DeepFocus: Deep Encoding Brainwaves and Emotions with Multi-Scenario Behavior Analytics for Human Attention Enhancement

Min Chen, Yong Cao, Rui Wang, Yong Li, Di Wu, and Zhongchun Liu

Abstract

Evaluation of the degree of attention has very important practical significance and application prospects in our lives. It plays considerably important functions in the fields of education, medical treatment, and automatic driving, and assists people in assessing psychological states automatically and gives early warning. However, there are some problems in the existing research, such as simplex data modality and simple modeling, which make the algorithm difficult to deploy in real circumstances. To solve these problems, a new system of attention degree evaluation, called DeepFocus, is put forward in this article. We propose an attention evaluation method based on multi-modal data and multi-scenario modeling for the first time. In addition, the relationship between emotional data and attention is analyzed in depth, and labels are corrected with emotional data. Based on the DeepFocus system, a completely new attention enhancement system is constructed, and the algorithm is deployed in a practical application for students to perceive and enhance their attention. It can be foreseen that our algorithm could explore people's inward world and assess users' attention degree accurately and comprehensively to help people work, study, and live better with higher efficiency in the near future.

Introduction

Attention means the cognitive process of selectively focusing on executing a certain task and ignoring other things in a certain environment. It plays the leading role in various cognitive activities. In some special fields, serious accidents can be caused by inattention. As for children, they can suffer from attention deficit hyperactivity disorder (ADHD), which is associated with a serious lack of attention. The worldwide prevalence of ADHD is about 4.5 percent [1]. In many special situations, such as doctors' operations, driving, and engineers' precise instrument maintenance, accidents could be avoided if they were able to be attentive all the time. Therefore, attention plays an important role in our lives. Evaluating a human's attention degree accurately can be helpful for finding problems quickly and solving them. For example, people tend to judge the degree of a student's attention to studying in class by observing their facial expression and movement, and quickly remind distracted students to concentrate on their studies or adjust the contents and method of teaching to attract the students' attention [2, 3]. However, appearance observation cannot be the only basis of judging students' attention degree. Therefore, many scholars have further focused on the evaluation of attention degree and have achieved some research findings.

At present, the existing research on the evaluation of people's attention is only to study superficial physiological data. In [4], the author tries to recognize people's emotion by combining EEG data and physiological data, but the designed experiment is relatively simple and could not lead participants into a certain emotional state accurately. The idea and implementation method of the conventional attention degree evaluation system is shown in Fig. 1. Research on conventional attention degree evaluation is done in three layers. In the first layer, it is evaluated according to users' behaviors; some researchers assess it aimed at a computer desktop scenario [5] and a driving scenario. In the second layer, it is evaluated according to users' physiological feature indicators [6]; it is mainly used in the scenario of teaching [7]. In the third layer, the source of corresponding activities is sought, internal physiological features are explored, and EEG is used for evaluation [8, 9]; it is mainly applied in lab scenarios.

The main undesirable qualities are listed below:

- The modality of data is single. The above-stated one or two data types are not enough for judging attention degree accurately. The data mentioned above have appeared in previous studies, but each researcher only focuses on one or two data types for judging the degree of attention. Attention is the orientation and focus of mental activity toward a certain object. It is a psychological feature accompanying sensation, consciousness, memory, thinking, and imagination. These are a series of reactions based on EEG in a certain scenario, and the reaction presents people's physiological indicators and behaviors.
- The present scenario is simplex. Prior research focuses on some special scenarios while ignoring daily fields. As mentioned above, the previous studies tend to focus on driving, class teaching, and remote online

Min Chen, Yong Cao, and Rui Wang are with the School of Computer Science and Technology, Huazhong University of Science and Technology; Yong Li is with Tsinghua University; Di Wu is with Sun Yat-sen University and Guangdong Key Laboratory of Big Data Analysis and Processing; Zhongchun Liu (corresponding author) is with the Department of Psychiatry, Renmin Hospital of Wuhan University.

Digital Object Identifier: 10.1109/MNET.001.1900054

teaching. However, evaluation of attention should involve most daily situations. The points involved in the two aspects are too narrow, and a person's attention cannot be cognized comprehensively. Therefore, the subsequent reminding and enhancement cannot be completed.
- The research methods are simple, and the algorithms are not mature [6–8]. Most of them are only for academic research without any corresponding application. Previous studies were based on users' behaviors and physiological features. Most algorithms of attention degree evaluation are basic processing algorithms and simple machine learning algorithms. With these methods it is difficult to deeply extract features relevant to attention. Therefore, it is more difficult to assess the degree of attention in complex and changeable environments based on the dynamic change of users' behaviors and physiological data.

To solve the above-stated problems, the DeepFocus system is put forward in this article. DeepFocus uses EEG, video data, and eye movement data to evaluate attention in four scenarios [10], namely studying, working, amusement, and relaxing. The core of this article is a method of researching attention degree. The thought of algorithm implementation is expounded in detail. In addition, a testbed is constructed, and an application for attention detection and enhancement is given.The remainder of this article is structured as follows. We describe the evolution process of the DeepFocus system, mainly involving multi-modal data collection, multi-scenario behavior analysis, and the deep encoding of brainwaves and emotional data. We expound the implementation scheme and algorithm model of the DeepFocus system. We describe a testbed of the DeepFocus system. Based on the testbed, the scenario classification and attention evaluation by EEG and video are realized. We realize the application of attention detection and enhancement based on the proposed scheme. Finally, we summarize the article.

Evolution of the DeepFocus System

The conventional attention evaluation system has many problems. Therefore, we put forward a completely new framework based on current research on historical attention evaluation systems and research findings. Input to the evaluation system is not single-modal data, and the evaluation model is not simplex. Deeply integrated multiple models are needed for comprehensive evaluation.

In the evolution from the conventional attention evaluation system to the DeepFocus system, the following aspects are greatly changed and optimized.

Multi-Modal Data Collection

Until now, most research has been based on single-modal data. Some studies have used eye movement data. Other researchers carry out work between attention and EEG, whose main works only focus on selecting and extracting data features [11].Therefore, we consider integrating all the above information and collect data from the face, behavioral performance data, eye movement data, and EEG data, as shown in Fig. 2a. Users' data from the three modes is collected simultaneously based on an eye tracker, a brain-wearable device, and a video camera.

- The existing eye tracker can directly track users' pupils and sight line, and obtain the data relevant to the movement of eyes.
- The brain-wearable device can read users' EEG signal directly. EEG is the general reflection of the physiological activity of brain nerve cells in the cerebral cortex or on the scalp.
- A video camera is used to capture users' facial expression features and behavior information. The key points on the face and behavior are extracted to obtain relevant feature vectors for modeling and analysis.

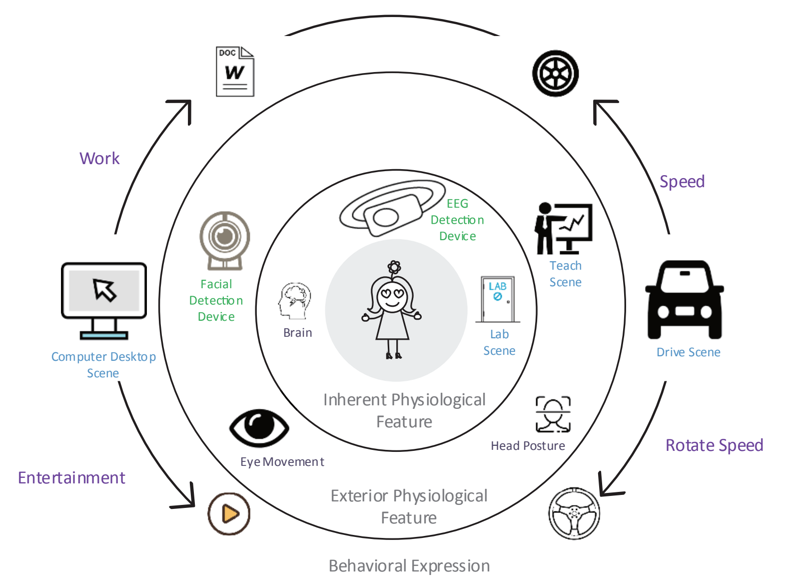

FIGURE 1. Conventional attention degree evaluation system.

Multi-Scenario Behavior Analysis

Most existing attention evaluation systems design an algorithm in certain specific experiments based on the data collected, such as a keyboard click experiment, a mathematical calculation experiment, or a reaction speed test [12]. They almost deviate from practical life without considering practical application scenarios and algorithm deployment in the real world. Meanwhile, it would be better to monitor attention for an entire day and give suggestions and guidance.

As shown in Fig. 3, to analyze users' attention at different times comprehensively and reasonably, users' attention state through a whole day is divided into four scenarios: working, studying, amusement, and relaxing. The four scenarios contain almost all possible states for a user in an entire day. As for different life scenarios, a different dataset is built. When building a dataset, different tasks should be allocated to different scenarios. The model is built and analysis is made for the dataset of each scenario to build four attention evaluation models.

It is necessary to classify users' life scenarios to match the users' evaluation model in real time. Therefore, a scenario classification algorithm should be designed. Second, for a specific scenario, the user's behavior analysis algorithm can be designed to analyze the user's behavior and obtain the relevant feature vector.

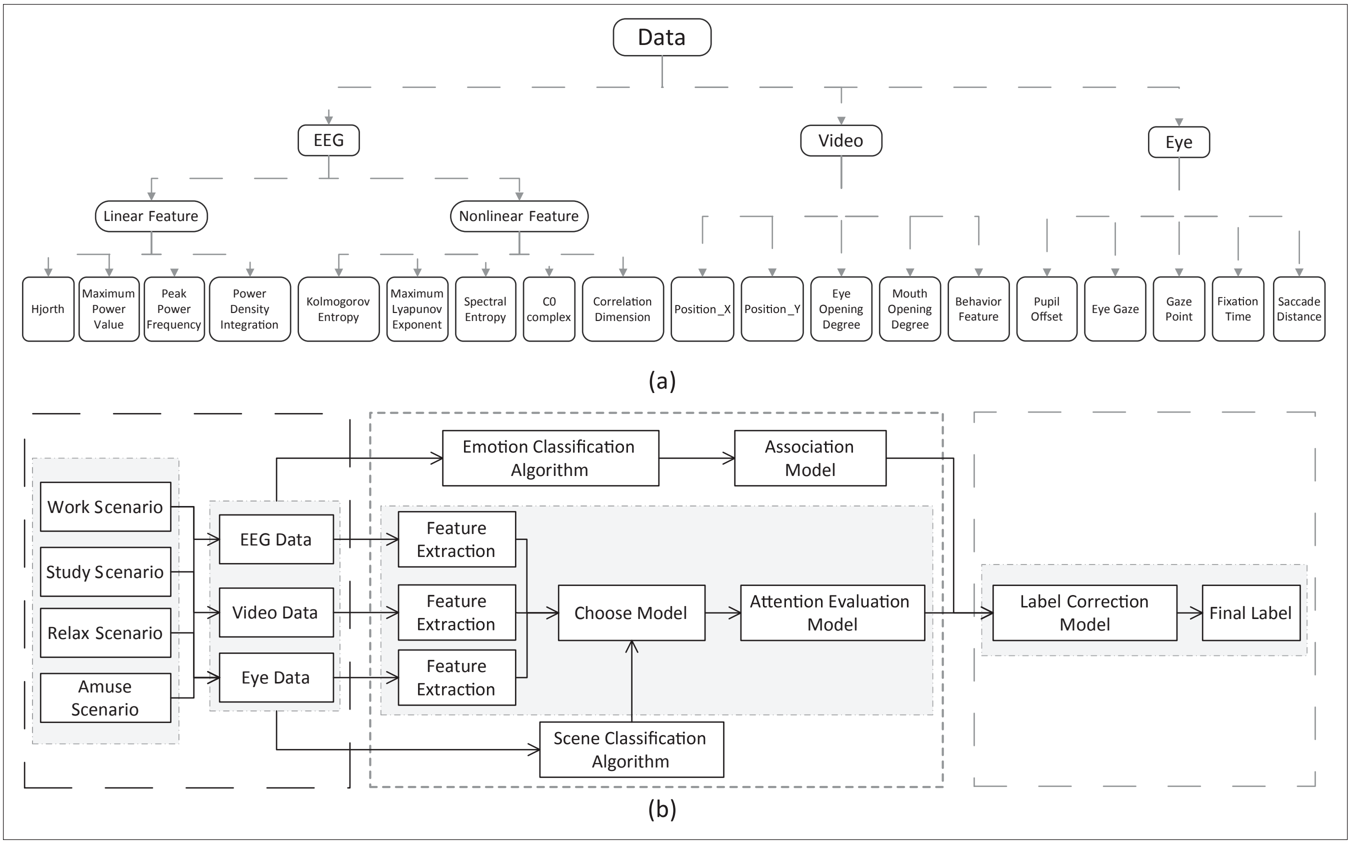

FIGURE 2. Data taxonomy and data process model of DeepFocus: a) dataset composition; b) framework of the attention evaluation system.

Deep Encoding Brainwaves and Emotions

Here, we propose the concept of deep encoding brainwaves and emotional data, that is, based on user brainwaves data encoding, the user's attention label is obtained through traditional feature extraction and a deep convolutional neural network, and then the user's attention label is obtained.

Regarding emotional data coding, until now, there has been no relevant study that mentions the relations between emotion and attention degree. In some brain science research, people's degree of attention varies with different emotional states. For example, it is difficult for a person to calm down to study when he or she is excited, and the attention degree is high when a person is angry.

Emotion is people's subjective feeling toward the value of objective things. The cognitive process is the basis of emotion and guides the development of emotion [13]. In current scientific research on the brain, there are some findings of emotion classification based on brain signals. A frequently used method is to extract rhythm wave features and nonlinear features. After collecting users' EEGs, their emotions can be analyzed [14]. Finally, the model of correlation between users' emotions and attention degree can be built to make an emotion label map to the label of users' attention degree. It plays the function of assistant correction.

In summary, compared to conventional attention evaluation systems, the improvements in our system are as follows:

- A multi-scenario and multi-modal attention evaluation system is put forward. The system first judges the scenario independently according to the data collected. After confirming the scenario, the corresponding model is chosen for evaluating the degree of attention. Multi-dimensional data are used in the model for multi-modal attention degree evaluation.
- The degree of attention is evaluated by combining the multi-modal dataset. Aimed at the above attention evaluation system, a multi-scenario and multi-modal attention-related data collection scheme should be built. Considering typical real-life scenarios, we collect multi-modal data simultaneously aimed at the four scenarios: working, studying, relaxing, and amusement. Concretely, corresponding to the above classification of scenarios, data is collected based on numbers handwriting, mathematical questions and answers, video guidance, and a difference search from pictures.
- Our system explores the deep relations between attention degree and multi-modal data, emotion, and different scenarios. In the whole evaluation system, the model is built between different modal data and attention degree, between emotion and attention degree, and between scenarios and attention degree. By referring to relevant parameters in the model, the internal relations between attention degree and modal data, and emotion and different scenarios can be explored.

Architecture for DeepFocus System

Figure 2b shows the complete process model of the DeepFocus system. It mainly contains of dataset collection, the scenario classification model, the four-class attention evaluation model, the emotion classification model, the emotion-atten-

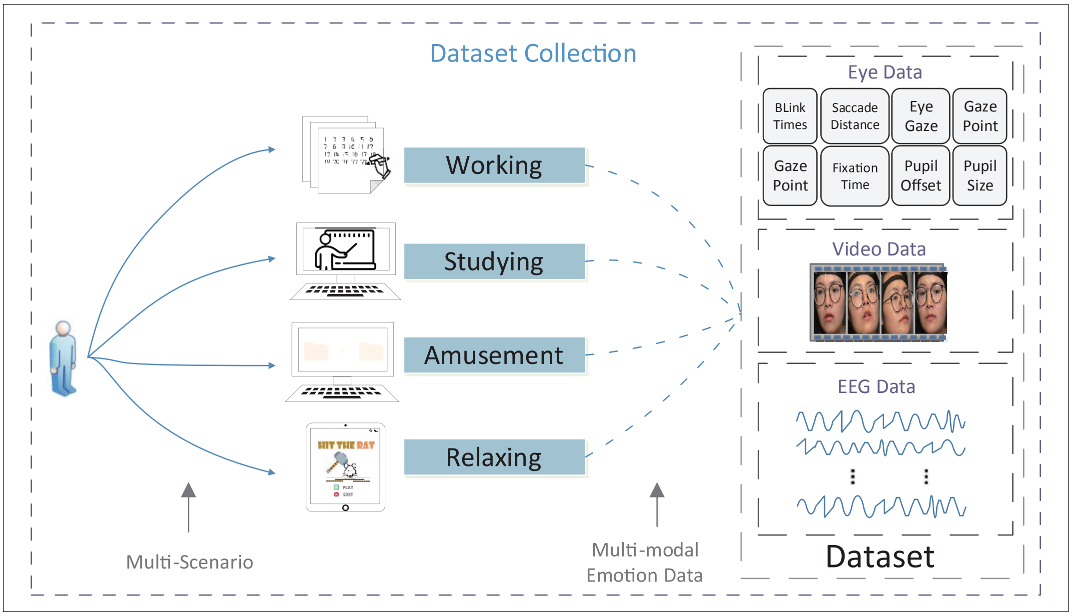

FIGURE 3. Typical application scenarios of DeepFocus: working, studying, amusement, and relaxing.

tion correlation model, and the attention correction model. Constituent parts of the framework are introduced below.

Dataset collection: Experimental equipment includes a computer, video camera, brain-link head band, the brain hoop of NeuroSky, and a 1-route electrode cap. The scenarios include:

- Working scenario — numbers handwriting. The subject is required to write numbers for six minutes from one. If an error occurs, the handwriting should restart from one.
- Studying scenario — reading learning materials. Learn relevant materials and complete corresponding exercises. The subject completes the learning and corresponding exercises within a designated time. The duration of the experiment is 20 minutes. The first 15 minutes are for learning, and the last five minutes are for completing exercises.
- Relaxing scenario — video guidance. In the scenario of relaxing, a comfortable video is chosen as scenario guidance. The duration is also 20 minutes.
- Amusement scenario — amusement game. The subject plays a game such as Mole Attack on a computer. The accuracy of game playing is calculated. The difficulty of the game is different, and the duration is 20 minutes.

Finally, video stimulus is used to enlighten the experimenter's emotion. Songs of different moods are joined to prominently guide a change of the subject's emotion.

Label of attention degree: For attention degree, we use 0–100 to evaluate the degree of attention. There are three degrees of attention according to the score:

1. Very attentive: 80–100
2. Attentive: 60–80
3. Inattentive: 0–60

The tester's activities will be recorded by video. Several evaluators make a label every 1.5 s by watching video. At the same time, the subject wears the brain hoop of NeuroSky during the experiment. The attention label of the NeuroSky product is added into the label data. In addition, the self-reporting of the subject was very important. The subject evaluates his or her attention in retrospect according to the video.

Scenario classification model: We use multi-scenario data for modeling. The difficulty is how to judge the scenario of the users. One is to judge the scenario based on users' background; the other is to judge the scenario based on the similarity of data collected. Both of them need experiments to verify their validity. In the first scheme, users' backgrounds are shot by a camera, the background image is segmented, and objects are extracted and recognized to roughly judge users' scenarios according to the correlation of the objects. The key techniques used in this method are image segmentation, image recognition, and image correlation analysis. In the second scheme, by judging the similarity between the EEG signal collected and that of the four scenarios in the database, the scenario of the tester is defined to be the one with the largest similarity. The usable classification methods can be *K*-nearest neighbor (KNN), clustering, or others. The method of defining the similarity can be the Euclidean distance or cosine distance.

EEG-based attention evaluation model:

1. Data collection: A brain-wearable device, which is a 1-route EEG collector, is used to collect data from the four scenarios.
2. Data preprocessing: Data is down sampled, and a band-pass filter is used to filter the low-frequency component and high-frequency component; artifacts are removed (interference of electro-oculogram, ECG, and EMG).
3. Feature extraction: Short-time Fourier transform (STFT) is used to transfer the EEG from time domain to frequency domain; rhythm waves (β, θ, α, γ, δ) are extracted, and the average energy and power spectrum of the rhythm waves are calculated to obtain the linear and nonlinear features.

4. Attention grading model:
   –The basic method of machine learning data analysis: applicable to circumstances in which data quantity is not enough and the quantity of features is small. Features are chosen artificially, such as the extraction of rhythm waves, the calculation of the power spectrum, and the calculation of the average energy.
   –Multiple linear regression: Linear iteration, segmented regression, or the least square method is used to solve. After features are chosen, support vector regression (SVR) is used to classify them.
   –Deep learning data analysis method: More than 10,000 parameters should be trained, and the data that is several times as many as the parameters is needed. The workload of the data collection is heavy, and there is no need to choose features artificially. STFT can be used to make the data into an image. Features are extracted automatically using convolution. The usable networks can be long short-term memory (LSTM), random forest (RF), and convolutional neural network (CNN) [4, 6].

Video and eye tracker data: Based on the attention evaluation model, the video signal is first divided into frames, and the size of window is set to be 1.5 s. Face images are preprocessed, the facial features are extracted, and the sizes of the eyes, the opening of the mouth, the width of the mouth, the opening degree of the eyes, the opening degree of the mouth, and the direction of the mouth corner are calculated based on facial features to finally obtain the 160-dimension facial expression eigenvector X. An eye tracker can be used to directly measure the position of pupils, eyes' staring area, position of eyes, times of blinks, point of gaze, the duration and times of gazes, the distance of the twitching of the eyelid, the size of pupil, and so on. After obtaining the above-stated data features, the support vector machine (SVM) model can be directly used to classify and get the attention labels.

Emotion data model: Emotion data can be used to rectify attention labels. In relevant research, the attention of people is largely different in different emotional states. Therefore, the relation between emotion and attention degree should be explored. First, emotion classification based on EEG was researched. The detailed algorithm is described below:
1. Regarding the EOG signal as a noise reference signal, Independent Component Correlation Algorithm (ICA) is used to denoise the EEG signal.
2. Band-pass filtering is used to obtain a signal between 0.5 Hz and 50 Hz.
3. Five rhythm waves are extracted: β, θ, α, δ, and γ.
4. Features extraction: The differential entropy of the five rhythm waves is first extracted, and then a deep belief network (DBN) is used for dimension transformation, features learning, and obtaining the final features.
5. Model: DBN and hidden Markov model.
6. Label: Obtain the discrete emotion label. It is assumed that the final emotion label is recorded as $A = [A_1, A_2, ..., A_n]$. The model of relevance between emotion and degree of attention is built. $Q = a_0 + a_1A_1 + ...+ a_nA_n$, namely $P_{emotion} = e^Q/1 + e^Q$. The value $P_{emotion}$ is the attention degree based on emotion.

Attention degree correction model: After the training of the attention evaluation model and emotion data model, the attention label of the three models can be obtained. They are assumed to be A_{11}, A_{12}, A_{13}, A_{21}, A_{22}, A_{23}, A_{31}, A_{32}, and A_{33}. Finally, after receiving the data of the EEG signal, facial image, eye movement, and emotion, multiplex logistic regression can be used for modeling and obtaining the final attention label. The formula is shown below. ω_{ij} is the weight allocated to each kind of label: $A_{score} = \omega_{11}A_{11} + \omega_{12}A_{12} + \omega_{13}A_{13} + ... + \omega_{33}A_{33}$.

Testbed for the DeepFocus System

In this part, we briefly describe the experiment of the DeepFocus system based on multi-scenario and multi-modal data and the analysis of the experiment results. The experimental equipment mainly includes:
1. Wearable brain equipment for collecting user's EEG signal
2. A 24-frame camera, for recording user's face image information in the experimental process
3. Synchronization clock, to ensure that the collected multi-modal data are synchronized
4. Computer, for playing guiding audio
5. Paper material, for completing data collection tasks

The experiment is divided into two parts: one to train the scenario classification model, the other to train the attention evaluation model.

Training Scenario Classification Model

The key element of scenario classification is to get the scenario of the tester according to things, the environment, and items in the video screen, such as office, bedroom, or study, and then map the actual scenario to the four divided scenarios, working, studying, amusement, and relaxing. First, the large-scale scenario category dataset Place205 [15] is used to train ResNet18, and the corresponding scenario is marked as the label of the four scenarios. After the training is completed, the parameters in the model are obtained so that the image scenario can be predicted and described to scenario information in the image. The experiment results are shown in Fig. 4a. We can see that the model can accurately identify some typical scenarios. The recognition accuracy of four kinds of scenarios is 0.981, 0.413, 0.878, and 0.984, respectively. Therefore, the model can be deployed to the DeepFocus system to accurately judge the scenarios where users are located and match the corresponding model parameters.

Training Attention Evaluation Model

Datasets are collected in four kinds of scenarios, such as studying scenarios, where the tester needs to complete the task of handwritten numerals. The detailed process is: first, let the tester be in a resting state for 1 min; let the tester write numerals for 6 min from the beginning. In this process, the tester is disturbed by dialogue with others, playing music, and other forms while recording the tester's state and writing result. We then remove the

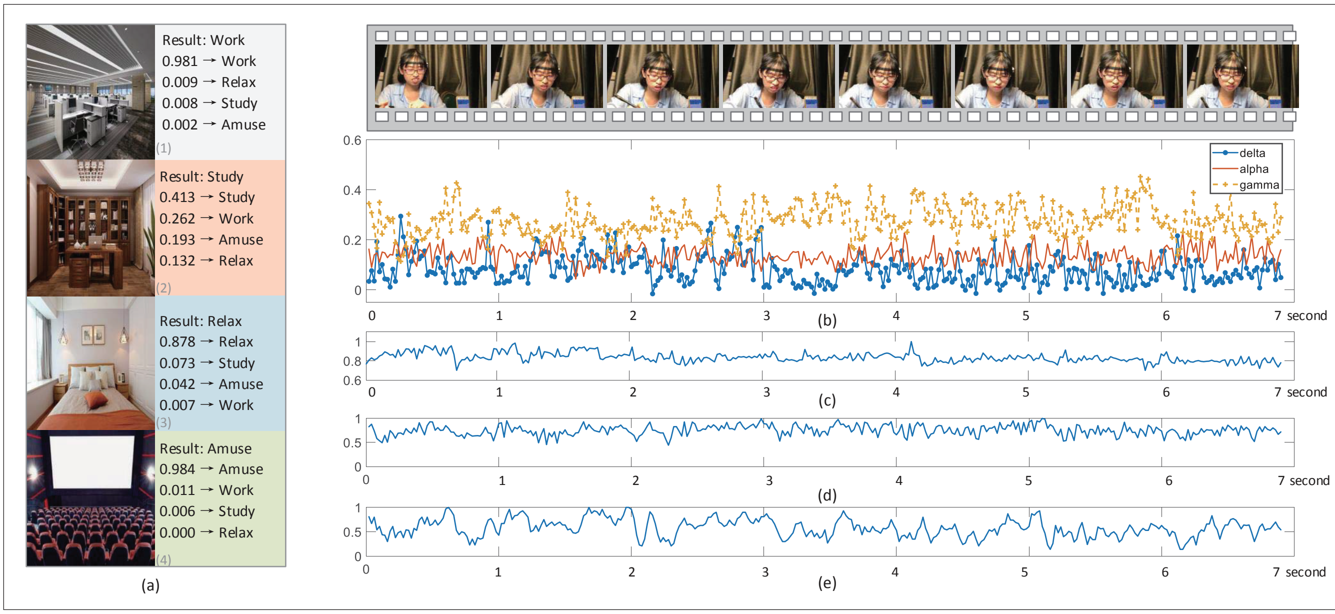

FIGURE 4. The result of scenario classification and attention evaluation: a) result of scene classification model; b) attention curve based on video data; c) attention curve based on EEG signal; d) attention curve of final label.

EEG data of the first 9 s and the last 9 s to prevent the initial unstable data from being retrieved, and label it once per second. It can be judged by the change of the tester's action in the video. If there is a pause or a writing error, the tester is judged to be inattentive; when speaking, skewing, or writing slowly, the tester's attention is also judged to be weak. Five people are selected for evaluation. Finally, there are 114 valid data samples for a person. Each tester's data is collected again every two to three days, so there are three testing times for each tester. Eight testers were selected for the test, and a total of 8208 samples are obtained.

The attention evaluation model is based on EEG data and video data (including face data and eye movement data). First, the EEG signal is filtered to retain the spectrum components below 50 Hz. Then the original signal is converted to the frequency domain by wavelet transform to extract all kinds of rhythmic waves. Several important rhythmic waves are extracted, including δ wave, α wave, and γ wave, as shown in Fig. 4b. The short-term energy of all kinds of rhythmic waves and the nonlinear characteristics are calculated as shown in Fig. 2a. Finally, SVM is used to classify the attention level. After the model training is completed, the attention of each video frame is evaluated, and the result is shown in Fig. 4c. As for the video signal, first, the video data is framed, and the window size is 1 s. Then video frames that do not contain face data or cannot capture faces are removed. Using Dlib and Opencv open source libraries to locate the key points of face data, 68 key points of face are labeled. The labeled key points of a human face are used to calculate some typical features, as shown in Fig. 2a, such as mouth opening size, face offset, and eye opening size. The traditional image processing algorithm is used to judge the degree of attention of the locator. The processing results are shown in Figs. 4c and 4d. Finally, the fusion of the two types of tags is completed, and weighted average is used to get the final result. The final label is shown in Fig. 4e.

The Attention Monitoring and Enhancement System Based on DeepFocus

In this section, we put forward an attention detection and enhancement system based on DeepFocus. In the system, the attention degree of users in the four scenarios of studying, working, relaxing, and amusement can be recorded. Based on this, the user can allocate and coordinate the times and do the appropriate tasks at the proper time.

Architecture

Figure 5 shows the architecture of the attention detection and enhancement system based on DeepFocus. Existing communication technology is used in the architecture to increase the speed of data transmission, containing the brain-wearable device, photographic device and eye tracker, smart terminal, and data center. Each of these are introduced below.

Brain-wearable device: The brain-wearable device is mainly used to collect users' EEG. The brain hoop of NeuroSky is used to collect the EEG signal of one route. Once a user is wearing the brain-wearable device, EEG can be detected at any time and sent to a data center in real time for processing. EEG can be used to obtain and analyze the scenario, attention degree, and emotion state of a user.

Photographic device and eye tracker: Generally speaking, users participate in some activities at fixed sites. For example, they work in an office, study in a quiet room at home, entertain themselves sitting on a sofa in the living room, and relax in the bedroom. Therefore, some photographic devices and eye trackers can be installed at the sites where users appear often. This is convenient for observing external physiological features, including the opening degree of the mouth and eyes, the staring area of eyes, the times of blinks, and so on.

Smart terminal: After the data processing system analyzes the data on which users concen-

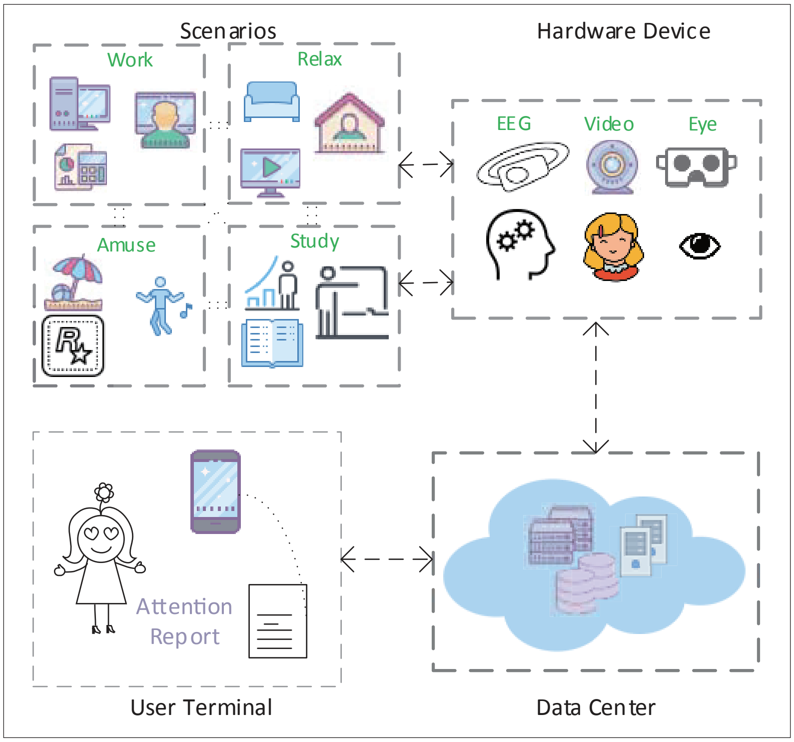

FIGURE 5. The architecture of an attention monitoring and development system.

trate, an analysis report is sent to the users' smart terminals. The users can know their attention in real time and properly adjust their activities at a proper time based on that.

Data center: This is the core of the attention detection and enhancement system. The data center's main function is to store, calculate, and transfer data. After receiving the EEG, external physiological features of the users, and behavior data in multiple scenarios, the algorithm deployed in the data center analyzes and processes the data. After evaluating the attention degree of users, the data center stores relevant key data as historical data so as to provide security for comprehensive and accurate evaluation of a user's general attention in the future. In addition, the data center sends an evaluation report to the user's smart terminal so that the user can view it.

Pilot Application

Based on the DeepFocus system, the student's attention detection and enhancement system is proposed. As shown in Fig. 6, the focus is on the following two contents:

Detect students' attention: Generally speaking, the degree of attention directly affects working efficiency. Therefore, it is necessary to evaluate the attention of students. The daily activities of students are small and controllable, which provides the possibility to collect real-time and attention-related data, and to comprehensively and accurately assess a student's attention. The student's attention evaluation is based on four basic scenarios: studying in class, independent study, relaxing, and amusement. The collected data types extend from internal and external physiological characteristics to external behavioral performance. In the above four scenarios, the brain wearable device, the eye tracker, and the camera are used to collect the physiological characteristic data and the external behavior performance data in real time. At the same time, the data is sent to the data center for analysis through the communication module. After the processing result is obtained, the student's attention report is sent to the smart terminal of the student, the parent, and the teacher. It will greatly promote the advance of students by comprehensively knowing their attention situation.

Enhance students' attention: Attention enhancement, which is completed via a smartphone application, is the goal of attention detection. To enhance students' attention, the main interaction method is an application on a smart terminal. First, students are guided to set a fixed time to study, relax, and have fun, and make a plan for the specific goals in each time period. In the beginning, some music that stimulates alpha brain waves is played to guide immersive work. When doing things, if the system detects that the user is inattentive, it uses a bell vibration to remind them. For some users who have obstacles to being attentive, some attention training methods can be set. For example, for users who are unable to concentrate due to excessive stress, the system can use the "breathing method" feature to help users calm down and improve their attention. There are also many progressive regular attention training tasks to increase attention during work. Through these methods, students' attention can be improved, in terms of study, relaxation, and amusement.

Open Issues and Future Work

In future work, the following two problems should be solved to enhance the system's performance.

Construction of a Big Data Health Platform: In future work, we need to collect a large number of users' multi-modal data and optimize the model based on big data. Second, we need to establish a user group psychological evaluation management platform, which stores all user data and analysis results in the cloud and provides more comprehensive services to user groups.

User Behavior Analysis: To further enhance the accuracy of the model, it is necessary to analyze the typical behavior of user groups and add user behavior data into the model as a new data resource. For example, if the user is detected writing, the degree of attention can be properly improved. If the user is in a daze, the degree of attention will be reduced.

Conclusion

In this article, completely new attention evaluation architecture is put forward. We analyze the problems of a conventional attention evaluation system in detail and put forward an improvement system. In the process of evolution from the historical system to DeepFocus, multi-modal data is introduced for integrated modeling. A different evaluation model is matched in different scenarios. Besides, we analyze the advantages of building such a multi-modal and multi-scenario evaluation model. The evaluation method is closer to the source of people's inward world. We collected EEG signals, facial images, eye tracker data, and extracted feature vectors of the three kinds of data. Finally, we use the logistic model to analyze the data and use the weight function to calculate the attention label. We believe that our architecture can evaluate users' attention status comprehensively and accurately.

Acknowledgments

This work is supported by the National Key R&D Program of China (2018YFC1314600), the National Natural Science Foundation of China (81771472), the Innovation Fund of WNLO, and NSFC No. 61821003. Prof. Di Wu's work was supported by the National Natural Science Foundation of China under Grant 61572538, Guangdong Special Support Program under Grant 2017TX04X148, and the Fundamental Research Funds for the Central Universities under Grant 19LGZD37. Prof. Zhongchun Liu is the corresponding author.

References

[1] G. V. Polanczyk *et al.*, "Annual Research Review: A Meta-Analysis of the Worldwide Prevalence of Mental Disorders in Children and Adolescents," *J. Child Psychology and Psychiatry*, vol. 56, no. 3, 2015, pp. 345–65.

[2] J. Zaletelj, "Estimation of Students' Attention in the Classroom from Kinect Features," *IEEE 2017 10th Int'l. Symp. Image and Signal Processing and Analysis*, Ljubljana, Slovenia, 2017, pp. 220–24.

[3] M. S. Hossain, and G. Muhammad, "Emotion Recognition Using Deep Learning Approach from Audio-Visual Emotional Big Data," *Info. Fusion*, vol. 49, 2019, pp.69–78.

[4] J. Lopez-Gil *et al.*, "Method for Improving EEG Based Emotion Recognition by Combining It with Synchronized Biometric and Eye Tracking Technologies in a Non-Invasive and Low Cost Way," *Frontiers in Computational Neuroscience*, 2016.

[5] J. Pacheco, "Analysis of Interaction Patterns-Attention," Universidade do Minho: RepositoriUM, 2014.

[6] C. Cimpanu *et al.*, "A Comparative Study on Classification of Working Memory Tasks Using EEG Signals," *IEEE Int'l. Conf. Control Systems and Computer Science*, 2017.

[7] J. Eriksson and L. Anna, *Measuring Student Attention with Face Detection: Viola-Jones versus Multi-Block Local Binary Pattern Using OpenCV*, Dissertation, 2015.

[8] N. Liu, C. Chiang, and H. Chu, "Recognizing the Degree of Human Attention Using EEG Signals from Mobile Sensors," *Sensors*, vol. 13, no. 8, 2013, pp. 10,273–86.

[9] J. Yiend, "The Effects of Emotion on Attention: A Review of Attentional Processing of Emotional Information," *Cognition and Emotion*, vol. 24, no. 1, pp. 3–47, 2010.

[10] M. Chen *et al.*, "Wearable Affective Robot," *IEEE Access*, vol. 6, 2018, pp. 64,766–76.

[11] C. Chen, J. Wang, and C. Yu, "Assessing the Attention Levels of Students by Using a Novel Attention Aware System Based on Brainwave Signals," *British J. Educational Technology*, vol. 48, no. 2, 2017, pp. 348–69.

[12] J. Olfers and G. Band, "Game-Based Training of Flexibility and Attention Improves Task-Switch Performance: Near and Far Transfer of Cognitive Training in an EEG Study," *Psychological Research*, 2017.

[13] M. Chen *et al.*, "Cognitive Information Measurements: A New Perspective", *Info. Science*, vol. 505, 2019, pp. 487–97.

[14] M. S. Hossain and G. Muhammad, "Emotion-Aware Connected Healthcare Big Data Towards 5G," *IEEE Internet of Things J.*, vol. 5, no. 4, 2018, pp. 2399–2406.

[15] B. Zhou *et al.*, "Learning Deep Features for Scene Recognition Using Places Database," *Advances in Neural Info. Processing Systems*, 2015.

Biographies

Min Chen [SM'09] (minchen2012@hust.edu.cn) has been a full professor in the School of Computer Science and Technology at Huazhong University of Science and Technology (HUST), Wuhan, China, since February 2012. He is Chair of the IEEE Computer Society STC on Big Data. His Google Scholars Citations reached 18,800+ with an h-index of 68. He received the IEEE Communications Society Fred W. Ellersick Prize in 2017 and the IEEE Jack Neubauer Memorial Award in 2019. His research focuses on cyber physical systems, IoT sensing, 5G networks, SDN, healthcare big data, and more.

Yong Cao (yongcao_epic@hust.edu.cn) is currently a Ph.D. student at the Embedded and Pervasive Computing (EPIC) Laboratory in the School of Computer Science and Technology, HUST. He received his Bachelor's degree in communication engineering from the College of Electrical Information at Sichuan University, China, in 2018. His research interest is focused on cognitive computing, machine learning, signal processing, and other topics.

Rui Wang (ruiwang2018@hust.edu.cn) is currently a Master's student at EPIC Laboratory in the School of Computer Science and Technology, HUST. She received her Bachelor's degree in computer science and technology from the College of Information Science and Engineering at Lanzhou University, China, in 2018. Her research interest is focused on cognitive computing and big data analysis among other areas.

Yong Li [M'09, SM'16] (liyong07@tsinghua.edu.cn) received his B.S. degree in electronics and information engineering from HUST in 2007, and his Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2012. During July to August 2012 and 2013, he was a visiting research associate with Telekom Innovation Laboratories and Hong Kong University of Science and Technology, respectively. During December 2013 to March 2014, he was a visiting scientist at the University of Miami. He is currently a faculty member of the Department of Electronic Engineering, Tsinghua University. His research interests are in the areas of networking and communications.

Di Wu [M'06, SM'17] (wudi27@mail.sysu.edu.cn) received his B.S. degree from the University of Science and Technology of China in 2000, his M.S. degree from the Institute of Computing Technology, Chinese Academy of Sciences, in 2003, and his Ph.D. degree in computer science and engineering from Chinese University of Hong Kong in 2007. He is a professor and the assistant dean of the School of Data and Computer Science at Sun Yat-sen University, Guangzhou, China. During 2007–2009, he worked as a postdoctoral researcher in the Department of Computer Science and Engineering, Polytechnic Institute of NYU, advised by Prof. Keith W. Ross. He was the co-recipient of the IEEE INFOCOM 2009 Best Paper Award.

Zhongchun Liu (zcliu6@whu.edu.cn) is currently the Luojia Scholar Distinguished Professor, chair of the Psychiatric Department,, Remmin Hospital of Wuhan University. He is Secretary General of the Chinese Psychiatrist Association (CPA) and Chair of the Precision Medicine Committee of CPA. He is mainly engaged in the diagnosis, treatment, and bio-marker research of mood disorders, and has been awarded five items for the National Key R&D Program of China and the National Natural Science Foundation.

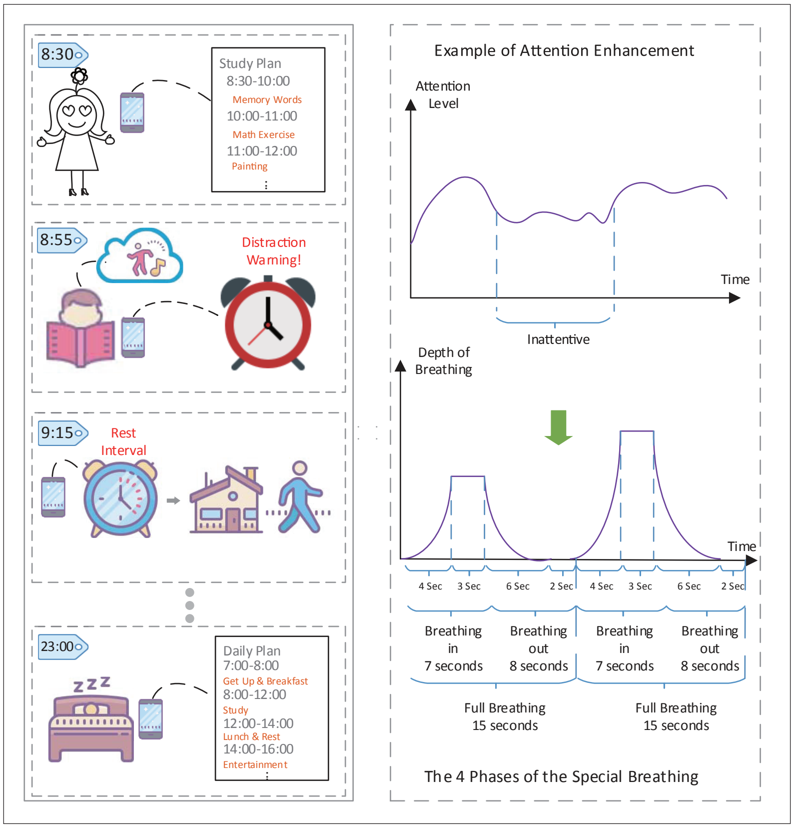

FIGURE 6. Attention enhancement case of a student.