



Research Article

Detection of Lying Behaviors Based on Facial Micro-Expressions Using Artificial Neural Network

Wei Ci Gan ¹, Gwo Chin Chung ^{2,*}, It Ee Lee ², Kah Yoong Chan ², Soo Fun Tan ³

¹MS Supply Chain Solutions, Jalan P10/10, Kawasan Perindustrian Miel, 43650 Bandar Baru Bangi, Selangor, Malaysia

²Faculty of AI & Engineering, Multimedia University, Jalan Multimedia, 63100 Cyberjaya, Selangor, Malaysia

³Preparatory Centre for Science and Technology, University Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia

*Corresponding author: gcchung@mmu.edu.my; Tel.: +60383125565

Abstract: Lying is the deliberate provision of false information with the intent to deceive others by concealing the truth. Traditionally, lie detection has been conducted using a polygraph machine, which measures physiological responses to identify signs of lying. However, it might not be accurate because people can undergo training to hide these behaviors detected by the polygraph. A modern way to address this problem is by observing the facial micro-expression that is unintentionally produced by the facial muscles. In this research, the main objective is to develop a deep learning model to detect possible lying behaviours based on facial micro-expressions. A facial behavior analysis toolkit, named OpenFace 2.0, was used to extract different categories of facial muscles into Action Units (AU) from the video frames. The data obtained were trained by Artificial Neural Network (ANN) to classify AU, which has the possible lying frames. Different case studies were also made based on different hyperparameters for performance evaluation. For all case studies, the training and testing accuracies can achieve performances of approximately 80%-90%, and the prediction on unseen data has a record of 55%-70% accurate prediction. Therefore, the detection of lying behaviors based on facial micro-expressions using deep learning is possible, and the result obtained from this research is crucial for the development of a more complete and advanced lying detection system to assist authorities in fighting crimes.

Keywords: Action unit; Artificial neural network; Facial micro-expressions; Lying; OpenFace 2.0

1. Introduction

According to the Oxford English Dictionary, the word "lie" is defined as a false statement made with the intent to deceive. When an individual attempts to conceal the truth, cognitive processes often guide them to manipulate both verbal statements and facial expressions in order to convince others of a falsehood. Detecting deception is critical, as undetected lies may lead to more serious criminal behavior. Traditionally, a polygraph is used to conduct polygraph tests, which consist of a physiological recorder that assesses three indicators, such as heartbeat rate/blood pressure, skin conductivity/sweat, and respiration (National Research Council Committee, 2003).

However, in the 21st century, an increasing number of individuals have learned to effectively conceal deceptive behavior by undergoing specialized training aimed at countering polygraph detection, primarily through the regulation of physiological responses such as heart rate and respiration (Rad et al., 2023). Therefore, researchers have developed many techniques to detect lying individuals based on their facial emotions (Canal et al., 2022). According to the results, seven

basic emotions can be observed from the face, such as happy, fear, sadness, contempt, neutral, anger, disgust, and surprise (Revina and Emmanuel, 2021). Expression of fear can be one of the features of lying detection, as reported in (Shen et al., 2021), but it also leads to false statements. Under high-stakes conditions, emotional responses, particularly fear can often be observed through facial expressions, as the stress of the situation induces nervousness and anxiety, even in individuals who are telling the truth. The disadvantage of the proposed method in this research is low flexibility because the liars not only show fear when they are lying but sometimes show happy emotions to hide the truth. Although using high-speed hardware from Owayjan et al. (2012) was able to provide a higher resolution of video and higher accuracy of model performance, this increased the expenses of the system significantly. Additionally, Chavali et al. (2014) provided an idea that used the Histogram of Oriented Gradients (HOG) features as the feature extraction and observed the changing of the emotion density graph through every frame of the videos to classify whether individuals were lying or not. However, this method is less reliable because the machine can only detect the changes in the emotions associated with individuals who are lying (Sharifnejad et al., 2021).

Although words can occasionally deceive others, facial expressions are one of the most reliable ways to indicate an individual's genuine intentions (Ekundayo and Viriri, 2021). This is due to the ease with which one can observe and identify emotions from facial expressions, as these factors represent a change in muscle activity on the face. In 2005, Ekman and Rosenberg (2005) developed the Facial Action Coding System (FACS), a worldwide known comprehensive system for evaluating facial expressions of the human face. FACS categorizes each of the hundreds of possible facial expressions formed by combining one or more facial muscles(s). According to this research, facial expression can be divided into 2 main categories, such as macro-expression and micro-expression. The macro-expressions are typically visible for 3/4th to 2 seconds, and there are six universal forms of macro-expressions, such as anger, disgust, fear, happiness, sadness, and surprise. The micro-expressions are more difficult to detect by bare eyes because the action is very fast, and it lasts for 1/25 second to 1/5 second. Facial micro-expressions are an essential behavior source for detecting aggressive intent and danger disposition (Ben et al., 2021; Li et al., 2019).

Micro-expression recognition systems are sometimes used as a backup authentication mechanism (Zhou et al., 2021; Takalkar et al., 2018). Human eyes can barely observe the micro-expressions due to the time taken being extremely short. Because micro-expressions can aggregate to form the six universally recognized macro-expressions, researchers are able to train systems to detect and analyze these subtle facial movements for the purpose of emotion and deception recognition (Mehendale et al., 2020; Zahara et al., 2020). Since micro-expressions are very small, Ekman and Rosenberg (2005) created the Micro-Expression Training Tool (METT) to educate humans on how to recognize and respond to them. To imitate the training tool, a lie detector with facial expressions using deep learning can be developed with various machine learning methods (Aftab et al., 2023). However, most of the studies mentioned used the convolutional neural network (CNN) method to perform the lying detection with the images or video frames (Ma et al., 2019). This method can be time-consuming, which can take up to several hours if the videos are too long (Abdullah et al., 2023) or the images are too large (Hor et al., 2022). This challenge becomes more complicated for real-time detection as reported by (Tran-Le et al., 2021). However, an ANN has been introduced to simplify the process, as reported by (Nurçin et al., 2017) for lie detection on pupil size. However, there are researchers who proposed a combination of other factors such as skin and voice (Rahman et al., 2023) or temperature (Aranjo et al., 2021), but it is still not reliable. To address this problem, the extraction of Action Unit (AU) with OpenFace (Nadeeshani et al., 2020) can be used to collect data on the facial expression of a human before training the machine learning models. According to the results, the general performance of OpenFace was better than FaceReader and Py-Feat (Namba et al., 2021).

Table 1 shows the advantages and disadvantages of several related studies. Facial micro-expressions provide better accuracy performance in detecting lying behaviours when compared to

macro-expressions. By applying machine learning models, the prediction can be improved with a lesser hardware cost setup. However, due to the complex background and diverse facial features, the amount of data that needs to be processed by the machine learning models is huge and requires high processing power or long processing time. Although more advanced models such as zero-shot learning have been deployed (Barakat et al., 2023), it is mainly dependent on the accuracy of the extracted facial micro-expression. Therefore, the research gaps addressed focus on the development of more effective facial expression methods for lying detection and more appropriate machine learning methods to improve the accuracy and speed performance of lying detection.

Table 1 Summary of literature review

| Reference Paper | Feature | Advantage | Limitation |
|---|--|--|--|
| Reference (Shen et al., 2021) | Use fear's AU as a feature of lying behaviours | Simple | Low flexibility |
| Reference (Owayjan et al., 2012) | High-speed hardware for detection | Improved accuracy | High complexity and costly |
| Reference (Chavali et al., 2014) | Use HOG for micro-expression detection | Able to clarify mini changes on the face and emotion | Less reliability due to the limit on detecting emotions only |
| Reference (Abdullah et al., 2023; Ma et al., 2019) | Use CNN for model training. | Improved accuracy | Time-consuming for training |
| Reference (Nadeeshani et al., 2020; Avola et al., 2019) | Use OpenFace 1.0 to extract AUs | Utilize AUs information to detect facial micro-expression | Lesser features of AU as compared to OpenFace 2.0 |
| Proposed System | Use OpenFace 2.0 to extract AUs and ANN for model training | More features of AU being used and faster training time than CNN | Lack of unseen data to train the model for better accuracy |

Machine learning models produce promising high-accuracy performance in facial expression detection, as discussed earlier. Their ability to adapt to varying conditions, such as complex environments and diverse facial features, allows lying detection to be applied in real-time applications. In this research, a new machine learning-based system was proposed that detects lying behaviors by extracting the features of facial micro-expressions. OpenFace 2.0 software was used to obtain AU from the video as the features for the prediction to improve the accuracy of the collected information over OpenFace 1.0 used in the previous investigations. Furthermore, the deep learning method, ANN, was selected because the features of the videos were in the form of values but not in pixels to perform prediction with a faster pace and higher learning rate compared to CNN (Dapito et al., 2024). The performance of the proposed system was evaluated in terms of loss and accuracy with different hyper-parameters such as the activation function, number of neurons, dropout layer, dropout rate, optimizers, batch size, and epochs. Three case studies were conducted, where case study 1 was to determine suitable parameters of the ANN model, case study 2 was to examine the performance of hyperparameter fine-tuning with Bayesian optimization, and case study 3 was the modification of ANN model based on case studies 1 and 2 for performance improvement.

The novelty of this research is to provide a comprehensive investigation of the performance of the ANN model using the extraction of AU with OpenFace 2.0 for lying detection as a significant reference for other researchers to continuously develop a more advanced and improved lying detection system in real-life applications. The goal of this research meets the requirement of

Sustainable Development Goal (SDG) 11, titled "Sustainable Cities and Communities," which was established by the United Nations General Assembly (United Nations, 2024).

This research is organized as follows: Section 2 presents the research methodology by illustrating the details of the overall architecture. This also introduces the usage of OpenFace 2.0 software. Section 3 discusses the results of lying detection using the ANN deep learning model, which includes performance comparison and analysis of the model with different hyper-parameters. Lastly, Section 4 consists of a summary of the lying detection through micro-facial expression using the proposed deep learning model as well as future recommendations.

2. Research Methodology

2.1. System Architecture

This research aims to develop a new deep-learning solution for lying detection behaviors from video frames. The overall architecture of the proposed system is shown in Figure 1. The video database was taken from an online source, thereby the size of the input is enough for the model training. Before the model training, the videos were edited to crop and store the main subjects' faces which are lying or telling the truth, to prevent the other subjects' faces appearing in the frame from interfering with the result. Additionally, the videos were processed with OpenFace 2.0 to analyze the information of AU from the video frame. The video frame was maintained at 30 frames per second and 480-pixel resolutions during the video extraction. AU was then stored in the Excel CSV file, and only the important features of AU were filtered and used for model training. In the CSV file, the confidence level must be more than 70% and the success value was set as '1' to indicate the subjects' faces are being fully detected by the software. Then, the ANN deep learning model (Abdolrasol et al., 2021) was used to train CSV files with different hyper-parameters, and the performance of the model was tested under different case studies. Finally, the prediction of lying actions was performed with unseen online sources except for the database to evaluate the performance of the system under different scenarios.

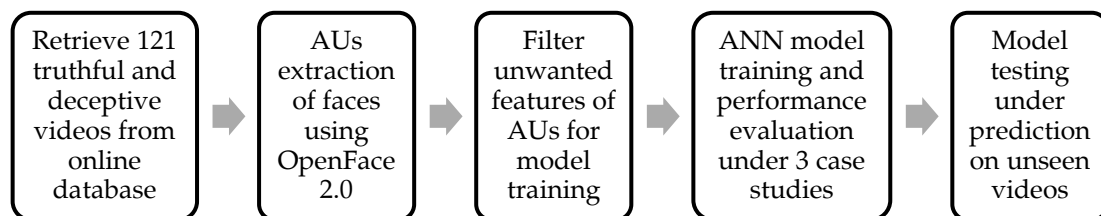


Figure 1 Block diagram of the proposed system

2.2. Dataset

In this research, the dataset was obtained from an online source (Sen et al., 2020) that showcases real-life trial videos with different kinds of individuals and races. The videos were classified as truthful and deceptive. To improve the performance accuracy, some features in the videos were required to be maintained, such as the viewing angle of the individual's face, the resolution of the video quality, the frame rate, and the background noise of the video. The source contains 121 videos, 61 of which were deceptive and 60 of which were accurate. The sample had 21 unique female speakers and 35 unique male speakers, ranging in age from 16 to 60 years. The videos required a frame rate of 30 frames per second and had a resolution of 480 pixels. In the videos, several faces appeared that were not the primary subjects of interest. Therefore, it was necessary to apply preprocessing methods to isolate and focus on the relevant facial regions. For instance, if the subjects' faces were not shown in the video frame or were not clear or covered by some objects, the frame part was cropped and removed from the dataset. Videos that had occurred in many subjects were removed from the dataset too to maintain one subject at a time (Avola et al., 2019) as separate subjects. In addition, all the filtered video frames were used to train the model.

In addition to the dataset mentioned above, this research also obtained some test videos from the online source. They were unseen in the training dataset and did not possess the same characteristics as testing samples, to evaluate the performance of the training model. The test videos were extracted from real-life movies or videos and went through the same operations as mentioned above to make sure the subjects were clear for AU extraction later.

2.3. OpenFace 2.0

OpenFace 2.0 is a more advanced version of the OpenFace 1.0 toolkit that can handle more circumstances due to the addition of a new CNN-based face detector and an optimized facial landmark identification algorithm (Baltrusaitis et al., 2018). In OpenFace 2.0, the program can detect facial landmarks, monitor head poses, recognize AU, and estimate eye focus. This research focuses on the method of using facial expression recognition and providing the facial AU intensity and presence, depending on a recent AU recognition framework reported by Baltrusaitis et al. (2015). Figure 2a shows the list of AU detected in OpenFace 2.0. When the face is detected by OpenFace 2.0, the subject's face will be covered by the 3D facial landmark, the eye gaze estimation, head pose estimation, and appearance extraction face alignment to obtain all this information (Kreiensteck et al., 2023) as shown in Figure 2b.

AU data obtained from OpenFace 2.0 was divided into two parts, such as the presence (0 or 1) and the intensity (ranging from 0 to 5). The value of the intensity was multiplied by the value of the presence to improve the accuracy of the training. The confidence level and the success value depend on the information that can be collected from the subject's face through the software. If the information can be obtained, the confidence level would be higher than 0.7 and the success value was set to 1. On the other hand, if the information cannot be obtained, the confidence level would be lower than 0.7 and the success value would be set to 0. After the dataset was processed into a truth and lie file, both data were concatenated and processed with the Min-Max scaler to normalize the input features of the data due to the inconsistent behavior of the subject's face. The equation of the scaler with x feature value is reported by Sharma (2022):

$$X_{scaled} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (1)$$

The data was separated into 75% of training data and 25% of testing data for ANN deep learning model training and unseen data prediction.

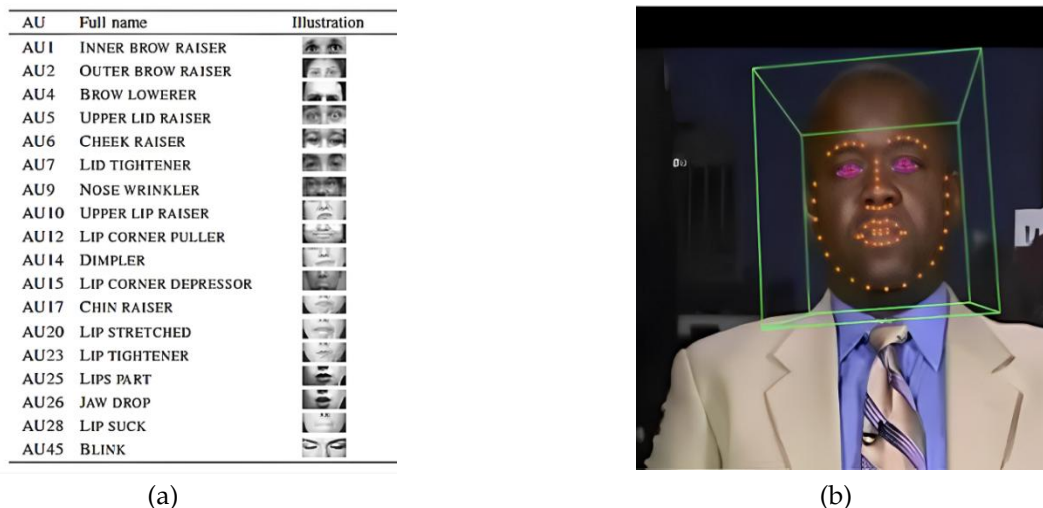


Figure 2 (a) List of AUs detected in OpenFace 2.0 and (b) face detected in OpenFace 2.0

2.4. Artificial Neural Network

ANN is a class of deep learning models designed to simulate the structure and functions of biological neural networks. Every building block in ANN is an artificial neuron, as shown in Figure

3 (Abdolrasol et al., 2021). At the neuron's entry, the inputs are weighted, which means that each input value is multiplied by its weight. All weighted inputs, W , and biases, b , are added using the sum function in the artificial neuron's central section. At the neuron's exit, the total of previously weighted inputs and bias travels through an activation function. On the other hand, hidden layers are added in the network in between the input and output layers to learn complex patterns. This is the concept of multi-layer perceptron (MLP). Dropout can be implemented in these hidden layers to thin the network or lower capacity during training by randomly subsampling the outputs of a layer. Finally, optimizers are algorithms for altering the characteristics of a neural network, such as weights and learning rate, to minimize losses.

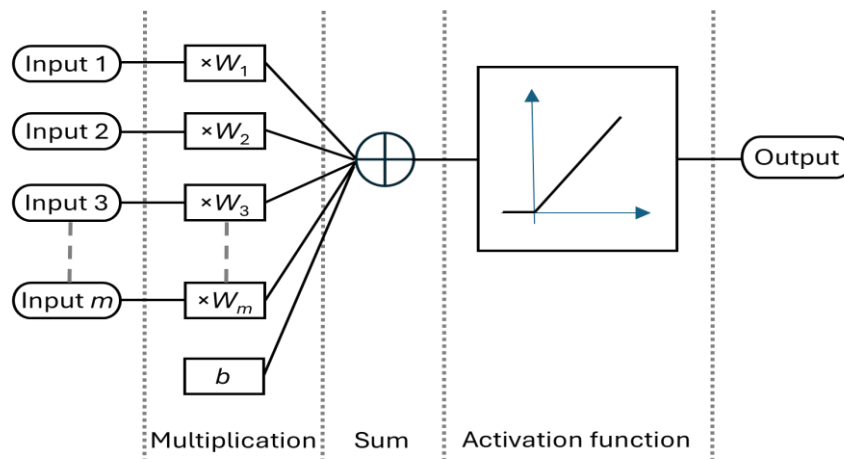


Figure 3 Working principle of ANN

In this research, a supervised classification type of ANN was implemented with a dropout method. The default optimizer is Adam and the learning rate is 0.001. The overall flowchart is presented in Figure 4.

3. Results and Discussion

3.1. ANN Model Training and Testing

There are three different case studies that have been performed in this research. The purpose of case study 1 was to determine suitable activation function, number of neurons, dropout layer, dropout rate, optimizers, batch size, and epochs and obtain a better training accuracy. Case study 2 was to examine the performance of hyperparameter fine-tuning with Bayesian optimization. Case study 3 was the modification of the activation function, optimizer, epochs, and batch size based on case studies 1 and 2 to improve the performance.

In case study 1, the model was tested with different activation functions, batch sizes, and epochs to obtain the best hyperparameters that fit the dataset. Figure 5 shows the loss plot of the training model with different activation functions of ReLU, tanh, and sigmoid (Rasamoelina et al., 2020). According to the observation, the ReLU function has a better learning curve compared to the other two functions since it has the lowest cost value after 50 epochs. Therefore, the ReLU function was used to test the model's accuracy performance with several epochs ranging from 1 to 500 and batch sizes ranging from 4 to 512 as shown in Figure 6. As a result, the batch size of 128 and the epochs of 200 were selected to achieve the best performance of all.

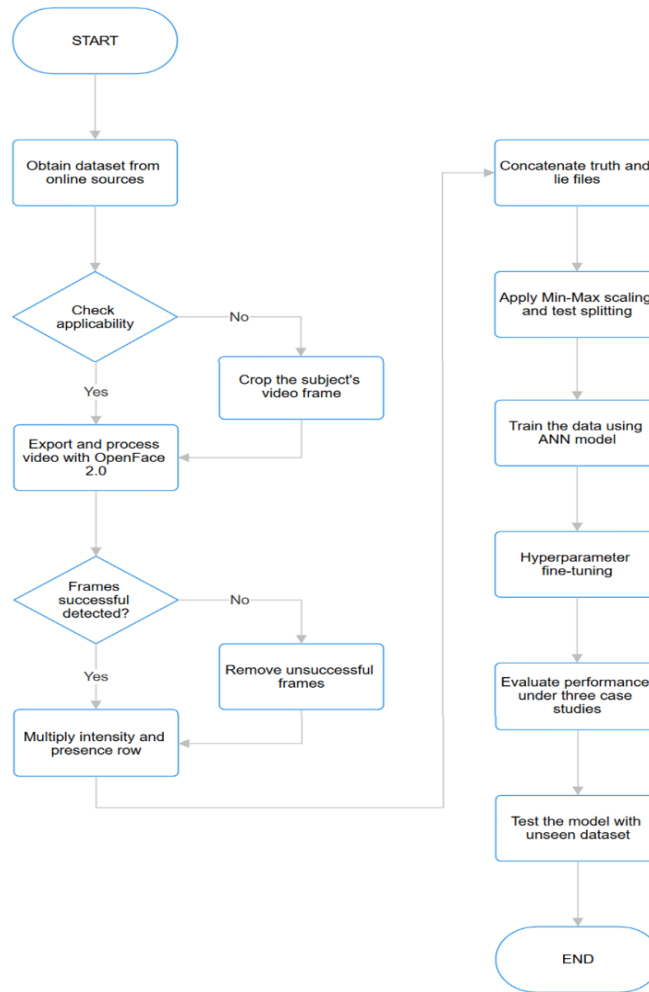


Figure 4 Flowchart of the proposed system

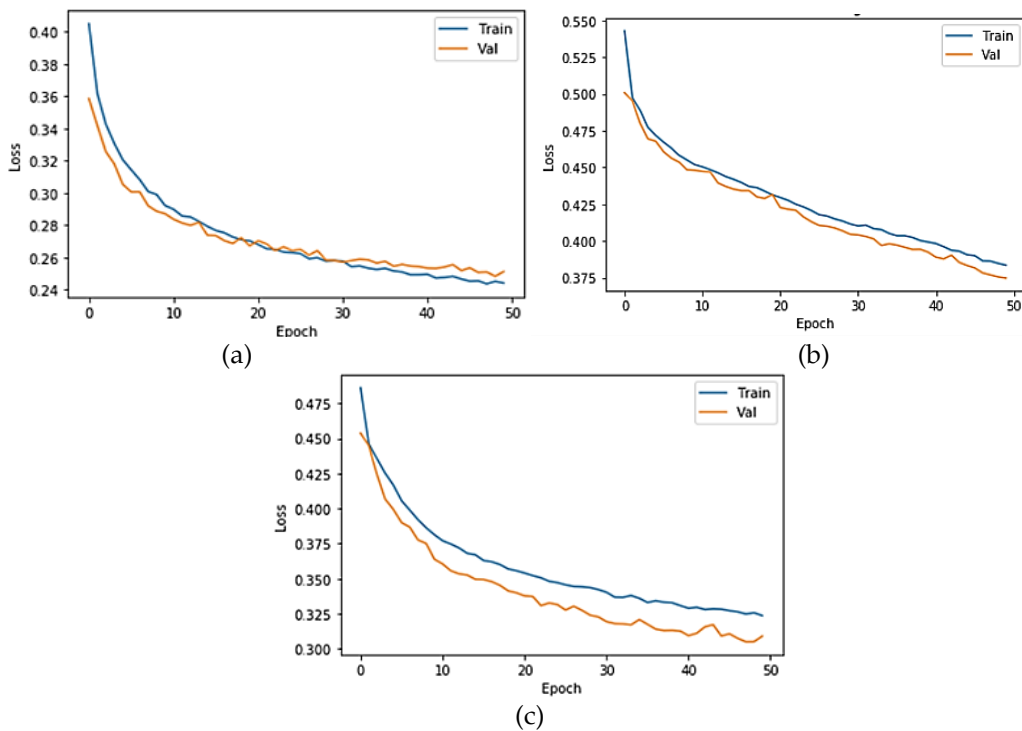


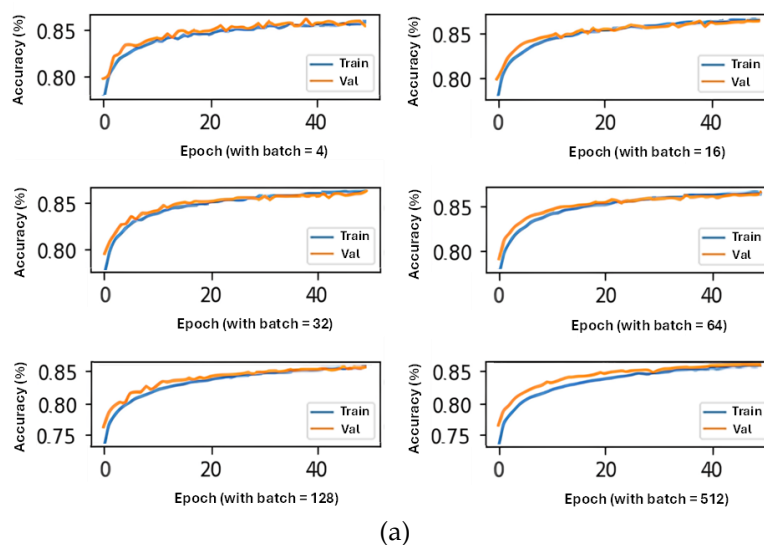
Figure 5 Loss plot of (a) ReLU, (b) tanh, and (c) sigmoid activation functions

In case study 2, the method of hyperparameter tuning with Bayesian optimization has been applied (Victoria and Maragatham, 2021), which is summarized in Table 2. These hyperparameters are a standard set of parameters used in Bayesian optimization. All ranges of parameters' values were used to run the model. In the end, a set of optimum hyperparameters has been selected which will be summarized later. Figure 7 shows the loss and accuracy plots of the model in case study 2. The accuracy of the training is about 82%, which is almost the same as the testing (validation). The loss value is about 0.32 for both training and validation, which means the dataset is fully fitted for both training and testing sets. However, the learning curves of the testing set in the accuracy and loss plots have some noise, and the loss value is less than the result in case study 1. Therefore, case study 3 has been proposed to integrate both cases to obtain a better performance.

Table 2 The hyperparameters to be tuned in case study 2

| Hyperparameters | Types/Values |
|----------------------------|--|
| 1. Optimizer | SGD, RMSprop, Adadelata, Adagrad, Adam, Nadam, Adamax, Ftrl |
| 2. Activation function | ReLU, tanh, sigmoid, Softsign, Softplus, Selu, Elu, exponential, and LeakyReLU |
| 3. Number of neurons | range from 10 to 100 |
| 4. Batch sizes/Epochs | range from 200 to 1000/range from 20 to 100 |
| 5. Learning rate | range from 0.01 to 1 |
| 6. Dropout | present or not (dropout rate range from 0 to 0.3) |
| 7. Batch normalization | present or not |
| 8. Number of hidden layers | range from 1 to 3 |

In case study 3, the number of neurons in layers and the structure of the neural layer, such as the batch normalization layer, the dropout layer, and the dropout rate, were maintained as in case study 2, but the other hyperparameters, such as the activation function, has changed from Softplus to ReLU, and the optimizer has changed from AdaGrad to Adam. The batch size and the number of epochs have changed back to the same value as in case study 1. Figure 8 shows the loss and accuracy plots of the model in case study 3. From the loss plot, the training loss is about 0.23 and the testing loss is about 0.26, which is the condition of overfitting, but it is acceptable. The accuracy plot shows that the training accuracy reaches approximately 90%, while the testing accuracy is around 87%. Although the learning curves exhibit minor fluctuations, these are less pronounced compared to those observed in case study 2. Table 3 shows the summary of hyperparameters in all case studies.



(a)

Figure 6 Relationship between (a) epochs with accuracy and (b) batch with accuracy

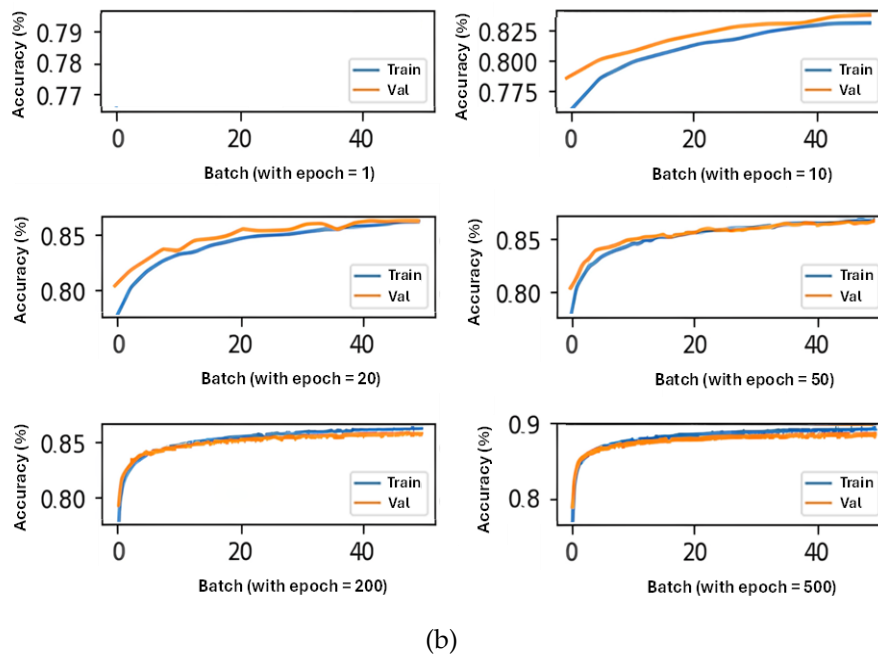


Figure 6 Relationship between (a) epochs with accuracy and (b) batch with accuracy (cont.)

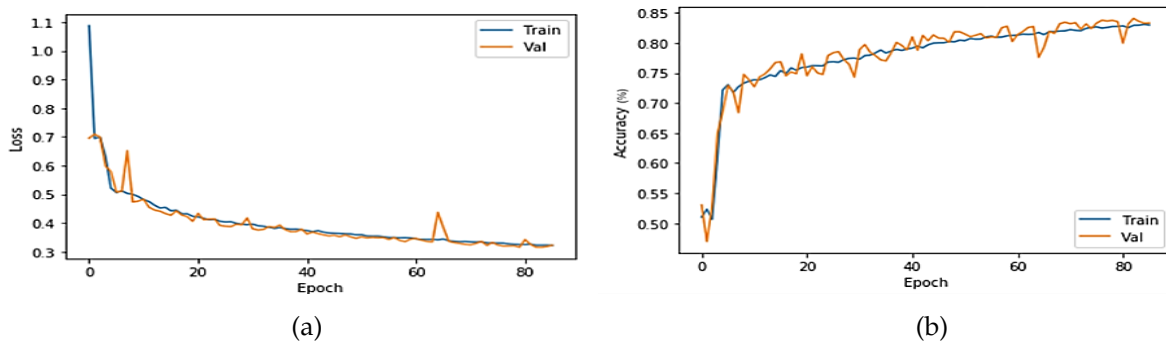


Figure 7 (a) Loss plot and (b) accuracy plot of case study 2

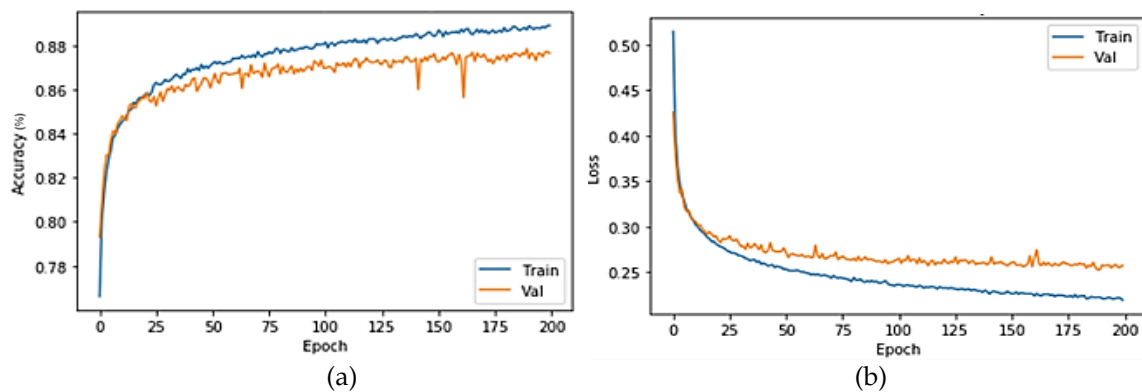


Figure 8 (a) Loss plot and (b) accuracy plot of case study 3

3.2. Prediction with Unseen Data

To test the performance of the proposed lying detection models, the prediction of the unseen data was conducted. The characteristics of the testing videos are one of the concerns that will affect the prediction result. Therefore, the subject's face inside the videos has to be fully shown in the frames. The confidence level obtained from OpenFace 2.0 must always be above 70%, and the frames of the video must be 30 frames per second and 480p resolution. To evaluate the performance

of each case study, the accuracy result was evaluated by testing 10 sample videos obtained from online YouTube sources, each class has 5 videos, and the accuracy is the average of the total number for each class. Table 3 shows the accuracy performance of the training and testing models as well as the performance of the prediction on unseen data under all case studies.

Table 3 Summary of hyperparameters in different case studies

| Case Study 1 | Case Study 2 | Case Study 3 |
|---|---|---|
| 1 input layer, 256 neurons, ReLU activation function, HeNormal weight initializer, | 1 input layer, 93 neurons, Softplus activation function, | 1 input layer, 93 neurons, ReLU activation function, |
| 1 st hidden layer, 256 neurons, ReLU activation function, HeNormal weight initializer, | 1 batch normalization layer after input layer, | 1 batch normalization layer after input layer, |
| 2 nd hidden layer, 128 neurons, ReLU activation function, HeNormal weight initializer, | 1 st hidden layer, 93 neurons, Softplus activation function, | 1 st hidden layer, 93 neurons, ReLU activation function, |
| Dropout layers after the input layer and hidden layers with a dropout rate of 20%. | Dropout layer after 1st hidden layer with 11% of dropout rate, | Dropout layer after 1st hidden layer with 11% of dropout rate, |
| 1 output layer, 1 neuron, sigmoid activation function, | 2 nd hidden layer, 93 neurons, Softplus activation function, | 2 nd hidden layer, 93 neurons, ReLU activation function, |
| 128 batch size and 200 epochs. | 3 rd hidden layer, 93 neurons, Softplus activation function, | 3 rd hidden layer, 93 neurons, ReLU activation function, |
| | 1 output layer, 1 neuron, sigmoid activation function, | 1 output layer, 1 neuron, sigmoid activation function, |
| | 921 batch size and 86 epochs. | 128 batch size and 200 epochs. |

From Table 4, it can be observed that Case Study 1 and Case Study 2 have the best comparable accuracy performances, with almost 90% of training accuracy and 88% of testing accuracy. For the prediction on unseen data, case study 3 achieved the highest truth prediction score of 69% and lie prediction score of 59%. However, all the predictions are not as accurate as in the training and testing stages. This is mostly due to the lack of unseen data collected from the online video resources, and most of the sources are not reliable and consistent.

Table 5 shows the results of this study for detecting lying behaviour using the ANN model. Hyperparameter tuning is a machine-based trial-and-error process based on the calculation made on the machine. Although this can as a determinant factor for better parameters to improve performance, and can also cause the model training to be overfitted. Therefore, manual adjustment helps to resolve the overfitting problem, and it is more adaptable to various conditions.

Table 4 Performance comparisons of different case studies

| Proposed Model | Accuracy Performance | | Prediction Performance | |
|----------------|----------------------|---------|------------------------|-------|
| | Training | Testing | Truth | Lie |
| Case study 1 | 90.15% | 87.95% | 64.8% | 54.4% |
| Case study 2 | 82.39% | 81.61% | 57.2% | 59.2% |
| Case study 3 | 90.31% | 87.65% | 69% | 59% |

Table 5 Results of different case studies

| Case Study 1 | Case Study 2 | Case Study 3 |
|---|--|--|
| ReLU activation function has a better performance than tanh and sigmoid based on loss plot, | Softplus activation function is chosen by Bayesian optimization to obtain optimum performance, | ReLU activation function, batch size of 128, and epochs of 200 are selected to achieve the best performance while maintaining the other optimum hyperparameters tuned in case study 2, |
| A batch size of 128 and epochs of 200 achieve the best performance, | 921 batch size and 86 epochs are selected, | |
| Loss performance is better than in case study 2. | Learning curves of the testing set in the accuracy and loss plots have some noises, | Overfitting has reduced as compared to case study 2, |
| | Fine-tuning results without regularization for the model may cause training to be overfitted. | Noises in learning curves are less than case study 2, |
| | | Performance is almost the same as case study 1 under testing, but better under unseen data prediction. |

4. Conclusions

In conclusion, a deep learning model has been applied in this project to predict the possible lying behaviours extracted from the videos. The dataset for the training was obtained from an online source, and every video was edited to make sure the subject's face appeared in the frames. OpenFace 2.0 was used to extract AU information from the video frames and convert the data into a CSV file. The unwanted data was then filtered out in a CSV file to prevent the training accuracy from being disrupted. With the aid of the ANN method, a huge dataset of micro-expressions obtained from OpenFace 2.0 has been trained and tested in three different case studies with different hyperparameters. Case study 3 has the best performance for both training and testing accuracy. To assess the performance of the case studies, additional unseen data obtained from online sources was used for prediction. Case study 3 achieved the highest prediction accuracy. Hyperparameter optimization remains a valuable method for refining neural network parameters to further enhance model accuracy. The performance of the optimization may be worse than the default parameter model due to factors such as the parameter range being too narrow and the model not covering the regularization layer, which causes the model to be overfit. Therefore, further modifications need to be carried out to fine-tune the parameters according to the requirements and conditions. For future recommendations, more reliable unseen data must be collected to evaluate and improve the deep learning models. Image pre-processing, such as annotation using Roboflow, can be applied to analyze the performance of the models under challenging scenarios. Other features such as head positioning, eye gaze positioning, HOG features can be used, or possible lying actions.

Author Contributions

Conceptualization, Gwo Chin Chung and Soo Fun Tan; Methodology, Gwo Chin Chung and It Ee Lee; Software, Wei Ci Gan; Investigation, Wei Ci Gan and Gwo Chin Chung; Writing – Original Draft, Wei Ci Gan and Gwo Chin Chung; Writing – Review & Editing, Gwo Chin Chung and Kah Yoong Chan; Supervision, Gwo Chin Chung. All authors have read and approved the final version of the manuscript.

Conflict of Interest

No potential competing interest was reported by the author(s).

References

- Abdolrasol, MG, Hussain, SMS, Ustun, TS, Sarker, MR, Hannan, MA, Mohamed, R, Ali, JA, Mekhilef, S & Milad, A 2021, 'Artificial neural networks based optimization techniques: A review', *Electronics*, vol. 10, no. 21, article 2689, <https://doi.org/10.3390/electronics10212689>
- Abdullah, MSN, Karim, HA & AlDahoul, N 2023, 'A combination of light pre-trained convolutional neural networks and long short-term memory for real-time violence detection in videos', *International Journal of Technology*, vol. 14, no. 6, pp. 1228-1236, <https://doi.org/10.14716/ijtech.v14i6.6655>
- Aftab, F, Bazai, SU, Marjan, S, Baloch, L, Aslam, S, Amphawan, A & Neo, T-K 2023, 'A comprehensive survey on sentiment analysis techniques', *International Journal of Technology*, vol. 14, no. 6, pp. 1288-1298, <https://doi.org/10.14716/ijtech.v14i6.6632>
- Aranjo, S, Kadam, MA, Sharma, MA & Antappan, MA 2021, 'Lie detection using facial analysis electrodermal activity pulse and temperature', *Journal of Emerging Technologies and Innovative Research*, vol. 8, no. 5, pp. 1-13,
- Avola, D, Cinque, L, Foresti, GL & Pannone, D 2019, 'Automatic deception detection in RGB videos using facial action units', *In: Proceedings of the 13th International Conference on Distributed Smart Cameras*, pp. 1-6, <https://doi.org/10.1145/3349801.3349806>
- Baltrusaitis, T, Mahmoud, M & Robinson, P 2015, 'Cross-dataset learning and person-specific normalisation for automatic action unit detection', *In: Proceedings of the 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG 2015)*, pp. 1-6, <https://doi.org/10.1109/FG.2015.7284869>
- Baltrusaitis, T, Zadeh, A, Lim, YC & Morency, LP 2018, 'OpenFace 2.0: Facial behavior analysis toolkit', *In: Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, pp. 59-66, <https://doi.org/10.1109/FG.2018.00019>
- Barakat, M, Chung, GC, Lee, IE, Pang, WL & Chan, KY 2023, 'Detection and sizing of durian using zero-shot deep learning models', *International Journal of Technology*, vol. 14, no. 6, pp. 1206-1215, <https://doi.org/10.14716/ijtech.v14i6.6640>
- Ben, X, Ren, Y, Zhang, J, Wang, SJ, Kpalma, K, Meng, W & Liu, YJ 2021, 'Video-based facial micro-expression analysis: A survey of datasets, features and algorithms', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5826-5846, <https://doi.org/10.1109/TPAMI.2021.3067464>
- Canal, FZ, Müller, TR, Matias, JC, Scotton, GG, de Sa Junior, AR, Pozzebon, E & Sobieranski, AC 2022, 'A survey on facial emotion recognition techniques: A state-of-the-art literature review', *Information Sciences*, vol. 582, pp. 593-617, <https://doi.org/10.1016/j.ins.2021.10.005>
- Chavali, GK, Bhavaraju, SKN, Adusumilli, T & Puripanda, V 2014, 'Micro-expression extraction for lie detection using Eulerian video (motion and color) magnification', *Dissertation*, Blekinge Institute of Technology, Sweden
- Dapito, JL & Chua, AY 2024, 'Coefficient of performance prediction model for an on-site vapor compression refrigeration system using artificial neural network', *International Journal of Technology*, vol. 15, no. 3, pp. 505-516, <https://doi.org/10.14716/ijtech.v15i3.6728>
- Ekman, P & Rosenberg, E 2005, *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS)*, Oxford University Press, Cary, NC, United States
- Ekundayo, OS & Viriri, S 2021, 'Facial expression recognition: A review of trends and techniques', *IEEE Access*, vol. 9, pp. 136944-136973, <https://doi.org/10.1109/ACCESS.2021.3113464>
- Hor, SL, AlDahoul, N, Karim, HA, Lye, MH, Mansor, S, Fauzi, MFA & Wazir, ASB 2022, 'Deep active learning for pornography recognition using ResNet', *International Journal of Technology*, vol. 13, no. 6, pp. 1261-1270, <https://doi.org/10.14716/ijtech.v13i6.5842>
- Kreiensteck, E, Ai, Y & Zhang, L 2023, 'A comprehensive evaluation of OpenFace 2.0 gaze tracking', *In: Proceedings of the International Conference on Human-Computer Interaction*, pp. 532-549, https://doi.org/10.1007/978-3-031-35596-7_34
- Li, J, Soladie, C & Segulier, R 2019, 'A survey on databases for facial micro-expression analysis', *In: Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pp. 241-248, <https://doi.org/10.5220/0007309202410248>
- Ma, C, Chen, L & Yong, J 2019, 'AU R-CNN: Encoding expert prior knowledge into R-CNN for action unit detection', *Neurocomputing*, vol. 355, pp. 35-47, <https://doi.org/10.1016/j.neucom.2019.03.082>
- Mehendale, N 2020, 'Facial emotion recognition using convolutional neural networks (FERC)', *SN Applied Sciences*, vol. 2, no. 3, article 446, <https://doi.org/10.1007/s42452-020-2234-1>

- Nadeeshani, M, Jayaweera, A & Samarasinghe, P 2020, 'Facial emotion prediction through action units and deep learning', *In: Proceedings of the 2nd International Conference on Advancements in Computing*, vol. 1, pp. 293-298, <https://doi.org/10.1109/ICAC51239.2020.9357138>
- Namba, S, Sato, W & Yoshikawa, S 2021, 'Viewpoint robustness of automated facial action unit detection systems', *Applied Sciences*, vol. 11, no. 23, article 11171, <https://doi.org/10.3390/app112311171>
- National Research Council Committee 2003, *The polygraph and lie detection*, National Academies Press, Washington DC, United States
- Nurçin, FV, Imanov, E, Işın, A & Ozsahin, DU 2017, 'Lie detection on pupil size by back propagation neural network', *Procedia Computer Science*, vol. 120, pp. 417-421, <https://doi.org/10.1016/j.procs.2017.11.258>
- Owayjan, M, Kashour, A, Al Haddad, N, Fadel, M & Al Souki, G 2012, 'The design and development of a lie detection system using facial micro-expressions', in *Proceedings of the 2nd International Conference on Advances in Computational Tools for Engineering Applications*, Beirut, Lebanon, pp. 33-38, <https://doi.org/10.1109/ICTEA.2012.6462897>
- Rad, D, Paraschiv, N & Kiss, C 2023, 'Neural network applications in polygraph scoring—A scoping review', *Information*, vol. 14, no. 10, article 564, <https://doi.org/10.3390/info14100564>
- Rahman, HA, Shah, MQ & Qasim, F 2023, 'Artificial intelligence based approach to analyze the lie detection using skin and facial expression', *Pakistan Journal of Emerging Science and Technologies*, vol. 4, no. 4, pp. 1-16, <https://doi.org/10.58619/pjest.v4i4.159>
- Rasamoelina, AD, Adjailia, F & Sinčák, P 2020, 'A review of activation function for artificial neural network', *In: Proceedings of the 18th World Symposium on Applied Machine Intelligence and Informatics, IEEE 18th World Symposium on Applied Machine Intelligence and Informatics*, pp. 281-286, <https://doi.org/10.1109/SAMI48414.2020.9108717>
- Revina, IM & Emmanuel, WS 2021, 'A survey on human face expression recognition techniques', *Journal of King Saud University—Computer and Information Sciences*, vol. 33, no. 6, pp. 619-628, <https://doi.org/10.1016/j.jksuci.2018.09.002>
- Şen, MU, Perez-Rosas, V, Yanikoglu, B, Abouelenien, M, Burzo, M & Mihalcea, R 2020, 'Multimodal deception detection using real-life trial data', *IEEE Transactions on Affective Computing*, vol. 13, no. 1, pp. 306-319, <https://doi.org/10.1109/TAFFC.2020.3015684>
- Sharifnejad, M, Shahbahrami, A, Akoushideh, A & Hassanpour, RZ 2021, 'Facial expression recognition using a combination of enhanced local binary pattern and pyramid histogram of oriented gradients features extraction', *IET Image Processing*, vol. 15, no. 2, pp. 468–478, <https://doi.org/10.1049/ipr2.12037>
- Sharma, V 2022, 'A study on data scaling methods for machine learning', *International Journal for Global Academic & Scientific Research*, vol. 1, no. 1, pp. 31-42, <https://doi.org/10.55938/ijgasr.v1i1.4>
- Shen, X, Fan, G, Niu, C & Chen, Z 2021, 'Catching a liar through facial expression of fear', *Frontiers in Psychology*, vol. 12, article 675097, <https://doi.org/10.3389/fpsyg.2021.675097>
- Takalkar, M, Xu, M, Wu, Q & Chaczko, Z 2018, 'A survey: Facial micro-expression recognition', *Multimedia Tools and Applications*, vol. 77, no. 15, pp. 19301-19325, <https://doi.org/10.1007/s11042-017-5317-2>
- Tran-Le, MT, Doan, AT & Dang, TT 2021, 'Lie detection by facial expressions in real time', *In: Proceedings of the International Conference on Decision Aid Sciences and Application*, pp. 787-791, <https://doi.org/10.1109/DASA53625.2021.9682390>
- United Nations 2024, Make cities and human settlements inclusive, safe, resilient and sustainable, Department of Economic and Social Affairs 2024, viewed 1 October 2024 <https://sdgs.un.org/goals/goal11>
- Victoria, AH & Maragatham, G 2021, 'Automatic tuning of hyperparameters using Bayesian optimization', *Evolving Systems*, vol. 12, no. 1, pp. 217-223, <https://doi.org/10.1007/s12530-020-09345-2>
- Zahara, L, Musa, P, Wibowo, EP, Karim, I & Musa, SB 2020, 'The facial emotion recognition (FER-2013) dataset for prediction system of micro-expressions face using the convolutional neural network (CNN) algorithm based Raspberry Pi', *In: Proceedings of the 5th International Conference on Informatics and Computing*, pp. 1–9, <https://doi.org/10.1109/ICIC50835.2020.9288560>
- Zhou, L, Shao, X & Mao, Q 2021, 'A survey of micro-expression recognition', *Image and Vision Computing*, vol. 105, article 104043, <https://doi.org/10.1016/j.imavis.2020.104043>