Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 6, pp. 4844–4855, DOI 10.12694/scpe.v25i6.2844

APPLICATION OF LSTM-BASED BODY FATIGUE DETECTION ALGORITHM IN TAI CHI TRAINING

ZHEHUA FAN* AND JINMAO TONG†

Abstract. Tai Chi training sessions are often lengthy, and athletes are prone to experiencing fatigue during the process. Timely detection of body fatigue can help athletes prevent injuries caused by excessive fatigue. This study combines Long Short-Term Memory (LSTM) networks with facial muscle activity detection models to propose a novel fatigue detection algorithm. In this algorithm, the limitations of LSTM networks in capturing future information are addressed by introducing an improved LSTM network model and combining attention mechanisms to highlight the important features of physical fatigue. This study utilizes hyperspectral imaging technology to extract real-time muscle fatigue signals from the faces of subjects in the dataset. Performance validation of the proposed model shows that it effectively extracts facial fatigue features with a detection accuracy of 97.67% and a recall rate of 96.78%, outperforming other existing models in this field. The model constructed through research has excellent performance and has broad application prospects in current and future technological development due to its high flexibility and adaptability, providing support and innovation momentum for different industries.

Key words: LSTM; Body Fatigue Detection; Tai Chi Training; FMAD; Attention Mechanism

1. Introduction. Tai Chi is a traditional sport in China with a wide audience demand. It is characterized by gentle movements and slow speed, and the training duration is usually long [1]. Fatigue is inevitable in daily Tai Chi training. The state of exercise fatigue not only fails to achieve the goal of strengthening the body, but also gradually deteriorates the individual's physical health [2]. In recent years, rapid detection of athletes' states has become a key focus in the field of sports. Continuing to engage athletes in high-intensity training tasks in a fatigued condition can have adverse effects on their athletic performance and physical health [3]. Timely assessment of athletes' physical fatigue status can prevent injuries caused by excessive fatigue [4]. In recent years, fatigue detection and diagnostic technologies have acted as an increasingly crucial factor in the construction of sports teaching and experimental centers [5]. Current detection technologies have drawbacks such as complexity, difficulty in implementation, and low accuracy. This study combines Long Short-Term Memory (LSTM) networks and Facial Muscle Activity Detection Model (FMAD) to propose an FMAD-Bi-LSTM-Attention model for body fatigue detection. The aim of this study is to provide relevant technical support for body fatigue detection in Tai Chi training. The innovations of this study are as follows: (1) Introducing the Bi-LSTM model to address the limitation of LSTM in capturing future information. (2) Incorporating the Attention mechanism in Bi-LSTM to highlight important features and make the model overall performance better. (3) Proposing the FMAD algorithm for facial muscle activity detection and extracting real-time muscle fatigue signal features using hyperspectral imaging technology from the faces of subjects in the dataset. The structure of this study consists of four parts: the first part introduces the development status of the required technologies in the related work section, the second part establishes the main detection model of this study in the model construction section, the third part validates the performance of the model, and the final part provides a summary and outlook for the entire paper.

2. Related Works. Currently, there have been discussions among scholars on fatigue detection algorithms. Fatima B et al. proposed a driver fatigue detection method based on microsleep patterns [6]. This method captures the driver's state image through a camera and uses a deep learning model combining SVM and Ad boost to classify the driver's mental state. The proposed model achieves an average detection accuracy of 98.7% among the participants and has potential applications. Ansari S et al. proposed a fatigue detection method

*Physical Education College, Putian University, Putian, 351100, China (ptxyfzh@163.com)
†College of General Education, Fujian Chuanzheng Communications College, Fuzhou, 350007, China (tjm2021@163.com)

based on motion capture systems [7]. This method monitors the driver's head posture movements using the XSENS system to determine the driver's mental fatigue state. The study conducted experiments on 15 healthy participants, and the outcomes displayed that the proposed model can accurately identify the driver's activity, fatigue, and excessive states. Li X et al. addressed the lack of facial detection functionality in existing visual fatigue detection methods and developed a driver fatigue detection system based on CNN [8]. This system uses a face detection network to locate the face and classify the located face into "normal" or "distracted attention" states. The comprehensive evaluation outcomes showed that the mean detection accuracy of the system is 89.55%, indicating room for improvement. Li F et al. discussed the detection of human fatigue states in traffic control and proposed an interpolation method based on eye tracking to assess fatigue indicators [9]. This method can extract fatigue indicators based on eye tracking from low-quality eye tracking data and adaptively classify missing gaze points. Two types of simulation experiments demonstrate that the proposed method has good detection performance and contributes to the application of eye tracking data in human fatigue detection. Zhang S et al. proposed an intelligent method combining CNN and LSTM to determine the fatigue status of medical staff, addressing the difficulties and complexities of traditional manual feature extraction methods [10]. The validation results on the dataset showed that the model could quickly learn the temporal information of time series and achieve high classification accuracy. The techniques used in this study also have certain inspiring significance for the research topic.

LSTM networks have been favored by scholars from various fields due to their practicality, and discussions on them have always been highly active. Zhang J et al. constructed a human activity recognition system based on dense LSTM models and WIFI networks [11]. The proposed system synthesized variant activity data using eight channel state information transformation methods to mitigate the impact of activity inconsistency and specific subject issues. The system achieved 90% accuracy and demonstrated good robustness in adapting to small-scale data. Amin J et al. approached from the perspective of human gait recognition and built a convolutional bidirectional LSTM model [12]. The model used CNN to extract human gait features and used them as inputs to LSTM to provide distinguishable temporal information. Additionally, the proposed model identified human gait with predicted scores using tinyYOLOv2. Experimental results validated the model's good recognition accuracy. Pan C et al. conducted research on driver action recognition from the perspective of traffic safety [13]. They proposed a human motion recognition model based on graph convolution and LSTM. The model first used graph convolution for spatial structural feature inference and then used LSTM for temporal motion feature learning within sequences. Experiment outcomes displayed that the proposed model achieved a recall rate of 8.24% for 88 driving activities and could meet practical application requirements. Xu S et al. conducted research on action recognition based on 3D skeletal sequences and proposed an attention-based multi-level co-occurrence graph convolution LSTM model [14]. This model can utilize body structural information from the skeleton to enhance multi-level co-occurrence feature learning. The spatial attention module in the model can be used to enhance features of key joints in the skeleton input. Simulation experiments validated the superior performance of the proposed model. Li X et al. developed a novel LSTM model for human motion recognition [15]. This model uses a bottom-up approach to identify human body key points in images and then combines multiple joints as nodes in the system. The advantage of LSTM is that it can recognize actions in different regions without human identification. Three sets of cross-validation experiments demonstrated the model's ability to extract deep human motion features.

In summary, there have been some research achievements in fatigue detection methods and LSTM, but few scholars have proposed more innovative approaches. Additionally, most algorithms for fatigue detection are focused on areas such as traffic driving, and there is still room for enrichment in the field of sports. Therefore, this study proposes an FMAD-Bi-LSTM-Attention model for body fatigue detection, aiming to provide relevant technical support for body fatigue detection in Tai Chi training.

3. Body Fatigue Detection Method Based on LSTM. This section first improves the LSTM model by constructing Bi-LSTM-Attention. Then, the FMAD algorithm is introduced, and finally, the FMAD-Bi-LSTM-Attention model for body fatigue detection is constructed.

3.1. Body Fatigue Detection Model Based on LSTM and its Improvement. Long Short-Term Memory networks (LSTM) are a type of recurrent neural network (RNN) that belong to the class of gated algorithms and are suitable for processing time series data [16]. LSTM uses three different types of gates

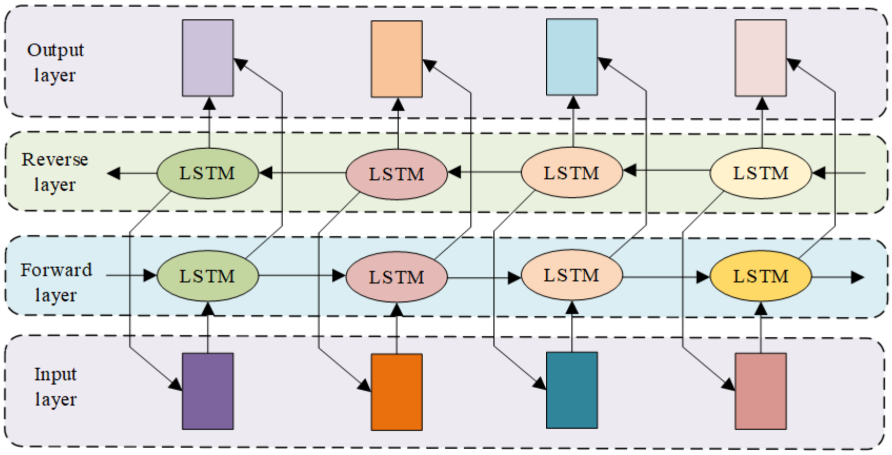

Fig. 3.1: Basic structure of Bi-LSTM model

to control the internal state and recurrent output units. Specifically, the input gate determines whether the current internal state should be updated based on the input at the current time step and the output from previous time steps. The forget gate influences the current internal state based on the previous internal state, determining whether to discard or keep historical data. The output gate determines the output based on the input at the current time step and the internal state of the system [17]. Conventional LSTM models can only make predictions in a forward direction in the sequence, ignoring future information. However, the relevant features of body fatigue are not only related to the preceding sequence but also closely related to the succeeding sequence [18]. Therefore, in this study, a Bi-LSTM model is chosen for training the body fatigue detection model, as shown in Figure 3.1. In the Bi-LSTM model, the current input not only depends on the preceding information but also on the succeeding information, allowing for a comprehensive consideration of the temporal information before and after video frames.

To better extract the feature information from motion videos and enhance the LSTM's ability to learn temporal features, this study proposes a Bi-LSTM-Attention model. This model first extracts deep features of Tai Chi movements, then feeds the corresponding feature vectors into the Bi-LSTM network to learn the temporal sequence features between frames comprehensively. Next, the feature vectors are passed to the Attention layer to adaptively perceive the network weights that have a significant impact on the recognition results, allowing certain features to receive more attention. Finally, the classification results are obtained by connecting the fully connected layer to the classifier, which is used for detecting body fatigue.

Let w_i represent the weights from one unit layer to another unit layer; x_t represent the extracted feature vectors; h represent the feature sequence input from forward to backward; h' represent the feature sequence input from backward to forward; w_i represent the output results of the Bi-LSTM network. At time t, the feature vector input from forward to backward can be represented by Equation (3.1).

$$h_t = \text{sigmoid}\left(w_1 x_t + w_2 h_{t-1} + b_t^{(1)}\right) \tag{3.1}$$

In Equation (3.1), h_{t-1} represents the output of the previous feature vector; $b_t^{(1)}$ represents the bias term for the control gates of the t-th feature vector; *sigmoid* represents the activation function. At time step t, the feature vector input from backward to forward can be represented by Equation (3.2).

$$h_t^{'} = \text{sigmoid}\left(w_3 x_t + w_5 h_{t+1}^{'} + b_t^{(2)}\right) \tag{3.2}$$

In Equation (3.2), $h_{|_{t+1}}^{'}$ represents the output of the subsequent feature vector; $b_t^{(2)}$ represents the bias term for the control gates of the t-th feature vector. The feature vector output from forward to backward at timet in the Bi-LSTM unit can be represented by Equation (3.3).

$$o_t^{'} = \tanh\left(w_4 h_t + b_t^{(3)}\right) \tag{3.3}$$

In Equation (3.3), $b_t^{(2)}$ represents the bias term for the control gates of the t-th feature vector. The feature vector output from backward to forward at time t in the Bi-LSTM unit can be represented by Equation (3.4).

$$o_t^{''} = \tanh\left(w_6 h_t^{'} + b_t^{(4)}\right) \tag{3.4}$$

In Equation (3.4), $b_t^{(4)}$ represents the bias term for the control gates of the t-th feature vector. The final output vector is obtained by summing and averaging the obtained $o_t^{'}$ and $o_t^{''}$, as shown in Equation (3.5).

$$o_t = \frac{o_t^{'} + o_t^{''}}{2} \tag{3.5}$$

Afterwards, the obtained feature vectors are input into the attention mechanism for network weight perception. Compared to traditional LSTM, Bi-LSTM can learn both past and future information simultaneously, resulting in more robust temporal information. Attention is mechanism for signal processing, which weights the features at different time points in Bi-LSTM and represents the salient features, thereby improving the overall performance of the network. When classifying body fatigue using this method, it can first make a preliminary prediction to narrow down the recognition range, and then adjust the weights based on the correlation between behaviors to achieve more accurate recognition. Let o_t represent the t-th feature vector output from Bi-LSTM, which is passed to the attention model. The initial input state vector s_t is acquired via the attention model hiding layer. The weight coefficient α_t is vector proportion. The final output vector Y is obtained by summing the product of the initial input state vectors s_t and the weight coefficients α_t. The energy value is shown in Equation (3.6).

$$e_t = \tanh\left(w_t s_t + b_t\right) \tag{3.6}$$

In Equation (3.6), e_t represents the energy value; b_t represents the energy bias term. Based on Equation (3.6), the expression for the weight coefficient can be obtained, as shown in Equation (3.7).

$$\alpha_t = \frac{\exp\left(e_t\right)}{\sum_{j=0}^{t} e_j} \tag{3.7}$$

In Equation (3.7), $\exp\left(e_t\right)$ represents the exponentiation of the energy values with e as the base; $\sum_{j=0}^{t} e_j$ represents the cumulative sum of the energy values from the previous parts. By comparing the two, the weight coefficients that affect the detection results can be acquired, thereby achieving the transformation from the initial state to the attention state. Thus, the final state vector is shown in Equation (3.8).

$$Y = \sum_{t=0}^{n} \alpha_t s_t \tag{3.8}$$

Using the formula shown in Equation (3.8), the final output vector Y can be obtained. With the final classification result, it is integreated. The overall Bi-LSTM-Attention framework is shown in Figure 3.2.

3.2. Construction of Body Fatigue Detection Model by Combining Bi-LSTM-Attention and FMAD. Before using the Bi-LSTM-Attention model for body fatigue detection, the feature signals of body fatigue need to be extracted. When a person is in a fatigued state, there are significant differences in facial muscles compared to normal conditions. Therefore, the FMAD method is well adapted for fatigue detection in the human body. In this study, the FMAD model is applied to extract fatigue signals [19]. FMAD is a method that combines non-invasive body fatigue detection with deep learning, which can effectively improve the accuracy of body fatigue detection.

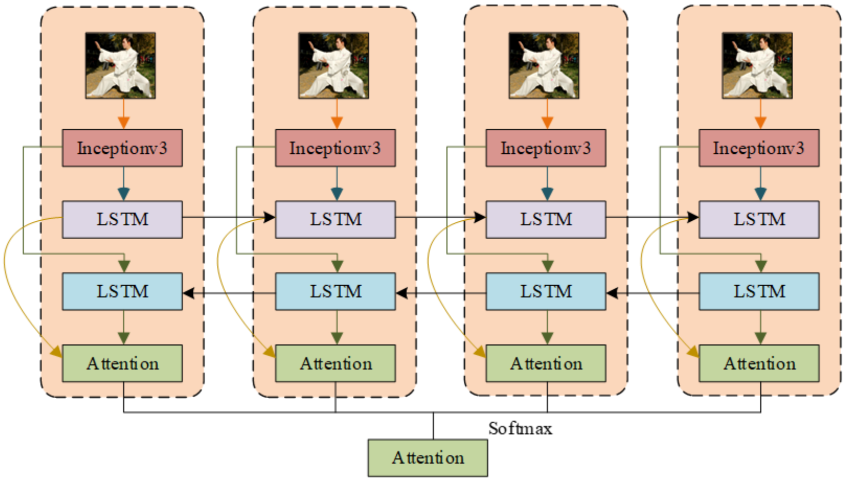

Fig. 3.2: Bi-LSTM-Attention framework

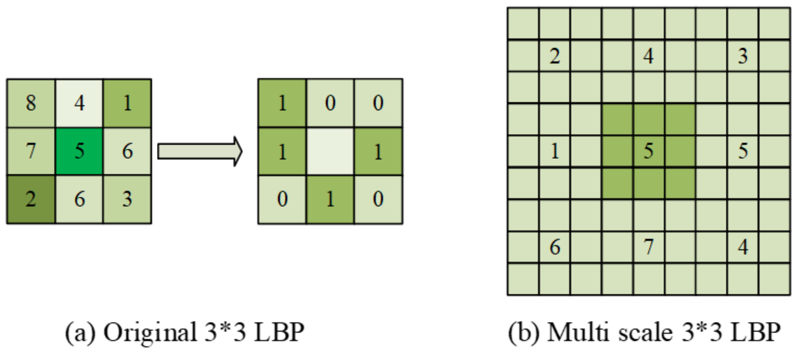

Fig. 3.3: Structure of LBP and MB-LBP

Due to the large amount of system errors and noise originated from path and scattering effects, it becomes very difficult to obtain body fatigue features. At the same time, interference of external factors is also capable of affecting the data. To make the noise not affect facial images' quality, this study uses Multi-Block Local Binary Patterns (MB-LBPs) for facial feature point extracting. It can be recognized as an LBP extension. As shown in Figure 3.3(a), the LBP operator uses the grayscale value of the pixel at the center of its neighborhood as the threshold, compares the grayscale values of the adjacent 8 pixels with the grayscale value of the center pixel, and obtains the LBP value of the center pixel [20]. As shown in Figure 3.3(b), MB-LBP divides the image region into multiple sub-blocks at appropriate scales, and further divides the sub-blocks into smaller regions. The LBP feature is acquired via comparison between the small region's grayscale value and the grayscale values of the surrounding small regions. The MB-LBP algorithm makes the facial extracting increased; it as well makes the robustness to image noise more reliable [21]. The face informative extracting degree varies under MB-LBP of various scales. With multiple experiments, this study uses MB-LBP with a scale of 4*4 to extract and filter facial signals.

After denoising the images, this study conducted a detailed analysis of the facial signal features in the state of fatigue. Before further identifying body fatigue, it is necessary to determine a suitable Region of Interest (ROI). Selecting the appropriate ROI is to identify which areas of the face of different subjects are most sensitive and representative of the feature signals in the state of fatigue. By comparing and analyzing a large number of experimental research results, an ROI with high sensitivity to body fatigue signals is selected, in order to dispel traditional methods dependence of the baseline data. First, the regions sensitive to the feature signals in the state of fatigue are defined as ROIs at different positions. Typically, there are four possible ROI positions on the face, namely the left cheek, right cheek, left corner of the mouth, and right corner of the mouth. Therefore, this study needs to analyze each of these four regions to determine which region is sensitive to body fatigue and can extract good signal features.

After determining the facial ROIs of the subjects, the movement changes of their facial muscles need to be tracked [22]. In this study, the Lucas-Kanade optical flow method is used for this process. The Lucas-Kanade method calculates the movement of each pixel from time t to $t + \alpha t$ between two frames. It is based on the Taylor series of the image signal, also known as the differential method, which takes partial derivatives with respect to spatial coordinates and that of time. Equation (3.9) displays the image constraint.

$$I(x, y, z, t) = I(x + \delta x, y + \delta y, z + \delta z, t + \delta t) \tag{3.9}$$

In Equation (3.9), $t + \alpha t$ represents the pixel at that point in the stereo image. Assuming that the object's motion is small enough, Taylor series can approximate the equation, as shown in Equation (3.10).

$$\frac{\partial I}{\partial x} V_x + \frac{\partial I}{\partial y} V_y + \frac{\partial I}{\partial z} V_z + \frac{\partial I}{\partial t} V_t = 0 \tag{3.10}$$

In Equation (3.10), V_x, V_y, V_z represent the optical flow vectors of x, y, z; $\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}, \frac{\partial I}{\partial z}, \frac{\partial I}{\partial t}$ represent the partial derivatives of the pixel (x, y, z, t) in the image. Thus, Equation (3.10) can be rewritten as Equation (3.11).

$$I_x V_x + I_y V_y + I_z V_z = -I_t \tag{3.11}$$

To solve the problem of over-determination, this study uses the least squares method to obtain the motion positions. In the determined ROI, three points are randomly selected in this study, and the Lucas-Kanade method is used to calculate the motion trajectories of these three points, S_1, S_2, S_3 as shown in Equation (3.12).

$$\begin{aligned} S_1 &= \left[\left(x_1^a, y_1^a\right), \ldots, \left(x_r^a, y_r^a\right), \ldots, \left(x_R^a, y_R^a\right)\right] \\ S_2 &= \left[\left(x_1^b, y_1^b\right), \ldots, \left(x_r^b, y_r^b\right), \ldots, \left(x_R^b, y_R^b\right)\right], \ r = 1, 2, 3, \ldots, R \\ S_3 &= \left[\left(x_1^c, y_1^c\right), \ldots, \left(x_r^c, y_r^c\right), \ldots, \left(x_R^c, y_R^c\right)\right] \end{aligned} \tag{3.12}$$

In Equation (3.12), x, y represents the position of the point; a, b, c represent feature points; R represents the frame number. According to Equation (3.12), the centroid of the three points is obtained as S_0,reference points that are fixed. The Euclidean distance between the three feature points and S_0 is acquired to obtain three sets of feature sequences D_1, D_2, D_3. The final feature sequence D, as shown in Equation (3.13).

$$D = \sum_{i=1} D_i / 3 \quad i = 1, 2, 3 \tag{3.13}$$

In Equation (3.13), D represents the signal output. This item expresses the mean distance of the sequences Therefore, the high-frequency jitter signal in the state of body fatigue can be extracted and used as a feature for classification training in the Bi-LSTM-Attention model. The main steps of the FMAD algorithm are shown in Figure 3.4.

4. Performance Verification of the FMAD-Bi-LSTM-Attention Body Fatigue Detection Model. This section first tests the performance of the FMAD component in the FMAD-Bi-LSTM-Attention body fatigue detection model, and then experimentally verifies the training process and detection accuracy of the overall FMAD-Bi-LSTM-Attention model.

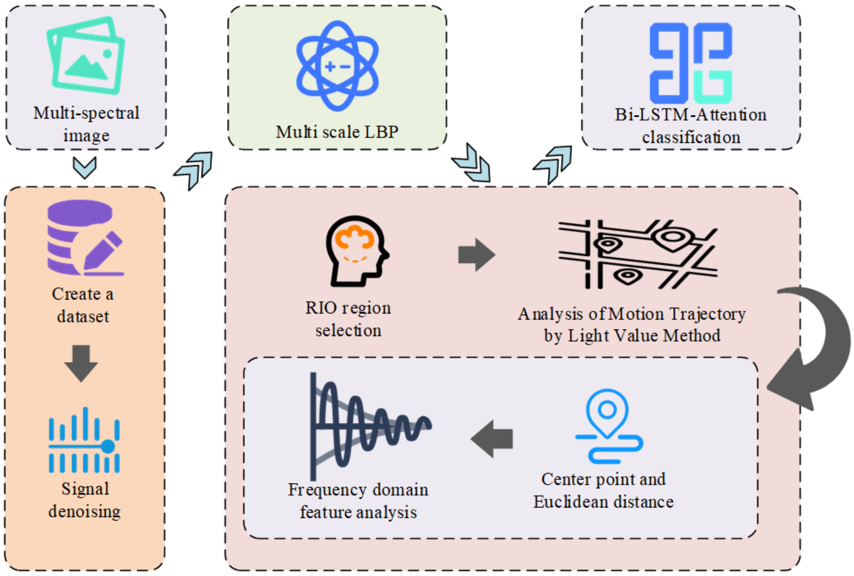

Fig. 3.4: Overall operation process of MAD-Bi-LSTM-Attention

4.1. Performance Testing of the FMAD Model. Currently, there is no available reference data for human fatigue detection experiments both domestically and internationally. Therefore, this study recruited some participants to provide experimental data. Most of the participants were recruited from job advertisements in newspapers, and 37 healthy objects take part in in the experiment. Among them, the data of 25 individuals were processed to train the algorithms, and the data of 12 individuals were testing data. The experiment device used to collect data from the subjects was mainly a visible-near-infrared multispectral imaging system in the 450-800 nanometer wavelength range. For the region of interest, an orange light source with high sensitivity was selected and applied. Finally, the data obtained from each participant were 120-second videos. The use of 120 seconds as a baseline is based on the ability of all test subjects' physical conditions to stabilize back to the baseline state during this time period, ensuring data consistency and reliability. If a test object requires a longer time to recover to baseline, it will not affect the validity of the dataset. The experimental design has taken into account baseline state uniformity and controlled individual recovery differences through standardized processing to ensure data consistency and scientific results. Since within 120 seconds, the physical condition of all individuals would return to near the baseline, the frames in the videos were segmented into a series of images, and these images were further segmented into groups as the processed dataset. In the preparation of these two datasets, all participants underwent three main experiments: firstly, each person wore a chest strap heart rate monitor and a finger pulse oximeter to accurately measure heart rate; secondly, the subjects were brought into a well-lit room and comfortably seated, and after appropriate rest, photographs were taken of the subjects to obtain baseline data; thirdly, the subjects were asked to perform some Tai Chi exercises to induce physical fatigue, and then they were asked to sit down and photographs were taken to show the subjects' Tai Chi movements. The hardware and software config for the best possible experiment environment is displayed Table 4.1.

This study extracted and analyzed the motion characteristics of the four ROI regions of the face in the state of body fatigue from 12 participants. In Figure 4.1, the frequency domain of the motion characteristic signals extracted from these four ROI regions is shown as the experimental result. The waveforms in the spectrum of the left cheek and right cheek regions are prominent peaks, indicating a strong sensitivity of these two regions to facial muscle motion characteristic signals in the state of body fatigue. The main reason is that the muscle groups in the cheek area participate more in facial expression changes and muscle activity during fatigue, making their movement characteristic signals more prominent and concentrated in the frequency domain. On

Table 4.1: Experimental hardware and software environment configuration

Experimental environment		Disposition
Software	System	Ubuntu 16.04
	Dependency library	Opencv, Protobuf, Lmdb, Hdf5
	Language	Shell, Python
	Deep learning framework	Caffe1.0
Hardware	GPU	NVIDIA GeForce GTX 1070 Graphics card
	CPU	Intel i3-7100

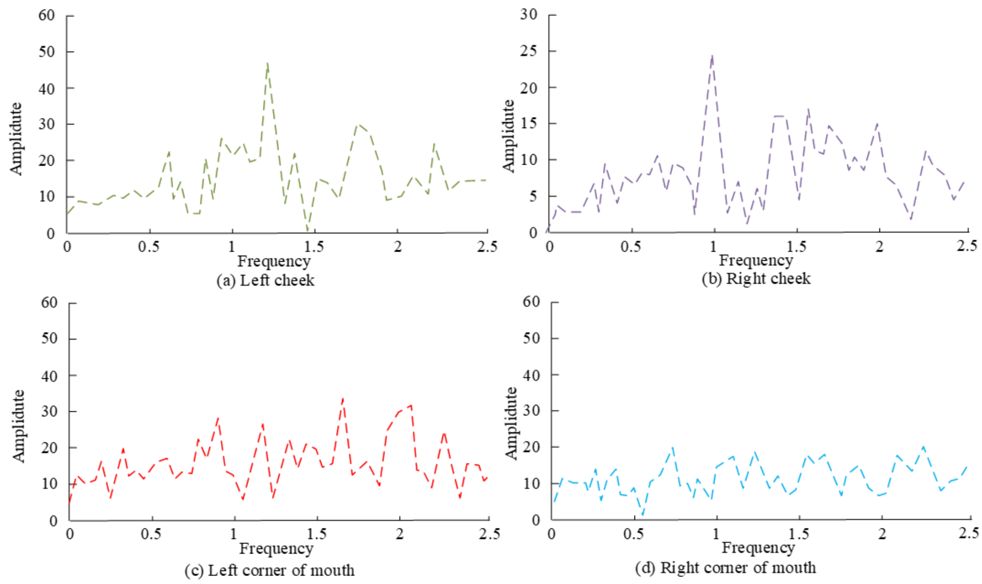

Fig. 4.1: Motion feature signals of four ROI regions under fatigue condition

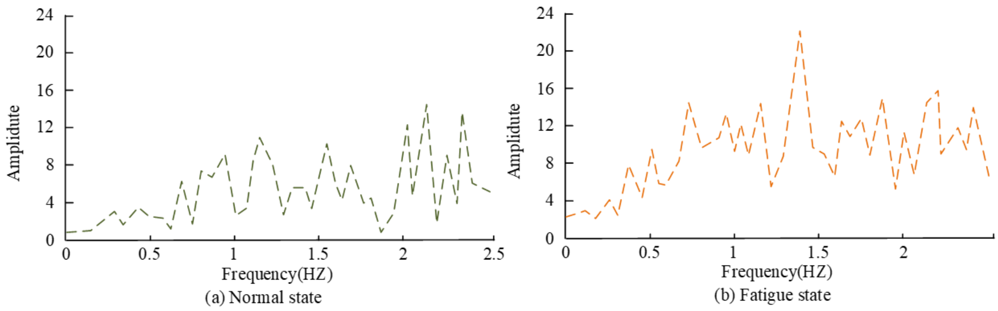

Fig. 4.2: Spectrum of characteristic signal under two states

the other hand, the waveforms in the spectrum of the left corner of the mouth and right corner of the mouth regions are chaotic, indicating a lower sensitivity of these two regions to facial muscle motion characteristic signals in the state of body fatigue.

By analyzing the sample data of the 12 participants, good recognition results were obtained. Figure 4.2

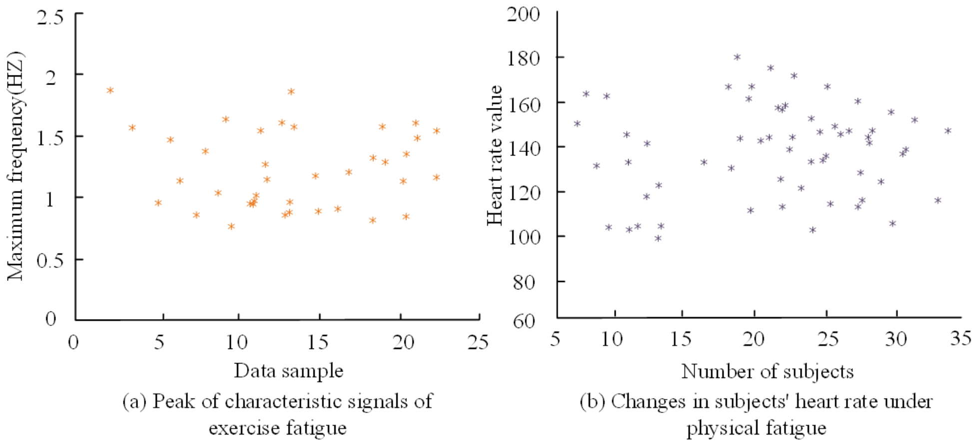

Fig. 4.3: Changes of subjects' physical indexes

represents a typical spectral characteristic signal obtained through fast Fourier transform (FFT). Through the spectral analysis of the baseline and fatigue states, it was found that the FFT curve during the baseline showed no significant peaks and had irregular and messy waveforms. However, during the fatigue state, a significant peak at 1.5 Hz appeared on the FFT curve.

Additionally, this study summarized the range of peak frequencies for all subjects. In Figure 4.3(a), muscle tremors within the high-frequency range above 1 Hz were observed in the experimental subjects. Each individual had a different frequency at which their muscle activity peaks occurred during body fatigue, indicating that the muscle motion characteristics are sensitive to physiological fatigue. Figure 4.3(b) shows the reference indicator for body fatigue, which is the heart rate monitor. The experiment recorded the heart rate of all participants in the state of body fatigue and recorded their maximum heart rate. The results were consistent with the results presented by the muscle activity peaks.

4.2. Performance Analysis of FMAD-Bi-LSTM-Attention Model in Tai Chi Training. Each time, 10% images are picked stochastically from the training set to train the FMAD-Bi-LSTM-Attention model for detecting eye fatigue. From the experimental analysis, when the raining iterations goes up, the accuracy increases synchronously. However, the accuracy on the test set displays a initially going up hen going down trend, from underfitting to convergence and then to overfitting. Therefore, it is necessary to find an appropriate number of training iterations. In this study, training for 3000 iterations yielded the best results. Figure 4.4 displays the training, via which a observation can be made that the model converges after 3000 iterations.

Similarly, each time, 10% images are picked stochastically from the training set to train the FMAD-Bi-LSTM-Attention model for detecting mouth fatigue. From the experimental analysis, when the raining iterations goes up, the accuracy increases synchronously. However, the accuracy on the test set displays an initially going up hen going down trend, from underfitting to convergence and then to overfitting. When the number of training iterations is less than 500, the accuracy is less than 90%. But when the number of training iterations exceeds 5000, the training accuracy reaches 99.9%, while the test accuracy is only 88.21%. Therefore, it is necessary to find an appropriate number of training iterations, and in this case, training for 3000 iterations yielded the best results. Figure 4.5 displays the training, via which an observation can be made that the model converges after 3000 iterations.

Figure 4.6 shows the probability values of nodding intervals for a participant when they are drowsy. When people are fatigued, their attention decreases, and their control over their head significantly decreases, causing the head to droop. The occurrence of nodding indicates that the athlete is in a fatigued state. Calculating the nodding frequency of the athlete is an important factor in fatigue detection. According to the FMAD-Bi-LSTM-Attention algorithm, which has good real-time performance and high accuracy, the head pose and its changes

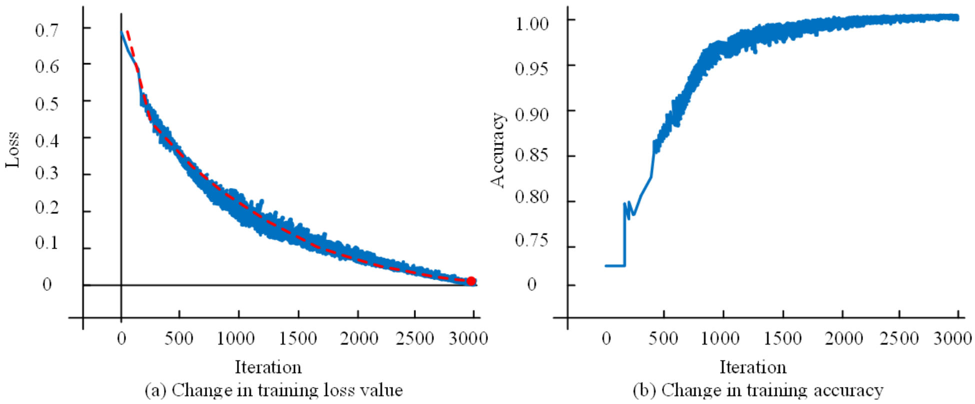

Fig. 4.4: Training process of FMAD-Bi-LSTM-Attention to detect eye fatigue state of subjects

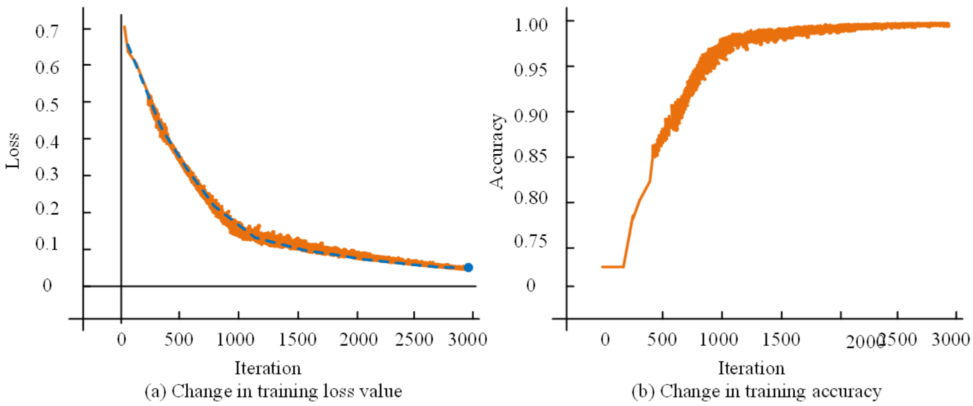

Fig. 4.5: Training process of FMAD-Bi-LSTM-Attention to detect mouth fatigue state

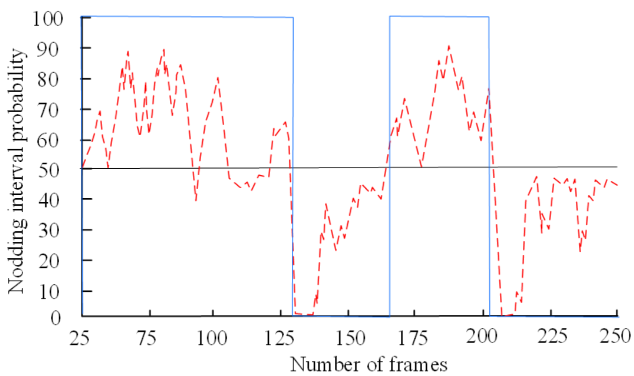

Fig. 4.6: Probability of nodding interval in sleepy state

Table 4.2: Performance comparison results of the four models

Model	Accuracy rate/%	Recall rate/%
3D-DCNN	93.81	90.63
CNN-LSTM	92.52	91.22
MSTN	93.31	92.25
FMAD-Bi-LSTM-Attention	97.67	96.78

during each frame can be obtained. Based on the head pose parameters, the probability value of whether the participant is in a nodding interval can be determined.

In order to highlight the superior performance of the FMAD-Bi-LSTM-Attention model, this study selected some currently popular body fatigue detection algorithms for comparison and verification, and the results are shown in Table 4.2. The detection accuracy of the FMAD-Bi-LSTM-Attention model is 97.67%, and the recall rate is 96.78, which is higher than the 3D-DCNN, CNN-LSTM, and MSTN models. Therefore, this result verifies the superior performance of the FMAD-Bi-LSTM-Attention model in the field of body fatigue detection and has practical significance.

5. Conclusion. Tai Chi, as a common form of physical exercise, is characterized by slow movements and long training durations. Therefore, athletes are prone to experience physical fatigue during training. In order to timely detect the physical fatigue of athletes and prevent injuries caused by excessive fatigue, this study proposed the FMAD-Bi-LSTM-Attention model. The performance of this model was verified. The muscle groups in the cheek area of 12 participants participated in more facial expression changes and muscle activity during fatigue, so the waveforms in the spectrograms of the left and right cheek areas were very significant peaks, indicating that these two areas have strong sensitivity to the characteristic signals of facial muscle movement under physical fatigue, verifying the effectiveness of FMAD. Training the FMAD-Bi-LSTM-Attention model, it converges after around 3000 iterations. The FMAD-Bi-LSTM-Attention model can effectively determine the probability value of whether the participant is in a nodding interval based on the head pose parameters. The final detection accuracy of the FMAD-Bi-LSTM-Attention model is 97.67%, and the recall rate is 96.78, which is higher than the 3D-DCNN, CNN-LSTM, and MSTN models. Therefore, the FMAD-Bi-LSTM-Attention model has practical significance. The limitation of this study is that it did not extend the experimental environment to align the algorithm with practical applications, which can be a direction for future research.

REFERENCES

[1] C. Goumopoulos and N. Potha, “Mental fatigue detection using a wearable commodity device and machine learning,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 8, pp. 10103-10121, 2023.
[2] L. Zeng, K. Zhou, Q. Han, Y. Wang, G. Guo, and L. Ye, “An fNIRS labeling image feature-based customized driving fatigue detection method,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 9, pp. 12493-12509, 2023.
[3] B. Akrout and W. Mahdi, “A novel approach for driver fatigue detection based on visual characteristics analysis,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 1, pp. 527-552, 2023.
[4] J. Lin, H. Li, N. Liu, J. Gao, and Z. Li, “Automatic lithology identification by applying LSTM to logging data: A case study in X tight rock reservoirs,” *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 8, pp. 1361-1365, 2020.
[5] J. Ma, H. Liu, C. Peng, and T. Qiu, “Unauthorized broadcasting identification: A deep LSTM recurrent learning approach,” *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 5981-5983, 2020.
[6] B. Fatima, A. R. Shahid, S. Ziauddin, A. Safi, and H. Ramzan, “Driver fatigue detection using Viola-Jones and principal component analysis,” *Applied Artificial Intelligence*, vol. 34, no. 6, pp. 456-483, 2020.
[7] S. Ansari, F. Naghdy, H. Du, and Y. Pahnwar, “Driver mental fatigue detection based on head posture using new modified ReLU-BiLSTM deep neural network,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 10957-10969, 2021.
[8] X. Li, J. Xia, L. Cao, G. Zhang, and X. Feng, “Driver fatigue detection based on convolutional neural network and face alignment for edge computing device,” *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 235, no. 10-11, pp. 2699-2711, 2021.
[9] F. Li, C. H. Chen, G. Xu, and L. Khoo, “Hierarchical eye-tracking data analytics for human fatigue detection at a traffic control center,” *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 5, pp. 465-474, 2020.

[10] S. Zhang, Z. Zhang, Z. Chen, S. Lin, and Z. Xie, “A novel method of mental fatigue detection based on CNN and LSTM,” *International Journal of Computational Science and Engineering*, vol. 24, no. 3, pp. 290-300, 2021.
[11] J. Zhang, F. Wu, B. Wei, Q. Zhang, H. Huang, S. Shah, and J. Cheng, “Data augmentation and dense-LSTM for human activity recognition using WiFi signal,” *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4628-4641, 2020.
[12] J. Amin, M. Anjum, M. Sharif, S. Kadry, Y. Nam, and S. Wang, “Convolutional Bi-LSTM based human gait recognition using video sequences,” *Computational Materials and Continua*, vol. 68, no. 2, pp. 2693-2709, 2021.
[13] C. Pan, H. Cao, W. Zhang, X. Song, and M. Li, “Driver activity recognition using spatial-temporal graph convolutional LSTM networks with attention mechanism,” *IET Intelligent Transport Systems*, vol. 15, no. 2, pp. 297-307, 2021.
[14] S. Xu, H. Rao, H. Peng, X. Jiang, Y. Guo, X. Hu, and B. Hu, “Attention-based multilevel co-occurrence graph convolutional LSTM for 3-D action recognition,” *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15990-16001, 2020.
[15] X. Li and X. Cao, “Human motion recognition information processing system based on LSTM Recurrent Neural Network Algorithm,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 7, pp. 8509-8521, 2023.
[16] R. Vincent, A. Wagadre, A. Sivaraman, and M. Rajesh, “Human Activity Recognition Using LSTM/BiLSTM,” *International Journal of Advanced Science and Technology*, vol. 29, no. 4, pp. 7468-7474, 2020.
[17] R. Huang, Y. Wang, Z. Li, Z. Lei, and Y. Xu, “RF-DCM: multi-granularity deep convolutional model based on feature recalibration and fusion for driver fatigue detection,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 630-640, 2020.
[18] E. Wu, C. Lin, L. Zhu, Z. Tang, Y. Jie, and G. Zhou, “Fatigue detection of pilots’ brain through brains cognitive map and multilayer latent incremental learning model,” *IEEE Transactions on Cybernetics*, vol. 52, no. 11, pp. 12302-12314, 2021.
[19] A. Ahmadi, H. Bazregarzadeh, and K. Kazemi, “Automated detection of driver fatigue from electroencephalography through wavelet-based connectivity,” *Biocybernetics and Biomedical Engineering*, vol. 41, no. 1, pp. 316-332, 2021.
[20] X. Zhang, D. Lu, J. Pan, J. Shen, M. Wu, X. Hu, and B. Hu, “Fatigue detection with covariance manifolds of electroencephalography in transportation industry,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 5, pp. 3497-3507, 2020.
[21] M. Ünver, M. Olgun, and E. Türkarslan, “Cosine and cotangent similarity measures based on Choquet integral for spherical fuzzy sets and applications to pattern recognition,” *Journal of Computational and Cognitive Engineering*, vol. 1, no. 1, pp. 21-31, 2022.
[22] Y. Yang and X. Song, “Research on face intelligent perception technology integrating deep learning under different illumination intensities,” *Journal of Computational and Cognitive Engineering*, vol. 1, no. 1, pp. 32-36, 2022.

Edited by: Zhengyi Chai
Special issue on: Data-Driven Optimization Algorithms for Sustainable and Smart City
Received: Nov 14, 2023
Accepted: Aug 6, 2024