Undisturbed Mental State Assessment in the 5G Era: A Case Study of Depression Detection Based on Facial Expressions

Minqiang Yang, Yu Ma, Zhenyu Liu, Hanshu Cai, Xiping Hu, and Bin Hu

Abstract

5G technology brings a comprehensive improvement in the network layer, which meets real-time, high-efficiency, and stability requirements in medical scenarios to a large extent, such as remote diagnosis and surgery. The heavy burden and severe impact of mental disorders make it desirable to find quantitative and automatic assessment approaches for early-stage detection of mental disorders. Facial expressions contain abundant emotional information, which may reflect abnormal mental states like anxiety and depression. With low latency and high bandwidth, 5G makes real-time monitoring of mental health feasible. In this article, a novel undisturbed mental state assessment prototype is proposed, which uses facial video streaming collected with 5G terminals to assess the mental state of a user in real time. A case study of depression detection using facial expressions has been developed based on the prototype. As a study case, we collected facial expression data from patients with depression and healthy people as control subjects. We extracted the transitional optical flow under stimulus feature and used the decision tree for classification. Results show that our depression assessment model is effective, and also reflect the feasibility and validity of our prototype.

Introduction

Applications based on fast, real-time, and reliable communication links provided by 5G technology, such as remote diagnosis, remote surgery, and remote monitoring, can bring tremendous innovation in the medical field. Its enhanced mobile broadband (eMBB) characteristic can meet the needs of 8K high-definition audio and video. Ultra-reliable low-latency communication (URLLC) and massive machine type communication (mMTC) could provide remote operation in real time and reliable feedback. Mobile healthcare shows significant influence and vitality, which plays an essential role in promoting the development of the healthcare industry. The continuous development of mobile and wearable devices in terms of capacity, performance, and intelligence provides new possibilities and opportunities for emotion awareness. Various sensors are embedded in smartphones and wearable devices, which can detect users' movements, capture images and videos, and record surrounding sounds without causing discomfort or intrusion to the users. With the development of mobile Internet, more and more devices are connected to the mobile network, and new services and applications are emerging. The surge of mobile data traffic, especially high-definition video, brings challenges to the bandwidth of the cellular network. 5G technology meets the growing demand for mobile traffic well, can improve the user experience, and facilitate applications that require real-time feedback with small latency.

As one of the most common healthcare issues, mental disorders have become a vital problem threatening human health and affecting people's thinking, emotion, and behavior. It is a general term used to describe a group of diseases mainly manifested as cognitive, emotional, volitional, and behavioral disorders. Most mental illnesses, such as schizophrenia, depression, anxiety, and obsessive-compulsive disorder, do not improve or heal without treatment, and may produce a series of adverse effects on people's physical and mental health, such as language barriers and interpersonal difficulties. For some safety-critical professions, such as bus drivers and air traffic controllers [1], a healthy mental state is essential and crucial during working hours. Drivers with abnormal mental states such as depression have a higher possibility of traffic accidents, which always cause injuries and deaths. It is necessary to assess and provide feedback on the mental state of bus drivers in real time to avoid disastrous consequences.

The clinical diagnosis of mental disorders mostly depends on extensive participation of experts, who will conduct a mental examination of the patient, combining some necessary auxiliary examinations and psychological tests. Depression is a kind of mental disorder that can severely disturb daily life and work, and sometimes has harmful effects on society. Most people with depression spend their whole lives in panic; some of them end their lives prematurely because of unbearable psychological pain and emotional disorder. At present, the assessment methods of depression are mainly self-report scale and clinician rating scale. Clinicians' scoring scales, such as the Hamilton Rating Scale for Depression (HRS-D), depend on clinical skills and professional knowledge. Still, the results

Digital Object Identifier:
10.1109/MWC.001.2000394

Minqiang Yang, Yu Ma, Zhenyu Liu, Hanshu Cai, Xiping Hu (corresponding author), and Bin Hu (corresponding author) are with Lanzhou University.

are easily influenced by subjective factors and cannot fully support the diagnosis of depression. Early intervention and effective management are crucial to patients, which always have a better curative effect than in later periods. The heavy burden and severe consequence of depression make a quantitative, accurate, objective, and automatic screening approach for early-stage detection even more urgent than it was before. The quantitive assessment of mental disorders has widely attracted researchers based on physiological signals and behavioral data. An improved algorithm for emotion-aware smart systems [2] is proposed to help medical staff effectively manage the treatment of patients who may develop postpartum depression through real-time data analysis.

In daily communication, only 7 percent of the information is transmitted through language, while 55 percent of the information is conveyed through facial expressions [3]. The typical symptoms of depression can be well described by a series of different manifestations of patients, such as disordered changes in facial expressions and changes in voice. These changes can reflect irritability, slow speech, and monotonous disorders of patients with depression. There are studies that focus on the feasibility of automatic depression detection with visual-based nonverbal behavior [4]. The severity of depressive symptoms is related to facial expressions over time, in [5], the authors proposed a method based on machine learning to measure depression symptom severity, which uses spoken language and 3D facial expressions as input data and achieves good results.

Facial expression data can be collected with a camera in a proper manner, which is a convenient way of assessing mental state that is worthy of investigation. The forthcoming 5G network applies new possibilities in terms of mental state healthcare, empowering relevant healthcare applications and assessment systems to monitor numerous users and process large amounts of data in real time. The demand for health information is growing fast in big data scenarios; with the development and applications of 5G technology in medical treatment, we can effectively collect a large amount of multi-modal health data such as video, audio, and eye movements.

In this article, a prototype of undisturbed mental health assessment in 5G scenarios is proposed, and based on the proposed prototype, a case study of depression detection using facial expressions has been investigated. The experimental paradigm was designed to collect facial videos of patients with depression and control groups. We extracted the transitional optical flow under stimulus (TOFS) feature to identify depression. The xperiment achieves good performance with recall of 0.93, which shows the validity and feasibility of the proposed prototype.

Mental Healthcare: Data-Driven and 5G

Data-Driven Approach to Mental State Assessment

The data-driven approach allows researchers to use massive, multi-dimensional data to establish a more comprehensive evaluation system. In [6], the authors introduced the concept of precision medicine, which is the basis for planning large cohort studies using genomics and phenotypes (physical and behavioral traits) to improve diagnosis and treatment in the medical application area. The idea is to combine a large amount of clinical data with other information on patients to improve the accuracy of the patient classification and achieve better treatment effectiveness. In addition to quantitative and automatic screening methods for mental disorders based on physiological signals such as electroencephalogram (EEG), electromyogram (EMG), and electrocardiogram (ECG), massive behavioral data like facial expressions and eye movements are also used to assess the mental state. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), one of the most common manifestations of depression is a persistent or cyclic depressed mood, which can be reflected by facial expressions. Video-based facial expression analysis can capture the nuances of facial expressions, and can be used in clinical studies of mental diseases. As it contains abundant emotional information and is easy to obtain, facial-expression-based affective computing has attracted many researchers to investigate topics such as emotional expression and micro-expression. In [7], the depression recognition based on visual cues and decision strategies of visual feature extraction, dimensionality reduction, classification, as well as regression methods are summarized. Also, they point out that the development of automated, objective depression assessment methods may be of value to study and apply in clinical practice. In [8], the authors proposed a computational framework of creating probabilistic expression profiles for video data, which may be of help to quantify facial expression nuances automatically between patients with neuropsychiatric disorders and healthy controls.

5G-Based Mobile Healthcare

The latest advances in wireless networks and big data technologies, as well as development in wearable computing and artificial intelligence, can effectively improve the computer network capabilities of the healthcare industry and ensure the smooth development of links, coordination, command, and scheduling. Technologies such as 5G, cloud computing, and the Internet of Things make real-time smart healthcare possible. Important health signals can be monitored in real time during daily human activities. The wireless body area networks (WBANs) in telemedicine systems can efficiently transmit critical medical information detected by body nodes [9]. The mobile application health programs have emerged with the development of intelligent sensing devices, which conduct real-time perception of the user's status and regular evaluation of the user's health status as well as feedback of relevant suggestions [10]. This type of application can be more intuitive for human-computer interaction and is simple to operate. It can also stimulate the enthusiasm of users and help users develop good living habits. In [11], the authors proposed a novel cooperative strategy application for mobile medical services, which increases the overall network throughput and improves the possibility of service delivery. 5G interconnects everything and provides the interoperability between networks and devices, which can ensure higher efficiency and securi-

The latest advances in wireless networks and big data technologies, as well as development in wearable computing and artificial intelligence, can effectively improve the computer network capabilities of the healthcare industry and ensure the smooth development of links, coordination, command, and scheduling.

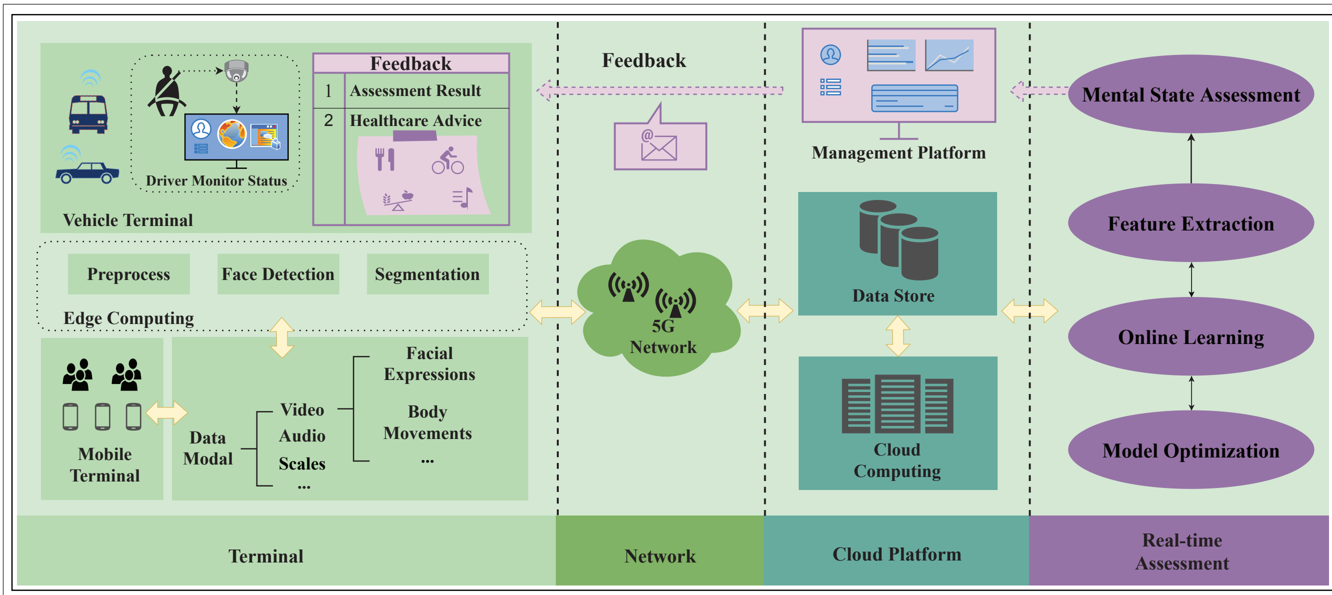

FIGURE 1. Prototype of mental healthcare in 5G scenarios.

ty, and thus increase the transmission rate and reduce latency. Some representative medical scenarios that could benefit from 5G technology were identified in [12]. Meanwhile, the article discussed some key technologies of 5G, such as medical network slicing, mobile edge computing, and heterogeneous network management by introducing typical cases of networked ambulances. An emotion-aware cognitive system (ECS) is presented in [13], which aims to provide a systematic approach to facilitate the different emotion-aware mobile applications in multi-channel cognitive radio ad hoc networks (mCRAHNs).

A Prototype of Mental Healthcare

As physiological data require expensive professional equipment and are time-consuming, it is more practical to assess mental state by visual tracking. We proposed a prototype (Fig. 1) of undisturbed mental state assessment in 5G scenarios, which is mainly composed of a terminal, a 5G network, a cloud platform, and an assessment model.

Facial expression contains rich emotional information that can effectively reflect the mental state of the subject in a natural state. It does not require the cooperation of the user when collecting the facial video stream, so an undisturbed mental state assessment can be realized based on the facial expression in the 5G scenario. Data acquisition terminals would be smart mobile devices or safety auxiliary devices of the vehicle, which can obtain a user's mental health information without disturbance by high-resolution cameras and transmit the information to a remote management platform by wireless technology. The main tasks of the terminal are collecting data, preprocessing, edge computing, and receiving information feedback of real-time assessment of the users' mental state, as well as some advice and necessary intervention such as some suggestions on mental health. For users of smartphones or wearable devices, multi-modal behavior data such as video, audio, and scales, which record users' facial expressions, body movements, and other information, are collected undisturbingly by the sensors with a healthcare application equipped on the smartphones. In our prototype, we specifically design a real-time status assessment scheme for drivers, which will continuously record a driver's facial expression video without disturbance through an auxiliary device like a driver monitor status (DMS) system and provide feedback from the cloud platform in time to improve the safety of driving. Edge computing provides useful computing resources on the edge of the network to maintain low-latency and real-time computing, and thus can process huge amounts of data collected from smart devices.

With the high bandwidth and low latency of 5G, it is possible to send high-definition facial expression video streams to the cloud platform and receive feedback from the cloud platform with small latency. It would meet the need for some scenarios that require real-time feedback of a user's mental state, such as a driver on the road.

On the cloud platform, effective features of multi-modal data such as video, audio, and scale are obtained by cloud computing. Those data flows are stored in the data center at the same time, which also helps the evaluation system conduct real-time evaluation of users' mental states and provide feedback consisting of assessment results and some suggestions like keeping a healthy and balanced diet for reasonable nutrition, listening to music to enhance relaxation, as well as developing a habit of regular moderate exercise and rest. The feedback will be sent to users' smart devices so as to effectively remind them to release pressure, reduce security risks, and improve the quality of life and work.

The system regularly pushes mental health assessment scales to users for data labeling. Classification is conducted in the long term with a large number of samples, which helps the evaluation model to be continuously optimized via online learning. With a large number of multi-modal data collected from smart devices uploaded to the cloud platform, the evaluation model we built makes use of those data to improve the assessment capability. It is worth noting that the mental state assessment of drivers needs to be carried out in real time because they need to concentrate at all times. The cloud computing platform needs to focus on the user's mental state data in real time to avoid safety incidents effectively. The mental health assessment scales can be collected through smart terminals, and long-term facial state monitoring can be performed. As the experiment progresses, the volume and modal of data will be expanded to ensure that our mental health assessment model achieves higher reliability.

Case Study: Depression Recognition

The prototype we propose describes mental state assessment in the 5G era from four stages in which the most critical part is the mental state assessment model. We need to figure out features that describe mental state as well as optimal parameters to facilitate the model which determines the accuracy of mental assessment in real time. We can carry out a series of case studies based on our prototype, such as the real-time assessment of the driver's mental state by using facial expressions, and the assessment of the public's mental state by using smartphones and other devices. The following section introduces our first case study to detect depression using video of facial expressions. We designed an experimental paradigm, and recruited depressed patients and healthy control groups according to inclusion criteria. Features of their facial expressions were extracted, and the assessment model was built to detect whether a subject is depressed, and experiment results showed the validity of our mental state assessment.

Paradigm

DSM-V believes that the existence of depression can be inferred from facial expressions and behaviors, which has attracted the attention of researchers to identify depression through facial expressions. The change of an individual's mental state can be reflected in his/her expressions, which stay natural for a long period; thus, emotional stimulus is needed in order to obtain transitory expression nuances. The method for selecting stimulus material in our experimental paradigm referred to the design process of the Database for Emotion Analysis using Physiological Signals (DEAP) and SJTU Emotion EEG Dataset (SEED). We choose fragments from high arousal Chinese movies containing emotional emphasis as stimulus materials to design experiments. Each stimulus video is about 2 minutes; at the beginning of experiment, there was a voice prompting the subject to be ready to watch the next video. After watching the video, the subject had 2 minutes to rest and prepare for the next one. This period was also used for the subject to adjust the emotion. The experiment was performed in a quiet and clean room, and lasted about 10 minutes for one subject. All three experimental tasks were completed in a comfortable chair, and videos of the facial expressions of subjects were recorded with Logitech C1000E cameras while watching the movie clips.

Subjects

Written informed consent was obtained from all subjects before the experiment. The local ethics committee approved consent forms and studies designed for biomedical research at the Second People's Hospital of Gansu Province, following the Code of Ethics of the World Medical Association (Declaration of Helsinki). All patients with major depressive disorder (MDD) received a structured Mini-International Neuropsychiatric Interview (MINI) that met the diagnostic criteria for major depression based on the DSM-IV. For all subjects, the inclusion criteria were that the age should be between 18 and 55 years old, with primary or higher education level, and the exclusion criteria were alcohol or psychotropic drug abuse or dependence in the past year, and women who were pregnant, lactating, or taking birth control pills. For MDD patients, the inclusion criteria were the diagnostic criteria of MINI for depression, the Patient Health Questionnaire-9 item (PHQ-9) score of subjects was greater than or equal to 5, and no psychotropic drug treatment performed in the last two weeks. For MDD patients, the exclusion criteria were having mental disorders or brain organ damage, having a severe physical illness, and extreme suicidal tendencies. For the control group, the exclusion criteria were personal or family history of mental disorders. Sixty-six subjects with 34 outpatients diagnosed with depression by at least one clinical psychiatrist from The Second People's Hospital of Gansu Province, China, and 32 healthy controls were recruited. All subjects had normal or corrected-to-normal vision. Excluding some invalid data such as part of the face disappearing in the video for a long time during the acquisition process and frequent body movements covering the subject's face, we have 30 outpatients (10 males, 20 females; 18–55 years old) with depression, as well as 30 healthy controls (16 males, 14 females; 18–55 years old).

Preprocess and Locating Effective Video Segment

Each video clip we acquire from a subject lasts about 2 minutes and records the subject's facial expressions during the experimental process. Each video has been cut into an image sequence only containing the entire face of the subject; images caused by sudden face movement are deleted. With the continuous affective video stimulus, subjects show different expression changes and emotional intensity. The change of facial expression does not show up obviously in part of the video because it takes some time to make an impression and empathy with the subjects by watching the emotional stimulus movie. It is meaningless to directly extract optical flow features from the whole video because a long-term accumulative optical flow contains too much noisy information, which inspires us to find significant short video clips. This process is shown in Fig. 2. We use sliding windows to find more effective video segments, and set the size of the sliding window at 100 frames and each video segment at 300 frames.

Feature Extraction and Classification

The optical flow uses the time domain changes of pixels in image sequence and the correlation between adjacent frames to find the correspondence between the previous frame and the current frame, and calculate the motion information of an object. We obtain the MDMO [14] feature of our video clip, which, based on the face action

The prototype we proposed describes mental state assessment in the 5G era from four stages in which the most critical part is mental state assessment model. We need to figure out features that describe mental state as well as optimal parameters to facilitate the model which determines the accuracy of mental assessment in real time.

The high recall rate indicates that the proposed approach could significantly reflect the facial change of the depressive subjects. In addition, we found that among the positive, neutral, and negative stimuli, the classification model performed best under the negative stimulation, which reflects the negative emotional attention bias of depressed people.

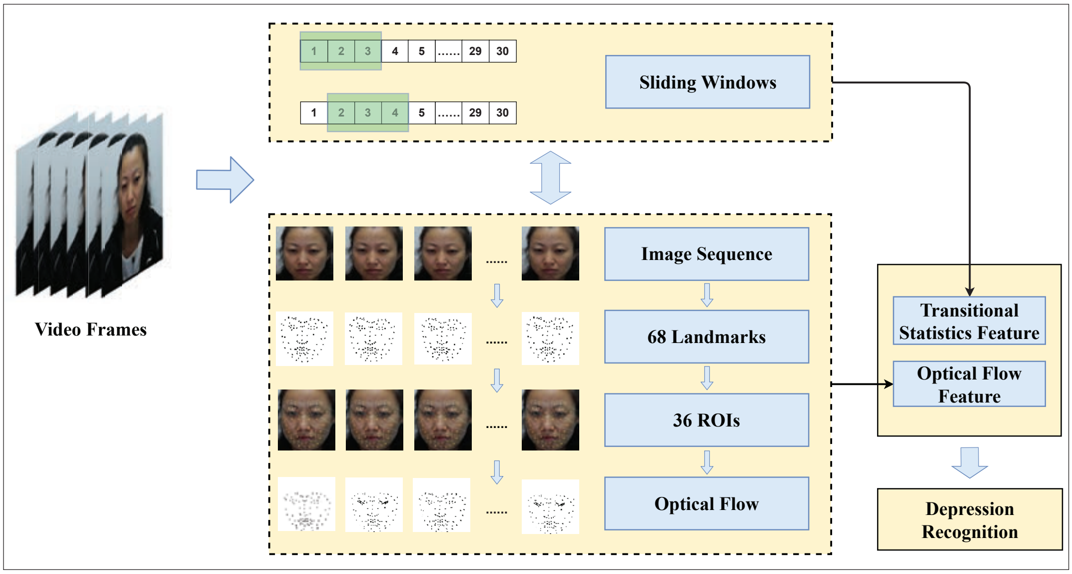

FIGURE 2. Feature extraction and depression recognition.

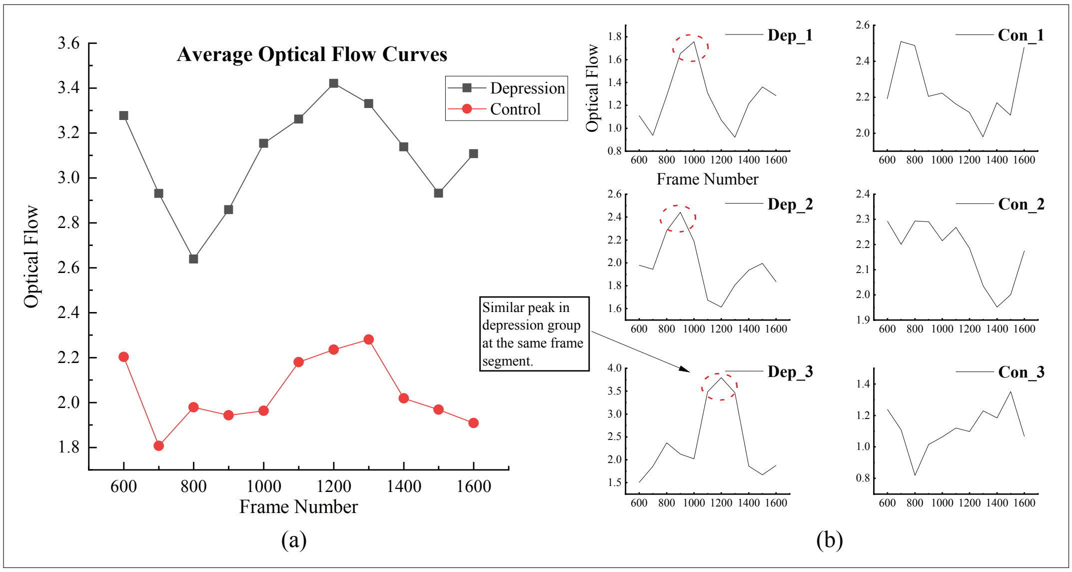

FIGURE 3. a) Average optical flow curves of all subjects from the depression group as well as the control group; b) some of the subjects' average optical flow curves in two groups.

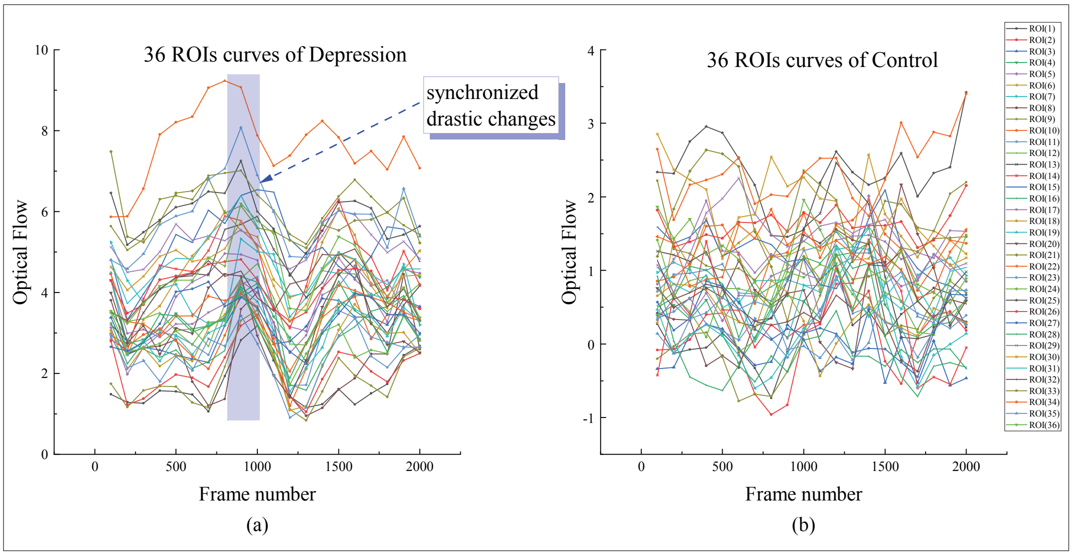

FIGURE 4. a) A depressed subject's optical flow curves of 36 ROIs; b) a subject's same curves from the control group.

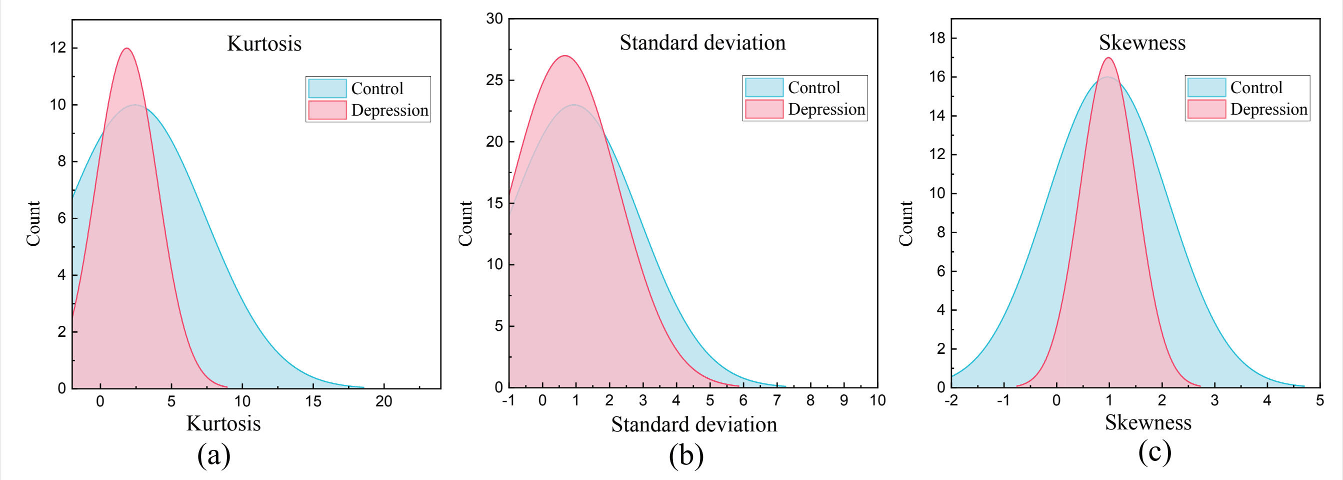

FIGURE 5. Statistics features: a) kurtosis; b) standard deviation; c) skewness.

coding system, uses 66 face landmarks to divide the face area of each frame into 36 regions of interest (ROIs), and calculate the optical flow of every point between each frame (Fig. 2). The feature can effectively reflect the change of facial ROIs between two frames at the pixel level, which include the magnitude and direction of flow, so as to determine the most effective frames of our video clips where a subject's expression changes significantly. Moreover, statistical characteristics of sliding windows in the time-frequency domain — mean, variance, standard deviation, skewness, and kurtosis — are extracted and used to measure the symmetry and distribution of optical flow in the whole process.

We obtain the TOFS feature based on statistical characteristics of the time-frequency domain as well as optical flow in effective video frames, and use decision tree for depression recognition. The grid search method is used to obtain the optimal parameters, and leave-one-out cross-validation is conducted to assess the prediction performance of our classification model. We achieved the best performance of depression recognition by setting the value of max_depth, min_samples_split, and min_samples_leaf as 3, 11, and 2.

Patterns in the Feature

Figure 3a shows average optical flow curves of all subjects from the depression group and the control group, and some of the subjects' average optical flow curves in two groups are separately shown in Fig. 3b. We also randomly selected one subject from the depression group and one from the control group and drew the sliding window curves of the facial optical flow of their 36 ROIs (Figs. 4a and 4b). From these two sets of images, significant changes of optical flow around frame 1000 occurred in the depression group, which are not obvious in the control group.

We calculate the time and frequency domain characteristics of these curves (Fig. 5) to measure the changing trend of optical flow. As can be seen from the figure, the standard deviation, skewness, and kurtosis of the depression group are all smaller than those of the control group, which indicates that patients with depression show more pronounced and similar facial expressions changes under stimulus than the control group.

Feature	Accuracy	Recall	Precision	F1 score
TOFS	0.7833	0.9333	0.7179	0.8115
MDMO	0.6500	0.6000	0.6667	0.6315
LBP	0.5500	0.6667	0.5405	0.5970
HOG	0.6167	0.7000	0.6000	0.6461
CNN	0.6000	0.5217	0.7058	0.5999

TABLE 1. Accuracy, recall, precision, and F1 score of TOFS, MDMO, LBP, HOG, and CNN.

To evaluate the proposed TOFS feature and compare it with other methods, we also select the classical Local Binary Pattern (LBP), Histograms of Oriented Gradients (HOG) feature, MDMO feature, as well as convolutional neural networks (CNNs) for comparison. We obtain the features of the same video clips for these algorithms and seek the corresponding optimal parameters. The best performance of depression recognition was achieved by setting the value of max_depth, min_samples_split, and min_samples_leaf as: LBP (3,5,10), HOG (3, 2, 15) and MDMO (3, 7, 15). The performance metrics are shown in Table 1.

As can be seen from the performance metrics, our evaluation finally reaches a recall rate of 0.93, which significantly surpasses the precision and might be competent in the preliminary screening of depression. The high recall rate indicates that the proposed approach could significantly reflect the facial change of depressive subjects. In addition, we found that among the positive, neutral, and negative stimuli, the classification model performed best under the negative stimulation, which reflects the negative emotional attention bias of depressed people [15].

Research Challenge and Open Issue

We have conducted a case study using our prototype and have some interesting findings, such as that subjects in the depression group have a regular trend of optical flow curves, and the TOFS feature we extracted from videos of facial expressions can effectively detect depression. In combination with four stages in our prototype and the whole experimental process of the first case

The finding from our study that patients with depression have an attention bias toward the negative stimulus is consistent with the earlier study. With the technical feasibility and potential needs, it is worthy to explore and implement the proposed prototype. However, more open issues and challenges need to be solved in the future.

study, we believe the following aspects are worthy of further exploration:

Robustness of Assessment Model

The evaluation results of our depression classification model achieve relatively high recall. There are several causes from the perspective of experiment process and algorithm. One is that some subjects in the control group show similar reactions to the experimental stimulus as depressed patients in the frame segment, and the optical flow curves peak is similar to those of depression group, which could lead the classifier to misjudge these individuals as depressed patients. The other is that our algorithm may be interfered with by some subjects' head movements, which may lead the TOFS feature to not effectively capture the nuances of facial expressions under experiment stimulus. Concerning the relatively low precision, we need to improve the algorithm and explore more specific features to facilitate our assessment model in the future.

Paradigm-Free Environment

Mobile medical devices or wearable diagnosis systems provide new opportunities for more convenient healthcare. We conducted a case study of depression detection using facial expressions, and our experimental result demonstrated the feasibility of the proposed prototype in practical scenarios. Considering building an effective mental state assessment model with general applicability, it is vital to create a paradigm-free environment that could be used to collect data without simulation continuously. In addition, the body movement, environmental brightness, and angle of visibility may lead to data inconsistency, which needs to be calibrated. This will also bring new challenges to be addressed.

Multi-Modal Behavioral Data

A large number of sensor systems, smartphones, and wearable devices can passively collect useful data streams related to the human environment and behavior, which can assess people's current mental state efficiently. How to extract effective features from these raw data streams is a fundamental problem to resolve despite many earlier studies investigating this topic. One more challenge would be the fusion of multi-modal heterogeneous data, such as data related to movement and social activities, namely video, audio, operating habits, GPS data, inertial sensor data, and other data.

Privacy in the Era of 5G Healthcare

As a new direction in the development of the medical industry, 5G healthcare has successively launched remote diagnosis, remote surgery, and 5G medical ambulances through cooperation with operators and equipment vendors, bringing new medical experiences and services to the public. In our proposed prototype, lots of private data will be collected, including psychological data, physiological data, and behavioral data, which brings legal and ethical challenges in the privacy of patients. A series of problems require our attention, such as how to protect the privacy of patients, including the entire life cycle of data acquisition, communication, process, and storage.

Conclusion

Facial expression data contains rich emotional information, but changes in facial actions are usually subtle and difficult to capture with computer vision. To meet the requirements of self-healthcare and mental state management for safety-critical positions, we propose a prototype of undisturbed mental state assessment in 5G scenarios, which can assess users' mental state and provide feedback in real time. We introduce the use cases of real-time mental state monitoring of drivers and mental state assessments based on smart mobile terminals in the prototype and carry out a case study of using facial expressions to detect depression. As a case study, the experimental paradigm of depression recognition is designed to acquire data from recruited subjects, and we finally achieved a depression assessment model with good performance. The finding from our study that patients with depression have an attention bias toward negative stimuli is consistent with [15]. With itds technical feasibility and potential needs, it is worth exploring and implementing the proposed prototype. However, more issues and challenges need to be solved in the future.

Acknowledgments

This work was supported in part by the National Key Research and Development Program of China (Grant No. 2019YFA0706200) and in part by the National Natural Science Foundation of China (Grant No.61632014, No.61627808).

References

[1] Z. Ning *et al.*, "A Cooperative Quality-Aware Service Access System for Social Internet of Vehicles," *IEEE Internet of Things J.*, vol. 5, no. 4, Aug. 2018, pp. 2506–17.

[2] Mário W. L. Moreira *et al.*, "Postpartum Depression Prediction through Pregnancy Data Analysis for Emotion-Aware Smart Systems," *Info. Fusion*, vol. 47, May 2019, pp. 23–31.

[3] A. Mehrabian and J. A. Russell, *An Approach to Environmental Psychology*, MIT, 1974.

[4] L. Wen *et al.*, "Automated Depression Diagnosis Based on Facial Dynamic Analysis and Sparse Coding," *IEEE Trans. Info. Forensics Security*, vol. 10, no. 7, July 2015, pp. 1432–41.

[5] A. Haque *et al.*, "Measuring Depression Symptom Severity from Spoken Language and 3D Facial Expressions," arXiv preprint arXiv:1811.08592, 2018.

[6] T. R. Insel and B. N. Cuthbert, "Brain Disorders? Precisely," *Science*, vol. 348, no. 6234, 2015, pp. 499–500.

[7] A. Pampouchidou *et al.*, "Automatic Assessment of Depression Based on Visual Cues: A Systematic Review," *IEEE Trans. Affective Comp.*, vol. 10, no. 4, 2019, pp. 445–70.

[8] P. Wang *et al.*, "Automated Video-Based Facial Expression Analysis of Neuropsychiatric Disorders," *J. Neuroscience Methods*, vol. 168, no. 1, Feb. 2008, pp. 224–38.

[9] E. Ibarra *et al.*, "QoS-Aware Energy Management in Body Sensor Nodes Powered by Human Energy Harvesting," *IEEE Sensors J.*, vol. 16, no. 2, 2016, pp. 542–49.

[10] B. M. Silva *et al.*, "SapoFitness: A Mobile Health Application for Dietary Evaluation," *Proc. 13th IEEE Int'l. Conf. e-Health Net. Appl. Serv.*, June 2011, pp. 375–80.

[11] B. M. C. Silva *et al.*, "A Novel Cooperation Strategy for Mobile Health Applications," *IEEE JSAC*, vol. 31, no. 9, supplement, Sept. 2013, pp. 28–36.

[12] G. Cisotto, E. Casarin, and S. Tomasin, "Requirements and Enablers of Advanced Healthcare Services over Future Cellular Systems," *IEEE Commun. Mag.*, vol. 58, no. 3, Mar. 2020, pp. 76–81.

[13] X. Hu *et al.*, "Emotion-Aware Cognitive System in Multi-Channel Cognitive Radio Ad Hoc Networks," *IEEE Commun. Mag.*, vol. 56, no. 4, Apr. 2018, pp. 180–87.

[14] Y. Liu *et al.*, "A Main Directional Mean Optical Flow Feature for Spontaneous Micro-Expression Recognition," *IEEE Trans. Affective Comp.*, vol. 7, no. 4, 2016, pp. 299–310.

[15] I. H. Gotlib *et al.*, "Attentional Biases for Negative Interpersonal Stimuli in Clinical Depression," *J. Abnormal Psychol.*, vol. 113, no. 1, 2004, pp. 121–35.

Biographies

Minqiang Yang [M] (yangmq@lzu.edu.cn) is an engineer and Ph.D. candidate with the Gansu Provincial Key Laboratory of Wearable Computing in the School of Information Science and Engineering, Lanzhou University, China. His current research interests include affective computing, image processing, machine learning, and automatic depression detection.

Yu Ma (mayu2019@lzu.edu.cn) received her B.S. degree from Nankai University, Tianjin, China, in 2019. She is currently pursuing an M.S. degree in computer technology with the School of Information Science and Engineering, Lanzhou University. Her research interests include affective computing, machine learning, and automatic depression detection.

Zhenyu Liu [M] (liuzhenyu@lzu.edu.cn) received his Ph.D. degree from Lanzhou University in 2017. He is a lecturer in the School of Information Science and Engineering, Lanzhou University. His current research interests include affective computing, image processing, speech processing, machine learning, and automatic depression detection.

Hanshu Cai [M] (hcai@lzu.edu.cn) is an associate professor at the Gansu Provincial Key Laboratory of Wearable Computing, School of Information Science and Engineering, Lanzhou University. His research interests include affective computing, neurofeedback, semantic web, ontologies, and computerized cognitive behavioral therapy.

Xiping Hu [M] (huxp@lzu.edu.cn) received his Ph.D. degree from the University of British Columbia, Vancouver, Canada. He is currently a professor at Lanzhou University. He has more than 100 papers published and presented in prestigious conferences and journals. He was the co-founder and CTO of Bravolol Ltd., Hong Kong, a leading language learning mobile application company with over 100 million users listed as the second language education platform globally.

Bin Hu [SM] (bh@lzu.edu.cn) received his Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Science in 1998.Since 2008, he has been a professor in the School of Information Science and Engineering, Lanzhou University. He was a guest professor at ETH Zurich, Switzerland, in 2011.His research interests include pervasive computing, computational psychophysiology, and data modeling.