International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 14, No.1 (Oct-23)

http://dx.doi.org/10.12785/ijcds/1401109

Optimizing ANN-Based Lyapunov Stability for Facial Expression Recognition as A Base Monitoring Neurological Disorders

Ahmad Ilham*[1,2], Hadiyanto Hadiyanto[1,3,4] and Catur Edi Widodo[1,4,5]

[1]*Doctoral Program of Information Systems, Diponegoro University, Semarang, Indonesia*
[2]*Department of Informatics, Universitas Muhammadiyah Semarang, Semarang, Indonesia*
[3]*Center of Biomass and Renewable Energy (C-BIORE), Diponegoro University, Semarang, Indonesia*
[4]*School of Postgraduate Studies, Diponegoro University, Semarang, Indonesia*
[5]*Department of Physics, Diponegoro University, Semarang, Indonesia*

Received 8 Jul. 2023, Revised 12 Oct. 2023, Accepted 13 Oct. 2023, Published 15 Oct. 2023

Abstract: This study emphasizes the importance of facial expression recognition in identifying neurological problems in individuals with limited verbal communication abilities. Current evaluation models are time-consuming and expensive, hindering medical professionals. To address these limitations, we present an improved artificial neural network based on lyapunov stability theory (ANN-LST). By combining these models, we tackle convergence issues while encountering overfitting problems with high-dimensional data, thereby affecting prediction and analysis. Our approach employs principal component analysis (PCA) for feature reduction and extraction to effectively solve overfitting problems. The proposed model was evaluated using the Japanese female facial expression (JAFFE) database and the ahmad ilham simple face database (AIsFD) as our own database, with accuracy (ACC) as the evaluation metric. The results demonstrate higher recognition rates and faster training speeds owing to the adaptive learning rate parameters and the extraction of relevant feature information. The proposed system achieved a 13% higher success rate than face recognition systems using raw images alone. Overall, this model represents a significant advancement and offers promising applications for facial expression recognition in patients with neurological disorders.

Keywords: Face Expression Recognition, Feature Reduction and Extraction, Overfitting, Neurological Disorders

1. Introduction and Related Works

Facial expressions play a crucial role in identifying and understanding human emotions and mentality, as evidenced by psychological research [1]. Language accounts for only 7% of human communication, whereas speech and facial expressions account for 38% and 55%, respectively. Facial expressions, characterized by fleeting and involuntary movements of facial muscles, provide genuine insight into hidden emotions [2]. This significance extends to the field of neuroscience in psychiatry, where facial expressions serve as valuable tools for diagnostic and therapeutic approaches targeting the dopaminergic system in patients with Parkinson's disease [3] and psychiatric disorders [4]. Due to the impact of dopaminergic dysfunction on physiological processes, such as heart rate and body temperature, as well as behavioral issues, such as anxiety and depression [4], techniques for facial expression recognition have become particularly valuable in understanding patients' emotional states. Automated expression detection has been widely adopted in critical healthcare domains, including clinical diagnosis of mental disorders, interrogation, interviews, and lie detection [5]. Clinically, interpreting patients' facial expressions enables the comprehension of their genuine emotions and facilitates the development of appropriate therapeutic strategies to support the healing process of patients with neurological disorders.

Extensive discussions and investigations have been conducted on strategies for treating patients with neurological disorders. Traditional approaches, including a two-week interview period to observe changes in the patient's mood, cognition, and neurovegetative functions (such as vasomotor-hyper sympathicotonia), have been implemented [6]. Nevertheless, these strategies have demonstrated limited success in managing patients with mild mental symptoms, with over 20% of cases exhibiting non-response to standard

E-mail address: ahmadilham@unimus.ac.id, hadiyanto@live.undip.ac.id, catur.ediwidodo@gmail.com*

interventions [7]. Moreover, technology has been explored extensively, particularly in severe cases of mental illness, through the introduction of Deep Brain Stimulation (DBS) as a potentially effective diagnostic and treatment approach for individuals with treatment-resistant depressive symptoms [8]. Additional strategies targeting severe depression include pharmacological interventions, such as transcranial magnetic stimulation and electroconvulsive therapy [9]. Notably, recent advancements have incorporated diverse recovery processes to assess the severity of depression in patients, considering the variability in emotional sensitivity and reactivity, both of which can be evaluated through the analysis of facial expressions [4].

However, recognizing emotions through facial expressions poses a significant challenge for researchers in the field of computer vision. This challenge arises because of the brief duration of facial expressions, typically lasting between 100 and 500 ms, and the subtle intensity of facial muscle movements during these expressions, as individuals often conceal their emotions and suppress their facial cues. Moreover, only a subset of distinctive facial muscle movements associated with facial expressions has been observed [10]. As stated in [11], even experts in the field can identify only five types of expressions with an accuracy as low as 47%. Consequently, there is a critical need to develop automated systems that can accurately detect facial expressions. This need is particularly pressing in light of advancements in non-face-to-face interview techniques leveraging artificial intelligence [12], as well as the utilization of computer-assisted interventions for mood disorders [13]. Hence, the demand for precise techniques for gradually recognizing facial expressions is steadily increasing.

Numerous studies have been conducted on automatic facial analysis using various techniques to examine facial expressions. These techniques include geometric features [14], facial emotions [15], facial landmarks [16], the facial unit action system (FACS) [6], and various other behavioral features such as the mouth and eye deformation [17], gaze direction [18], and pupil dilation [19]. These features possess inherent interpretative qualities and exhibit a close relationship with mood, rendering them potential biomarkers for diagnosing depression. Emotional features, including smile intensity, anger, happiness, neutrality, disgust, contempt, and variability, have also received extensive attention because of their statistical significance in discerning differences in depression [5]. Additionally, facial action units (FAUs) have been identified as robust indicators of depression discrepancies [20]. Nonetheless, the FAUs annotation process requires human judgment, which introduces subjectivity in identifying and assessing certain action units, thereby resulting in discrepancies between different annotators [21]. In this study, we focused on the FER techniques.

Artificial neural networks (ANN) are widely utilized in facial expression recognition problems in facial emotion techniques owing to their effective integration of statistical and structural information, resulting in notable performance advancements compared to rule-based systems [22]. However, during the training phase, challenges such as slow convergence and frequent entrapment in the local minima of the error performance surface often arise [23]. Various models have been proposed to expedite convergence in the training phase, including momentum terms, standard optimization techniques, and adaptive learning rates (e.g., iterative least-squares, conjugate gradient, quasi-Newton, and Levenberg–Marquardt) [24]. Despite their faster convergence rates compared to traditional gradient-based ANNs, these models encounter significant issues such as extensive storage requirements, high memory demands, intricate computational complexity, and reliance on heuristic knowledge [24].

The Lyapunov stability theory (LST) [25] has garnered considerable attention due to its reported ability to effectively address the aforementioned issue by identifying the global minimum point along the error performance surface of an artificial neural network (ANN) through weight coefficient updates, thereby achieving heightened stability. The adaptive gain-ratio parameter is a critical factor that influences the convergence rate. However, it is worth noting that fixed values were assigned to the candidate Lyapunov function and the adaptive gain-ratio parameter for this model.

Despite the enhanced training speed achieved by the utilization of ANN and LST (ANN-LST), the challenge of overfitting arises in the handling of high-dimensional data [26], [27], [28]. Overfitting arises because of the excessive complexity of the ANN model, causing it to closely adhere to the training data and subsequently perform poorly in predicting unseen data, thereby posing difficulties in stability analysis. To mitigate this, Principal Component Analysis (PCA) can be employed [29]. PCA reduces high-dimensional data to a lower-dimensional feature space by calculating the eigenvalues and eigenvectors of the covariance matrix.

The primary aim of this study was to reduce the overfitting of high-dimensional data and enhance exercise recognition speed through the utilization of PCA on an ANN-LST. Through the application of PCA to reduce data dimensionality, essential information can be preserved while simultaneously reducing the occurrence of overfitting. Consequently, the effectiveness and accuracy of the proposed model were improved, making it more proficient and precise in yielding results and accelerating the training process.

2. Proposed Model and System

The proposed model is illustrated in Fig. 1.

A. *Data Gathering and Schematic Partition*

First, we used the Japanese female facial expression (JAFFE) database [30], which comprises 213 images portraying seven distinct expressions. Among these, six were

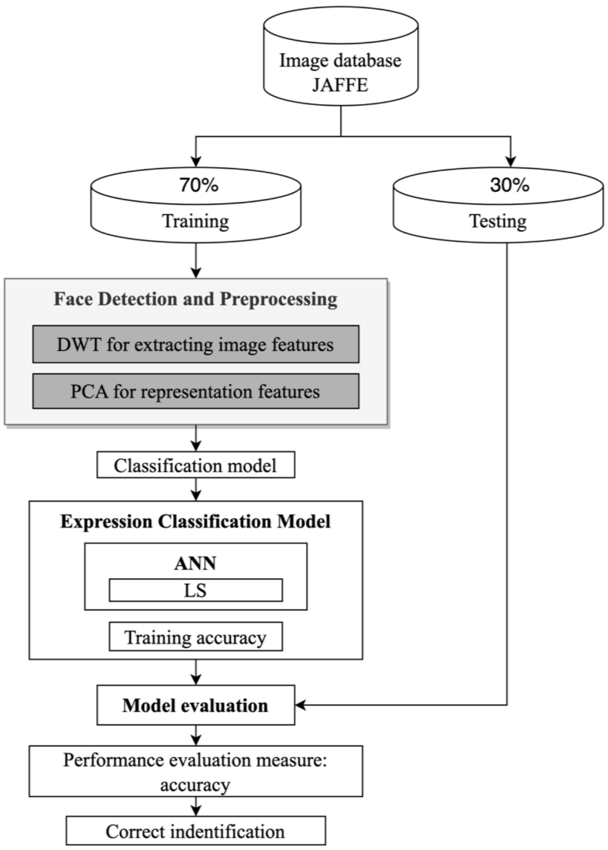


Figure 1. Proposed model for facial expression recognition system as a base for monitoring neurological disorders

considered fundamental facial expressions, while the remaining one represented a neutral facial expression. This database encompasses images generated by ten Japanese female models, each with a resolution of 256 × 256 pixels. To evaluate the emotional content of each image, 60 Japanese individuals assessed them using six adjectives denoting emotions. In this study, all images from the database were preprocessed to obtain facial expression images with normalized intensities, sizes, and shapes. Preprocessing involved eliminating lighting and shading effects following the methodology introduced in [31]. The initial step in the preprocessing phase involves the automatic detection of facial features, including the eyes, nose, and mouth, with the objective of implementing histogram equalization to eliminate lighting variations. The processed and resized data samples are depicted in Figure 2, serving as the independent variable used to examine the response of patients with neurological disorders [32].

Second, we used our own database, named the Ahmad Ilham Simple Face Database (AIsFD) [33]. This database was created by capturing our own face using a 13-megapixel camera on a Vivo 2007 mobile phone. We took pictures of facial expressions while facing the camera, maintaining a

Figure 2. Facial expression samples from the JAFFE database

distance of 40 cm between the camera and the face. A total of eleven images were collected, representing seven distinct expressions: surprised (1), sad (1), afraid (2), happy (1), disgust (2), angry (2), and neutral (2).

To comprehensively analyze the behavior of the proposed model, we meticulously segmented it into four distinct experimental stages. Each stage involved altering the composition of both the training and testing sets at various levels to ensure rigorous evaluation. A schematic of data partitioning for each experiment is presented in Table 1.

B. *Face Detection and Pre-Processing*

In the second stage, discrete wavelet transform (DWT) [34] and principal component analysis (PCA) [35], [36] were employed to extract and represent features from the input database images.

The database image features were extracted using DWT, a two-dimensional (2D) transform. This model encompasses two essential functions: a low-pass filter and a high-pass filter, which are employed to decompose the original image. A low-pass filter generates an approximate image, whereas a high-pass filter generates a detailed image. Based on specific applications, an approximate image can be further divided into deeper levels of approximation and detail. A comprehensive description of the DWT process can be found in [34]. Fig. 3 illustrates a two-level DWT applied to a "happy" facial expression image from the JAFFE database.

Next, the relevant image features were represented using PCA. This model was inspired by [35] and [34]. For instance, a facial image, represented by a 2D matrix, exhibits intensity values with a size of $q \times q$. To facilitate the training process of the ANN model, which serves as the classifier in the face recognition system, the facial image was initially converted into a 2-dimensional vector. Consequently, the training set comprising q^2 facial images can be defined as $Z = z_1, z_2, ...z_p \subset \Re^{q^2} \times p$, where each vector signifies a class to be classified by the face recognition system. Fulfilling this criterion establishes the initial prerequisite for defining the covariance matrix of PCA, as presented in (1).

$$C = \frac{1}{p} \sum_{i=1}^{p} \mathbb{Z}_i - \bar{\mathbb{Z}} * \mathbb{Z}_i - \bar{\mathbb{Z}}^T \quad (1)$$

TABLE I. Database partitioning schematic in each experiment

Experiments	Training %	Testing %
1	80	20
2	70	30
3	60	40
4	50	50

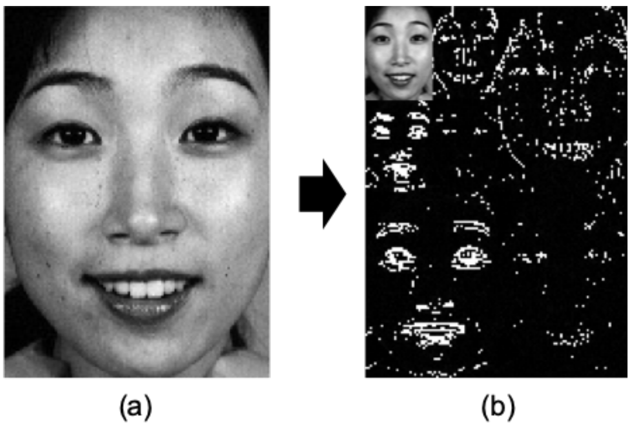


Figure 3. Feature extraction sample using DWT of "happy" face expression image: (a) original image and (b) decomposed subimages

where $\frac{1}{p}\sum_{i=1}^{p}$ is the average image vector.

Subsequently, the eigenvalues and eigenvectors of the covariance matrix C are computed, and a matrix r, consisting of r eigenvectors associated with the largest eigenvalues, is defined as $U = u_1, u_2, \ldots, u_r \in \Re^{q^2 \times p}\,(r < p)$. The eigenface-based features of the original face image, $\mathbb{Z} \subset \Re^{2\times p}$, will be used in the training process, which is obtained by projecting onto the corresponding eigenface space $\mathbb{Z}$, as defined in (2).

$$X = U^T \mathbb{Z} \tag{2}$$

If the system is tested with a new face image $\mathbb{Z}_{new}$, it will be projected into the eigenface space as per Equation (3).

$$X_{new} = U^T\left(z_{new} - \bar{z}\right) \tag{3}$$

where X_{new} represents the eigenface-based features of the new image applied to the system for testing. For more information on the PCA model, please refer to [35], [36].

C. Classification Model

At this stage, we adopted an artificial neural network classification model based on the Lyapunov stability theory (ANN-LST) proposed by [26] as a classification model. The utilized ANN structure accounts for the input-output relations from $\Re^n$ to $\Re^m$, where n and m represent the number of inputs and outputs, respectively. The M-dimensional weight vector contains the weight coefficients $w_t(k)$ of the ANN structure defined in (4).

$$\mathbf{w}_t(k) = \begin{bmatrix} \mathbf{w}_t^{(2,1)}(k) \\ \mathbf{w}_1^{(1,0)}(k) \\ \vdots \\ \mathbf{w}_u^{(1,0)}(k) \end{bmatrix} \tag{4}$$

where u represents the number of neurons in the hidden layers, $M = u(n+1) + (u+1)$ and $t = 1, 2, ..., m$. The equations of state and desired outputs of the ANN structure are formulated in (5) and (6), respectively.

$$w_t = w_1(k-1) + v_t(k) \tag{5}$$

$$d_t(k) = f(w_t(k-1)) + c_t(k) \tag{6}$$

where w_t and $d_t(k)$ are expressions representing random Gaussian processes with zero means.

Similar to the definitions provided above, the structure of the ANN undergoes linearization. Therefore, the linearized output of the ANN is formulated as Equation (7). Comprehensive information can be found in Ref. [26].

$$\tilde{y}_t(k) = h_t^T(k)\, w_t(k) \tag{7}$$

where $h_t(k)$ is the column vector of the variable t obtained from the Jacobian matrix in the linearization process. The linearization process is formulated in (8).

$$\mathbf{h}_t(k) = \begin{bmatrix} \frac{\partial y_t(k)}{\partial \mathbf{w}_t^{(2,1)}(k)} \\ \frac{\partial y_t(k)}{\partial \mathbf{w}_1^{(1,0)}(k)} \\ \vdots \\ \frac{\partial y_t(k)}{\partial \mathbf{w}_u^{(1,0)}(k)} \end{bmatrix} = \begin{bmatrix} G'\left(z_t^{(2,1)}(k)\right)\mathbf{s}^{(2,1)}(k) \\ G'\left(z_t^{(2,1)}(k)\right) w_{t,1}^{(2,1)}(k) F_1'\left(z_1^{(1,0)}(k)\right)\mathbf{s}^{(1,0)}(k) \\ \vdots \\ G'\left(z_t^{(2,1)}(k)\right) w_{t,u}^{(2,1)}(k) F_u'\left(z_u^{(1,0)}(k)\right)\mathbf{s}^{(1,0)}(k) \end{bmatrix} \tag{8}$$

where, $G'\left(z_t^{(2,1)}(k)\right) = 1 - \left(z_t^{(2,1)}(k)\right)^2$ and $F_u'\left(z_u^{(1,0)}(k)\right) =$

$1 - \left(z_u^{(1,0)}(k)\right)^2$ are input vectors from the input layer to the hidden layer where $s^{(1,0)}(k)$ is the input vector from the hidden layer to the output layer. Functions $z_t^{(1,0)}(k)$ and $z_u^{(1,0)}(k)$ are the summed outputs of each neuron, function $F() = .G() = .\left(1 - e^{-\beta(.)}\right)$ is the hyperbolic tangent sigmoid function, and $F'(.)$ and $G'(.)$ represent the derivatives of functions $F(.)$ and $G(.)$.

After the linearization process, the function $V_t(k) = a_t^k e_t^2(k)\,(a_t > 1)$ will be chosen as the candidate Lyapunov function for the output of the k-th ANN. Subsequently, a constrained optimization problem is constructed using the strict negativity condition of the LST, as a constraint function formulated in (9).

$$Argmin\left[\frac{1}{2}\delta w_t^T \delta w_t\right] subjectto \left(a_t^k e_t^2(k) - a_t^{k-1} e_t^2(k-1)\right) < 0, \forall k \tag{9}$$

where $\delta w_t = w_t(k) - w_t(k-1)$ and $e_t(k) = d_t(k) - h_t^T(k)w_t(k)$ represent the error signals of the ANN. Due to the Lyapunov stability of the optimization process, as determined by the cost function formulated in (9), stability is always guaranteed. Solving Equation (9) using the Lagrange multiplier theorem yields the following weight vector update rule for the ANN, as formulated in (10).

$$w_t(k) = w_t(k-1) + \left(1 - \frac{|e_t(k-1)|}{(a_t(k-1))^{\frac{k}{2}}\, a_t(k)}\right) a_t(k) \tag{10}$$

where $a_t(k) = d_t(k) - h_t^T(k)w_t(k)w_t(k-1)$ and $a_t(k) = 1 + \frac{e_t^2(k-1)}{e_{t,}^2(k)}$ represent the prediction errors and rate of adaptation gain, respectively. We note that the updated adaptation gain ratio $a_t(k)$ depends on consecutive training error levels, thus significantly improving the training process. This finding was also reported in [26]. Additionally, this parameter is set to $a_t(0) > 1$ to ensure the asymptotic convergence of the Lyapunov training error $e_t(k)$ is always achieved.

Furthermore, the potential singularity issue formulated in (11) can be resolved by introducing an appropriately small positive variable, denoted as λ, to maintain optimal performance.

$$w_t(k) = w_t(k-1) + \frac{h_t(k)}{\lambda + \|h_t(k)\|^2} \times \left(1 - \frac{|e_t(k-1)|}{\lambda + (a_t(k-1))^{\frac{k}{2}} |a_t(k)|}\right) a_t(k) \tag{11}$$

D. Model Evaluation

To evaluate the model, the accuracy score is formulated in Eq. (12).

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \tag{12}$$

3. Results and Discussion

All models were executed on a 2.5 GHz Intel Core i5 Dual-Core CPU with 16 GB of RAM and were simulated using Python 3.10.6, which incorporates several modules for accuracy calculation and visualization, such as SciKit-Learn, Matplotlib, and Seaborn. Additionally, Pandas was utilized for database processing, and NumPy handled scientific calculations. Python was used to construct the sample application system.

To ensure a fair comparison, we set specific parameters as follows: (i) The adaptation gain of the proposed face recognition system ANN algorithms was set to 0, while the adaptation gain of the other ANN algorithms was set to 1.01. (ii) The initial value of the error signal for both ANN algorithms was set to 0. (iii) The learning rate of the gradient-based ANN model was set to 0.95.

In the experimental tables, the highlighted bold values indicate good performance, while non-bold values indicate poor performance.

Tables 2, 3, 4, and 5 present the results of the experiments (Exp. 1, 2, 3, and 4). As shown in Tables 2, 3, 4, and 5, the proposed model exhibited outstanding performance across all PCA features. Specifically, when utilizing 40 features out of all available, the proposed model achieved the highest performance compared to the other models, with accuracy values of 90.00% in Exp. 1, 92.01% in Exp. 2, 89.62% in Exp. 3, and 84.12% in Exp. 4.

Tables 2, 3, 4, and 5 present the results of the four experiments, highlighting the optimal PCA feature count as 40. This finding serves as the basis for analyzing its impact on the model's training cycle speed. In this study, we investigated the effects of different cycle values ranging from 10 to 210 while using 40 features as a reference. Determining the ideal cycle value for adjusting the model's cycle configuration during training is a topic of ongoing debate in the literature. Therefore, in this study, we adopted a trial-and-error approach.

Figure 4 displays the experimental outcomes, showing that our proposed model achieved a significantly low recognition error rate of 0.101 with 30 cycles during training. This represents a 1% improvement compared to ANN+LST-adaptive [26], a 2% improvement over ANN+LST-non-adaptive [37], a 3% improvement over BP-adaptive, and a 5% improvement over the BP-non-adaptive model. The mean squared error (MSE) values for the three comparative models were 0.270, 0.310, 0.401, and 0.618, respectively.

These results indicate that implementing PCA as a feature extraction technique is sufficient to mitigate overfitting in the ANN+LST-adaptive model. The model's robustness

TABLE II. Model comparison of testing recognition rate in Exp. 1

Number of features	Testing recognition rate: 20%				
	BP-non-adaptive (η = 0.95)	**BP-adaptive** (η = 0.95)	**ANN+LST-non-adaptive [37]** ($\mathbf{a}_t$ = 101)	**ANN+LST-adaptive [26]** ($\mathbf{a}_t$ = 101)	**Proposed model** ($\mathbf{a}_t$ = 101)
20	68.30	75.64	73.63	77.31	85.13
40	76.54	84.31	80.19	86.03	**90.00**
60	79.12	85.00	81.26	85.63	85.97
80	78.49	81.45	82.15	85.17	88.32
100	76.18	80.12	75.13	81.39	88.49
120	75.43	80.00	72.00	80.42	88.16
180	75.89	79.89	61.42	75.51	86.73
240	74.75	79.38	60.19	72.11	85.60

TABLE III. Model comparison of testing recognition rate in Exp. 2

Number of features	Testing recognition rate: 30%				
	BP-non-adaptive (η = 0.95)	**BP-adaptive** (η = 0.95)	**ANN+LST-non-adaptive [37]** ($\mathbf{a}_t$ = 101)	**ANN+LST-adaptive [26]** ($\mathbf{a}_t$ = 101)	**Proposed model** ($\mathbf{a}_t$ = 101)
20	62.13	74.89	76.13	79.66	84.60
40	74.14	81.21	82.19	86.91	**92.01**
60	72.29	82.00	83.17	84.92	87.16
80	73.91	82.41	83.24	84.31	87.82
100	77.96	74.12	74.24	82.79	89.12
120	78.81	80.00	76.11	79.94	84.71
180	74.32	78.19	70.19	77.75	85.12
240	76.51	79.18	69.26	80.13	86.16

TABLE IV. Model comparison of testing recognition rate in Exp. 3

Number of features	Testing recognition rate: 40%				
	BP-non-adaptive (η = 0.95)	**BP-adaptive** (η = 0.95)	**ANN+LST-non-adaptive [37]** ($\mathbf{a}_t$ = 101)	**ANN+LST-adaptive [26]** ($\mathbf{a}_t$ = 101)	**Proposed model** ($\mathbf{a}_t$ = 101)
20	60.62	70.1	74.16	78.93	84.11
40	69.45	72.39	78.92	80.42	**89.62**
60	73.23	80.1	80.12	80.79	82.17
80	75.93	79.45	79.81	80.11	81.63
100	72.28	76.23	72.32	82.73	85.32
120	71.3	70.71	71.08	78.55	86.47
180	70.19	75.19	71.42	77.94	82.35
240	72.75	74. 83	58.43	79.44	82.07

TABLE V. Model comparison of testing recognition rate in Exp. 4

Number of features	Testing recognition rate: 50%				
	BP-non-adaptive ($\eta = 0.95$)	**BP-adaptive** ($\eta = 0.95$)	**ANN+LST-non-adaptive [37]** ($\mathbf{a}_t = 101$)	**ANN+LST-adaptive [26]** ($\mathbf{a}_t = 101$)	**Proposed model** ($\mathbf{a}_t = 101$)
20	60.11	69.66	70.52	74.10	80.11
40	68.04	70.82	75.12	76.17	**84.12**
60	71.56	78.42	80.02	81.09	82.98
80	74.19	73.81	74.32	78.33	80.11
100	70.48	73.11	70.41	77.81	83.53
120	70.15	70.02	70.08	75.35	82.21
180	70.04	73.52	71.38	72.23	80.61
240	70.91	72. 11	52.21	72.95	82.14

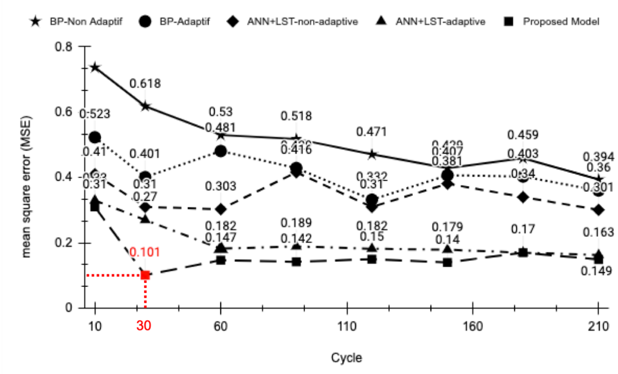


Figure 4. Comparison of the training recognition error rate in 40 features PCA on the JAFFE database

against overfitting can be attributed to the accurate representation of the true data distribution in the training database. Geometrically, this implies that all non-zero regions of the true (unknown) data distribution contain a corresponding number of training data points, which is typically a characteristic of a well-structured database.

Based on the experimental results obtained from the JAFFE database, we recommend using PCA in ANN+LST-based adaptive learning for high-dimensional databases.

To evaluate the effectiveness of the proposed model, we conducted tests on a facial expression recognition system using expression image data, both separately and in real-time. The results are presented in Tables 6 and 7. Table 6 demonstrates that the proposed model achieves superior accuracy when compared to the ANN+LST-non-adaptive [37] and ANN+LST-adaptive [26] models. For nearly all randomly selected facial expression labels, the proposed model achieves high accuracy rates: 85.91% for "Neutral," 88.21% for "Sad," 82.63% for "Surprise," 86.13% for "Fear," 89.82% for "Happy," and 83.04% for "Angry." The only exception is the "Disgust" expression, where ANN+LST-adaptive outperforms both ANN+LST-non-adaptive and the proposed model. This can be attributed to the fact that ANN models generally require substantial training data to achieve optimal results. The use of more training data enables the neural network to effectively learn and comprehend patterns within the data, thereby allowing it to capture the variations and complexities inherent in the database.

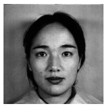
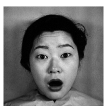
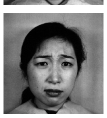
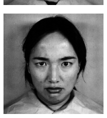
Table 7 presents the novel findings regarding real-time data usage. Adaptive-based models, including ANN + LST-adaptive [26] and the proposed model, exhibit significant improvements compared with experiments conducted on the JAFFE database. The average accuracy increased to 83.18% and 88.71% for the aforementioned models. Conversely, non-adaptive models such as ANN+LST-non-adaptive [37] experience a substantial decrease in average accuracy, Dropping to 71.80%, as depicted in Table 6. These results corroborate the findings of previous studies conducted by [36] and [29], which emphasized the enhanced performance achieved through the integration of feature extraction techniques using PCA and conventional machine learning models in facial expression recognition. Thus, this study provides additional empirical evidence to support the notion that the integration of robust feature extraction techniques can enhance the performance of the proposed face recognition system based on an adaptive model.

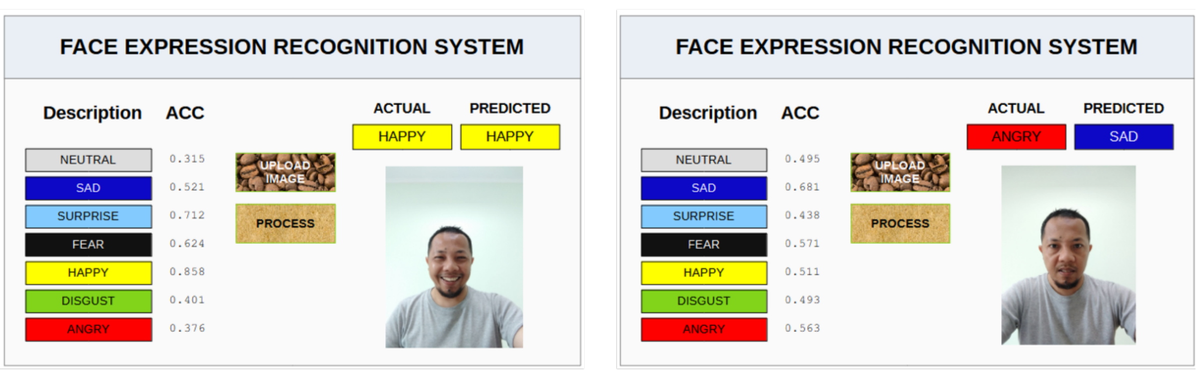
Moreover, the proposed facial expression recognition system is illustrated in Fig. 5 with two examples. In Fig. 5(a), the actual facial expression is 'Happy,' and the suggested model accurately predicts it. Conversely, in Fig. 5(b), the actual expression is 'Angry', while the predicted label based on JAFFE labels is 'Sad'.

4. Conclusions and Future Work

This study presents a novel approach that can be used to identify the expression and mentality of patients with neurological disorders, especially those with limited verbal communication skills. It is well known that current evaluation models have time and cost limitations that may hinder medical professionals from providing treatment. In this study, we successfully integrated PCA and ANN+LST-

TABLE VI. Comparison proposed model and previews studies for each randomly selected sample image expression from JAFFE

Sample Images	Descriptions	Predicted Expression (ACC %)		
		ANN+LST-Non-Adaptive [37]	ANN+LST- adaptive [26]	Proposed Model
	Neutral	76.83	80.14	**85.91**
	Sad	76.53	79.85	**88.21**
	Surprise	78.18	81.42	**82.63**
	Fear	74.08	79.64	**86.13**
	Happy	78.62	83.50	**89.82**
	Disgust	73.14	**85.22**	81.53
	Angry	71.06	76.60	**83.04**
	Average	75.49	80.91	**85.32**



(a) (b)

Figure 5. Proposed system interface. (a) Correct recognition predicted; (b) Incorrect recognition predicted

TABLE VII. Comparison proposed model and previews studies for each sample image expression for real-time image testing

Sample Images	Descriptions	Predicted Expression (ACC %)		
		ANN+LST-Non-Adaptive [37]	**ANN+LST- adaptive [26]**	**Proposed Model**
	Neutral	71.83	82.11	**87.62**
	Sad	69.75	80.24	**89.19**
	Surprise	74.33	82.11	**87.37**
	Fear	70.88	83.58	**88.51**
	Happy	73.49	85.14	**89.97**
	Disgust	71.52	89.43	**89.71**
	Angry	70.81	79.62	**88.58**
	Average	71.80	83.18	**88.71**

adaptive models. The findings of this study showed excellent results. PCA effectively mitigates the high-dimensional data problem by selecting relevant features, thereby overcoming convergence challenges and minimizing the impact of overfitting on ANN+LST-adaptive. The results demonstrated excellent accuracy compared to the other models used in the experiments. We also tested the model on systems with separate and real-time databases, and the results were quite promising compared to those of previous models. Therefore, it can be concluded that PCA can enhance the ANN+LST-adaptive model, which can be applied in FER systems.

For future research, we are currently working on another study to integrate VGG-16 as a pre-train on ANN+LST-adaptive for extracting informative features from an input image database, with the aim of improving FER accuracy. We will test this approach not only on the JAFFE database but also on several other databases, including RaFD [38], MUG [39], TFEID [40], FER2013 [41], and CK+ [42]. In future studies, we may utilize real data from patients with neurological disorders, as we are currently preparing a research funding proposal focused on collecting databases directly from patients in psychiatric hospitals. To the best of our knowledge, there is no similar database derived directly from patients.

Acknowledgement

We extend our gratitude to our supervisors, including Prof. Dr. Hadiyanto, S.T., M.Sc., and Dr. Drs. Catur Edi Widodo, M.T., from the Information Systems Doctoral Program at Diponegoro University. We also appreciate the funding support provided by Universitas Muhammadiyah Semarang throughout the research process.

References

[1] A. Mehrabian, "Communication without words," in *Communication theory*. Routledge, 2017, pp. 193–200.

[2] P. Ekman, W. V. Friesen, M. O'sullivan, A. Chan, I. Diacoyanni-Tarlatzis, K. Heider, R. Krause, W. A. LeCompte, T. Pitcairn, P. E. Ricci-Bitti *et al.*, "Universals and cultural differences in the judgments of facial expressions of emotion." *Journal of personality and social psychology*, vol. 53, no. 4, p. 712, 1987.

[3] U. S. Clark, S. Neargarder, and A. Cronin-Golomb, "Specific impairments in the recognition of emotional facial expressions in parkinson's disease," *Neuropsychologia*, vol. 46, no. 9, pp. 2300–2309, 2008.

[4] K. Grabowski, A. Rynkiewicz, A. Lassalle, S. Baron-Cohen, B. Schuller, N. Cummins, A. Baird, J. Podgórska-Bednarz, A. Pieniażek, and I. Łucka, "Emotional expression in psychiatric conditions: New technology for clinicians," *Psychiatry and clinical neurosciences*, vol. 73, no. 2, pp. 50–62, 2019.

[5] A. Ilham, S. Suryono, and C. E. Widodo, "A review of face expression recognition system recommendation for psychiatric disease," in *AIP Conference Proceedings*, vol. 2683, no. 1. AIP Publishing, 2023.

[6] M. Gavrilescu and N. Vizireanu, "Predicting depression, anxiety, and stress levels from videos using the facial action coding system," *Sensors*, vol. 19, no. 17, p. 3693, 2019.

[7] S. H. Kennedy, P. Giacobbe, S. J. Rizvi, F. M. Placenza, Y. Nishikawa, H. S. Mayberg, and A. M. Lozano, "Deep brain stimulation for treatment-resistant depression: follow-up after 3 to 6 years," *American Journal of Psychiatry*, vol. 168, no. 5, pp. 502–510, 2011.

[8] A. L. Crowell, S. J. Garlow, P. Riva-Posse, and H. S. Mayberg, "Characterizing the therapeutic response to deep brain stimulation for treatment-resistant depression: a single center long-term perspective," *Frontiers in integrative neuroscience*, vol. 9, p. 41, 2015.

[9] Z. J. Daskalakis, A. J. Levinson, and P. B. Fitzgerald, "Repetitive transcranial magnetic stimulation for major depressive disorder: a review," *The Canadian Journal of Psychiatry*, vol. 53, no. 9, pp. 555–566, 2008.

[10] W.-J. Yan, Q. Wu, J. Liang, Y.-H. Chen, and X. Fu, "How fast are the leaked facial expressions: The duration of micro-expressions," *Journal of Nonverbal Behavior*, vol. 37, pp. 217–230, 2013.

[11] 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 & Intelligent Hardware*, vol. 3, no. 1, pp. 1–17, 2021.

[12] J. Lee, P. H. Tung, E. Im, and G. Gim, "Ai recruitment system using eeg to explore the truth of interviewers," in *Emotional Artificial Intelligence and Metaverse*. Springer, 2022, pp. 71–84.

[13] J. H. Wright, M. Mishkind, T. D. Eells, and S. R. Chan, "Computer-assisted cognitive-behavior therapy and mobile apps for depression and anxiety," *Current psychiatry reports*, vol. 21, pp. 1–9, 2019.

[14] D. Ghimire and J. Lee, "Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines," *Sensors*, vol. 13, no. 6, pp. 7714–7734, 2013.

[15] J. Zhang, Z. Yin, P. Chen, and S. Nichele, "Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review," *Information Fusion*, vol. 59, pp. 103–126, 2020.

[16] Y. Wu and Q. Ji, "Facial landmark detection: A literature survey," *International Journal of Computer Vision*, vol. 127, pp. 115–142, 2019.

[17] Z. S. Syed, K. Sidorov, and D. Marshall, "Depression severity prediction based on biomarkers of psychomotor retardation," in *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge*, 2017, pp. 37–43.

[18] L. He, D. Jiang, and H. Sahli, "Automatic depression analysis using dynamic facial appearance descriptor and dirichlet process fisher encoding," *IEEE Transactions on Multimedia*, vol. 21, no. 6, pp. 1476–1486, 2018.

[19] K. M. Charan, A. K. Aswin, K. A. Varshini, and S. Kirthica, "Implementation of pupil dilation in ai-based emotion recognition," in *Artificial Intelligence and Technologies: Select Proceedings of ICRTAC-AIT 2020*. Springer, 2021, pp. 447–454.

[20] E. A. Clark, J. Kessinger, S. E. Duncan, M. A. Bell, J. Lahne, D. L. Gallagher, and S. F. O'Keefe, "The facial action coding system for characterization of human affective response to consumer product-based stimuli: a systematic review," *Frontiers in psychology*, vol. 11, p. 920, 2020.

[21] K. Belhouchette, "Facial action units detection to identify interest emotion: An application of deep learning," *Advances in Data Science and Adaptive Analysis*, vol. 14, no. 01n02, p. 2250006, 2022.

[22] O. N. Al-Allaf, "Review of face detection systems based artificial neural networks algorithms," *arXiv preprint arXiv:1404.1292*, 2014.

[23] F. Nonis, N. Dagnes, F. Marcolin, and E. Vezzetti, "3d approaches and challenges in facial expression recognition algorithms—a literature review," *Applied Sciences*, vol. 9, no. 18, p. 3904, 2019.

[24] L. N. Smith and N. Topin, "Super-convergence: Very fast training of neural networks using large learning rates," in *Artificial intelligence and machine learning for multi-domain operations applications*, vol. 11006. SPIE, 2019, pp. 369–386.

[25] J. A. Vargas, W. Pedrycz, and E. M. Hemerly, "Improved learning algorithm for two-layer neural networks for identification of nonlinear systems," *Neurocomputing*, vol. 329, pp. 86–96, 2019.

[26] N. Acir and E. C. Meng¨u, "Lyapunov theory based adaptive learning algorithm for multilayer neural networks," *Neural Network World*, pp. 619–636, 2014.

[27] D. Kim, J. Kim *et al.*, "A reliable, self-adaptive face identification framework via lyapunov optimization," *ACM Symposium on Operating Systems Principles (SOSP) Workshop on AI Systems (AISys)*, 2021.

[28] V. Naveen, Y. Mareedu, N. S. Mandava, S. Kaveti, and G. Krishna Kishore, "Performance-based adaptive learning rate scheduler algorithm," in *Applications of Artificial Intelligence in Engineering: Proceedings of First Global Conference on Artificial Intelligence and Applications (GCAIA 2020)*. Springer, 2021, pp. 417–433.

[29] J. Ma and Y. Yuan, "Dimension reduction of image deep feature using pca," *Journal of Visual Communication and Image Representation*, vol. 63, p. 102578, 2019.

[30] M. J. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba, and J. Budynek, "The japanese female facial expression (jaffe) database," in *Proceedings of third international conference on automatic face and gesture recognition*, 1998, pp. 14–16.

[31] H. Ali, M. Hariharan, S. Yaacob, and A. H. Adom, "Facial emotion recognition using empirical mode decomposition," *Expert Systems with Applications*, vol. 42, no. 3, pp. 1261–1277, 2015.

[32] G. Yolcu, I. Oztel, S. Kazan, C. Oz, K. Palaniappan, T. E. Lever, and F. Bunyak, "Facial expression recognition for monitoring neurological disorders based on convolutional neural network," *Multimedia Tools and Applications*, vol. 78, pp. 31 581–31 603, 2019.

[33] A. Ilham, "Ahmad Ilham Simple Face Database," dec 2022. [Online]. Available: https://zenodo.org/record/7437652

[34] B. R. Ilyas, B. Mohammed, M. Khaled, A. T. Ahmed, and A. Ihsen, "Facial expression recognition based on dwt feature for deep cnn," in *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*. IEEE, 2019, pp. 344–348.

[35] M. R. Mohammadi, E. Fatemizadeh, and M. H. Mahoor, "Pca-based dictionary building for accurate facial expression recognition via sparse representation," *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 1082–1092, 2014.

[36] S. Varma, M. Shinde, and S. S. Chavan, "Analysis of pca and lda features for facial expression recognition using svm and hmm classifiers," in *Techno-Societal 2018: Proceedings of the 2nd International Conference on Advanced Technologies for Societal Applications-Volume 1*. Springer, 2020, pp. 109–119.

[37] K. H. Lim, K. P. Seng, L.-M. Ang, and S. W. Chin, "Lyapunov theory-based multilayered neural network," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 56, no. 4, pp. 305–309, 2009.

[38] O. Langner, R. Dotsch, G. Bijlstra, D. H. Wigboldus, S. T. Hawk, and A. Van Knippenberg, "Presentation and validation of the radboud faces database," *Cognition and emotion*, vol. 24, no. 8, pp. 1377–1388, 2010.

[39] N. Aifanti, C. Papachristou, and A. Delopoulos, "The mug facial expression database," in *11th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 10*. IEEE, 2010, pp. 1–4.

[40] L.-F. Chen and Y.-S. Yen, "Taiwanese facial expression image database," *Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan*, 2007.

[41] I. J. Goodfellow, D. Erhan, P. L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee *et al.*, "Challenges in representation learning: A report on three machine learning contests," in *Neural Information Processing: 20th International Conference, ICONIP 2013, Daegu, Korea, November 3-7, 2013. Proceedings, Part III 20*. Springer, 2013, pp. 117–124.

[42] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," in *2010 ieee computer society conference on computer vision and pattern recognition-workshops*. IEEE, 2010, pp. 94–101.

Ahmad Ilham is an Indonesian national. He completed a bachelor's degree in Information Systems in Computer Science at Universitas Al Asyariah Mandar (UNASMAN) and a master's degree in Intelligent Systems in Informatics Engineering at Universitas Dian Nuswantoro (UDINUS). Since 2019, he has worked as a faculty member in the Department of Informatics at Universitas Muhammadiyah Semarang. He is also a member of the International Association of Engineers (IAENG). Currently, he is pursuing a Doctoral (Ph.D.) degree in Information Systems at Diponegoro University. His research interests include data science, machine learning, and intelligent information processing.

Hadiyanto Hadiyanto is a Senior Lecturer in the Doctoral Program of Information Systems at Diponegoro University, Semarang, Indonesia. He obtained his Ph.D. from Wageningen University in the Netherlands. In May 2017, he was appointed as the head of a master's program in environmental science within the School of Postgraduate Studies and was also promoted to the rank of full professor in the field of bioprocess engineering. His research interests include process modeling, bioprocess, and process optimization.

Catur Edi Widodo is working as a Senior Lecturer in the Doctoral Program of Information Systems at Diponegoro University, Semarang, Indonesia. He obtained his Ph.D. from Universitas Gajah Mada, Indonesia. His research interests include computational physics, machine learning, and artificial intelligence.