IAES International Journal of Robotics and Automation (IJRA)
Vol. 12, No. 1, March 2023, pp. 20~28
ISSN: 2722-2586, DOI: 10.11591/ijra.v12i1.pp20-28

Microexpression recognition robot

Yi-Chang Wu, Yao-Cheng Liu, Chieh Tsao, Ru-Yi Huang
Department Forensic Science Division, Investigation Bureau, Ministry of Justice, Taipei, Taiwan

Article Info

Article history:

Received Feb 17, 2022
Revised Oct 3, 2022
Accepted Nov 18, 2022

Keywords:

Artificial intelligence
Deception detection
Microexpression
Real-time recognition

ABSTRACT

Following the development of big data, the use of microexpression technology has become increasingly popular. The application of microexpressions has expanded beyond medical treatment to include scientific case investigations. Because microexpressions are characterized by short duration and low intensity, training humans to recognize their yields poor performance results. Automatically recognizing microexpressions by using machine learning techniques can provide more effective results and save time and resources. In the real world, to avoid judicial punishment, people lie and conceal the truth for a variety of reasons. In this study, our primary objective was to develop a system for real-time microexpression recognition. We used FaceReader as the retrieval system and integrated the data with an application programming interface to provide recognition results as objective references in real-time. Using an experimental analysis, we also attempted to determine the optimal system configuration conditions. In conclusion, the use of artificial intelligence is expected to enhance the efficiency of investigations.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Yi Chang Wu
Forensic Science Division, Investigation Bureau, Ministry of Justice
No. 74, Zhonghua Rd., Xindian Dist., New Taipei City 231, Taiwan
Email: shintenwu@gmail.com

1. INTRODUCTION

Many methods exist for determining whether someone is lying. Physiological responses such as body temperature, blood pressure, heart rate, and respiratory rate, which are detectable by a lie detector, and speech, tone, facial expressions, and body language can all serve as a basis for judgment. However, the polygraph recording process is an invasive procedure, and its actual application is limited by the consent of the person concerned and its immediate effect. When a person lies, the brain sends signals to the facial muscles, which makes lying difficult without giving away hints [1], [2]. Therefore, microexpressions have been viewed as a possible indicator of lying [3]. Two psychologists, Ekman and Friesen, suggested that microexpressions occur when people attempt to conceal their emotions, which is why they proposed the facial action coding system (FACS), a system for coding facial actions [1], [4].

The majority of microexpressions are characterized by subtle changes in facial features and muscles that are rapid, difficult to control, and difficult to feign [5]. Therefore, microexpressions are regarded as crucial indicators for lie detection [6]. During an interrogation, the personnel of the Ministry of Justice Investigation Bureau typically evaluate the credibility of statements by observing the interrogee's face and body movements. However, the duration of microexpressions ranges from approximately 1/25 to 1/3 s [7], and the precise definition of such duration varies [8]. In addition, the variations of microexpressions are extremely small and require careful observations to discern [9]. Although training can improve the ability to detect microexpressions [10], the recognition success rate is low [7], [8]

While investigating a crime or collecting evidence, the police may contact a suspect for questioning. Nevertheless, the method of questioning used must adhere to the applicable regulations. Illegal and improper interrogation techniques, such as threatening, coercion, deception, and illegal detention, should be avoided lest the obtained evidence be rendered useless. The purpose of such interrogation is to coax the interrogee into telling the truth and to glean information from their statements to further clarify the case or uncover clues. The key to solving many cases lies in the statements of the person being questioned. Therefore, interrogation is regarded as one of the most crucial methods used by investigators to discover the facts of a crime.

If the investigators are unable to confirm the veracity of a statement or determine what the interrogee is thinking about, the investigation may be jeopardized or may steer into the wrong direction, which not only wastes time and resources but also allows criminals to escape justice. Therefore, under the premise of legal law enforcement, investigators must seek effective interrogation techniques. The rest of this paper is organized as. Section 2 describes our research methodology, including how we collected the data and extracted useful information for further analysis. Section 3 presents our findings and discussion. Finally, section 4 shows our conclusions.

2. METHOD

The FACS uses 44 action units (AUs) to describe facial expressions [11]. Multiple types of AUs can be combined freely, allowing for a variety of expressions. For instance, the combination AU4 (lower eyebrow) + AU5 (upper eyelid) + AU24 (lips pressed together) corresponds to the emotional state of anger. The facial motion coding system can objectively, precisely, and comprehensively describe facial expressions. In the past, the high cost of labeling has severely hampered the progress of AU identification research. This is because different AUs are distributed in various regions of the face, expressing subtle changes of different intensities. Currently, facial recognition and model learning technologies based on artificial intelligence are used to instantly label and publish AU units, allowing for the recognition and evolution of microexpressions. Both emotional facial action coding system (EMFACS) [11] and facial action coding system affect interpretation dictionary (FACSAID) [12] consider only emotion-related facial actions. Table 1 depicts the seven combinations of emotional AUs, and Table 2 depicts the facial muscles responsible for microexpressions [13].

Table 1. Facial expression calculation from single AUs

Emotion	Action units
Happiness	6+12
Sadness	1+4+15
Surprise	1+2+5B+26
Fear	1+2+4+5+7+20+26
Anger	4+5+7+23
Disgust	9+15+17
Contempt	R12A+R14A

Table 2. Main action codes

AU number	FACS name	Muscular basis
1	Inner brow raiser	Frontalis (pars medialis)
2	Outer brow raiser	Frontalis (pars lateralis)
4	Brow lowerer	Depressor glabellae, depressor supercilii, corrugator supercilii
5	Upper lid raiser	Levator palpebrae superioris, superior tarsal muscle
6	Cheek raiser	Orbicularis oculi (pars orbitalis)
7	Lid tightener	Orbicularis oculi (pars palpebralis)
9	Nose wrinkler	Levator labii superioris alaeque nasi
10	Upper lip raiser	Levator labii superioris, caput infraorbitalis
11	Nasolabial deepener	Zygomaticus minor
12	Lip corner puller	Zygomaticus major
13	Sharp lip puller	Levator anguli oris (also known as caninus)
14	Dimpler	Buccinator
15	Lip corner depressor	Depressor anguli oris (also known as triangularis)
16	Lower lip depressor	Depressor labii inferioris
17	Chin raiser	Mentalis
18	Lip pucker	Incisivii labii superioris and incisivii labii inferioris
20	Lip stretcher	Risorius w/ platysma
22	Lip funneler	Orbicularis oris
23	Lip tightener	Orbicularis oris
24	Lip pressor	Orbicularis oris
25	Lips part	Depressor labii inferioris, or relaxation of mentalis or orbicularis oris
26	Jaw drop	Masseter; relaxed temporalis and internal pterygoid
27	Mouth stretch	Pterygoids, digastric
28	Lip suck	Orbicularis oris

In this study, we focused on developing a microexpression recognition system for the personnel of the Ministry of Justice Investigation Bureau. The human face produces more than just microexpressions. When microexpressions occur, the human face may also change in a variety of ways. Therefore, isolating irrelevant information and extracting important information are essential [14], [15]. In addition, compared with manual detection, the advantage of using artificial intelligence (AI) to identify microexpressions is that as long as the camera objectively captures the identified object, the computer will have obtained the relevant data for processing [16]. In other words, by installing cameras, a certain level of recognition can be achieved [16], [17]. In this study, we used FaceReader, professional facial expression analysis software, for microexpression recognition.

In addition to the main recognition software, the software used in this study offers an expression item comparison analysis module, an expression muscle unit analysis module, and an expression remote Photoplethysmography (rPPG) module. The most recent version of the software also incorporates deep neural networks, meaning that a deep face model is added to the original active appearance model (AAM) [18]. To construct the face module, the AAM thoroughly considers the shape, texture, and local features of the face [15]. Subsequently, after face localization, the AAM starts operating and extracts feature data based on 500 feature points [19]. If the AAM is rendered inoperable because of obstacles such as face occlusion, the deep face model continues to function. Currently, this software can recognize eight basic facial expressions, facial changes, and physical features. The system also offers a variety of convenient functions such as the ability to present recognition results and relevant data in different forms depending on the requirements and to mark the selected categories. The open application programming interface facilitates the integration of the system data into the recognition system, and the rapid analysis function aids in the sample collection and labeling process, which is highly useful in research [20].

An AU labeling expert requires at least 30 minutes to analyze a one-minute video of a human face [21], [22]. The advantage of machine learning is that it can replace the time-consuming manual labor. The emotion recognition processing framework of this system consists of data collection (video), data input, data labeling (AUs), machine learning, and recognition output. A flowchart of the proposed system is depicted in Figure 1.

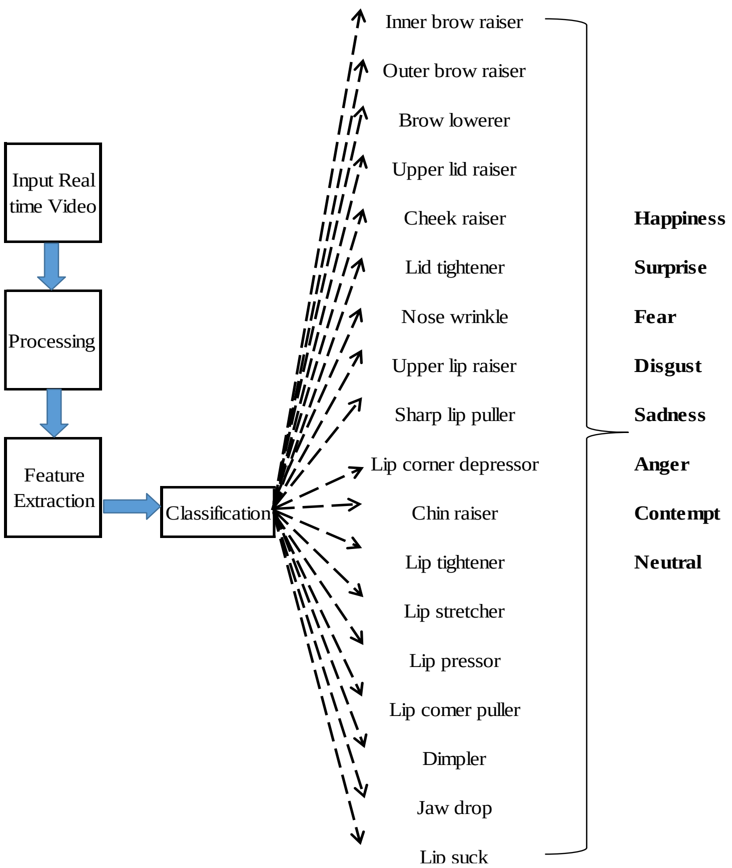

Figure 1. System flowchart

3. RESULTS AND DISCUSSION

The self-developed system proposed in this study is primarily intended to provide investigation bureau personnel with the instant emotional responses of the respondents, thereby enhancing the efficiency of case investigations. According to psychologists, when microexpressions occur, facial muscle movements are not directly related to individual characteristics, such as gender, age, or ethnicity [23], and the system is configured for use in an indoor confined space with no denoising concerns. Thus, during data labeling, only facial expressions are considered for emotional analysis. In addition, considering the usage sites, five photography conditions are planned to determine the optimal analysis configuration.

3.1. Real-time facial expressions and action recognition system

Given the convenience and practicability of actually using the recognition system, we built a three-part system consisting of a testing theme documentation system, a dynamic microreaction recording system, and a retrospective analysis system.

3.1.1. Validation of the theme documentation system

The system can be used to perform theme addition, modification, and deletion as shown in Figure 2. The file creation date is automatically generated by the system and cannot be manually entered or altered. The ideation and design mode of the questions is based on a single respondent as shown in Figure 3. The user can predict the respondent's emotional response toward a question, and multiple selections can be made.

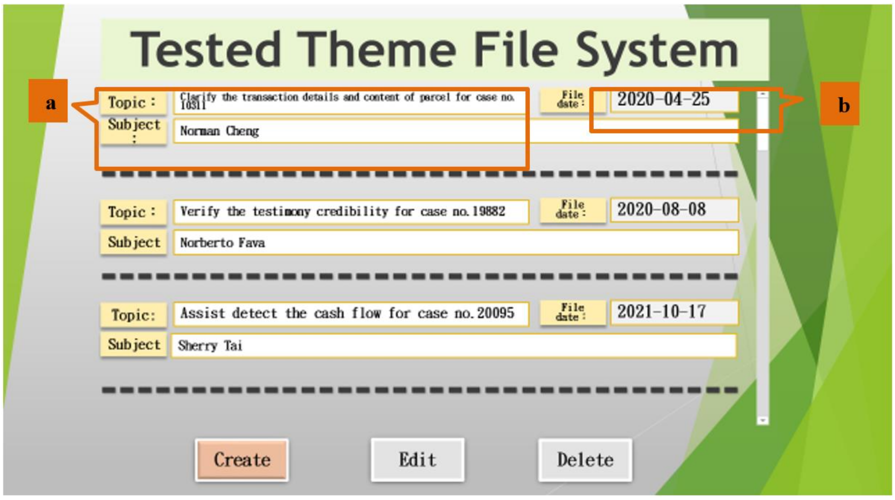

Figure 2. Tested theme file system

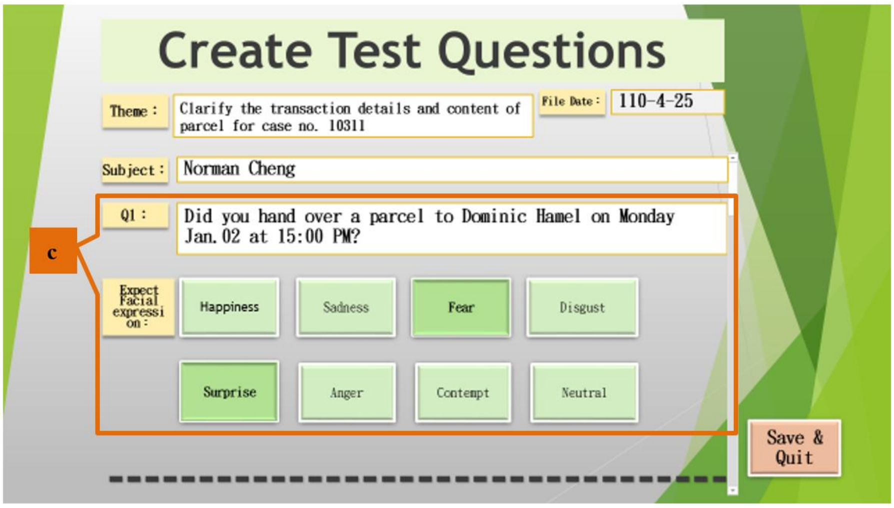

Figure 3. Creating test questions (create/edit)

3.1.2. Dynamic microreaction recording system

When a user accesses the system interface as shown in Figure 4 and clicks the button in the upper-right corner, the system starts recording and analyzing emotional responses and body movements. In the bottom-right corner of the screen, the user can pause or terminate the recording of the current test item. The test time is automatically generated by the system and cannot be manually entered or modified. The system logs the interrogation start time and the duration of the current interrogation. The purpose is to remind the user to consider the physical and mental effects of interrogation duration on the interrogee. The left-hand side records the interrogee's current emotional response and its intensity as shown in Figure 5. The right-hand side displays the interrogee's predicted emotional response on the basis of predefined questions. When the system detects a microexpression, a red frame appears around the label and a warning sound is played. The left-hand side indicates the number of times that the microexpression has occurred, and the expression indicator light lasts for 3 s. When unexpected expressions appear, a special warning is displayed in the form of a blue frame and a unique sound effect.

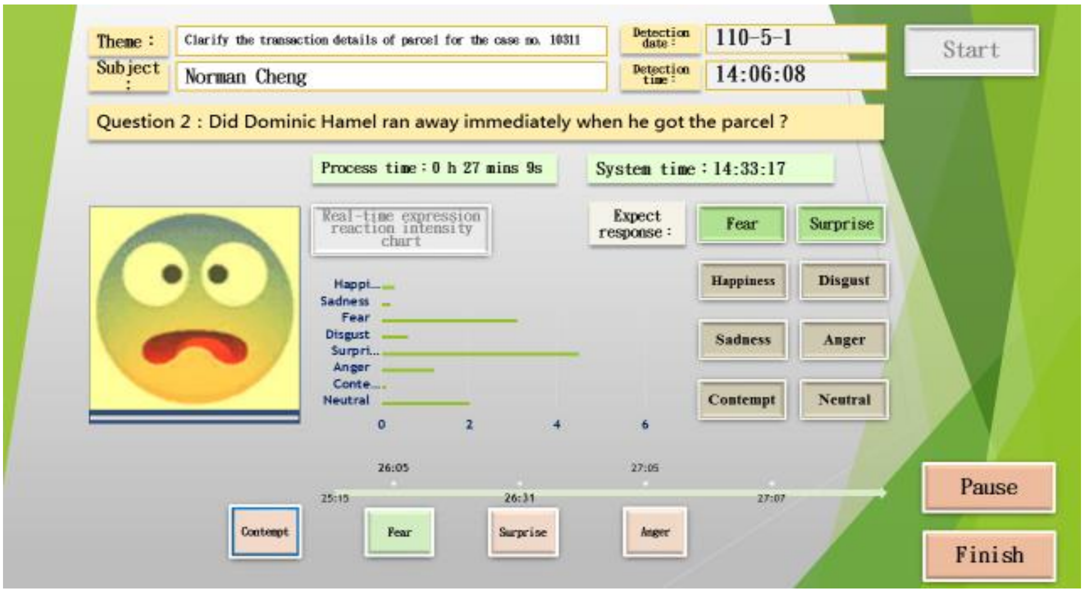

Figure 4. Screenshot of the real-time analysis interface

Figure 5. Analysis results of real-time expression reaction intensity

3.1.3. Review and analysis system

In Figure 6, The upper-left corner displays the frequency and time points at which the emotion has appeared. If the user clicks on the time stamp, the video skips to the corresponding section (5 s of an introduction clip). The bottom-left corner displays the current playback speed of the video (×0.5, ×1, or ×2). The right-hand side displays the predicted emotional response of the interrogee on the basis of predefined questions.

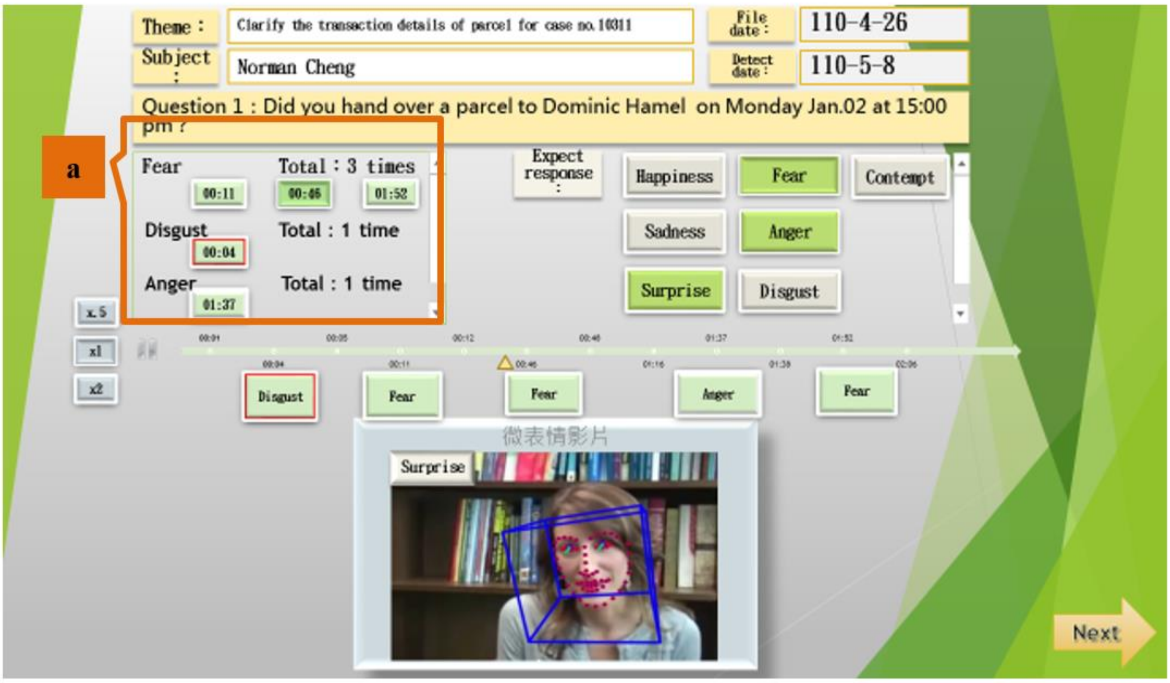

Figure 6. Recording analysis system

3.2. Analysis of the system configuration conditions

Regarding the interrogation room, we identified five system configuration factors that may influence the microexpression analysis results. These experimental factors included the resolution, shooting angle, shooting height, shooting distance, and frame rate (frames per second/FPS) of the camera. The results of the experimental tests are as.

3.2.1. Analysis of the recognition results at different resolutions

After the captured videos were analyzed using Facereader, editing software was used to adjust the video resolution to 600p (800×600) and 480p (640×480). Facereader was then used for analysis and comparison as shown in Figure 7. The results indicated that the recognition performance of the 600p videos were considerably higher than those of the 480p videos.

3.2.2. Analysis of the recognition results at different shooting angles

FaceReader was used to analyze the original videos, and the frontal and side videos (45° and −45°) were selected for analysis and comparison. As shown in Figure 8, the recognition results of the three shooting angles did not differ considerably.

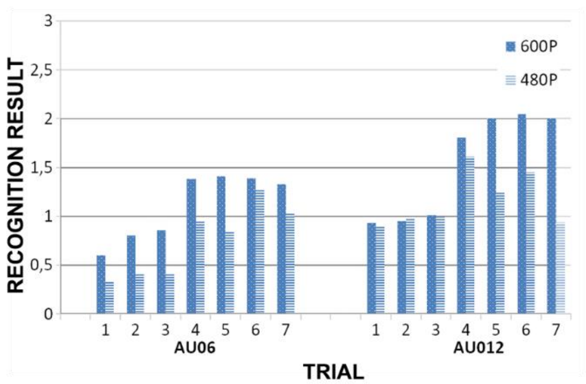

Figure 7. Recognition results of AU6 and AU12 at different resolutions

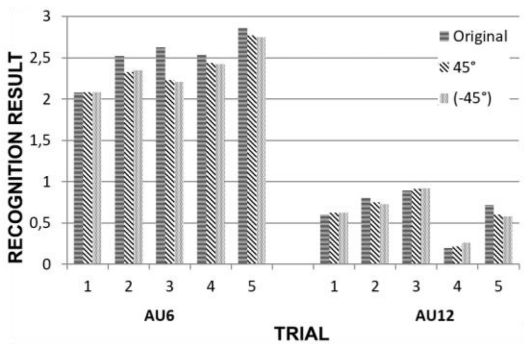

Figure 8. Recognition results of AU6 and AU12 at different shooting angles

3.2.3. Analysis of the recognition results at different shooting heights

FaceReader was used to analyze the original video, and the frontal and elevation videos (45°) were selected for comparison. As shown in Figure 9, the recognition results of the two shooting heights did not differ considerably.

3.2.4. Analysis of the recognition results at different shooting distances

The video was played back on a laptop, and the camera was positioned 30, 100, 200, and 300 cm away from the laptop for analysis. As shown in Figure 10, the recognition results obtained at 30, 100, and 200 cm did not differ considerably. Nonetheless, at a distance of 300 cm, the large gap hindered the recognition of microexpressions.

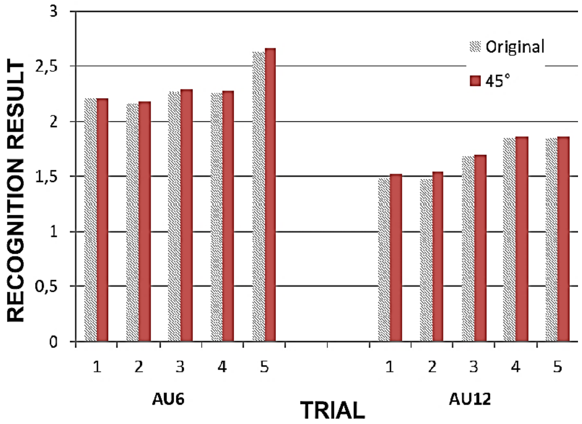

Figure 9. Recognition results of AU6 and AU12 at different shooting heights

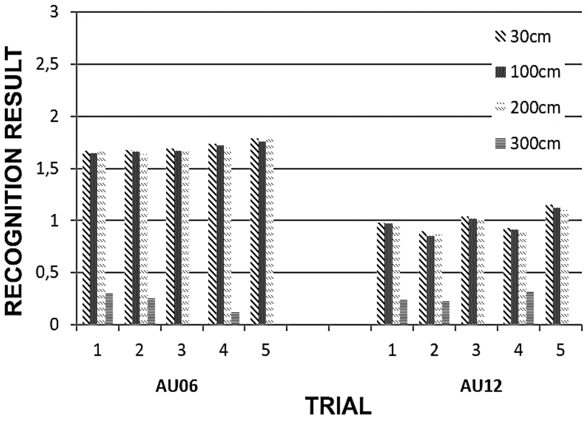

Figure 10. Recognition results of AU6 and AU12 at different shooting distances

3.2.5. Analysis of the recognition results at different frame rates

As shown in Figure 11, the recognition capacity of the high-speed camera (190 FPS) in recognizing happy emotions (AU6+AU12) was higher than that of the conventional camera (30 FPS). This is because FaceReader captured clearer microexpressions during video output.

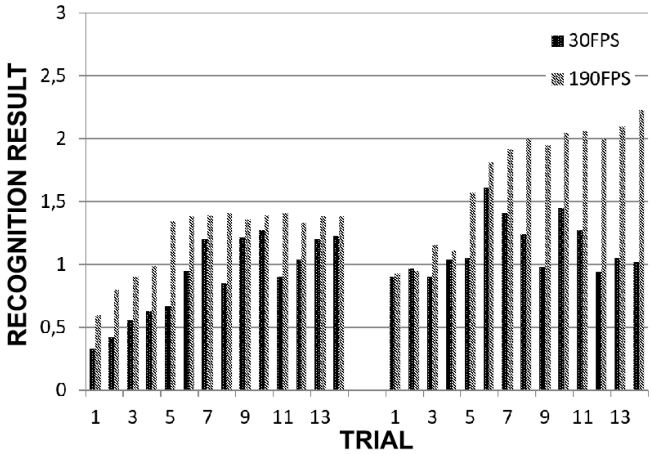

Figure 11. Recognition results of AU6 and AU12 at different frame rates

4. CONCLUSION

In today’s world of rapid technological advancement, an increasing number of criminal cases are being investigated with the aid of scientific methods as auxiliary tools. In this study, we proposed a real-time microexpression recognition system for a specific purpose. This system employs FaceReader to gather information for real-time analysis and aid law enforcement officers in their investigations. To ensure the accuracy of analysis, the system was evaluated under various conditions. However, the duration of microexpressions is too short, and the intensity of changes is low. When these characteristics are combined with the complexity of human emotions, this leaves large room for exploration in the presentation of microexpressions. Currently, the number of videos available for training and analysis is relatively small, and the uneven distribution of video lengths, sample frame rates, and sample types affects the recognition of microexpressions [23]. Therefore, the system must be continuously updated in the future.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the Ministry of Justice for the financial support through the Science and Technology Project (110-1301-10-17-03 and 111-1301-10-28-01).

REFERENCES

[1] P. Ensari, "How to improve emotional intelligence and social skills among adolescents: the development and test of a new microexpressions training," *Journal of Behavioral and Brain Science*, vol. 7, no. 5, pp. 211–225, 2017, doi: 10.4236/jbbs.2017.75016.

[2] T. S. Gunawan *et al.*, "Development of video-based emotion recognition using deep learning with Google Colab," *Telecommunication Computing Electronics and Control (TELKOMNIKA)*, vol. 18, no. 5, Oct. 2020, doi: 10.12928/telkomnika.v18i5.16717.

[3] E. A. Clark *et al.*, "The facial action coding system for characterization of human affective response to consumer product-based stimuli: a systematic review," *Frontiers in Psychology*, vol. 11, May 2020, doi: 10.3389/fpsyg.2020.00920.

[4] P. Ekman and W. V Friesen, "Measuring facial movement," *Environmental Psychology and Nonverbal Behavior*, vol. 1, no. 1, pp. 56–75, 1976, doi: 10.1007/BF01115465.

[5] E. A. Haggard and K. S. Isaacs, "Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy," in *Methods of Research in Psychotherapy*, Boston, MA: Springer US, 1966, pp. 154–165. doi: 10.1007/978-1-4684-6045-2_14.

[6] J. Li, T. Wang, and S.-J. Wang, "Facial micro-expression recognition based on deep local-holistic network," *Applied Sciences*, vol. 12, no. 9, May 2022, doi: 10.3390/app12094643.

[7] F. Xu and J.-P. Zhang, "Facial microexpression recognition: A survey," *Zidonghua Xuebao/Acta Automatica Sinica*, vol. 43, no. 3, pp. 333–348, 2017.

[8] T. Pfister, L. Xiaobai, G. Zhao, and M. Pietikainen, "Recognising spontaneous facial micro-expressions," in *2011 International Conference on Computer Vision*, Nov. 2011, pp. 1449–1456. doi: 10.1109/ICCV.2011.6126401.

[9] Y.-H. Oh, J. See, A. C. Le Ngo, R. C.-W. Phan, and V. M. Baskaran, "A survey of automatic facial micro-expression analysis: databases, methods, and challenges," *Frontiers in Psychology*, vol. 9, Jul. 2018, doi: 10.3389/fpsyg.2018.01128.

[10] G. Warren, E. Schertler, and P. Bull, "Detecting deception from emotional and unemotional cues," *Journal of Nonverbal Behavior*, vol. 33, no. 1, pp. 59–69, Mar. 2009, doi: 10.1007/s10919-008-0057-7.

[11] E. Friesen and P. Ekman, "Facial action coding system: a technique for the measurement of facial movement," *Palo Alto*, vol. 3, no. 2, 1978.

[12] P. Ekman, E. Rosenberg, and J. Hager, "Facial action coding system affect interpretation dictionary (FACSAID)," *In.*, 1998.

[13] P. Ekman and W. V. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. California: Consulting Psychologists Press, 1978.

[14] S. P. Namboodiri and D. Venkataraman, "A computer vision based image processing system for depression detection among students for counseling," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 14, no. 1, pp. 503–512, 2019.

[15] A. Al-Dhamari, R. Sudirman, N. H. Mahmood, N. H. Khamis, and A. Yahya, "Online video-based abnormal detection using highly motion techniques and statistical measures," *Telecommunication Computing Electronics and Control (TELKOMNIKA)*, vol. 17, no. 4, Aug. 2019, doi: 10.12928/telkomnika.v17i4.12753.

[16] P. Ghuli, B. N. Shashank, and A. G. Rao, "Development of framework for detecting smoking scene in video clips," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 13, no. 1, pp. 22–26, Jan. 2019, doi: 10.11591/ijeecs.v13.i1.pp22-26.

[17] J. J. Ali, N. M. Shati, and M. Talib Gaata, "Abnormal activity detection in surveillance video scenes," *Telecommunication Computing Electronics and Control (TELKOMNIKA)*, vol. 18, no. 5, Oct. 2020, doi: 10.12928/telkomnika.v18i5.16634.

[18] J. Hamm, C. G. Kohler, R. C. Gur, and R. Verma, "Automated facial action coding system for dynamic analysis of facial expressions in neuropsychiatric disorders," *Journal of Neuroscience Methods*, vol. 200, no. 2, pp. 237–256, Sep. 2011, doi: 10.1016/j.jneumeth.2011.06.023.

[19] K. Rangasamy, M. A. As'ari, N. A. Rahmad, N. F. Ghazali, and S. Ismail, "Deep learning in sport video analysis: a review," *Telecommunication Computing Electronics and Control (TELKOMNIKA)*, vol. 18, no. 4, Aug. 2020, doi: 10.12928/telkomnika.v18i4.14730.

[20] S. B. Jadhav, "Convolutional neural networks for leaf image-based plant disease classification," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 8, no. 4, pp. 328–341, Dec. 2019, doi: 10.11591/ijai.v8.i4.pp328-341.

[21] K. Zhao, W.-S. Chu, and A. M. Martinez, "Learning facial action units from web images with scalable weakly supervised clustering," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 2090–2099. doi: 10.1109/CVPR.2018.00223.

[22] D. Cannata, S. Redfern, and D. O'Hora, "OpenFaceR: developing an R Package for the convenient analysis of OpenFace facial information," 2020.

[23] H. Pan, L. Xie, Z. Wang, B. Liu, M. Yang, and J. Tao, "Review of micro-expression spotting and recognition in video sequences," *Virtual Reality and Intelligent Hardware*, vol. 3, no. 1, pp. 1–17, Feb. 2021, doi: 10.1016/j.vrih.2020.10.003.

BIOGRAPHIES OF AUTHORS

Yi-Chang Wu works for the Ministry of Justice Investigation Bureau. He is also a doctoral student of electronic and computer engineering at the National Taiwan University of Science and Technology. He can be reached at shintenwu@gmail.com.

Yao-Cheng Liu works for the Ministry of Justice Investigation Bureau. He holds a master's degree from the Department of Computer and Communication, Jinwen University of Science and Technology. He can be reached at 40981t1371@mail.kyu.edu.tw.

Chieh Tsao works for the Ministry of Justice Investigation Bureau. She obtained her bachelor's degree in tourism from Ming Chuan University. She can be reached at vanessatsao650425@gmail.com.

Ru-Yi Huang works for the Ministry of Justice Investigation Bureau. She obtained her bachelor's degree in tourism from Ming Chuan University. She can be reached at a082544@gmail.com.