on Cloud Computing and AlexNet.

As shown in Fig. 5, this model adopts Cloud Computing, and Communication technology are to obtain human health data. Then, the improved AlexNet extracts and classifies food nutrition data. The human health and food nutrition data obtained in the real space are mapped to the virtual space via DTs technology. The obtained data results are finally evaluated through data fusion and classification transmission to determine the healthy recipes suitable for human health.

When using Cloud Communication technology to acquire and analyze human health data, the proposed time switching protocol constructs the overall throughput objective function of multipoint WBAN. It presents the objective function of STM optimization. The overall throughput of the system is:

$$(\text{P1}): \max_{\tau_0,\tau_i} \sum_{i=1}^{K} R_i^{ab}\left(\tau_0,\tau_i\right) = \max_{\tau_0,\tau_i} \frac{1}{2}\sum_{i=1}^{K}\tau_i \log_2\left(1+\frac{X_i^{ab}\frac{\tau_0}{\tau_i}}{Y_i^{ab}\frac{\tau_0}{\tau_i}+Z_i^{ab}}\right) \tag{32}$$

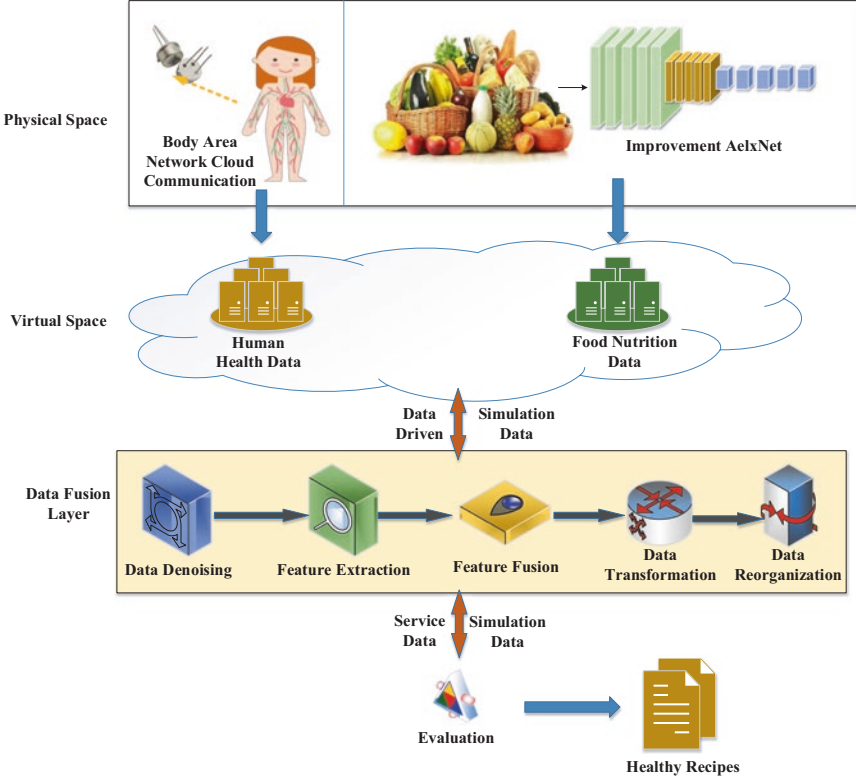$$s.t.\sum_{i=1}^{K}\tau_i \le 1, \tau_i \in \left[0,1\right], i = 0,1,\ldots,K \tag{33}$$

Fig. 5 Food nutrition and health DTs evaluation model based on Cloud Computing and AlexNet

where R_i^{ab} represents the throughput from the sensor S_i to the SN P under abnormal conditions; X_i^{ab}, Y_i^{ab}, and Z_i^{ab} are different parameters in the objective function, respectively; τ_0 refers to the initial energy transmission time slot; τ_i stands for the energy transmission time slot at time i.

In this model, a large amount of food image data is used to improve AlexNet for self-learning to accurately learn the personalized preferences matching the nutritional needs of users with the nutrition of agricultural products. When using the enhanced AlexNet for training, the Weighted Cross-Entropy is used as the cost function to optimize the model training. Denote $z_k(x, \theta)$ as the unnormalized logarithmic probability value of the pixel x of the k-th category of the food image with a specific parameter θ. The Softmax function $p_k(x, \theta)$ is decided as Eq. (34).

$$p_k\left(x,\theta\right) = \frac{\exp\left\{z_k\left(x,\theta\right)\right\}}{\sum\limits_{k'}^{K}\exp\left\{z_{k'}\left(x,\theta\right)\right\}} \tag{34}$$

In Eq. (34), K represents the amount of food categories. In the prediction stage, when Eq. (34) takes the maximum value, the pixel x is marked as the k-th class, i.e., $k^* = \arg\max\{P_k(x,\theta)\}$. The log-normalized probability value is abbreviated as p_{ik}. Thus, the training focuses on locating the best network parameters θ^* by minimizing the weighted CELF $\ell(x,\theta)$, i.e., $\theta^* = \min_{\theta} \ell(x,\theta)$. The Weighted Loss Function for food picture integration is defined as Eq. (35).

$$\ell(x,\theta) = -\sum_{i=1}^{N}\sum_{k=1}^{K} w_{ik} q_{ik} \log p_{ik} \tag{35}$$

In Eq. (35), $q_{ik} = q(y_i = k|x_i)$ refers to the actual label distribution of the pixel x_i of the k-th category; w_{ik} represents the weighting coefficient. In the training process, the calculation strategy shown in Eq. (36) is adopted.

$$w_{ik} = \frac{1}{\ln(c + p_{ik})} \tag{36}$$

In Eq. (36), c stands an extra super parameter, empirically set to 1.11 during this experiment.

Figure 6 indicates the training process of the improved AlexNet algorithm applied to the food image DTs model

3.5   Experimental Testing and Evaluation

This section evaluates the performance of the constructed food nutrition and health DTs evaluation model based on Cloud Computing combined with AlexNet. Firstly, the analysis effect of human health data under Cloud Communication technology is evaluated. The human body path loss model is chosen to be located on the torso of the human body. The channel reciprocity between uplink and downlink makes the channel coefficients the same for uplink and downlink. The objective function optimized by the STM algorithm reported here is compared with the Equal Time Allocation (ETA) algorithm [28] and the Minimum Throughput Maximization (MTM) algorithm [29]. These algorithms are compared from the changing curve of the overall throughput of the system with the transmission power (TP) P_a of the source node (SN) P, the number of sensors S, and the energy conversion efficiency η. The changing trend of the energy transmission time slot τ_i with the number of sensors is analyzed. In addition, various algorithms are analyzed in terms of fairness index.

The improved AlexNet proposed here is also evaluated from the food nutritional characteristics data. This study collects about 1284 kinds of food nutrition components from the *Chinese Food Nutrient Composition Table* as the data of this experiment. After collecting the nutritional data of agricultural products, the data is

1	**start**
2	Input: Food image nutrition data matrix D
3	Output: Improved AlexNet food nutrition data matrix Z_i
4	**For** $D:[1,\cdots,X_i], i=1,2,...,N$ **do**
5	Searching for the optimal network parameter $\theta*$ by CELF $\ell(x,\theta)$;
6	$\theta* \leftarrow \min_\theta \ell(x,\theta)$
7	$\ell(x,\theta) \leftarrow \sum_{i=1}^{N}\sum_{k=1}^{K} w_{ik}q_{ik}\log p_{ik}$
8	$q_{ik} \leftarrow q(y_i = k\mid x_i)$
9	$w_{ik} \leftarrow \dfrac{1}{\ln(c+p_{ik})}$
10	$c \leftarrow 1.10$
11	$z_t \leftarrow \mathrm{Re}lu\left(W_z h_t + b_z\right)$ //Fully connected layer
12	$prediction \leftarrow Soft\max\left(W_{sft}z_t + b_{sft}\right)$ // Classification and prediction of food nutrition
13	$Z_i \leftarrow Z_i \cup z_t$
14	**return** Z_i
15	**end**

Fig. 6 Training flow of the improved AlexNet algorithm applied to the food image DTs model

processed to remove some irrelevant data and parameters. Besides, the food data is divided into five categories (staple food, vegetables, meat, aquatic products, and fruits) according to the nutritional balance guidelines. There are about 200 pieces of data for the experiment after processing. A unique Identity Document is assigned to each data to be distinguished by food category. For the neural network, the following hyperparameters need to be set: the network iterates for 120 times, the simulation time is 2000 s, and the Batch Size is 128. The Poly learning rate adjustment method is adopted as the learning rate update strategy using Polynomial Decay, expressed as $init\_lr\times\left(1-\dfrac{epoch}{\max\_epoch}\right)^{power}$. The initial learning rate $init\_lr$ is 0.0005 (or 5e^{-4}), and the power is set to 0.9.

The algorithm reported here is compared with the algorithms applied by other scholars in related fields for the performance analysis, including Long Short-Term Memory (LSTM) [30], CNN [31], Recurrent Neural Network (RNN) [32], AlexNet [33], and Multilayer Perceptron (MLP) [34]. The classification accuracy is compared and analyzed from the perspectives of Accuracy, Precision, Recall, and F1 value. The time required by different models is also examined.

The specific experimental configuration primarily considers hardware and software, as summarized in Table 1.

Table 1 Experimental configuration

| Software			Hardware		
Operating System	Python	Development Platform	Central Processing Unit	Memory	Graphics Processing Unit
Linux 64bit	3.6.1	PyCharm	Intel core i7–7700@4.2GHz of eight cores	Kingston ddr4 2400 MHz 16G	Nvidia GeForce 1060 8G

4 Research Results

4.1 Comparative Analysis of Human Health Data Evaluation under Different Cloud Communication Technologies

The objective function optimized by the STM algorithm reported here is compared with the ETA algorithm and the MTM algorithm. This experiment compares the changing curves of the overall throughput of the system with the TP P_a of the SN P, the number of sensors S, and the energy conversion efficiency η. The changing trend of the energy transmission time slot with the change of the number of sensors is also discussed. In addition, various algorithms are analyzed from the perspective of fairness. The results are shown in Figs. 7, 8, and 9. Two repeated experiments for each algorithm are conducted separately.

As shown in Fig. 7a, as the TP of the SN P increases, the overall throughput of different algorithms grows. This is because the greater the transmit power, the more broadcast energy of the SN, resulting in a smooth rise in the overall throughput. In addition, according to the international standard of WBAN, the TP cannot exceed one mW; otherwise, it will cause damage to the human body. The TP of the SN ranges from 0 mW to 1 mW. The throughput performance of the STM algorithm reported here is significantly better than that of the MTM algorithm and the ETA algorithm. The throughput performance of the STM algorithm proposed here is considerably better than that of the MTM algorithm and ETA algorithm. The overall throughput of the ETA algorithm is lower, and the difference between them is no more than 2 bps/Hz. According to Fig. 7b, the overall throughput of different algorithms increases smoothly with the growth of energy conversion efficiency η. It is possible that the higher the energy conversion efficiency, the more energy the sensor can harvest for energy harvesting, resulting in greater throughput. In addition, it can be found that the throughput performance of the STM algorithm reported here is significantly better than that of the MTM algorithm and the ETA algorithm. The ETA algorithm has the lowest overall throughput. Besides, with the increase of energy conversion efficiency η, the overall throughput gap between the STM algorithm and ETA algorithm becomes larger. When η is 1, the throughput difference between the two is about 1.5 bps/Hz.

Figure 8**a** suggests that the overall throughput of each algorithm increases significantly as the number of sensors grows. In addition, the overall throughput
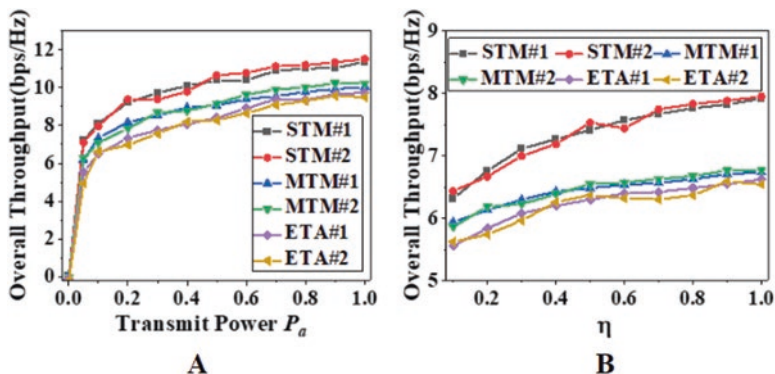
Fig. 7 Variation curves of the overall throughput varying with the TP (mW) of the SN and energy conversion efficiency η (**a**. the overall throughput varying with the TP of the SN; **b**. the overall throughput varying with the energy conversion efficiency η)
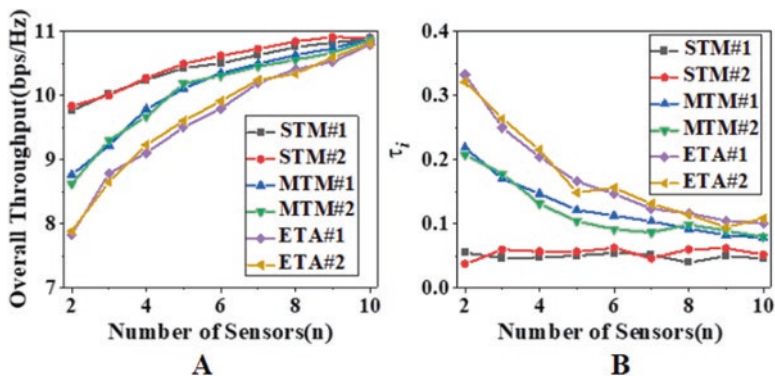
Fig. 8 Variation curves of overall throughput and energy transmission time slot with the number of sensors (**a**. The overall throughput changing with the number of sensors; **b**. the energy transmission time slot changing with the number of sensors)

performance of the STM algorithm reported here is substantially better than that of the MTM algorithm and the ETA algorithm. The overall throughput of the ETA algorithm is the lowest. With the increase of sensors, the overall throughput gap between STM algorithm, MTM algorithm, and ETA algorithm become smaller. When the number of sensors is 7, the throughput gap between the two does not exceed 0.5 bps/Hz. According to Fig. 8b, as the number of sensors increases, the energy transmission time slots τ_i of the MTM algorithm and the ETA algorithm show a downward trend, sharply affected by the number of sensors. However, the energy transmission time slot τ_i of the STM algorithm proposed reported here does not decrease significantly, almost stable within a specific range. The increase in the number of sensors implies that more information transmission time slots are required, reducing the energy transmission time slots within the normalized time
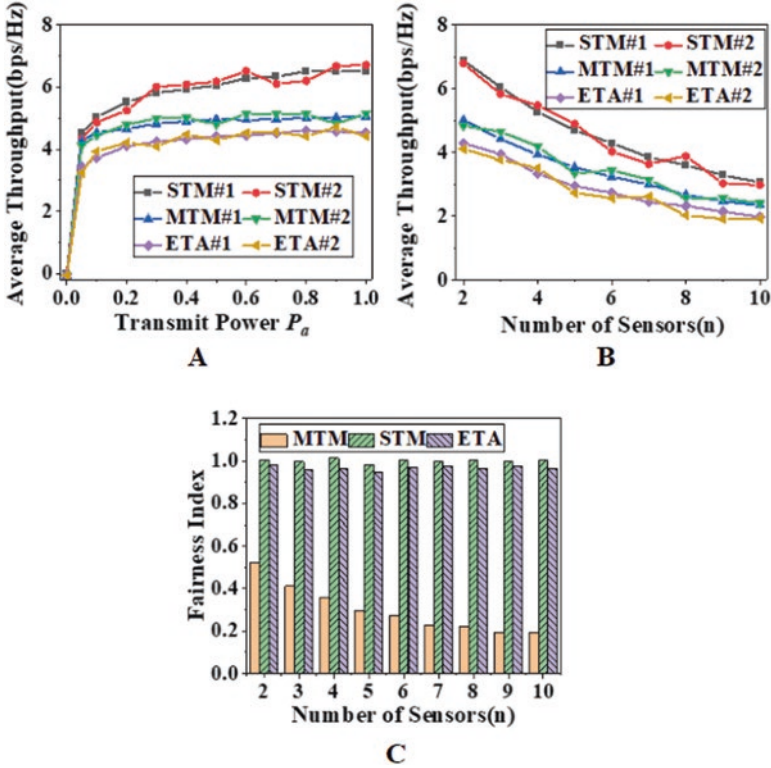
Fig. 9 Results of the average throughput varying with different factors (**a**. Average throughput varying with the TP of the SN; **b**. average throughput varying with the number of sensors; **c**. the fairness index varying with the number of sensors)

slot T. This explains why the energy transmission time slots τ_i for each algorithm are reduced. Therefore, the STM algorithm can be kept within a certain range (τ_i is around 0.05) to ensure the stability of energy transmission

As shown in Fig. 9a, with the increase of the TP of the SN, the average throughput of the system under different algorithms increases smoothly. This is consistent with the reason that overall throughput increases with transmit power. In addition, the STM algorithm reported here produces significantly better throughput than the MTM algorithm and the ETA algorithm. With the increase of the TP of the SN, the difference between the MTM and ETA algorithms gradually reduces. According to Fig. 9b, as the number of sensors increases, the average throughput decreases, which just verifies the phenomenon presented in Fig. 8b. The throughput of each sensor decreases due to the decrease of energy transmission time slots due to the increase of the number of sensors. In addition, the STM algorithm has a significantly higher average throughput than the MTM algorithm and ETA algorithm. The change of fairness index with the number of sensors is shown in Fig. 9c. It can be found that the fairness index of the MTM algorithm decreases gradually with the

uneven resource allocation caused by the addition of new sensors to the network. The fairness index of the STM algorithm is one and remains unchanged, making the resource allocation fair to obtain the same throughput to maximize the average throughput. The Equal Time Distribution Protocol assigns the same information and energy transmission time slot to all sensors. Therefore, the fairness index of the ETA algorithm decreases slightly with the increase of sensors, which is lower than that of the MTM algorithm. Thus, in terms of resource fairness, the STM algorithm is better than the ETA algorithm and MTM algorithm, which has better stability and better effect in analyzing human health data.

4.2 Prediction Performance Analysis of Food Nutrition Classification Under Different Algorithms

This test analyzes the food nutrition recognition accuracy of the food nutrition and health DTs evaluation model based on Cloud Computing combined with AlexNet constructed here. The recognition and prediction accuracy of the algorithm reported here is compared with LSTM, CNN, RNN, AlexNet, and MLP from Accuracy, Precision, Recall, and F1 value, respectively. The results are shown in Fig. 10. Figure 11 further compares the training and test time required by each algorithm.

Figure 10 shows the comparison results of the model constructed here and the models proposed by scholars in other related fields from the perspectives of Accuracy, Precision, Recall, and F1 value. It can be found that the Accuracy of this algorithm reaches 93.17%. Compared with other neural network algorithms, the Accuracy is increased by at least 2.01%. Meanwhile, it is evident that the algorithm reported here has the highest values of Precision, Recall, and F1. Compared with other algorithms, the difference is at least 3.24%. In conclusion, compared with other neural network algorithms, the algorithm used here to build a food nutrition and health DTs evaluation model based on Cloud Computing combined with AlexNet has excellent food nutrition recognition and prediction accuracy.

Figure 11 provides the results of the time needed by each algorithm in the test set and training set under different food image data amounts. It can be found that the optimized AlexNet algorithm is better than the algorithms used by other scholars. Under different food image data amounts, the time required for the food image data increases gradually. At the same time, the time required by various algorithms in the test set is slightly lower than the time needed for the training set. This phenomenon is since the model has not yet formed a nutritional identification and analysis path corresponding to the food image during the training process. Besides, the neural network learns autonomously through the training process. Afterward, it can precisely analyze the nutritional classification of many food images, significantly improving the efficiency. Therefore, in the comparative analysis of the experimental results of all methods, the food nutrition and health DTs evaluation model based on Cloud Computing combined with AlexNet performs the best and achieves the better food nutrition classification and identification effects.
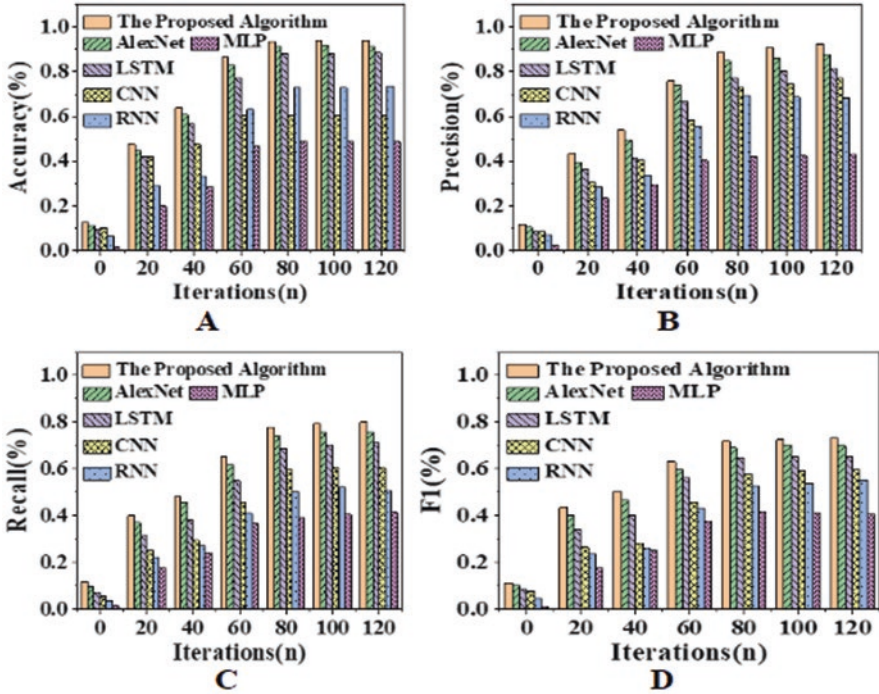
Fig. 10 Recognition results of different algorithms as the number of iterations increases (**a**. Accuracy; **b**. Precision; **c**. Recall; **d**. F1 value)
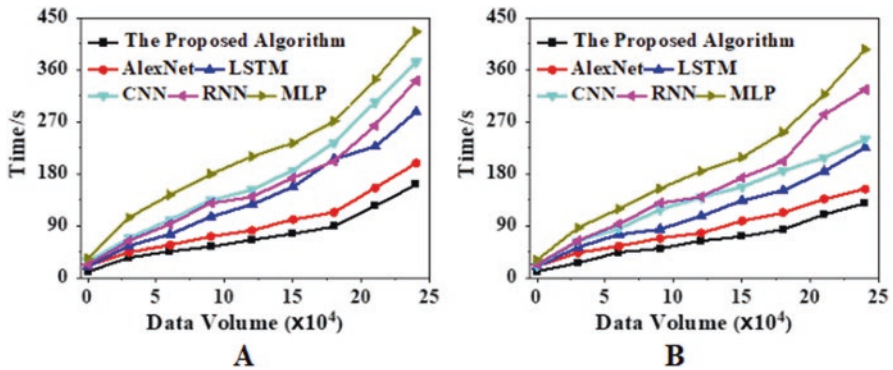

Fig. 11 Comparison of the running time of different algorithms under the training set and test set under different food image data amounts (**a**. training set; **b**. testing set)

5 Conclusion

Today, people pay increasing attention to balanced nutrition. Here, Cloud Communication technology, AlexNet, and DTs technology build a food DTs nutrition evaluation model. Through experiments, it is found that the throughput of the STM optimization algorithm proposed here is optimal, which can ensure the stability of the identification and processing of human nutrition and health data. At the same time, the improved AlexNet algorithm has a classification and recognition accuracy of 93.17% for food images, providing a direction fr the balanced intake and intellectual development of the follow-up dietary nutrition field. Still, the present work has some shortcomings. For example, the Body Area Network is cannot acquire all types of data on human nutrition and health. Therefore, future work will optimize the nutritional structure of users and study the impact of human social attributes on the correlation between food nutrition.

References[1]

1. Nasirahmadi, A., & Hensel, O. (2022). Toward the next generation of digitalization in agriculture based on digital twin paradigm. *Sensors, 22*(2), 498.
2. Hossain, M. S., Göbel, S., Xu, C., & El Saddik, A. (2020). IEEE access special section editorial: Mobile multimedia for healthcare. *IEEE Access, 8*, 153799–153803.
3. Chang, W. J., Chen, L. B., Lin, I. C., & Ou, Y. K. (2021). iBuffet: An AIoT-based intelligent calorie management system for eating buffet meals with calorie intake control. *IEEE Transactions on Consumer Electronics, 67*(4), 226–234.
4. Jiang, L., Qiu, B., Liu, X., Huang, C., & Lin, K. (2020). DeepFood: Food image analysis and dietary assessment via deep model. *IEEE Access, 8*, 47477–47489.
5. Park, J., Chung, H., & DeFranco, J. F. (2022). Multilayered diagnostics for smart cities. *Computer, 55*(2), 14–22.
6. Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., & Martynenko, A. (2020). IoT, big data and artificial intelligence in agriculture and food industry. *IEEE Internet of Things Journal*, 1–4.
7. Oteyo, I. N., Marra, M., Kimani, S., Meuter, W. D., & Boix, E. G. (2021). A survey on mobile applications for smart agriculture. *SN Computer Science, 2*(4), 1–16.
8. Alian, S., Li, J., & Pandey, V. (2018). A personalized recommendation system to support diabetes self-management for American Indians. *IEEE Access, 6*, 73041–73051.
9. Yunus, R., Arif, O., Afzal, H., Amjad, M. F., Abbas, H., Bokhari, H. N., et al. (2018). A framework to estimate the nutritional value of food in real time using deep learning techniques. *IEEE Access, 7*, 2643–2652.

---

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

10. Ghandar, A., Ahmed, A., Zulfiqar, S., Hua, Z., Hanai, M., & Theodoropoulos, G. (2021). A decision support system for urban agriculture using digital twin: A case study with aquaponics. *IEEE Access, 9,* 35691–35708.

11. Jia, W., Li, Y., Qu, R., Baranowski, T., Burke, L. E., Zhang, H., et al. (2019). Automatic food detection in egocentric images using artificial intelligence technology. *Public Health Nutrition, 22*(7), 1168–1179.

12. Liu, C., Liang, Y., Xue, Y., Qian, X., & Fu, J. (2020). Food and ingredient joint learning for fine-grained recognition. *IEEE Transactions on Circuits and Systems for Video Technology, 31*(6), 2480–2493.

13. Lu, Y., Stathopoulou, T., Vasiloglou, M. F., Christodoulidis, S., Stanga, Z., & Mougiakakou, S. (2020). An artificial intelligence-based system to assess nutrient intake for hospitalised patients. *IEEE Transactions on Multimedia, 23,* 1136–1147.

14. Rachakonda, L., Mohanty, S. P., & Kougianos, E. (2020). iLog: An intelligent device for automatic food intake monitoring and stress detection in the IoMT. *IEEE Transactions on Consumer Electronics, 66*(2), 115–124.

15. Arslan, B., Memis, S., Battinisonmez, E., & Batur, O. Z. (2021). Fine-grained food classification methods on the UEC Food-100 database. IEEE transactions on. *Artificial Intelligence,* 1–12.

16. Laamarti, F., Badawi, H. F., Ding, Y., Arafsha, F., Hafidh, B., & El Saddik, A. (2020). An ISO/IEEE 11073 standardized digital twin framework for health and well-being in smart cities. *IEEE Access, 8,* 105950–105961.

17. Wan, Z., Dong, Y., Yu, Z., Lv, H., & Lv, Z. (2021). Semi-supervised support vector machine for digital twins based brain image fusion. *Frontiers in Neuroscience, 802.*

18. Francisco, A., Mohammadi, N., & Taylor, J. E. (2020). Smart city digital twin–enabled energy management: Toward real-time urban building energy benchmarking. *Journal of Management in Engineering, 36*(2), 04019045.

19. Zhang, K., Cao, J., Maharjan, S., & Zhang, Y. (2021). Digital twin empowered content caching in social-aware vehicular edge networks. *IEEE Transactions on Computational Social Systems,* 1–13.

20. Van Der Werff, L., Fox, G., Masevic, I., Emeakaroha, V. C., Morrison, J. P., & Lynn, T. (2019). Building consumer trust in the cloud: An experimental analysis of the cloud trust label approach. *Journal of Cloud Computing, 8*(1), 1–17.

21. Bhatia, M., & Sood, S. K. (2019). Exploring temporal analytics in fog-cloud architecture for smart office healthcare. *Mobile Networks and Applications, 24*(4), 1392–1410.

22. Tao, Q., Ding, H., Wang, H., & Cui, X. (2021). Application research: Big data in food industry. *Food, 10*(9), 2203.

23. Arnaud, A., Marioni, M., Ortiz, M., Vogel, G., & Miguez, M. R. (2021). LoRaWAN ESL for food retail and logistics. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 11*(3), 493–502.

24. Kadam, P., Petkar, N., & Phansalkar, S. (2020). A systematic literature review with bibliometric meta-analysis of deep learning and 3D reconstruction methods in image based food volume estimation using scopus, web of science and IEEE database. *Library Philosophy & Practice,* 1–11.

25. Khanafer, M., & Shirmohammadi, S. (2020). Applied AI in instrumentation and measurement: The deep learning revolution. *IEEE Instrumentation & Measurement Magazine, 23*(6), 10–17.

26. Lo, F. P. W., Sun, Y., Qiu, J., & Lo, B. (2020). Image-based food classification and volume estimation for dietary assessment: A review. *IEEE Journal of Biomedical and Health Informatics, 24*(7), 1926–1939.

27. Chen, J., Zhu, B., Ngo, C. W., Chua, T. S., & Jiang, Y. G. (2020). A study of multi-task and region-wise deep learning for food ingredient recognition. *IEEE Transactions on Image Processing, 30,* 1514–1526.

28. Jiang, H., Starkman, J., Liu, M., & Huang, M. C. (2018). Food nutrition visualization on Google glass: Design tradeoff and field evaluation. *IEEE Consumer Electronics Magazine, 7*(3), 21–31.

29. Toledo, R. Y., Alzahrani, A. A., & Martinez, L. (2019). A food recommender system considering nutritional information and user preferences. *IEEE Access, 7*, 96695–96711.

30. Ozdemir, A., & Polat, K. (2020). Deep learning applications for hyperspectral imaging: A systematic review. *Journal of the Institute of Electronics and Computer, 2*(1), 39–56.

31. Siemon, M. S., Shihavuddin, A. S. M., & Ravn-Haren, G. (2021). Sequential transfer learning based on hierarchical clustering for improved performance in deep learning based food segmentation. *Scientific Reports, 11*(1), 1–14.

32. Iwendi, C., Khan, S., Anajemba, J. H., Bashir, A. K., & Noor, F. (2020). Realizing an efficient IoMT-assisted patient diet recommendation system through machine learning model. *IEEE Access, 8*, 28462–28474.

33. Lo, F. P. W., Sun, Y., Qiu, J., & Lo, B. P. (2019). Point2volume: A vision-based dietary assessment approach using view synthesis. *IEEE Transactions on Industrial Informatics, 16*(1), 577–586.

34. Ma, P., Lau, C. P., Yu, N., Li, A., Liu, P., Wang, Q., & Sheng, J. (2021). Image-based nutrient estimation for Chinese dishes using deep learning. *Food Research International, 147*, 110437.