

Research Article

Effect Analysis of Language Education Based on Computer and Network Sensors

Wang Dandan ¹, Wu Yufei ¹ and Zhu Yanwei²

¹School of Foreign Studies, Tangshan Normal University, Tangshan 063000, China

²School of Mathematics and Computational Sciences, Tangshan Normal University, Tangshan 063000, China

Correspondence should be addressed to Wang Dandan; 1212470926@st.usst.edu.cn

Received 8 September 2021; Accepted 20 October 2021; Published 17 November 2021

Academic Editor: Gengxin Sun

Copyright © 2021 Wang Dandan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Language talent education is an essential education in education, but today's language talent education needs to be improved because of the influence of teaching methods and other factors. This paper puts forward an idea of sensor-assisted education. The sensor is connected to the computer to help improve the language ability and information acquisition ability of the educated by means of network sensing and try to intelligently classify the relevant languages in language education, so as to reduce the time for the educated to process information, so as to realize the matching between language and reception. At the same time, the research also found that the sensors based on computers and networks have the function of intelligently strengthening the language, and the meaning of expression is simpler and more accurate. By studying the improvement of the effect of language education based on computer and network sensors, this paper provides a reference for the application of network sensors in the future.

1. Introduction

A sensor is a new breakthrough in the modernization of educational means. With the continuous updating of various instruments and technologies, people try to expand their perception range with its advantages. Sensors, like human "five senses," are responsible for information collection. In physical experiments, they can perfectly integrate sound, light, force, heat, etc. These perceived physical quantities with electricity and amplify, transmit, and store them in measurement. Applying the advantages of sensor technology in physics experiment teaching presents digital experimental means. The successful examples of physics experiment teaching can show the visualization of physical phenomena and laws. The key to success in the experiment is to let the students "see the phenomenon." Although physics is science based on experiments, it is impossible to demonstrate all the experiments in middle school laboratories. At this time, the advanced technical means provided by digital sensing information systems play a very important role. For example, the effect of interaction is a common phenomenon in life.

How to better understand the characteristics of interaction force from the law of physics is a teaching difficulty. Traditional experimental equipment can only measure the magnitude relationship of horizontal force with spring dynamometer, and the error of the result is relatively large. The advantage of digital sensors is shown at this time because it cannot provide students with valuable data for analysis. By introducing the sensor into the teaching of interaction force, students can clearly record the images of the two forces changing with time and the relationship between the two forces in the process of movement by pulling the force sensor on both sides in the experiment. With the organic combination of experimental technology and advanced experimental means, students not only see physical phenomena and summarize physical laws but also arouse their interest in physics, which is also the key factor of teaching.

The effect of language talent education pays more attention to the environment. Language teaching will also be related to learners' listening. A noisy environment will seriously affect the effect of language talent education. Not only is the language teaching of teachers easy to be affected but

also the information received by students is incomplete, which is not conducive to the communication between teachers and learners.

At present, intelligent sensor technology is applied in many aspects, and the characteristic of intelligent sensor technology is to retain information to the greatest extent. This feature makes the intelligent sensor be applied to the analysis of the effect of language talent education. Recently, the intelligent sensor has derived the function of strengthening information, which makes educators spend less energy on communication. With the combination of sensor and network, language talent education can also be applied to online teaching.

The first part is the introduction of the research background. The second part is the research on network sensors. The third part introduces the related algorithms based on computer and network sensors. The fourth part is experimental verification. The fifth part is the summary.

2. Related Work

The research of computer and network sensors has been very extensive, among which the most important is the transmission function of network sensors. S-mad introduces the idea of periodic sleep into wireless sensor network transmission for the first time. At the beginning of the protocol execution stage, nodes are divided into different virtual clusters and enter the sleep state according to different virtual clusters [1]. After the time synchronization phase, the node preempts the channel in the way of carrier listening, and the node that obtains the channel transmits data. However, because the node sleep time is determined by the virtual cluster, its sleep cycle is fixed. When the channel preemption is not successful, the node is still in the wake-up state, which increases the energy consumption of the network. Therefore, the T-MACt 22l protocol is improved. The node dynamically adjusts the length of wake-up time in the cycle according to the current network data volume, so as to reduce the idle listening time and avoid energy waste [2]. The above MAC protocols require nodes to perform synchronization. The synchronization mechanism generates not only unnecessary overhead but also periodic synchronization which will further increase the network delay. Therefore, the asynchronous low-power detection mechanism is proposed to solve the above problems. Its basic idea is that the node periodically wakes up and sends the detection frame, and the detection frame contains its own sleep plan. When a node receives the detection frame, it uses the obtained sleep plan to send data to the node, which will greatly reduce the waiting time of the sending node, so as to reduce energy consumption [3–5]. However, this method needs all nodes to send detection frames continuously, which increases the transmission cost of the whole network. At the same time, the detection frames sent by nodes will be received by multiple nodes. If these nodes wake up at the same time to send data, a large number of nodes wake up and seize the channel at the same time, resulting in low network channel utilization in the time domain. Therefore, the asynchronous low-power detection mechanism is only suitable for application scenarios with a

small amount of low-power data. In order to overcome the above problems, B-MAC proposes a low-power listening mechanism for the first time. Its basic idea is that when a node preempts the channel, it first compares the currently measured minimum received signal strength of the sender with all received signal strengths through exponential weighted average by evaluating the current channel quality. When the signal strength of the sender is large, it indicates that the channel quality is good [6]. Secondly, when channel competition occurs in the network, in addition to using the initial back-off algorithm, the node will evaluate the congestion state according to the number of currently competing channel nodes and select a specific node for transmission to reduce the degree of network congestion. Finally, when the node data transmission is completed, the receiving node returns ACK to increase the reliability of the network. However, the preamble sequence of B-MAC is too long, which increases the transmission delay. X-mac proposes to fragment the preamble sequence, and each fragment contains the destination address information. When the destination address of the preamble sequence fragment received by the receiving node is not its own, it will immediately enter the sleep state. The node sends the fragment preamble sequence at a certain time interval. If the receiving node matches the destination address, it immediately returns an ACK to inform the sending node. The sending node stops sending the check-in sequence and sends the data to the receiving node. In this way, the receiving node does not have to wake up frequently to listen to whether a node sends data. At the same time, the node adjusts its duty cycle according to the load in the current network to further reduce energy consumption and network delay [7]. However, the cost of channel detection is further increased because the transmitting node needs to continuously send preamble sequence fragments. Especially in low load wireless sensor networks, the transmission cost ratio is too small to meet the needs of applications. Therefore, b0x-mac replaces the preamble sequence fragment with the data packet actually sent by the node. When the receiving node receives the data packet, it returns an ACK to the sending node, and other nodes directly ignore the data packet and turn to sleep. Thus, the cost of channel competition is further reduced, and the overall transmission cost ratio of the system is improved [8].

Although the above MAC layer protocol reduces the transmission energy consumption of the node to a certain extent, the duty cycle of the node is usually dynamically adjusted according to a fixed value or according to the traffic situation. The duty cycle adjustment method of the MAC layer protocol is based on the way that the node can transmit as much as possible, and there is no further description of adjusting the duty cycle. Therefore, these methods can not only guarantee the end-to-end transmission delay of nodes but also increase the duty cycle of nodes, resulting in a waste of energy. Researchers began to pay attention to the data transmission mechanism with guaranteed delay and optimal energy consumption. It was first proposed in wireless networks. Assuming that the arrival information of all scheduling packets is known and has the same transmission deadline, researchers minimize the transmission energy

under the condition of meeting the transmission delay and give off-line algorithm and on-line approximation algorithm. Without assuming the network packet rate and channel statistical characteristics, the average data transmission energy is minimized under the constraint of ensuring the average queue waiting delay [9, 10]. However, wireless sensor networks are different from traditional wireless networks. Therefore, EDF designs an anticollision delay guaranteed transmission protocol according to the prediction of wireless sensor network traffic. Rap proposes a prioritization method based on packet deadline and distance to sink node. The end-to-end transmission delay of data packets is controlled by adjusting the node transmission rate [11, 12]. Speedpi ensures end-to-end transmission delay by implementing a unified transmission rate in the whole network [13]. El Khediri et al. proposed a data stream transmission control mechanism to ensure the transmission delay of the network [14]. Some scholars have proposed a lazy transmission scheduling mechanism to minimize data transmission energy in single-hop wireless sensor networks [15]. At the same time, many articles have studied how to minimize the transmission energy consumption of the node with the worst link quality in the transmission path by reasonably allocating the single-hop data transmission time under the condition of limited total transmission time based on dynamically adjusting the constellation point scale in multi-hop wireless sensor networks [16]. In recent years, some work has used graph theory to ensure transmission delay. For example, for minimizing the transmission energy consumption in the process of data collection in wireless sensor networks, the algorithm of constructing a data collection tree is based on a greedy algorithm to dynamically allocate channels [17]. Some scholars have proposed the transmission scheduling problem to minimize the cost of data collection and used the connected dominating set to solve the above problem [18]. Some scholars put forward the transmission proportion requirements based on data packets and dynamically adjust the value of α (ruler) through the greedy algorithm to obtain that all data packets meet their own delay requirements [19]. The node periodic sleep mechanism is still not considered in the delay control problem of wireless sensor networks. Therefore, the delay control protocol combined with the node periodic sleep mechanism has been studied. In this paper, some scholars use the sleep cycle of scheduling nodes to control the end-to-end transmission delay in low-power wireless sensor networks for the first time. It divides the problem into two subproblems. One is how to adjust the sleep cycle of nodes; the other is how to place sink nodes and the number of sink nodes [20].

Edge proposes an energy-efficient end-to-end delay guaranteed transmission protocol. The protocol is divided into two parts. Firstly, the minimum energy cost path or minimum delay path from the node to sink node is found through energy cost formula EEC and end-to-end delay formula EED. Then, the number of wake-up time slices in the node cycle is dynamically adjusted to make the single-hop delay meet the requirements [21]. Therefore, dutycon proposes a method to dynamically adjust the sleep time C in the node cycle to ensure that the end-to-end delay expecta-

tion of the node meets a certain limit [22]. Some researchers began to introduce the routing protocol mentioned in wireless mesh networks into wireless sensor networks [23]. On the whole, the research of network sensors is relatively diverse, but the focus is on transmission. Few sensors take into account the application of language teaching [24]. This paper is to explore the effect of language talent education based on computer and network sensors [25].

3. Method

This paper puts forward the idea of sensor-assisted teaching. The sensor is connected to the computer to help the educated improve their language ability and information acquisition ability through network perception and try to intelligently classify relevant languages in language education, so as to reduce the time for the educated to process information, so as to realize the matching between language and reception. By studying the improvement of language teaching effect based on computer and network sensors, this paper introduces the sensor from three aspects: (1) intelligent recognition algorithm of network sensor language, (2) sensing related technology based on computer and network sensors, and (3) introduction to multilayer sensing structure of network sensors.

3.1. Introduction of Intelligent Recognition Algorithm for Network Sensor Language. The principle of language intelligent recognition of network sensors is variance normalization. The parameter initialization method can well solve the problem of variance normalization. In this paper, a sensor decoder is constructed using a PIC microcontroller and assembly language. When configuring parameter initialization, the parameter initialization method is used to make the parameter initialization distribution obey Gaussian distribution, and the network model can be trained after actual operation.

As shown in formula (1), $x_i \in R^k$ represents the i th vocabulary with K dimension in the input statement. Of course, K is set to 200 in this paper, which means that the dimension of vocabulary in this paper is 200. After the words in the input statement are converted into word vectors, they are input to the input layer through superposition as an input matrix. The i value of x_i can be defined according to the actual situation. In this paper, this value is set to 50, which means that the input statement is allowed to include up to 50 Chinese characters:

$$x_{(1:n)} = x_1 \wedge x_2 \wedge \dots \wedge x_n. \quad (1)$$

As shown in equation (2), two parallel convolution calculations require a filter $y \in R^k$. In this way, when applied to a window with j words, a new feature will be generated, which can be calculated by

$$a_i = f(y \cdot x_{i:i+j-1} + b). \quad (2)$$

In equation (2), $b \in R$ is the preset offset value and f is the nonlinear activation function. In the sigmoid type

function and ReLU function, this paper selects the ReLU function as the activation function. In this way, for the input statement, after being processed by the convolution kernel filter, each window will get a feature map, as shown in

$$c = [c_1, c_2, \dots, c_{n-h+1}]. \quad (3)$$

It can be seen from equation (4) that $c \in R^{n-h+1}$; after that, this paper continues to maximize the pool of the obtained features, and $c' = \max(c)$ can be obtained as a new feature of the convolution kernel filter. The purpose of this is to facilitate the model to process input statements with different lengths. After the above steps, the input statement can be extracted into a group of features after convolution processing, and the extracted features become the basis for the final completion of intention recognition. In this process, the size of convolution kernels determines the window size. Generally speaking, small convolution kernels are selected, because compared with large convolution kernels, small convolution kernels can increase the complexity of the model and help to improve the accuracy of model training. Second, it can improve the network capacity and mine more hidden information of the input data. Third, it can reduce the number of convolution parameters. A cyclic neural network can be used in the sensor. The cyclic neural network uses the nonlinear function $f(\cdot)$ to convert the input statement sequence $x = \{x_1, x_2, \dots, x_N\}$ into a hidden state output $h = \{h_1, h_2, \dots, h_N\}$, where n is the sequence length of the input statement, and its network calculation unit is as shown in

$$h_t = f(x_t, h_{t-1}) = f(W_x x_t + W_h h_{t-1}). \quad (4)$$

The public Zi after heat treatment and the input query content instruction statement Qi are given, the public Zi is embedded in the matrix Zu, and the query content instruction statement Qi is embedded in the matrix Zq, so the public Zi and the statement Qi can be converted into a continuous hidden vector z_1^i and q_1^i . Finally, the coding is as follows:

$$h_{g,z}^i = \text{LSTM}(w_{g,0}^i, z_1^i), \quad h_{g,0}^i = 0, \quad (5)$$

$$h_{g,q}^i = \text{LSTM}(w_{g,2}^i, q_1^i), \quad h_g^i = h_{g,q}^i. \quad (6)$$

Formula (7) belongs to the further integration of formulas (5) and (6), and h_g^i is the output result obtained by the short-term and long-term memory network after processing the (z^i, q^i) sequence. On the basis that it can represent the user's query intention through the (z^i, q^i) sequence, the probability distribution of each alternative answer in the knowledge base to the intention can be obtained:

$$p^i = \text{softmax}(Z_g \cdot h_g^i + b_g). \quad (7)$$

In equation (8), Z_g represents the linear transformation

matrix and b_g represents the corresponding offset, which is mainly reflected in the matrix embedding of the input query statement which remains unchanged, but the invisible output h_g of the last layer is linearly transformed, and the results of the linear transformation are the initial input of the long-term and short-term neural network in this section:

$$q' = Z_g \cdot h_g + b_g. \quad (8)$$

In equation (8), Z_g is a linear transformation matrix and q' after linear change is more accurate than the original CC. In the training process of the long short neural network model, when using back propagation to calculate the gradient, h_g receives the gradient information from the global network for updating.

Equation (9) represents the calculation formula after the user input sentence is converted into a sentence vector. The sentence is represented by s , and W and V represent the words and word vectors that make up the sentence, respectively:

$$\vec{S} = \sum_{w \in S} V_w. \quad (9)$$

The similarity between the question sentences in the question answering database and the user input sentences can be calculated according to

$$\text{sim}(S_1, S_2) = \frac{\vec{S}_1 \cdot \vec{S}_2}{\|\vec{S}_1\| \cdot \|\vec{S}_2\|}. \quad (10)$$

3.2. Introduction of Sensing Related Technology Based on Computer and Network Sensors. Network sensing technology mainly involves the PMSM method, which is one of the common methods of network sensors. PMSM itself is a highly nonlinear structure, and the stator and rotor interact with each other, so the electromagnetic environment inside the motor is very complex. Since the magnetic circuit may have large current during operation, resulting in motor saturation, the sensing conduction equation of PMSM is

$$\begin{cases} u_a = R_s \cdot i_a + \frac{d}{dt} \psi_a, \\ u_b = R_s \cdot i_b + \frac{d}{dt} \psi_b, \\ u_c = R_s \cdot i_c + \frac{d}{dt} \psi_c, \end{cases} \quad (11)$$

where u_a, u_b, u_c is the phase voltage of three-phase winding, i_a, i_b, i_c is the current of three-phase winding, R_s is the equivalent resistance of three-phase winding, and ψ_a, ψ_b, ψ_c is the corresponding flux linkage of three-phase winding.

The stator flux linkage equation is

$$\begin{bmatrix} \psi_a \\ \psi_b \\ \psi_c \end{bmatrix} = \begin{bmatrix} L_{AA} & M_{AB} & M_{AC} \\ M_{AB} & L_{BB} & M_{BC} \\ M_{CA} & M_{CB} & L_{CC} \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} + \psi_f \begin{bmatrix} \cos \theta \\ \cos \left(\theta - \frac{2\pi}{3} \right) \\ \cos \left(\theta + \frac{2\pi}{3} \right) \end{bmatrix}, \quad (12)$$

where L_{AA}, L_{BB}, L_{CC} is the self-inductance of three-phase flux linkage, $M_{AB} = M_{BA}, M_{BC} = M_{CB}, M_{AC} = M_{CA}$ is mutual inductance between two-phase flux chains, ψ_f is the rotor flux linkage of the motor, and θ is the rotating electrical angle.

The electromagnetic torque equation is

$$T_c = -p_n \psi_f \left[i_a \sin \theta + i_b \sin \left(\theta - \frac{2}{3}\pi \right) + i_c \sin \left(\theta + \frac{2}{3}\pi \right) \right], \quad (13)$$

where p_n is the polar logarithm.

The motion equation of the network sensing factor is

$$J \frac{d\omega_m}{dt} = T_e - T_L - B\omega_m, \quad (14)$$

where ω_m is the speed of the network motor, J is the moment of inertia, B is the damping coefficient, and T_L is the load torque.

It can be seen that although the physical meaning of the mathematical model in the three-phase static coordinate system is easy to understand, it is not conducive to the application of the motor control algorithm due to the coupling of motor variables. Similarly, convert the ABC coordinate system to $\alpha\text{-}\beta$. The process of the coordinate system is called inverse Clarke transformation, and the transformation formula is as follows:

$$\begin{bmatrix} f_A & f_B & f_C \end{bmatrix}^T = T_{2S/3S} \begin{bmatrix} f_a & f_\beta & f_0 \end{bmatrix}^T. \quad (15)$$

3.3. Introduction of Multilayer Sensing Structure in Network Sensor. FCN is the most basic multilayer perceptron structure in sensor networks. First proposed a regression model based on FCN. The model uses FCN to reconstruct speech envelopes from EEG signals in the way of sample-wise prediction. Its network structure is shown in Figure 1. Considering that the system processing input speech into neural response is a causal system, and the input of stimulation reconstruction task is EEG signal and the output is speech time domain envelope, the reconstructed speech envelope at time 0 is only related to the neural response after time 0 + 1, but not to the neural response before time 0. Therefore, in order to predict the speech envelope value at time 0, FCN takes the EEG signal in a period of time after time 0 as the input. In the work of de Tailrez et al., the observation length of FCN is 27 sampling points, corresponding to 420 ms. After FCN expands all channels of EEG signals within 420 ms through one dimension (flatten) and activates two-

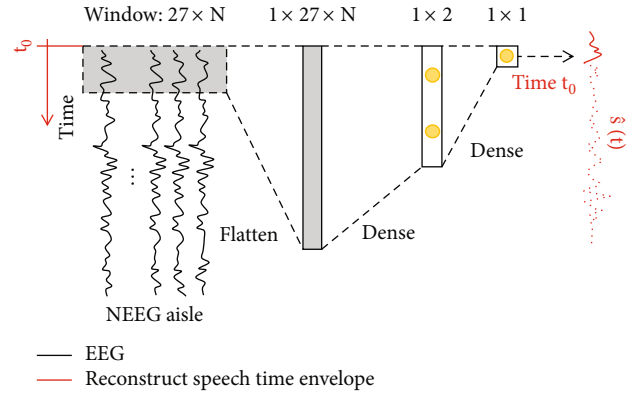


FIGURE 1: Flow chart of sensor's own sensing structure.

layer FCN (number of neurons: 2 and 1), the predicted value of the speech envelope at time 0 is obtained. When FCN slides continuously in the time dimension of EEG, the complete speech time domain envelope can be predicted (). Then, as in the second step of the linear decoding algorithm, the auditory attention object is determined by correlation analysis.

The authors conducted AAD experiments on data sets with only audio stimulation and spatial separation of sound sources (-45° and 45°). The results show that when the decoding window length is 10 s, compared with the linear decoding algorithm, the decoding algorithm based on the FCN regression model achieves higher accuracy (the accuracy is improved by about 20%). In addition, there are three findings in the experiment: (1) The accuracy of AAD using correlation coefficient loss function is higher than that using MSE loss function. This may be because the amplitude scale of the speech time domain envelope output by the network may not be fixed, the correlation value is independent of the signal amplitude, and the MSE value is related to the signal amplitude. Therefore, the loss function based on the correlation coefficient is better at capturing the consistency between the change trend of network output and the real speech envelope. (2) For the DNN decoder, the AAD accuracy when using broadband EEG is significantly higher than that when using narrowband EEG, while there is no significant difference in AAD accuracy when using two kinds of bandwidth EEG for the linear decoder. This result shows that compared with the linear decoder, the DNN decoder can use more information in the EEG signal and has stronger ability to describe the system. (3) After visualizing the weight of FCN, it is found that the neurons corresponding to the electrode channels of the temporal lobe (corresponding to the auditory cortex) have higher weight, which is similar to the temporal and spatial distribution patterns of the linear decoder. This result shows that the FCN regression model has certain interpretability. Ciccarelli et al. also used the above reconstruction model based on FCN regression. On the experimental data with only audio stimulation and single-channel presentation conditions, they also found that when the decoding window length is 10 s, the decoding accuracy of the FCN model (64%) is higher than that of the linear decoding algorithm (59%).

In addition to the above two studies, Nogueira and Dolhopiatenko proposed another decoding algorithm based on the FCN regression model. The difference from the previous model is that the causality between EEG and speech time domain envelope is realized by delaying EEG. FCN only observes the EEG signal at time $0 + \tau$ and then maps it to the predicted value of the speech envelope at time 0. Its network structure is shown in Figure 2. In the training, the author also uses the dropout strategy and the loss function based on the correlation coefficient. In addition, the operation of batch normalization (BN) is used to speed up the network convergence.

Compared with FCN, CNN is considered to have better spatial feature extraction ability. Nogueira and Dolhopiatenko proposed a decoding algorithm based on the CNN regression model, which uses the CNN structure to reconstruct the speech time domain envelope from EEG. Its network structure is shown in Figure 3. Similar to the FCN model proposed by the author, the causality between EEG and speech time domain envelope is realized by delaying EEG. In order to predict the speech envelope value at time 0, CNN takes the EEG signal in a period near time $0 + \tau$ as the input, which corresponds to the size of the convolution kernel in the time dimension. CNN used a size of $16 \times$ one-dimensional convolutions of N , corresponding to 250 ms. In the research using the linear decoding algorithm, it is found that the receptive field of this length can cover the most important delay range for speech envelope reconstruction; our results in this paper are also consistent with this conclusion. The EEG signal passes through single-layer one-dimensional CNN (number of convolution cores: 5, size: $16x$). And two-layer FCN (number of neurons: 5 and 1) are activated to obtain the predicted value of the speech envelope at the time. With the convolution kernel sliding continuously in the time dimension of EEG, the model can predict the complete speech time domain envelope, so it can also be regarded as a point-by-point prediction algorithm. Then, as in the second step of the linear decoding algorithm, the object is noticed through correlation analysis. In the model training, the author also uses the dropout and batch normalization strategy, as well as the loss function based on the correlation coefficient. The experimental results on the same data set show that when the decoding window length is 10 s, the decoding algorithm based on the CNN regression model only achieves about 50% accuracy, which is equivalent to the opportunity level. This shows that the model does not learn the mapping relationship between EEG and speech time domain envelope, which may be due to the unreasonable implementation of the causal system.

4. Results and Discussion

4.1. Results and Discussion of the Decoding of Language by Sensors. In fact, there are 6 stimulation conditions (3 spatial separations \times 2 auditory attention switching intervals). There is little difference between the decoders trained under the two auditory attention conversion intervals, so the AAD task uses the decoder jointly trained by all the data in the training set. At 90° , the masking release effect caused by

spatial separation is the strongest, so its reconstruction accuracy is the highest. However, for 30° condition, although it has weaker masking release effect than 60° , the reconstruction accuracy is higher. This may be because the subjects have less language behavior, less EMG artifacts, and less interference with AAD at 30° , which offsets the disadvantage of low masking release effect to a certain extent. Although the single factor RM-ANOVA shows that the main effect of the conversion interval condition is not significant ($f_{1215} = 2.139$, $P = 0.897$, $2 = 0.009$), we can still observe that the reconstruction accuracy (ATT = 0.0763) under the conversion interval condition of 30 s is higher than that of 15 s (ATT = 0.0729), which is consistent with the trend that the AAD accuracy decreases with the decrease of the decoding window length, as shown in Figure 4.

Then, we use the above decoder to reconstruct the speech envelope frame by frame (5 s frame length, 1 s frame shift) for each trial and determine whether the AAD result of each frame is correct. Figure 5 shows the continuous decoding results after intertrial average under six stimulation conditions. The vertical axis in the figure represents the probability of being determined as speaker 1. The closer the curve is to 1, the greater the probability that the result of AAD is speaker 1. The closer the curve is to 0, the greater the probability that the result of AAD is speaker 2. It can be seen that the change trend of AAD results after average between trials over time is roughly consistent with the conversion setting of the attention object in the experiment (marked by vertical black dotted lines), but there are also two problems. Firstly, the accuracy of the algorithm is low; that is, the curve is close to the position with the longitudinal axis of 0.5; secondly, the algorithm has a delay of about 5–10 s; that is, the algorithm can stably judge that the attention object of the speaker has changed after about 5–10 s. These experimental results show that in the auditory selective attention task with language matching conditions and attention conversion, the EEG-based AAD method can be used for language attention object detection, but the accuracy is low and has obvious algorithm delay (about 5–10 s). This is mainly because the “linear system” assumption based on the linear decoding algorithm is too simplified, and the modeling ability of the mapping relationship between EEG signal and speech time domain envelope is limited. Nonlinear modeling can be used to solve this problem.

4.2. Application of Classifier in Sensor. As mentioned above, since the amount of data with a label of 0° in the experimental data is much higher than that of the other 6 types of labels, the classification accuracy index will be biased towards the results of this kind of data. Therefore, we show the results of FCN and CNN classifiers under various input conditions through the classification result confusion matrix, as shown in Figure 6. It can be seen that under all conditions, the classification accuracy of 0° is the highest and significantly higher than other rotation angles, because the data of each sensor is relatively stable under this condition. The confusion mainly occurs in the same rotation direction and between adjacent rotation angles, especially

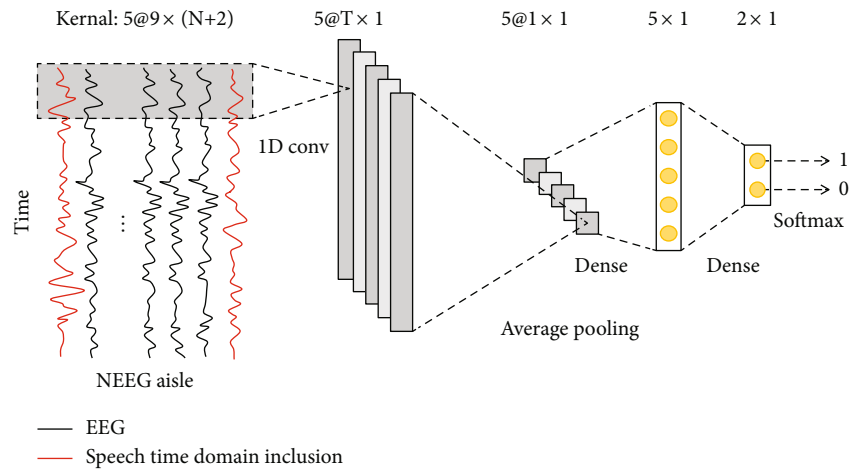


FIGURE 2: Flow chart of enhanced perception structure in network sensors.

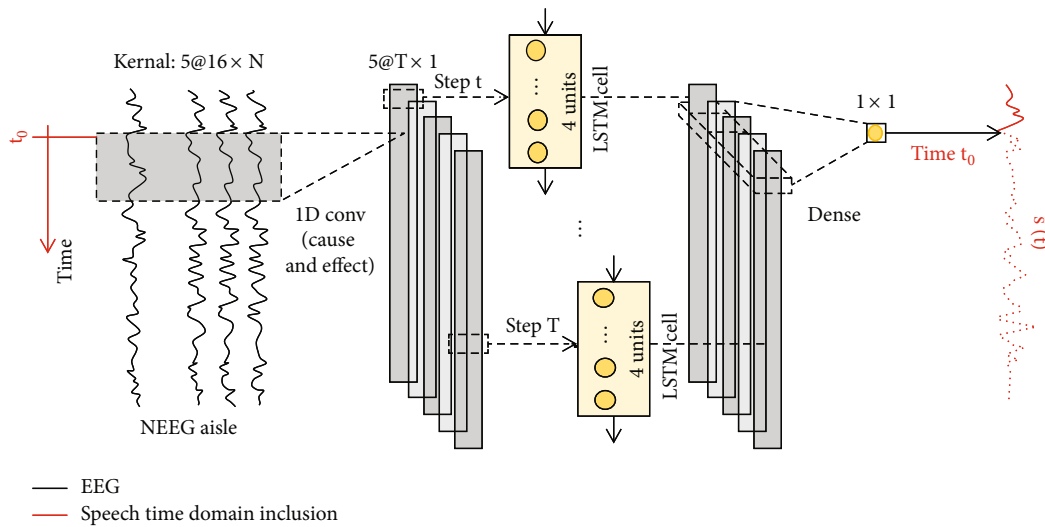


FIGURE 3: Flow chart of returning to the enhanced perception structure in network sensors.

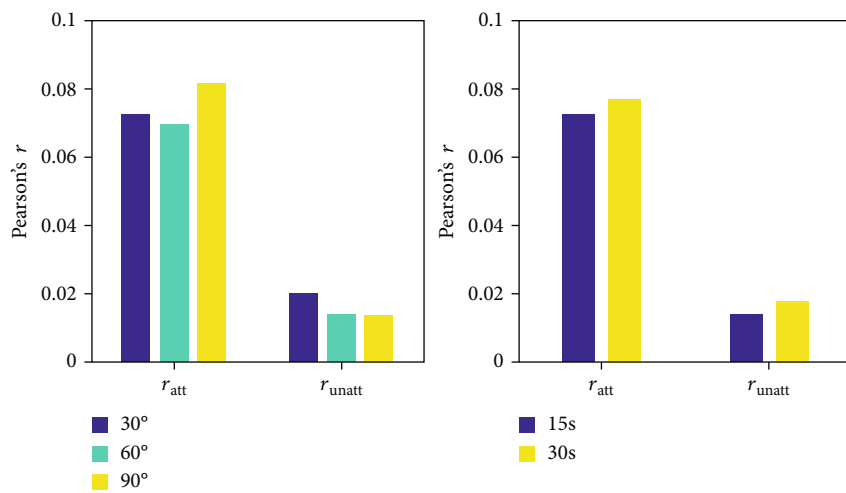


FIGURE 4: Comparison of reconstruction accuracy of network sensors under different language stimulation conditions.

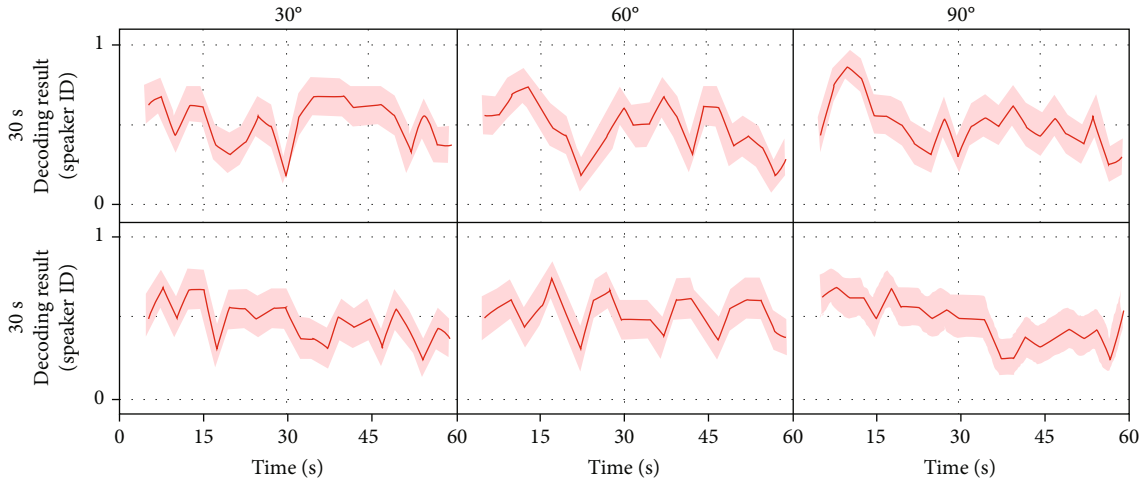


FIGURE 5: Continuous decoding results within a trial of the linear decoding algorithm in the network sensor.

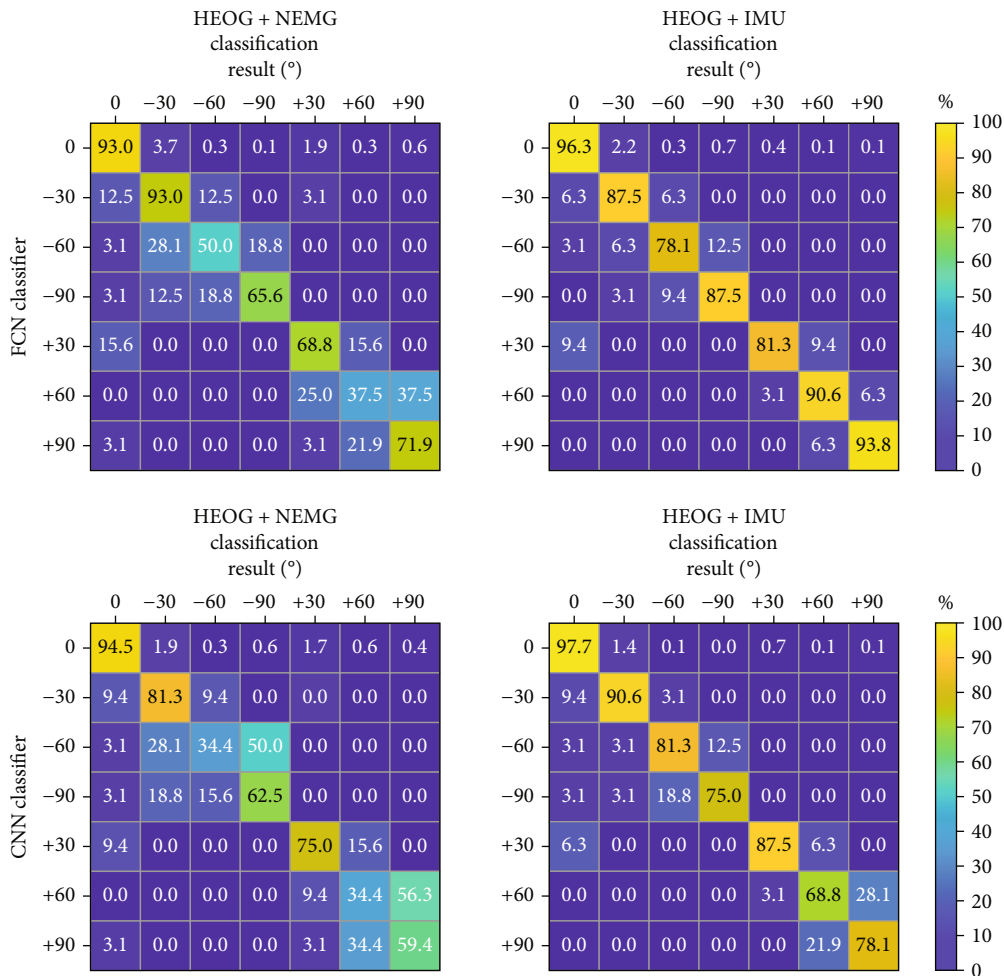


FIGURE 6: Confusion matrix of classification results of FCN and CNN classifiers in the sensor (%).

between $+60^\circ$ and $+90^\circ$ and -60° and -90° . For the two classifiers, the results of bivariate heog and NEMG are not as good as bivariate heog and IMU, which also shows that the effect of using NEMG to estimate head rotation is not as good as IMU.

We further calculated the continuous line of sight rotation estimation results for each trial in the test set. It should be noted that after the listener turns the line of sight, the sensor signal of several consecutive frames still fluctuates. Therefore, note that there is a certain oscillation in the

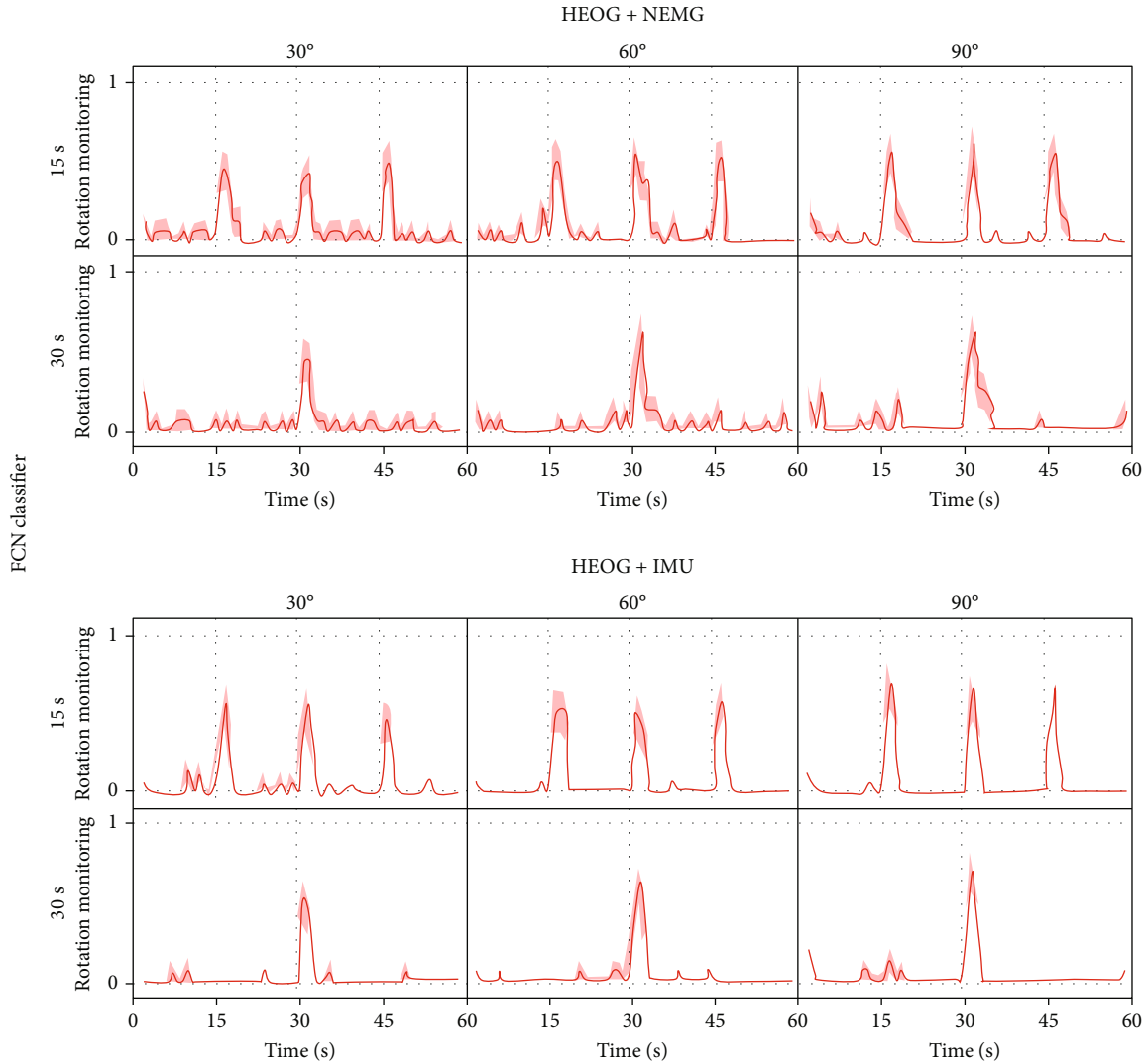


FIGURE 7: Estimation results of the rotation angle of the continuous line of sight within the trial.

estimation results of several consecutive frames after conversion. For example, as shown in Figure 6, after an impulse of a certain polarity occurs, the heog signal will gradually recover to the initial potential, and the latter process can also be regarded as a reverse impulse with a small amplitude. When we extract the signal waveform frame by frame, the forward and reverse impulse process will be divided into adjacent frames, so the algorithm may mistakenly classify the reverse impulse process as reverse line of sight rotation. Therefore, when calculating the estimation results of continuous line of sight orientation in the trial, when the estimation results of two consecutive frames are not 0° , we sharpen them. Set the smaller absolute value in the estimation result to 0° . Figure 7 shows the average continuous estimation results between trials under six stimulation conditions. Taking FCN classifier and bivariate heog and NEMG inputs as examples, it can be seen that the classifier can judge that the line of sight direction has rotated within 2 s after the attention conversion time (marked by vertical black dotted lines), and the estimated value output by the classifier increases with the increase of the spatial separation angle

set in the experiment. In addition, it can also be observed that at the nonattention conversion time (i.e., the rotation angle should be 0°), the classifier may also output a rotation amount of non- 0° (corresponding to the first row of the confusion matrix), which may be caused by the fluctuation in the sensor signal. The results using the CNN classifier show similar patterns. Under univariate (results not shown) and bivariate input conditions, the performance of the CNN classifier is slightly better than that of the FCN classifier, which is reflected in more accurate estimation of rotation angle, less misjudgment of nonrotation time, and smaller error interval. This is because the convolution kernel in the CNN structure can observe the signal waveform in a certain time window at one time and integrate the information in the window, which has stronger feature extraction ability and antinoise ability.

In general, these results show that under the condition of audiovisual matching, the Los rotation estimation based on heog and NEMG can more accurately reflect the information related to auditory attention conversion, such as rotation time and rotation angle. Considering that the AAD

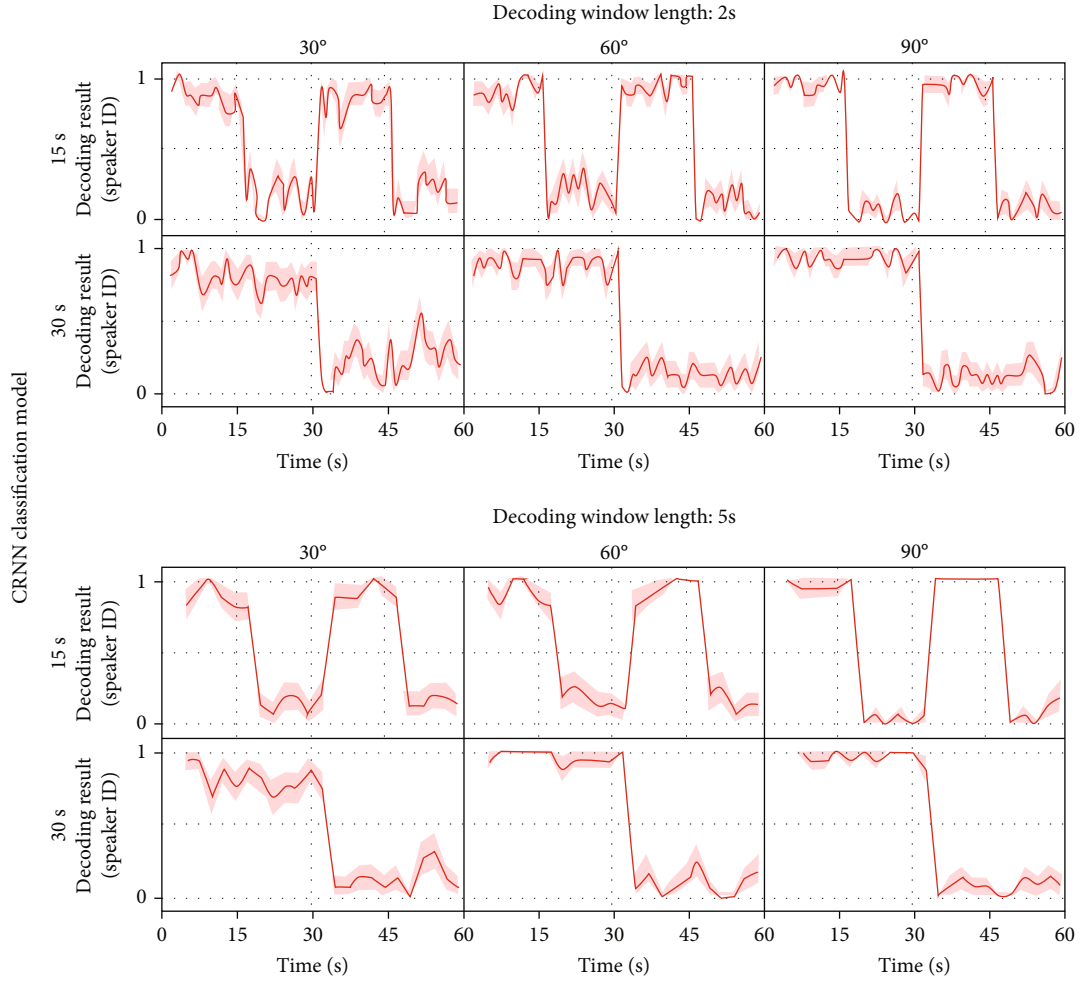


FIGURE 8: Continuous decoding results of the CRNN model in the sensor under different stimulus conditions.

task does not require accurate line of sight rotation angle information, we can reduce the line of sight rotation angle estimation to line of sight rotation detection; that is, the rotation label output by the classifier (Class 7: 0° , $\pm 30^\circ$, $\pm 60^\circ$, and $\pm 90^\circ$) is changed to rotation label (Class 2: rotation, no rotation), so that while meeting the requirements of AAD task, the advantage of low detection delay (2 s) is also retained. Based on this change, this section will continue to explore the feasibility of the AAD task based on line of sight rotation detection.

Based on the results in Figure 7, after remapping the output label of the classifier into rotation ($\pm 30^\circ$, $\pm 60^\circ$, and $\pm 90^\circ$) and nonrotation (corresponding to 0°), the experimental results of the AAD task can be obtained, and the confusion matrix is shown in Figure 6. It can be seen that when the bivariate heog and NEMG are input, FCN is easier to misjudge the rotation condition as no rotation than the CNN classifier (FCN: 6.3%, CNN: 3.1%). This missing alarm means that the AAD algorithm fails to guide the calculation of AAD, which will affect the language object detection. On the contrary, a false alarm has less impact, because the AAD algorithm can correct it. In order to further quantify the performance of each algorithm in the AAD task under various input conditions, we calculate three indicators: F1 value,

missed alarm rate, and false alarm rate according to the confusion matrix. F1 value is the harmonic average of accuracy rate (the proportion of samples divided into positive examples) and recall rate (the proportion of samples divided into positive examples). The model can be comprehensively evaluated. The missed alarm rate is the proportion of the missed positive cases in all the positive cases. The false alarm rate is the proportion of samples judged as positive cases, which are actually negative cases. It can be seen that the CNN classifier has better performance. It has a higher F1 value and lower false alarm rate than the FCN classifier. It is proved that in this task, the CNN network has stronger feature extraction ability and is more robust to noise interference. In addition, although the results when using bivariate heog and NEMG are still worse than bivariate heog and IMU on the whole, the gap is not large; especially when using the CNN classifier, the alarm leakage rate (both 3.1%) has almost no difference. The results show that the proposed AAD algorithm based on heog and NEMG is feasible.

The comparison and fusion of auditory attention decoding and auditory attention conversion detection methods combines the advantages and disadvantages of the AAD method based on the auditory selective attention neural mechanism (measured by EEG) and based on visual

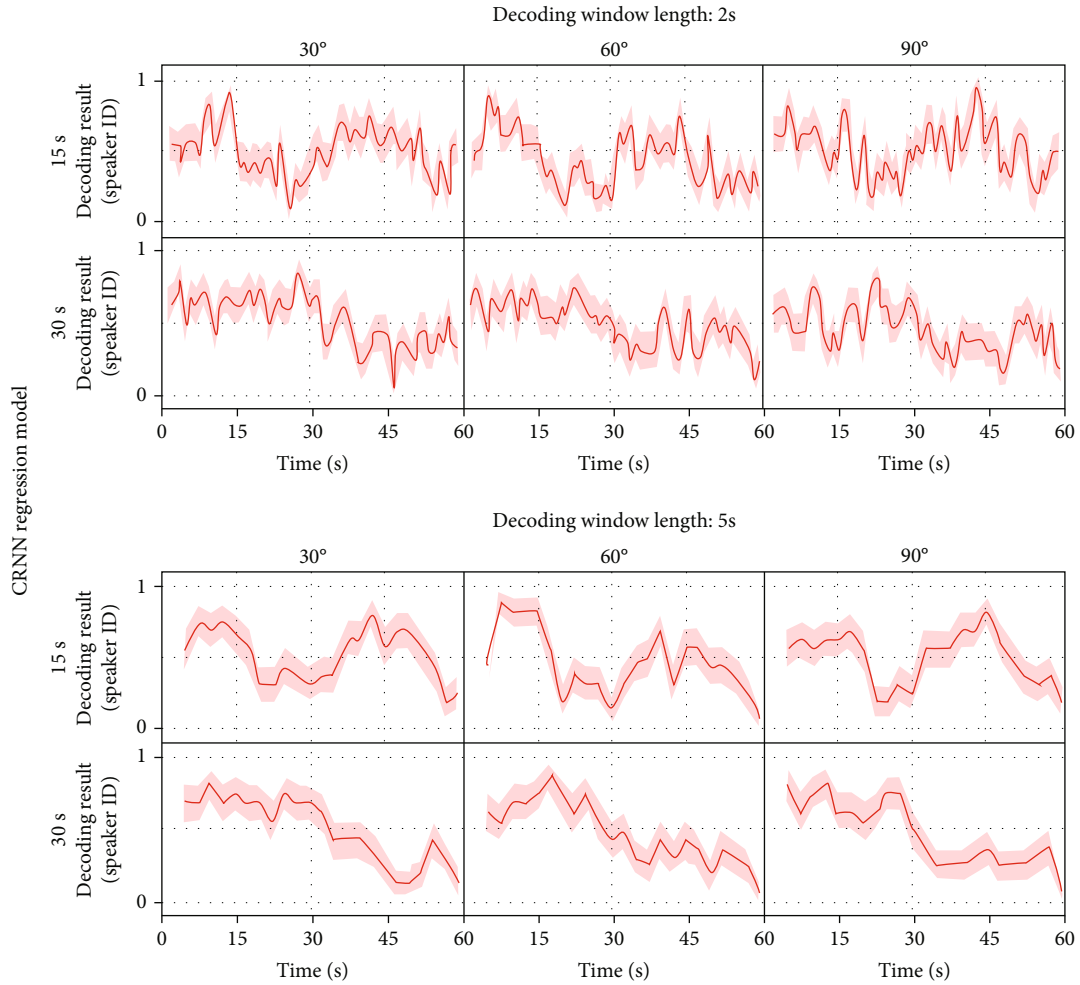


FIGURE 9: Continuous decoding results of the CRNN model in the enhanced sensor.

behavior (measured by heog and NEMG) and then puts forward the fusion strategy of the two kinds of methods. However, in the third chapter, due to the low accuracy of the linear decoding algorithm, the feasibility of this strategy is poor. Therefore, several AAD methods based on the DNN model are proposed to significantly improve the correct rate of understanding code, which makes the above fusion strategy possible. Therefore, in this section, we will recompare the performance of AAD methods and explore the feasibility of the fusion strategy proposed in this paper.

4.3. Research on Enhancement of Sensing Related Technology in Sensor Networks. Similar to Figure 6, we use the DNN decoder to perform AAD calculation frame by frame for each trial in the test set, where the frame length (i.e., decoding window length) is 2 s or 5 s and the frame shift is 50%. Considering that the classification and regression model based on the CRNN structure proposed in this paper is superior to other models in most cases, we will take these two models as examples for calculation. Figure 8 shows the continuous decoding results after intertrial average under six stimulation conditions. The vertical axis in the figure represents the probability of being determined as speaker 1. The closer the curve is to 1, the greater the probability that the

result of AAD is speaker 1. The closer the curve is to 0, the greater the probability that the result of AAD is speaker 2. It can be seen that compared with the linear decoding algorithm, the decoding performance of the two CRNN models is significantly improved, and the change trend of AAD results after average between trials is more consistent with the conversion setting of the object of attention in the experiment (marked by vertical black dotted lines). In addition, the CRNN classification model has better performance than the CRNN regression model, which is reflected in higher accuracy and smaller error interval. The decoding performance under the condition of 90° spatial separation is better than the other two angle conditions.

For the CRNN classification model, when the decoding window length of 2 s or 5 s is used, the model can detect the auditory attention conversion in time under each stimulus condition, and the algorithm delay is the decoding window length. For the CRNN regression model, when the decoding window length of 2 s or 5 s is used, the model can also detect auditory attention conversion in time. However, due to its large error interval, the model needs longer time to achieve stable results. In conclusion, compared with the linear decoding algorithm, the AAD algorithm based on the CRNN model has higher accuracy and lower algorithm

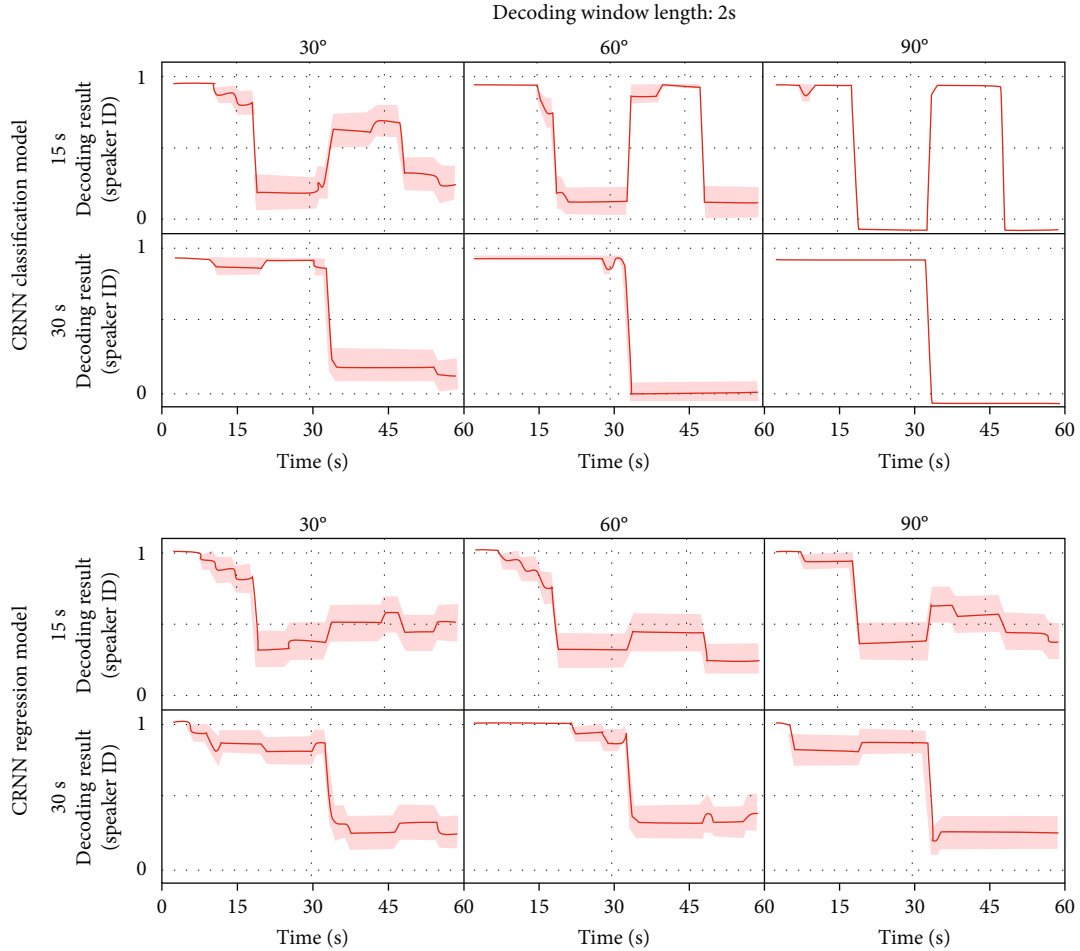


FIGURE 10: Continuous decoding results of the CRNN model in the sensor under the fusion strategy.

delay (about 2–5 s). Based on the above results, this paper summarizes the advantages and disadvantages of AAD, and the results are shown in Figure 9. Among them, the decoding window length is set to 2 s and the correction window length is set to 5 s. The AAD algorithm uses CNN classifier and bivariate heog and NEMG input. For the CRNN classification model, under the condition of 15 s attention conversion interval, the accuracy of auditory attention object detection after using the fusion strategy is reduced, and the accuracy of the initial part of the trial is higher than that of the second half, mainly because more attention conversion times will bring more cumulative errors (see Figure 10).

In addition, the decrease of this accuracy rate is greater than 60° and 90° under the condition of 30° spatial separation, because the accuracy of AAD itself is low under the condition of 30° and the algorithm cannot effectively correct the cumulative error. On the contrary, under the condition of 90° spatial separation, the AAD algorithm effectively corrects the cumulative error. For the CRNN regression model, under various stimulation conditions, the results after using the fusion strategy have been improved, and the model output can achieve stable results faster, reflecting the effectiveness of the fusion strategy. We can also observe that the result under the condition of 30 s attention conversion interval is significantly better than 15 s, which also indicates that

the model has the problem of cumulative error. The error correction ability of the CRNN regression model is weaker than that of the classification model, which is mainly due to its low decoding accuracy. In addition, we statistically find that the fusion strategy significantly reduces the amount of computation of the AAD algorithm (about 70%), and the results show that the effect of language talent education based on computer and network sensors is better.

5. Conclusion

The key of network sensors is information request and enhancement. Because the sensor method is based on cortical phase-locked response, and the establishment of phase-locked response takes a certain time, there is a high delay (e.g., 5–10 s) in the detection of attention conversion using the sensor method. Under the condition of large interference of environmental factors, the sensor can play a better role in information capture. The experiment explores the strengthening effect of network sensors on language talent education. In addition, the accuracy of using the CNN classifier in the sensor is higher than that of other classifiers, indicating that the CNN model in the sensor has stronger ability to extract features from multichannel input signals. These results preliminarily verify the effectiveness of this method.

Simultaneously interpreting the influence of sensor signal types and classification algorithms on the transmission modes of different sensors, based on this, this paper proposes a fusion strategy of various sensor conduction modes, which initially verifies the feasibility of the strategy in language talent education. However, the research in this paper only preliminarily verifies the effectiveness of this method. Because there is no more in-depth data verification on the ability of the CNN model to extract features from multi-channel input signals, it needs to be further discussed in future research.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

Acknowledgments

The study was funded by the following: (1) Research and Practice Project of English Teaching Reform in Higher Education of Hebei Province in 2020, Research on the Construction of Practical Curriculum System for Business English Major under the Background of Integration between Industry and Education (Project No.: 2020YYJG059), and (2) Project supported by Scientific Research Fund of Tangshan Normal University, Research on the Construction Path of English Cloud Platform of Jidong Culture under the Background of New Infrastructure Construction (Project No.: 2021B12).

References

- [1] D. Kandris, C. Nakas, D. Vomvas, and G. Koulouras, "Applications of wireless sensor networks: an up-to-date survey," *Applied System Innovation*, vol. 3, no. 1, p. 14, 2020.
- [2] A. Shahraki, A. Taherkordi, Ø. Haugen, and F. Eliassen, "Clustering objectives in wireless sensor networks: a survey and research direction analysis," *Computer Networks*, vol. 180, article 107376, 2020.
- [3] M. Numan, F. Subhan, W. Z. Khan et al., "A systematic review on clone node detection in static wireless sensor networks," *IEEE Access*, vol. 8, pp. 65450–65461, 2020.
- [4] D. K. Sah and T. Amgoth, "Renewable energy harvesting schemes in wireless sensor networks: a survey," *Information Fusion*, vol. 63, pp. 223–247, 2020.
- [5] H. Patel, D. Singh Rajput, G. Thippa Reddy, C. Iwendi, A. Kashif Bashir, and O. Jo, "A review on classification of imbalanced data for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 16, no. 4, Article ID 155014772091640, 2020.
- [6] S. Qu, L. Zhao, and Z. Xiong, "Cross-layer congestion control of wireless sensor networks based on fuzzy sliding mode control," *Neural Computing and Applications*, vol. 32, no. 17, pp. 13505–13520, 2020.
- [7] X. Liu, P. Lin, T. Liu, T. Wang, A. Liu, and W. Xu, "Objective-variable tour planning for mobile data collection in partitioned sensor networks," *IEEE Transactions on Mobile Computing*, vol. 1, pp. 1–1, 2020.
- [8] S. N. Mohanty, E. L. Lydia, M. Elhoseny, M. M. G. al Otaibi, and K. Shankar, "Deep learning with LSTM based distributed data mining model for energy efficient wireless sensor networks," *Physical Communication*, vol. 40, article 101097, 2020.
- [9] X. Liu, M. S. Obaidat, C. Lin, T. Wang, and A. Liu, "Movement-based solutions to energy limitation in wireless sensor networks: state of the art and future trends," *IEEE Network*, vol. 35, no. 2, pp. 188–193, 2021.
- [10] P. K. Singh, B. K. Bhargava, M. Paprzycki, N. C. Kaushal, and W.-C. Hong, *Handbook of Wireless Sensor Networks: Issues and Challenges in Current Scenario's*, Advances in Intelligent Systems and Computing, 2020.
- [11] R. Priyadarshi, B. Gupta, and A. Anurag, "Deployment techniques in wireless sensor networks: a survey, classification, challenges, and future research issues," *The Journal of Supercomputing*, vol. 76, no. 9, pp. 7333–7373, 2020.
- [12] B. D. Deebak and F. al-Turjman, "A hybrid secure routing and monitoring mechanism in IoT-based wireless sensor networks," *Ad Hoc Networks*, vol. 97, article 102022, 2020.
- [13] X. Fu, P. Pace, G. Aloï, L. Yang, and G. Fortino, "Topology optimization against cascading failures on wireless sensor networks using a memetic algorithm," *Computer Networks*, vol. 177, article 107327, 2020.
- [14] S. El Khediri, W. Fakhret, T. Moulahi, R. Khan, A. Thaljaoui, and A. Kachouri, "Improved node localization using K-means clustering for wireless sensor networks," *Computer Science Review*, vol. 37, article 100284, 2020.
- [15] K. Khujamatov, K. Ahmad, E. Reypnazarov, and D. Khasanov, "Markov chain based modeling bandwidth states of the wireless sensor networks of monitoring system," *International Journal of Advanced Science and Technology*, vol. 29, no. 4, pp. 4889–4903, 2020.
- [16] S. R. Mugunthan, "Novel cluster rotating and routing strategy for software defined wireless sensor networks," *Journal of ISMAC*, vol. 2, no. 3, pp. 140–146, 2020.
- [17] J. S. Raj, "Machine learning based resourceful clustering with load optimization for wireless sensor networks," *Journal of Ubiquitous Computing and Communication Technologies*, vol. 2, no. 1, pp. 29–38, 2020.
- [18] R. Lin, H.-J. Kim, S. Achavananthadith et al., "Wireless battery-free body sensor networks using near-field-enabled clothing," *Nature Communications*, vol. 11, no. 1, p. 444, 2020.
- [19] A. Ghosal, S. Halder, and S. K. Das, "Distributed on-demand clustering algorithm for lifetime optimization in wireless sensor networks," *Journal of Parallel and Distributed Computing*, vol. 141, pp. 129–142, 2020.
- [20] H. Landaluce, L. Arjona, A. Perillos, F. Falcone, I. Angulo, and F. Muralter, "A review of IoT sensing applications and challenges using RFID and wireless sensor networks," *Sensors*, vol. 20, no. 9, p. 2495, 2020.
- [21] A. Bulashenko, S. Piltyay, A. Polishchuk, and O. Bulashenko, "New traffic model of M2M technology in 5G wireless sensor networks," in *2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*, pp. 125–131, Kyiv, Ukraine, 2020.

- [22] R. Priyadarshi, B. Gupta, and A. Anurag, "Wireless sensor networks deployment: a result oriented analysis," *Wireless Personal Communications*, vol. 113, no. 2, pp. 843–866, 2020.
- [23] J. A. Alzubi, "Bipolar fully recurrent deep structured neural learning based attack detection for securing industrial sensor networks," *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 7, article e4069, 2021.
- [24] J. Singh, R. Kaur, and D. Singh, "A survey and taxonomy on energy management schemes in wireless sensor networks," *Journal of Systems Architecture*, vol. 111, article 101782, 2020.
- [25] B. Bhushan and G. Sahoo, "Requirements, protocols, and security challenges in wireless sensor networks: an industrial perspective," in *Handbook of Computer Networks and Cyber Security*, B. Gupta, G. Perez, D. Agrawal, and D. Gupta, Eds., pp. 683–713, Springer, Cham, 2020.