❏ 332

# The use of artificial intelligence in interrogation: lies and truth

**Yi-Chang Wu, Yao-Cheng Liu, Ru-Yi Huang**
Department Forensic Science Division, Investigation Bureau, Ministry of Justice, New Taipei City, Taiwan (R.O.C.)

## Article Info

## ABSTRACT

Following the development of artificial intelligence technology, a new trend has emerged in which this technology is increasingly used in case investigations. In this study, we developed a lie detection system that can instantly determine whether an interrogee is lying depending on their emotional responses to specific questions. Investigators then use these data, in addition to their personal experiences and case information, to adjust their interrogation strategies and techniques, thereby leading the interrogee to confess and accelerating the investigation process. Our system collects data using OpenFace and performs deep learning using gcForest. Deep learning training was performed using a real-life trial dataset, the Miami University Deception Detection Database, and a bag-of-lies dataset, and their corresponding trained systems achieved a detection accuracy of 95.11%, 90.83%, and 88.19%, respectively.

## Corresponding Author:

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

## 1. INTRODUCTION

During case investigations, officers may request criminal suspects to appear for an interrogation in order to investigate a crime or gather evidence. During these interrogations, officers are required to maintain an honest demeanor and are prohibited from using improper techniques, such as violence, coercion, bribery, fraud, or exhaustion. The purpose of these interrogations is to lead the interrogee into telling the truth and to glean information from their statements in order to further clarify the case or uncover new clues. Numerous cases are solved as a result of interrogee statements, indicating that interrogations are a crucial means for investigators to gather criminal facts. However, even if the interrogee is not the actual perpetrator, they may occasionally lie for a variety of reasons. In addition, to avoid punishment, actual criminals tend to deny or distort criminal facts. Therefore, if officers are unable to determine the veracity of a suspect's statements or decipher their intentions, then the investigation may become stalled or may proceed in the wrong direction, thereby wasting time and resources and allowing criminals to go unpunished. Hence, investigators must employ effective interrogation techniques while adhering to the law.

As a result of the development of hardware equipment and artificial intelligence (AI) in the area of image recognition, the traditional method of human supervision has evolved to include the integration of systems and information technologies to achieve smart image recognition. The computational and logical capabilities of various deep learning algorithms have also enabled the rapid processing of big data and the improvement of the learning and recognition efficiency [1]–[10]. Studies in this area have spanned a wide range of fields, including crime warning, license plate recognition, local disease recognition, and depression and dementia treatment [11]–[14]. Foreign law enforcement agencies are currently investigating the use of face recognition in interrogations in an effort to reduce prejudice against people of different races and genders [15].

The primary purpose of a case investigation is to obtain criminal evidence and clarify case-related facts. In this scenario, the investigation process is conducted in an interactive fashion. In each case, physical evidence, documentary evidence, and related parties are interconnected. To solve a crime, the investigators must rely on existing clues and evidence. Thus, interrogation is one of the most crucial means for investigators to clarify facts and obtain additional case-related information during case investigations.

Keli [16] examined the impact ratio (with vs. without impact) of the 17 reasons why 20 criminal suspects voluntarily confessed as shown in Table 1. Among these reasons, lies being exposed impacted the fifth largest number of suspects, indicating that if officers can immediately expose the interrogee's lies during interrogation, the interrogee becomes more likely to confess. The concept of facial microexpressions was first proposed in 1966 [17], and it refers to the subconscious responses that people exhibit when stimulated. The majority of these expressions consist of subtle changes in facial features and muscles. These changes are difficult to conceal and are prevalent across all racial and age groups [18]–[20]. In recent years, facial microexpressions have been used in studies to identify lying interrogees.

Table 1. Reasons why criminal suspects voluntarily confessed

| No. | Reasons | Impact Ratio | |
| --- | --- | --- | --- |
| | | Yes | No |
| 1 | The investigators have gathered all the criminal evidence, so resistance is futile. | 17 | 3 |
| 2 | Everybody involved has told the truth, so I did too. | 12 | 8 |
| 3 | To receive lighter punishment. | 18 | 2 |
| 4 | I have to be responsible for what I have done, after all. | 9 | 11 |
| 5 | My lies have been exposed. | 15 | 5 |
| 6 | It feels better to tell the truth than not. | 17 | 3 |
| 7 | To thank the investigators for their help. | 7 | 13 |
| 8 | To leave the interrogation room early. | 16 | 4 |
| 9 | At the time, I felt pressured to tell the truth. | 14 | 6 |
| 10 | I confessed to conceal other crimes, but I only told part of the truth. | 12 | 8 |
| 11 | It is only a matter of time before everything is exposed, so it is better to tell the truth early. | 10 | 10 |
| 12 | To show the interrogators that I was willing to cooperate. | 15 | 5 |
| 13 | I do not know why I confessed; I did it on impulse. | 4 | 16 |
| 14 | The facts will be found eventually, so I might as well tell the truth. | 14 | 6 |
| 15 | The investigators were very nice to me. | 6 | 14 |
| 16 | To be released on bail earlier. | 14 | 6 |
| 17 | The case-handling departments actually did not have sufficient criminal evidence, and I voluntarily confessed under the assumption that they did. | 7 | 13 |

In recent years, many units have developed vastly divergent perspectives on the use of polygraphs. Polygraphs are no longer used by law enforcement units as evidence in trials. Instead, they are used as interrogation aids. During an interrogation, the interrogee may lie or conceal the truth for a variety of reasons. Therefore, in this study, we developed a lie recognition system that uses AI deep learning to replace traditional lie detection techniques with a noncontact-based lie detection technique so that investigations are not delayed if the interrogee refuses to take the polygraph or is required to take the polygraph at another time. During interrogations, our system can objectively determine the emotional state and truthfulness of the interrogee. The results can then be combined with case information and the investigators' own experiences to modify the direction of the interrogation, thereby accelerating the investigation process.

The rest of this paper is organized as follows. Section 2 introduces our study methods, including how we collected the data and extracted useful information for subsequent analyses, section 3 presents our results and discussion, and section 4 concludes the study.

## 2.    METHODS

Because truth is essential in decision-making, detecting misleading information before such information is included in the decision-making process is crucial. To screen and clarify a large amount of case-related information within a short period of time, lowering the detection thresholds to avoid missing critical information is essential. Several lie detection methods are currently available, depending on the case, environment, and purpose. However, the most common lie detection method is polygraphs. Polygraphs rely on contact sensors and the analysis of physiological changes, such as heart rate fluctuations [21], to detect misleading behavior. After the emergence of the facial action coding system (FACS) and AI deep learning, changes have been observed in the traditional method for lie detection.

Deep neural networks (DNNs) have made considerable progress in image and sound processing. In addition, a number of deep learning techniques, such as DNNs, convolutional neural networks, deep belief

networks, and loop neural networks, have been applied to computer vision, speech recognition, natural language processing, audio recognition, and bioinformatics, with remarkable results.

Our proposed system captures facial action unit (AU) signals by using OpenFace [22]–[24], performs unsupervised learning by using gcForest, and determines whether an interrogee is likely to confess if their lies are exposed. The detection process of the proposed system is depicted in Figure 1. The system immediately performs detection and makes decisions during interrogations.
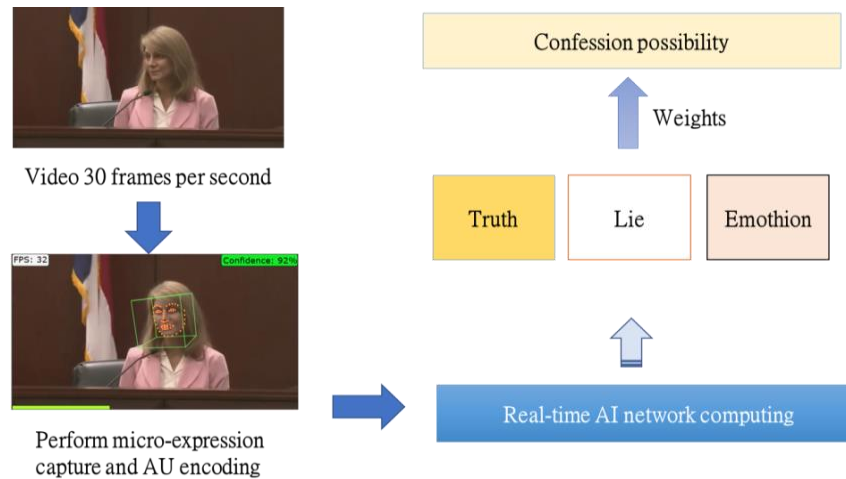


Figure 1. Detection process of the proposed system

## 2.1. Facial microexpressions

By using the human face anatomy and matching facial expressions to their meanings, the FACS defines 44 facial AUs. These AUs are used to describe local facial muscle movements and objectively, accurately, and precisely describe facial expressions [25]. By using AI face recognition and model learning technologies, AU units can be marked quickly in real-time, thereby allowing the FACS to improve its recognition capability and helping interrogators identify lying interrogees and expose their lies to increase the likelihood of confessing.

## 2.2. Deep learning

Deep learning is a branch of machine learning and is an algorithm based on artificial neural networks (ANNs) that learns data features. In the 1980s, a number of essential concepts related to associationism emerged, including distributed representation [26]. Distributed representation eliminates the need for users to manually extract features and enables computers to simultaneously learn how to extract and use features. Feature learning aims to discover superior representation methods, construct more robust models, and uncover representation methods from large amounts of unlabeled data. These representation methods are derived from neuroscience and are loosely based on the understanding of information processing and communication models similar to those of the nervous systems. For example, neural encoding attempts to define the relationships between pulling neuron responses and neuronal activities in the brain. Figure 2 shows the differences between a shallow neural network and a deep learning neural network model [27]. The shallow neural network has only one hidden layer, whereas the DNN has two or more hidden layers.
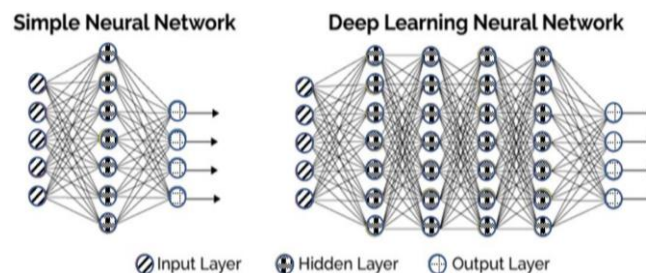


Figure 2. The difference between simple and deep learning neural networks

### 2.3. gcForest

Introduced by Zhou and Feng [28], multigrained cascade forest (gcForest) uses a cascading method to stack multilayer random forests in order to improve feature representation and learning performance. gcForest undergoes representation learning through random cascade forests, processes data hierarchically similar to deep learning networks (DNNs), and uses different forest types to create learning diversity and form waterfall-like structures. The model also uses sliding windows and multigrained scanning to preprocess input features, and it inputs the extracted feature vectors into cascade forests to train and splice the model repeatedly until the verification results converge. Compared with DNNs, gcForest requires considerably less training data to achieve satisfactory performance. In addition, because gcForest contains fewer hyperparameters, does not require the hyperparameter settings to be adjusted, and can control tree-like components through self-adaptation, it consumes only a few computational resources and samples, making its training relatively straightforward [29].

### 2.4. Training models

In this study, three video datasets were used for AI deep learning training: a real-life trial dataset [30], the Miami University Deception Detection Database [31], and a bag-of-lies (BOL) dataset [32]. The risk levels in the training videos were divided to reflect the liabilities and risks that the interrogees faced while lying. A high risk indicated that if the interrogee lied, they would face actual repercussions, such as criminal charges and imprisonment: i) the real-life trial dataset included high-risk videos of actual court proceedings; ii) the Miami University Deception Detection Database and BOL dataset included experimental laboratory-produced low-risk videos.

### 2.5. Effectiveness assessment

Confusion matrices are an essential instrument for evaluating the performance of classification models. They generate indicators (i.e., accuracy, precision, recall, and F1-score) that can be used to evaluate system detection results.
a. Confusion matrix is as follows.
- True positive (TP): Lied and tested positive for lying
- False positive (FP): Did not lie but tested positive for lying
- True negative (TN): Did not lie and tested negative for lying
- False negative (FN): Lied but tested negative for lying
b. Accuracy: The overall prediction accuracy is calculated as follows.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

c. Precision: Precision is the percentage of samples that are truly positive out of all samples predicted to be positive. It is calculated as follows.

$$p_{recision} = \frac{TP}{TP + FP}$$

d. Recall: Recall is the percentage of samples that are truly positive out of all positive samples. It is calculated as follows.

$$Recall = \frac{TP}{TP + FN}$$

e. F1-score: Precision and recall have an interdependent relationship. Optimally, they should both be high. However, in real-life scenarios, the higher one is, the lower the other one becomes. Therefore, precision and recall must be comprehensively assessed. To this end, the most common method is to use the F1-score as a comprehensive indicator as follows.

$$\frac{2}{F} = \frac{1}{P} + \frac{1}{R} \Rightarrow F_1 = \frac{2PR}{P + R}$$

f. Area under the receiver operating characteristic curve (AUC): The AUC is a common statistical value that represents the predictive ability of a classifier. The greater the area under the curve is, the stronger the predictive ability becomes:

- An AUC value greater than 0.5 indicates that the classification effect of the classifier is superior to random guesses, indicating that the model provides valuable predictions.
- An AUC value of 0.5 indicates that the classification effect of the classifier is equivalent to random guesses, indicating that the model does not provide valuable predictions.

## 3.    RESULTS AND DISCUSSION

During an interrogation, the interrogee may lie for a variety of reasons, including the avoidance of consequences or the protection of others. However, even lies can provide investigative leads. In this study, we developed a facial microexpression recognition system to immediately detect and expose lies, thereby compelling suspects to confess and assist with interrogations. Figure 3 depicts the DNN structure of the proposed system.
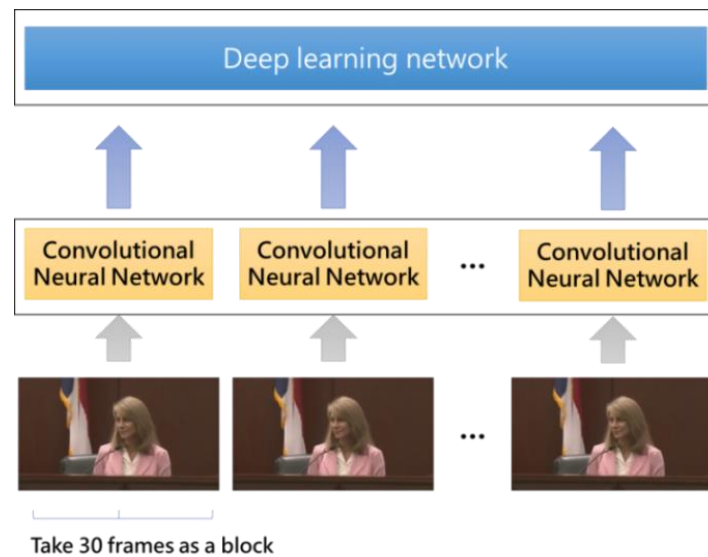


Figure 3. AI learning structure of the proposed lie detection system

Before the proposed lie detection system was actually used, it was first trained. The videos in the datasets were divided into 50% lies and 50% truths at 30 frames per second. The samples were then divided into two groups as follows: 70% for training and 30% for testing. To achieve facial marking and detection, a constrained local neural field was used, which provided over 700 features, of which 35 were associated with facial AUs. Because the p values of AU01_r, AU23_r, and AU17_C were too low, they were disregarded during the training process to increase the detection success rate of the training model [33].

Before detection was performed, the proposed system was trained using the real-life trial dataset, Miami University Deception Detection Database, and BOL dataset. The real-life trial dataset included high-risk videos of actual court proceedings, whereas the other two datasets included experimental laboratory-produced low-risk videos. Table 2 lists the detection results of the proposed system. We compared the accuracy, precision, recall, and F1-score of the proposed system trained using the three different datasets. The results indicated that the system produced the highest accuracy (95.11%) when trained using actual court videos and the lowest accuracy (88.19%) when trained using the BOL database. All other indicators exhibited scores over 80%.

Table 2. Assessment indicators of the effectiveness of the proposed lie detection system

| Data Set | Accuracy | Precision | Recall | F1-Score |
|----------|----------|-----------|--------|----------|
| TRIAL | 95.11% | 95.39% | 93.97% | 94.68% |
| MU3D | 90.83% | 88.99% | 81.27% | 84.96% |
| BOL | 88.19% | 88.64% | 89.06% | 88.85% |

Figure 4 depicts the system effectiveness indicators as measured by the AUC. When the proposed system was trained using the real-life trial dataset, the Miami University Deception Detection Database, and the BOL dataset, AUC values of 95.03%, 88.29%, and 88.14%, respectively, were obtained. Among the three datasets, the proposed system demonstrated optimal detection performance for all indicators when trained using actual court videos.
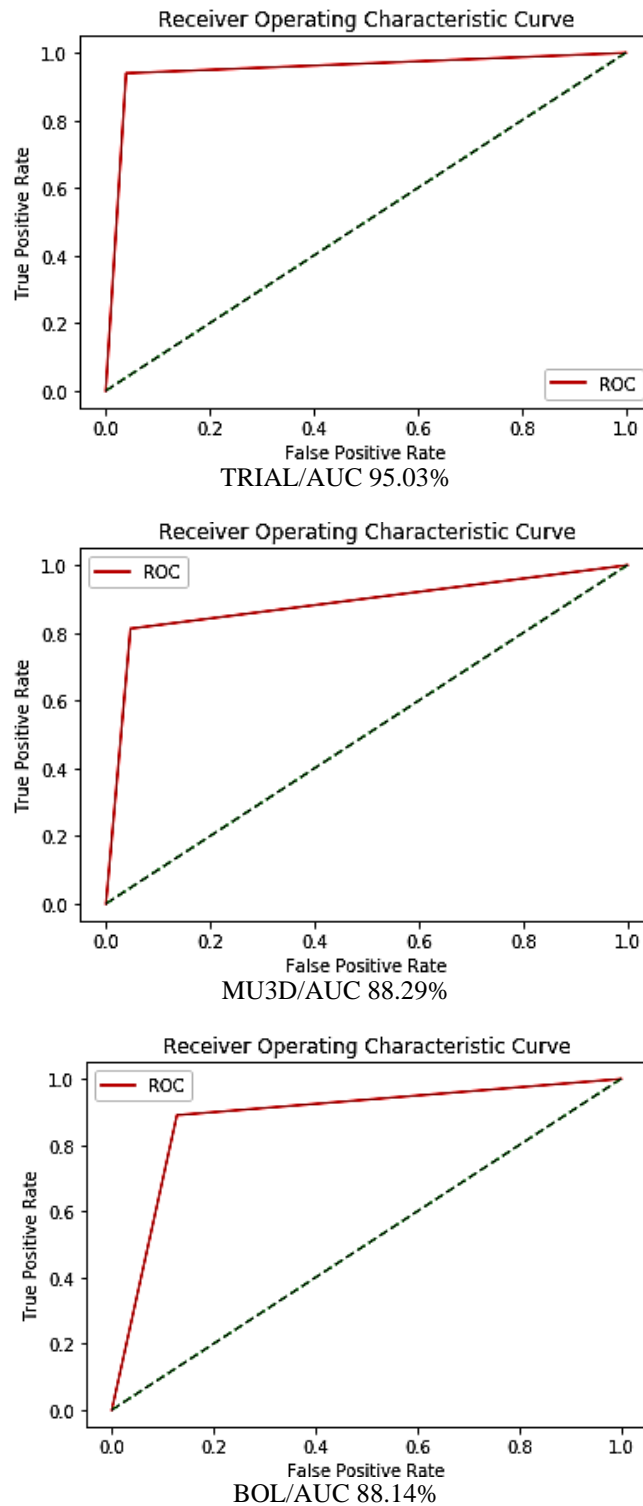


TRIAL/AUC 95.03%



MU3D/AUC 88.29%



BOL/AUC 88.14%

Figure 4. AUC values of the proposed lie detection system

Table 3 compares the effectiveness of the proposed system to other court video evaluation systems. The proposed system achieved the highest scores across all four indicators, with an accuracy and F1-score of 95.11% and 94.68%, respectively, indicating its superior detection rate stability and reliability.

Table 3. Comparison of system effectiveness indicators

| No | Methods | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|---|
| 1 | Feature Auto-Extraction+Fusion [34] | 66.36% | 43.57% | 49.08% | 43.43% |
| 2 | Emotion Transformation Feature for Detection [35] | 69.44% | 40.98% | 52.65% | 45.25% |
| 3 | Facial Affect Involved Method [36] | 72.95% | 41.47% | 55.00% | 46.49% |
| 4 | Multimodal Detection [37] | 76.29% | 69.73% | 63.64% | 60.54% |
| 5 | SVM/RF/NN+Fusion [38] | 77.12% | 58.51% | 64.64% | 59.26% |
| 6 | Multi-modal Neural Mode [39] | 77.92% | 71.43% | 71.58% | 67.74% |
| 7 | CNN+Fusion [40] | 83.84% | 74.94% | 73.86% | 73.00% |
| 8 | Face Focused Cross-Stream Net-work [41] | 85.19% | 79.82% | 7865% | 75.99% |
| 9 | GCFM [42] | 88.14% | 82.46% | 80.75% | 78.50% |
| 10 | Our Method | 95.11% | 95.39% | 93.97% | 94.68% |

## 4. CONCLUSION

Combining AI and image detection technologies has increased their applicability and development within the field of information engineering. Technology enables faster access to necessary information and has become an increasingly popular tool to enhance the effectiveness of investigations. In this study, we used existing imaging equipment in interrogation rooms and a noncontact lie detection method to expand the application of lie detection, determine the true emotions of interrogees without their knowledge, and provide investigators with objective detection results. Currently, the accuracy of our system exceeds 80%. To achieve the same accuracy in practice, additional related videos and videos of individuals of Asian descent must be added in future training scenarios to improve the feature extraction accuracy of our system.

## REFERENCES

[1] I. S. Hutomo and H. Wicaksono, "A smart door prototype with a face recognition capability," *IAES International Journal of Robotics and Automation (IJRA)*, vol. 11, no. 1, pp. 1–9, Mar. 2022, doi: 10.11591/ijra.v11i1.pp1-9.

[2] R. Thilahar C. and S. R., "Fuzzy neuro-genetic approach for feature selection and image classification in augmented reality systems," *IAES International Journal of Robotics and Automation (IJRA)*, vol. 8, no. 3, pp. 194–204, Sep. 2019, doi: 10.11591/ijra.v8i3.pp194-204.

[3] L. Liliana, J.-H. Chae, J.-J. Lee, and B.-G. Lee, "A robust method for VR-based hand gesture recognition using density-based CNN," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 761–769, Apr. 2020, doi: 10.12928/telkomnika.v18i2.14747.

[4] S. A. Baker, H. H. Mohammed, and H. A. Aldabagh, "Improving face recognition by artificial neural network using principal component analysis," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 6, pp. 3357–3364, Dec. 2020, doi: 10.12928/telkomnika.v18i6.16335.

[5] H. A. Al-Jubouri and S. M. Mahmmod, "A comparative analysis of automatic deep neural networks for image retrieval," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 3, pp. 858–871, Jun. 2021, doi: 10.12928/telkomnika.v19i3.18157.

[6] F. Martinez, C. Hernández, and F. Martínez, "Evaluation of deep neural network architectures in the identification of bone fissures," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 807–814, Apr. 2020, doi: 10.12928/telkomnika.v18i2.14754.

[7] H. M. Ariza, H. H. Martínez, and L. A. G. Roa, "Recognition system for facial expression by processing images with deep learning neural network," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 17, no. 6, pp. 2975–2982, Dec. 2019, doi: 10.12928/telkomnika.v17i6.12948.

[8] K. A. Lipi, S. F. K. Adrita, Z. F. Tunny, A. H. Munna, and A. Kabir, "Static-gesture word recognition in Bangla sign language using convolutional neural network," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 5, pp. 1109–1116, Oct. 2022, doi: 10.12928/telkomnika.v20i5.24096.

[9] M. Attamimi, R. Mardiyanto, and A. N. Irfansyah, "Inclined image recognition for aerial mapping using deep learning and tree based models," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 16, no. 6, pp. 3034–3044, Dec. 2018, doi: 10.12928/telkomnika.v16i6.10157.

[10] Z. N. Abdullah, Z. A. Abutiheen, A. A. Abdulmunem, and Z. A. Harjan, "Official logo recognition based on multilayer convolutional neural network model," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 20, no. 5, pp. 1083–1090, Oct. 2022, doi: 10.12928/telkomnika.v20i5.23464.

[11] P. Kulkarni and T. M. Rajesh, "Analysis on techniques used to recognize and identifying the Human emotions," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 3, pp. 3307–3314, Jun. 2020, doi:

10.11591/ijece.v10i3.pp3307-3314.

[12] A. Jahagirdar and R. Phalnikar, "Comparison of feed forward and cascade forward neural networks for human action recognition," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 25, no. 2, pp. 892–899, Feb. 2022, doi: 10.11591/ijeecs.v25.i2.pp892-899.

[13] M. Aouayeb, W. Hamidouche, K. Kpalma, and A. Benazza-Benyahia, "A spatiotemporal deep learning solution for automatic micro-expressions recognition from local facial regions," in *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, Oct. 2019, pp. 1–6, doi: 10.1109/MLSP.2019.8918771.

[14] K. V. Kumar and J. Harikiran, "Privacy preserving human activity recognition framework using an optimized prediction algorithm," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 1, pp. 254–264, Mar. 2022, doi: 10.11591/ijai.v11.i1.pp254-264.

[15] M. Noriega, "The application of artificial intelligence in police interrogations: An analysis addressing the proposed effect AI has on racial and gender bias, cooperation, and false confessions," *Futures*, vol. 117, Mar. 2020, doi: 10.1016/j.futures.2019.102510.

[16] W. Keli, "Interrogation under the camera: interrogation strategies and techniques under the recording and video recording of the whole process." China Legal Publishing House, 2016.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] M. Oswald, "Technologies in the twilight zone: early lie detectors, machine learning and reformist legal realism," *International Review of Law, Computers and Technology*, vol. 34, no. 2, pp. 214–231, May 2020, doi: 10.1080/13600869.2020.1733758.

[22] M. E. Kret, E. Prochazkova, E. H. M. Sterck, and Z. Clay, "Emotional expressions in human and non-human great apes," *Neuroscience and Biobehavioral Reviews*, vol. 115, pp. 378–395, Aug. 2020, doi: 10.1016/j.neubiorev.2020.01.027.

[23] T. Baltrusaitis, P. Robinson, and L.-P. Morency, "Constrained local neural fields for robust facial landmark detection in the wild," in *2013 IEEE International Conference on Computer Vision Workshops*, Dec. 2013, pp. 354–361, doi: 10.1109/ICCVW.2013.54.

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

[25] 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.

[26] G. E. Hinton, "Learning distributed representations of concepts," 1986.

[27] S. Larabi-Marie-Sainte, S. Ghouzali, T. Saba, L. Aburahmah, and R. Almohaini, "Improving spam email detection using deep recurrent neural network," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 25, no. 3, pp. 1625–1633, Mar. 2022, doi: 10.11591/ijeecs.v25.i3.pp1625-1633.

[28] Z.-H. Zhou and J. Feng, "Deep forest: towards an alternative to deep neural networks," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Aug. 2017, pp. 3553–3559, doi: 10.24963/ijcai.2017/497.

[29] A. Samat, E. Li, P. Du, S. Liu, and Z. Miao, "Improving deep forest via patch-based pooling, morphological profiling, and pseudo labeling for remote sensing image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 9334–9349, 2021, doi: 10.1109/JSTARS.2021.3110994.

[30] V. Pérez-Rosas, M. Abouelenien, R. Mihalcea, and M. Burzo, "Deception detection using real-life trial data," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, Nov. 2015, pp. 59–66, doi: 10.1145/2818346.2820758.

[31] E. P. Lloyd, J. C. Deska, K. Hugenberg, A. R. McConnell, B. T. Humphrey, and J. W. Kunstman, "Miami University deception detection database," *Behavior Research Methods*, vol. 51, no. 1, pp. 429–439, Feb. 2019, doi: 10.3758/s13428-018-1061-4.

[32] V. Gupta, M. Agarwal, M. Arora, T. Chakraborty, R. Singh, and M. Vatsa, "Bag-of-lies: a multimodal dataset for deception detection," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2019, pp. 83–90, doi: 10.1109/CVPRW.2019.00016.

[33] H. U. D. Ahmed, U. I. Bajwa, F. Zhang, and M. W. Anwar, "Deception detection in videos using the facial action coding system," *arXiv preprint arXiv: 2105.13659*, May 2021.

[34] J. Zhang, S. I. Levitan, and J. Hirschberg, "Multimodal deception detection using automatically extracted acoustic, visual, and lexical features," in *Interspeech 2020*, Oct. 2020, pp. 359–363, doi: 10.21437/Interspeech.2020-2320.

[35] J.-T. Yang, G.-M. Liu, and S. C.-H. Huang, "Emotion transformation feature: novel feature for deception detection in videos," in *2020 IEEE International Conference on Image Processing (ICIP)*, Oct. 2020, pp. 1726–1730, doi: 10.1109/ICIP40778.2020.9190846.

[36] L. Mathur and M. J. Matarić, "Introducing representations of facial affect in automated multimodal deception detection," in *Proceedings of the 2020 International Conference on Multimodal Interaction*, Oct. 2020, pp. 305–314, doi: 10.1145/3382507.3418864.

[37] Z. Wu, B. Singh, L. Davis, and V. Subrahmanian, "Deception detection in videos," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, Apr. 2018, doi: 10.1609/aaai.v32i1.11502.

[38] M. U. Sen, V. Perez-Rosas, B. Yanikoglu, M. Abouelenien, M. Burzo, and R. Mihalcea, "Multimodal deception detection using real-life trial data," *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 306–319, Jan. 2022, doi: 10.1109/TAFFC.2020.3015684.

[39] G. Krishnamurthy, N. Majumder, S. Poria, and E. Cambria, "A deep learning approach for multimodal deception detection," *arXiv preprint arXiv: 1803.00344*, Mar. 2018, [Online]. Available: http://arxiv.org/abs/1803.00344.

[40] M. Gogate, A. Adeel, and A. Hussain, "Deep learning driven multimodal fusion for automated deception detection," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Nov. 2017, pp. 1–6, doi: 10.1109/SSCI.2017.8285382.

[41] M. Ding, A. Zhao, Z. Lu, T. Xiang, and J.-R. Wen, "Face-focused cross-stream network for deception detection in videos," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 7794–7803, doi: 10.1109/CVPR.2019.00799.

[42] H. Zhang, Y. Ding, L. Cao, X. Wang, and L. Feng, "Fine-grained question-level deception detection via graph-based learning and cross-modal fusion," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 2452–2467, 2022, doi: 10.1109/TIFS.2022.3186799.

## BIOGRAPHIES OF AUTHORS

**Yi-Chang Wu** 🆔 [g] SC ◎ works for the Ministry of Justice Investigation Bureau. He is also a Ph.D. in electronic and computer engineering at the National Taiwan University of Science and Technology. He can be contacted at shintenwu@gmail.com.

**Yao-Cheng Liu** 🆔 [g] SC ◎ 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 contacted at 40981t1371@mail.kyu.edu.tw.

**Ru-Yi Huang** 🆔 [g] SC ◎ works for the Ministry of Justice Investigation Bureau. She obtained her bachelor's degree in tourism from National Chi Nan University. She can be contacted at a082544@gmail.com.