
www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 17(2), 2026, 296-311. ISSN: 2007-1558. https://doi.org/10.61467/2007.1558.2026.v17i2.1120

A New Hybrid Algorithm Proposed for the Analysis and Classification of Signals Obtained from Head Movements

Luis Alberto Hernández Montiel, Edmundo Bonilla Huerta, Roberto Morales Caporal, Carlos Bueno Avendaño[1]
TecNM, Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, México.
{D23370018, edmundo.bh, roberto.mc, carlos.ba}@apizaco.tecnm.mx

Abstract. In this paper, we present a new algorithm for the analysis and classification of signals obtained from the macro and micro head movements of participants. First, videos were presented to elicit the disgust emotion in multiple participants, and their reactions were recorded. Hu invariant moments were then applied to transform the head movements into one-dimensional signals. The resulting signals were preprocessed using polynomial regression and normalization techniques to standardize size and numerical scale. Finally, a fuzzy inference system was implemented to classify the signals. The results demonstrate high performance in detecting disgust-related head movements, capturing both subtle and pronounced reactions.

Keywords: Head movement analysis, Macro and Micro Movements, Preprocessing, Hu Moments, Fuzzy Logic, Video-based signal processing

Article Info
Received April 16, 2025
Accepted Dec 11, 2025

1 Introduction

Human emotions are expressed in a variety of ways throughout the day. Studying these expressions is a complex task because people react differently to the same situation: some are highly expressive, while others may experience emotions without showing visible reactions.

Emotion recognition is not limited to facial expressions. Other body parts, such as the torso, hands, eyes, skin temperature, and even feet, also provide valuable information for analyzing emotional states. These additional sources can complement facial analysis to improve the understanding of human affective responses.

In this work, we propose the analysis and classification of signals obtained from the macro and micro movements of a person's head to recognize the disgust emotion. Participants' reactions were recorded while being exposed to videos specifically selected to elicit disgust. The head movements were tracked using Hu invariant moments and then mapped into signals within a frequency plane. Signals were preprocessed using polynomial regression and normalization techniques and finally classified using fuzzy inference systems.

This approach allows precise characterization of head movement patterns, capturing both subtle and pronounced reactions, and provides a robust framework for emotion recognition. The following sections describe the experimental procedures, signal processing pipeline, and classification methodology used to evaluate the effectiveness of the proposed algorithm.

2 State of the art

A person's expression analysis can help create systems that can automatically recognize an emotion and generate more specific therapies or devices to improve people's quality of life (Shashidhar, 2022). Nevertheless, doing a comprehensive examination of an individual's face to discern an emotion requires scrutinizing several anatomical regions, not limited to the face. This requires the implementation of systems that use all the body's responses as a means of obtaining information to identify different emotions.

Following this approach, different research has been proposed where they propose multimodal systems that analyze different parts of the body to find a specific emotion (Yan, 2024; Cabada & Estrada, 2019) using machine learning techniques (Zhang et al, 2020; Paul et al, 2024). Other studies (Crenn, 2020; Xu, 2022) analyze the body movements that a person makes to detect emotion using neural networks (Tan et al, 2020; Mehendale, 2020) and deep learning (Mozaffari, 2023; Wang & Deng, 2021). Another part of the body that is analyzed is the eyes, specifically the person blinking makes when stimulated to generate some type of emotion (Antonio & Alessandro, 2019; Mahanama, 2022) .

An alternative approach to examining the eyes is by analyzing the eye movements or saccadic movements shown by individuals during emotional experiences (Demiral, 2023; Matthiopoulou, 2024). Certain individuals employ fuzzy inference algorithms to examine speech or facial expressions. Other studies analyze straightforward expressions in images with the infrared plane (Zhang et al, 2020; Assiri & Hossain, 2023). Despite various proposals in literature, a concrete solution remains elusive. That is why new algorithms come out every day to find a stable solution to this problem.

3 Experimental Procedures

Finding a person's emotion from recorded responses requires precise algorithms capable of extracting relevant patterns from video recordings. To address this challenge, we developed a method that integrates Hu invariant moments, polynomial regression, normalization techniques, and fuzzy logic.

This approach generates, analyzes, and classifies signals derived from both macro and micro head movements while participants are exposed to stimuli designed to elicit the disgust emotion. Figure 1 illustrates the overall workflow of the proposed hybrid algorithm.

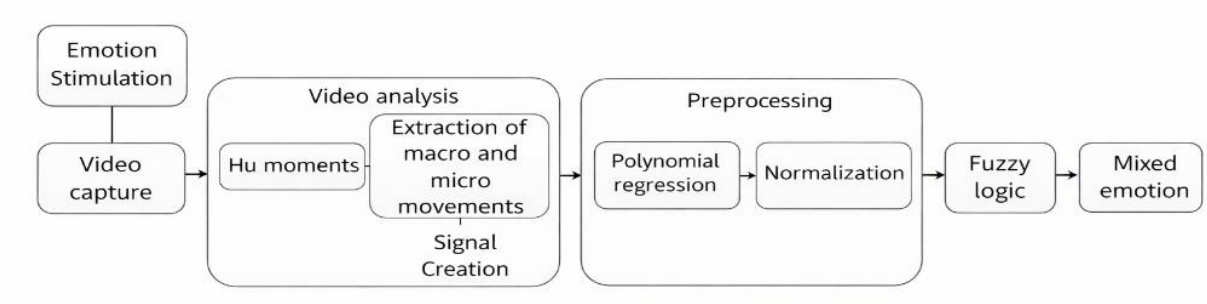

Fig.1. Overall procedure of the proposed hybrid algorithm for emotion recognition.

Figure 1 presents the main phases of the algorithm. The first step involves stimulating the participant and recording their reactions. Next, the videos are analyzed using Hu moments to extract macro and micro movements, which are then converted into one-dimensional signals. The signal undergoes preprocessing through polynomial regression and normalization techniques to standardize size and scale. Finally, the processed signals are classified using a fuzzy inference system. Each phase of the algorithm is explained in detail in the following subsections.

3.1 Stimulation and video capture

This phase involved stimulating a group of 24 participants (13 men and 11 women) from the Technological Institute of Apizaco, aged 18 to 27 years. Participants were selected based on a survey indicating high tolerance for descriptive images.

To elicit the disgust emotion, two video clips are projected, Planet Terror and The Human Centipede II, from the LATEMO-E database (Michelini et al., 2019), were presented. These clips elicited disgust in 79% and 83% of participants, respectively, and each had an average duration of 1–2 minutes. The videos were chosen according to two criteria: (1) the film's most frequently recommended, and (2) the consistency between the narrative content and the elicited emotion.

The videos were displayed on a Lenovo IdeaPad Slim 3 laptop with a 15.6-inch screen and its integrated speakers. Data collection took place over a two-week period, between 9:00 a.m. and 11:00 a.m., in a controlled environment measuring 2 × 4 meters to minimize ambient noise, excessive lighting, and drafts. The space was furnished with a table and a chair to ensure participant comfort.

A 1080p integrated webcam, positioned one meter from the participant, captured the upper part of the body. The OpenCV library (Hasan, 2021) was used to record the participants' head movements during stimulation. Figure 2 shows the sample collection workspace.

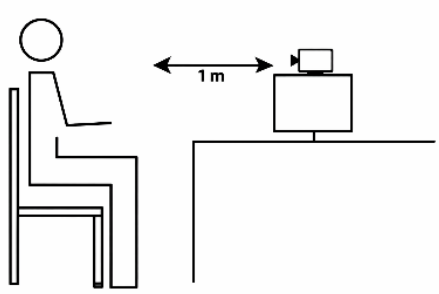

Fig. 2. Experimental setup used for recording participant videos.

3.2 Video analysis and signal creation

During this phase, the recorded videos are analyzed to generate signals representing participants' head movements. Head movements are particularly informative, as they are areas were people frequently express emotional reactions.

The first step involves applying the Mediapipe library (Lugaresi, Tang, & others, 2019), a machine learning–based framework for video processing. This library generates a dense mesh of points covering the face, enabling independent tracking of specific landmarks. Figure 3 shows the complete facial mesh.

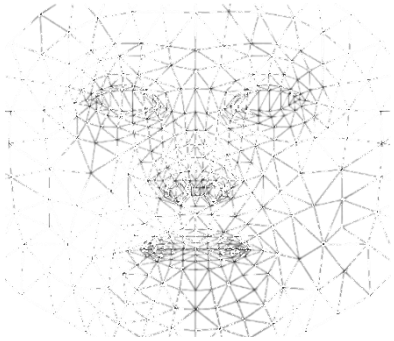

Fig. 3. Mediapipe *FaceMesh* generated around a participant's face.

Figure 3 shows the head movement analysis; five key landmarks were selected: forehead, beard, sides near the ears, and nose. Figure 4 shows the selected points.

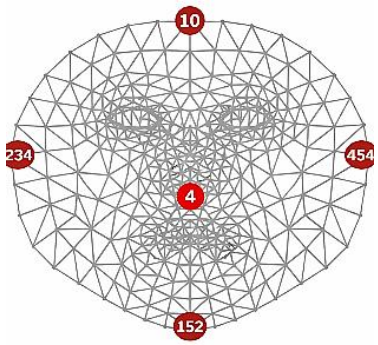

Fig. 4. Selected facial landmarks used for head movement tracking.

By selecting these five landmarks, the next step is to track the movements to transform them into a signal. The 2D coordinates of each landmark in a frame are computed as follows:

$$(\alpha)=LMs[i].x*w \quad , \quad (\beta)=LMs[i].y*h \tag{1}$$

Where *LMs[i]* is the *i-th* landmark generated by Mediapipe. *LMs[i].x* and *LMs[i].y* are the normalized horizontal and vertical coordinates in the range *[0,1]*. *w* and *h* are the width and height of the video frame (in pixels). α_i and β_i are the absolute *x* and *y* coordinates of the *i-th* landmark in the frame.

Once the coordinates are obtained, Hu invariant moments are computed to extract features from each landmark across frames. These moments provide rotation, scale, and translation-invariant characteristics of the landmark movements. The Hu moment of order $p+q$ is computed as (Sharma et al, 2022):

$$m_{pq}=\sum_{x}\sum_{y}(\alpha)^{p}(\beta)^{q}f(\alpha,\beta) \tag{2}$$

Where $m_{pq}$is the Hu moment of order $(p+q)$ for the landmark. $f(\alpha,\beta)$ represents the intensity or occupancy of the landmark at coordinates (α,β). p and q are the orders of the moment. In this work, first-order moments ($p + q = 1$) are used, since the generated signal is derived from a two-dimensional continuous function $f(\alpha, \beta)$ that represents the spatial distribution of head movement in the image plane.

To create a one-dimensional signal, the computed Hu moments for each landmark are concatenated across frames in temporal order. This process transforms the two-dimensional spatial movement information into a time-varying 1D signal, where:

- Macro movements correspond to high-amplitude variations.
- Micro movements correspond to subtle, low-amplitude fluctuations.

To identify the scenes that elicit the disgust emotion, time markers are introduced into the video recordings, allowing the detected signal perturbations to be temporally aligned with the corresponding emotional stimuli. This alignment ensures that each detected movement is associated with a specific moment in the video that corresponds to the elicitation of the disgust emotion.

By combining landmark tracking, Hu moments, and signal mapping, the 2D visual plane of the video is transformed into a frequency-based 1D signal, providing a reproducible and quantifiable representation of head movements for subsequent preprocessing and classification.

3.3 Signal preprocessing

The signal passes through two stages of preprocessing at this point. To normalize the signal to the same size, we employ a polynomial regression technique in the first phase. In the second stage, we implement three distinct normalization methods to transform the signal data to the same numerical scale: z-score, basic scaling, and min-max normalization.

The first phase is polynomial regression. In the previous step, Hu's invariant moments created different cues capturing the macro and micro head movements for each participant. Two factors contribute to the varying sizes and lengths of the signals generated in this step.

The first factor is that Hu's moments only capture the participant's movement, regardless of the video size. If there is no movement, the Hu moment captures nothing. The second factor arises from participants' varied responses to stimuli. Deviation from a normal distribution might cause overfitting in classification algorithms, leading to imprecise outcomes.

As a preprocessing step, we use polynomial regression to address this issue by matching all signals to a standard size. Polynomial regression, a specific instance of multiple linear regression, forecasts the reaction of the independent variable *X* to the dependent predictor variable *Y* by computing a polynomial using the following equation (Golondrino, 2020):

$$Y=\beta 0+\beta 1X+\beta 2X^{2}+\beta 3X^{3}+\ldots+BpX^{P}+\epsilon \tag{3}$$

where ϵ denotes the estimate error between the predicted and observed values. Equations 4 and 5 apply multiple linear regression to solve the polynomial that results from this equation:

$$X1=X, X2=X2,+\ldots+Xp=Xp \tag{4}$$

$$Y=\beta 0+\beta 1X+\beta 2X^2+\beta 3X^3+\beta pX^p+\epsilon \quad (5)$$

Equation 6 proposes the model from a matrix of *n* data samples to obtain multiple linear regression.

$$\begin{pmatrix} y1 \\ y2 \\ \vdots \\ yn \end{pmatrix}=\begin{pmatrix} 1x11\cdots x1p \\ 1x21\cdots x2p \\ \vdots\ \vdots\ \cdots\ \vdots \\ 1xn1\cdots xnp \end{pmatrix}\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}+\begin{pmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad (6)$$

With matrix notation, it is possible to express equation 6 more compactly as:

$$Y=X\beta+\in \quad (7)$$

The goal of matrix equation 7 is to determine the vector of coefficients β. It can be expressed as:

$$\beta=(XTX)-1XTY \quad (8)$$

The resulting vector β contains the coefficients β_0, β_1 ... β_p of the polynomial presented in equation 3. By applying polynomial regression, we can predict data that matches the shape of the source signal and supplement the signal to achieve uniform size.

To evaluate the performance of the polynomial regression, the Mean Squared Error (MSE) was employed as a quantitative metric to measure the approximation accuracy between the original head-movement signal and the polynomial-modeled signal. The MSE quantifies the average squared difference between the original signal and its estimated representation and is computed as follows (Hodson, 2022):

$$MSE=\frac{i}{N}\sum_{n=1}^{N}(x[n]-\hat{x}[n])^2 \quad (9)$$

where N represents the total number of signal samples, $x[n]$ denotes the value of the original signal at sample n, and $\hat{x}[n]$ corresponds to the estimated value obtained through polynomial regression. A lower *MSE* value indicates a closer approximation of the modeled signal to the original head-movement dynamics, whereas higher *MSE* values reflect larger deviations between both signals.

After transforming the signals to a standard size, the second step is to normalize the signal frequencies to the same numerical scale. Normalized signals prevent overadjustment during classification. Three standardization techniques are applied:

The min-max normalization method used the signal's highs and lows to transform the dataset values to a specific range, typically [0,1] (Henderi et al, 2021):

$$X'=\frac{(X-\min(X))}{(\max(X)-\min(X))} \quad (10)$$

Where *X* is the input signal, *min(X)* and *max(X)* are the minimum and maximum values in the signal. *X'* is the normalized signal. The z-score normalization calculates the signal mean around zero and a standard deviation of one. It is obtained with the following equation (Henderi et al, 2021):

$$X'=\frac{(X-\mu(X))}{\sigma(X)} \quad (11)$$

Where *X* is the input signal. $\mu(X)$ is the mean and $\sigma(X)$ is the standard deviation of the signal. *X'* are the normalized signals. Simple-scale normalization divides each value by the maximum value in the dataset (Izonin, 2022):

$$X'=\frac{X}{X_{max}} \quad (12)$$

Where X is the input signal. X_{max} is the maximum value within the signal base. X' are the normalized signals. The three methods have normalized all signals to the same numerical scale, preparing them for classification using fuzzy logic.

3.4 Fuzzy Logic Classification

This stage involves the implementation of a fuzzy inference algorithm to classify the signals obtained from the macro- and micro-head movements of the participants.

Fuzzy logic is used to handle fuzzy data collections, which allows the representation of inexact real-world patterns (Gupta, Gupta, Christopher, & Arunachalam, 2020). These sets consist of items that can take values within a real number range from 0 to 1. Each element of the set can be included with a certain degree of membership. Consequently, the closer the value is to one, the higher the probability that the element belongs to the set.

In this phase, a fuzzy inference system is implemented to classify the signals from the previous stages. The preprocessed signals first enter the fuzzification stage, where the system transforms crisp real values into fuzzy values. Then, an inference system evaluates these fuzzy values by applying fuzzy rules to the input data. Based on this analysis, the fuzzy data is subjected to a defuzzification procedure, transforming them back into crisp values. The output of the fuzzy system determines the signal relevance level associated with the unpleasant disgust emotion. Figure 5 illustrates this process.

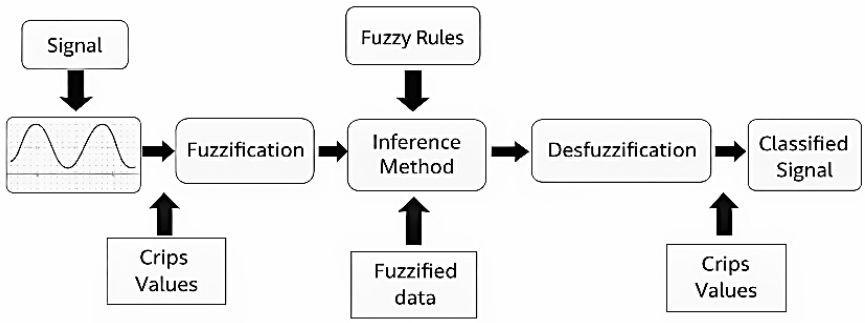

Fig. 5. Fuzzy inference system architecture used for signal classification.

Figure 5 shows the operation of the fuzzy system. The system works with a single input and a single output variable. Each variable uses multiple membership functions: five fuzzy sets for the input variable and two fuzzy sets for the output variable. Linguistic variables are applied to both input and output to fuzzification the data.

For the input variable, all signal disturbances are covered, corresponding to participants' head movements. For the output variable, the linguistic variables determine the relevance level of each data point in the signal.

To evaluate the signal data, the input and output membership functions are combined through a set of fuzzy IF-THEN rules, associating the input variable with the output variable. Figure 6 shows an example of input variables, a set of fuzzy rules, and output variables.

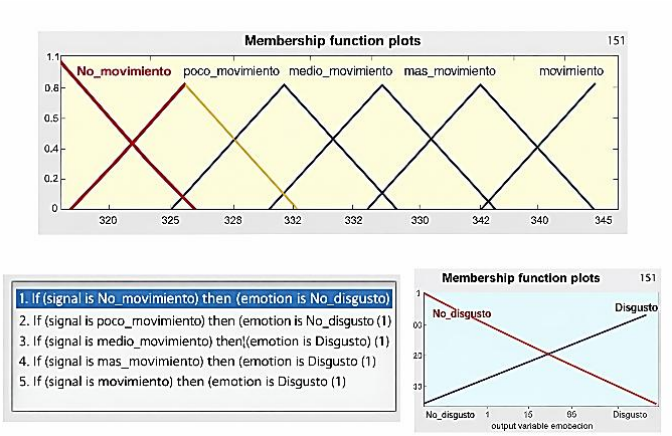

Fig. 6. Input and output variables with IF–THEN rules of the fuzzy system.

Figure 6 illustrates the fuzzy system's input variables (left), rule set (center), and output variables (right). The centroid method is used for defuzzification, which converts fuzzy values into crisp outputs. The inference method generates a fuzzy subset for each output variable, taking the maximum value of each label's membership function. Each subset determines the inferred certainty value, and the combined results are used to compute the center of gravity (COG) as follows (Golondrino, 2022):

$$COG=\frac{\int M(x)*xdx}{\int M(x)dx} \tag{13}$$

Where x is an element in the universe of discourse representing the possible values of the output variable. $M(x)$ is the membership function for element x, representing its degree of belonging to a fuzzy set within the output variable.

By combining these techniques, the algorithm analyzes and classifies various signals, allowing it to detect the disgust emotion. The algorithm's results are described below.

4 Experiments and Results

The experiments were conducted using the Mediapipe, OpenCV, and Scikit-learn libraries for signal analysis in Python 3.11.5. The fuzzy inference system was implemented in MATLAB R2017b. The algorithm was developed and tested on a Dell Inspiron 3000 laptop equipped with a Ryzen 7 processor and 24 GB of RAM. The best results are described below, highlighting the performance of the proposed hybrid algorithm in analyzing and classifying macro- and micro-head movements for the detection of disgust emotion.

4.1 Results

Two videos were recorded of each participant, resulting in a total of forty-eight videos (24 for each clip used to elicit disgust). These videos were transformed into a signal using the Mediapipe library to map landmarks around each participant's face, as shown in Figure 7.

Fig. 7. Facial landmarks selected using Mediapipe for head movement signal generation.

Figure 7 shows the five landmarks that cover the area of the head, beard, forehead, sides near the ears, and nose. After selecting these landmarks, the next step involves tracking the participant's movements during the emotion stimulation.

The coordinates of each landmark were obtained using Equation 1. Hu's invariant moments (Equation 2) transformed these coordinates into frequency values within a hyperplane. Disturbances in the signal correspond to the intensity of participant movements: large movements generate higher disturbances; minimal movements produce small disturbances. These disturbances generate a one-dimensional signal representing head movements.

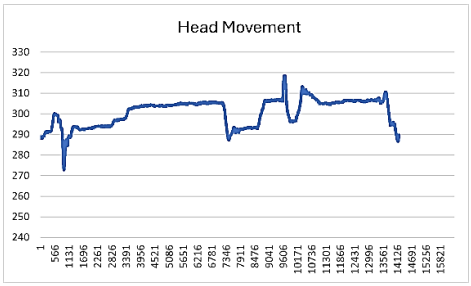

Fig. 8. Head movement signal obtained from Hu moments.

Figure 8 shows the signal obtained after applying Hu's invariant moments to the reference coordinates illustrated in Figure 7. The signal indicates that larger head movements generate more pronounced disturbances, whereas smaller movements result in lower-amplitude variations. Figure 9 presents the forty-eight signals generated from the participants' head movements, illustrating the variability in both the intensity and duration of the recorded responses.

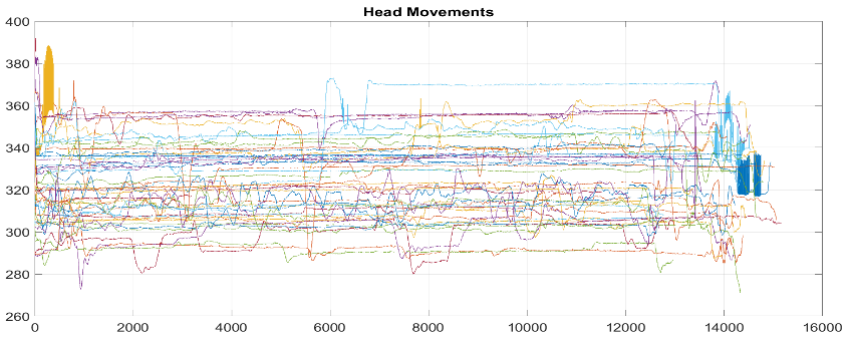

Fig. 9. Set of head movement signals for all participants.

Figure 9 shows the forty-eight signals generated from head movements. To create these signals, the proposed algorithm tracks only the participant, regardless of the video clip length. Consequently, when a participant exhibits multiple reactions, the algorithm captures these movements and transforms them into perturbations that form the resulting signal.

Due to variations in video duration and participant behavior, the resulting signals differ in length, leading to non-uniform signal sizes. Training a learning algorithm using signals of varying lengths may introduce overfitting and bias during the classification stage.

To address this issue, a two-phase preprocessing strategy is implemented. In the first phase, polynomial regression (Equation 3) is applied to model the overall trend of each signal while preserving its dynamic behavior. Polynomial regressions of different degrees were evaluated, and this process is illustrated in Figure 10.

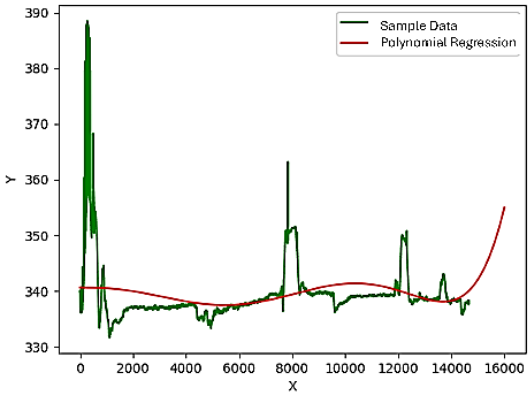

Fig. 10. Representative example of polynomial regression applied to a head movement signal.

Figure 10 presents a representative example of the polynomial regression applied during the preprocessing stage. This example illustrates the effect of applying a sixth-degree polynomial to a signal obtained from a head movement and is intended to demonstrate the general behavior of the regression process, without implying a criterion for selecting the polynomial degree.

During the evaluation phase, polynomial models of different degrees were analyzed, and third- and fourth-degree polynomials were identified as providing the best fit to the behavior of the generated signals. An example of the results obtained using these selected polynomials is shown in Figure 11.

To quantitatively support the selection of polynomial degrees, the approximation error between the original head-movement signals and the polynomial-modeled signals was evaluated using the Mean Squared Error (MSE). The analysis showed an average reduction of approximately 37% in the approximation error when using third- and fourth-degree polynomial models, indicating that these polynomial orders achieve a suitable balance between fitting precision and preservation of the signal's dynamic behavior.

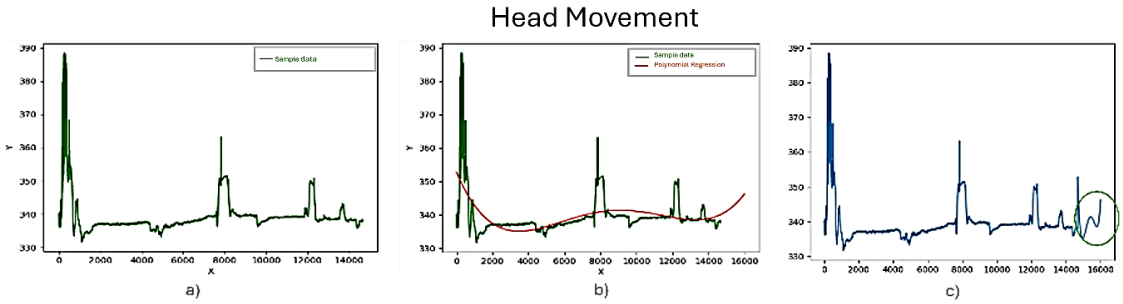

Fig. 11. Signal standardization using fourth-degree polynomial regression.

Figure 11 illustrates the process of data generation using polynomial regression with a fourth-degree polynomial. Subfigure (a) shows the original signal produced from Hu's moments. Subfigure (b) presents the curve obtained through fourth-degree polynomial regression, which closely follows the signal's behavior and enables the generation of data with characteristics like the original signal. Subfigure (c) shows the resulting extended signal, where the newly generated data are appended to the original signal (highlighted by the green circle), allowing all signals to be standardized to the same length.

With all signals standardized to the same length, the next preprocessing phase involves normalizing the signal amplitudes to a common numerical scale. This normalization is performed using three different methods, as described previously. Figure 12 presents the result of applying min–max normalization to the signal obtained from the fourth-degree polynomial regression.

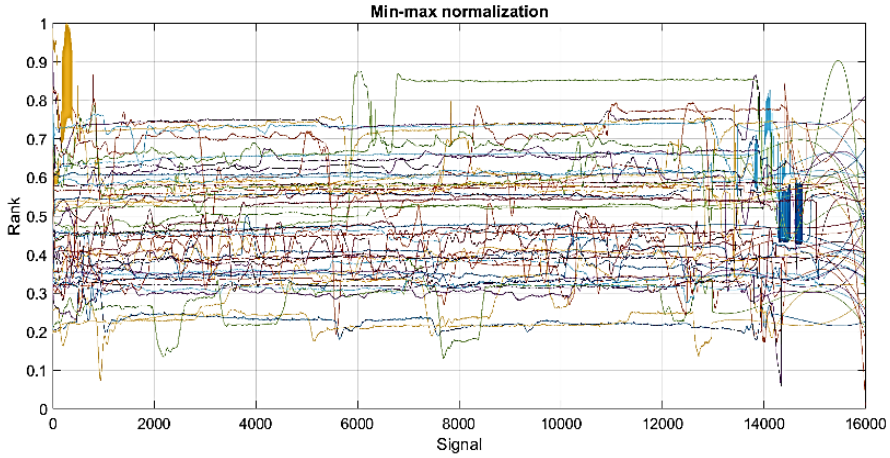

Fig. 12. Min–max normalized head movement signals.

Figure 12 presents the result of applying min–max normalization to the signals generated through fourth-degree polynomial regression. After the preprocessing stage, the next step consists of classifying the signals using a fuzzy inference system.

Figures 13 and 14 show the training results of the fuzzy inference system using four datasets and polynomial regressions of degree 3 and degree 4, respectively. Table 1 summarizes the performance of the fuzzy system, including the best achieved accuracy, as well as the mean and standard deviation obtained from four experimental runs for both polynomial degrees.

In Table 1, Column 1 indicates the method used to train the fuzzy system. Rows 2 and 3 report the best performance achieved for each polynomial degree (P3 for degree 3 and P4 for degree 4). Columns 4 and 5 present the mean and standard deviation of the performance obtained across the four tests for both polynomial regressions.

Figure 13 and figure 14 show the fuzzy system training results with four databases and degree 3 and 4 polynomials, respectively. Table 1 shows the best fuzzy system performance, as well as the mean and standard deviation for the four tests with the two polynomials as fallows: Column 1 shows the method with which the fuzzy system was trained; rows 2 and 3 show the best performance obtained by each proposed method for each polynomial (P3 grade 3 and P4 grade 4). Columns 4 and 5 show the mean and standard deviation of the performance obtained by the four tests and two polynomials.

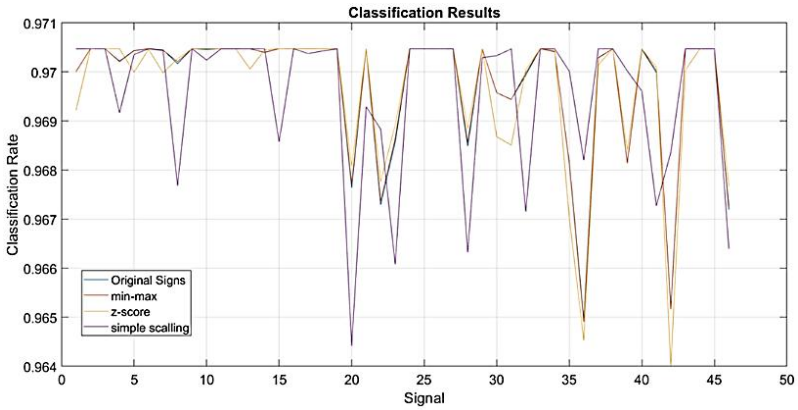

Fig. 13. Fuzzy system classification results using third-degree polynomial regression.

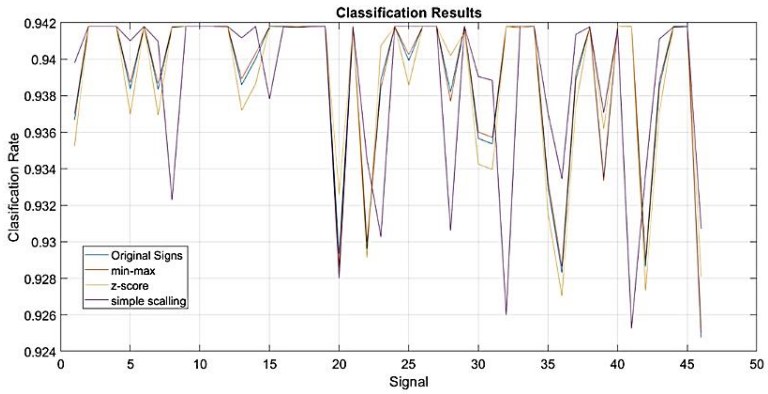

Fig. 14. Fuzzy system classification results using fourth-degree polynomial regression.

Table 1. Results of the fuzzy inference system.

Method	Best ranking (No. Signs)		Media (±*DS*)				
	P3	P4	P3		P4		
Original signs.	0.9418 (21)	0.9418(26)	0.7714 0.2325	±	0.9390 0.0045	±	
Min-max.	0.9400 (16)	0.9425(26)	0.7757 0.2320	±	0.9390 0.0045	±	
Z-score.	0.9399 (18)	0.9400(24)	0.7748 0.2327	±	0.9389 0.0044	±	
Simple scaling.	0.9435 (20)	0.9405(23)	0.8172 0.2103	±	0.9388 0.0048	±	

Figure 13 and Figure 14 show that the fuzzy inference system achieves classification accuracies above 90% across the four experimental databases when using signals generated by third- and fourth-degree polynomial regressions, respectively. As shown in Table 1, the best performance (94.35%) was achieved using simple scaling normalization with third-degree polynomial regression, while a comparable accuracy (94.25%) was obtained for fourth-degree polynomial signals using min–max normalization.

Overall, the fuzzy system demonstrates robust performance when trained with polynomial generated signals combined with different normalization strategies. The low standard deviation values reported in Table 1 indicate stable classification behavior across the evaluated preprocessing configurations, particularly for fourth-degree polynomial regression.

The consistently high performance observed can be explained by the fuzzy inference mechanism. In the evaluated databases, several signals reach maximum or near-maximum normalized values (e.g., values close to 1 under min–max normalization), which correspond to the highest membership degree of the fuzzy set representing high head movement. Consequently, during the defuzzification stage, similar output values are produced, leading to comparable classification results within the displeasure emotional category. An illustrative example of this behavior is shown in Figure 15.

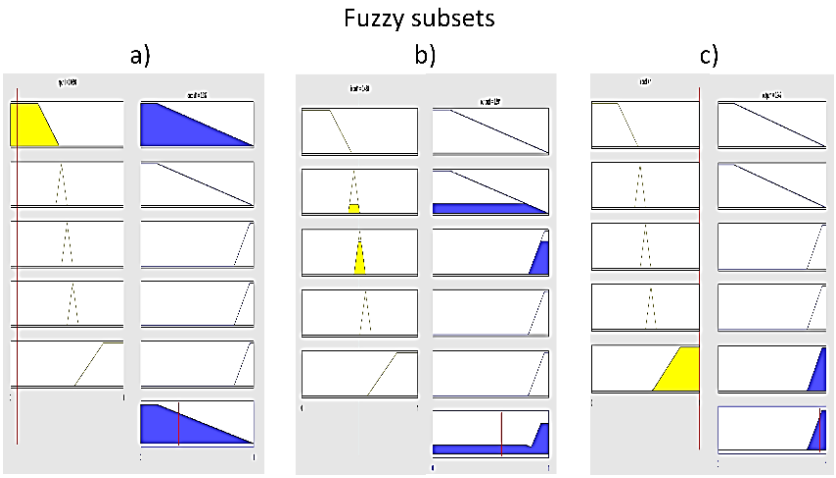

Fig. 15. Formation of fuzzy subsets during the defuzzification stage.

Figure 15 illustrates how different fuzzy subsets are activated during the evaluation of signal data. The figure is divided into three subfigures. In Figure 15(a), the activated fuzzy subsets correspond to signal values below 0.4. Figure 15(b) shows the subsets activated when signal values range between 0.41 and 0.6, while Figure 15(c) presents the subsets activated when signal values exceed 0.61. Despite the activation of different fuzzy subsets, the defuzzification process yields similar output values, leading to consistent classification results.

These results demonstrate that the proposed algorithm achieves a high classification rate by effectively handling variations in normalized signal values through fuzzy reasoning. To further validate the proposed approach, two validation methods are employed. The first method evaluates classification performance using confusion matrix–based metrics, including accuracy, precision, recall, and F1-score. Figure 16 presents the confusion matrix obtained from this evaluation.

		Real Values	
		Disgust	Not Disgust
Prediction values	Disgust	42	0
	Not Disgust	4	0

Fig. 16. Confusion matrix for the fuzzy system classification results.

Figure 16 presents a representative confusion matrix obtained for the fuzzy system using min–max normalization combined with third-degree polynomial regression (RP3) for the disgust emotional category. The matrix shows that most of the signals are correctly classified, with a small number of samples incorrectly classified as not disgusting. This behavior reflects the variability of head-movement signals and helps avoid overfitting effects.

The absence of false positive classifications indicates that the proposed fuzzy system does not incorrectly assign the disgust label to signals from other emotional states, while the few false negatives correspond to borderline cases characterized by weaker or less expressive motion patterns. These results are consistent with the performance trends observed in previous experiments and support the generalization capability of the proposed approach.

After obtaining the confusion matrix, the next step is to analyze the evaluation metrics of the proposed model. Table 2 shows these metrics for all evaluated configurations. Column 1 shows the proposed method (FS for fuzzy system), while Column 2 lists the applied normalization techniques, namely Original Signals (OS), Min–Max (M–M), Z-score (Z–S), and Simple Clipping (S–C). Columns 3 to 6 report the accuracy, precision, recall, and F1-score, respectively. Each metric is further divided into sub-columns labeled RP3 and RP4, corresponding to the polynomial regression degrees used during the preprocessing stage.

Table 2. Comparison of evaluation metrics.

Method.		Evaluation Metrics Acc.		Prec.		Sens.		F1.	
		RP3	RP4	RP3	RP4	RP3	RP4	RP3	RP4
F. S.	O. S.	0.9722	1.0	1.0	1.0	0.9725	1.0	0.9725	1.0
	M-M	0.9130	1.0	0.9999	1.0	0.9130	0.9999	0.9545	1.0
	Z-S	0.9999	1.0	1.0	1.0	0.9876	1.0	0.9876	1.0
	S. C.	1.0	0.999	0.999	0.999	0.9999	1.0	0.9545	1.0

Table 2 shows the evaluation metrics for the fuzzy system using third- (RP3) and fourth degree (RP4) polynomial regressions across different normalization methods. Both models exhibit high performance, confirming the robustness of the proposed approach. While RP4 consistently achieves near perfect or perfect scores, RP3 shows slightly lower values, reflecting a more conservative generalization behavior.

Accuracy for RP3 ranges from 0.913 to 1.0, compared to 0.999–1.0 for RP4. Precision remains at or near 1.0 for all models, indicating the fuzzy system's strong ability to correctly identify positive samples. Sensitivity for RP3 varies from 0.913 to 0.988, while RP4 reaches 0.999–1.0, demonstrating the higher responsiveness of the fourth-degree polynomial. F1-scores follow a similar trend, with RP3 ranging from 0.955 to 0.988 and RP4 achieving 1.0 across most normalization methods.

These results indicate that the fourth-degree polynomial consistently outperforms the third-degree polynomial, providing higher overall accuracy and a better balance between precision and sensitivity across all normalization strategies.

Table 3 summarizes the averaged evaluation metrics across all normalization methods, providing a concise overview of the system's stability and effectiveness regardless of the normalization strategy.

Table 3. Average values obtained for each metric.

Method	Evaluation Metrics Accuracy		Precision		Sensitivity		F1 score	
	RP3	RP4	RP3	RP4	RP3	RP4	RP3	RP4
Fuzzy System	0.971275	0.999975	0.99995	0.99995	0.96825	0.99995	0.967275	0.999975

The fuzzy system combined with the fourth-degree polynomial regression consistently outperforms the third-degree polynomial, as reflected in the averaged metrics across all normalization methods (Table 3). RP4 achieves near-perfect performance in all evaluation metrics, while RP3 shows slightly lower values, indicating a more conservative generalization behavior. These results confirm the robustness and stability of the proposed approach across different preprocessing configurations.

In the second comparison study, we evaluated the algorithm against results reported in the literature. Despite a thorough review, we did not find studies using head-movement-derived signals. Existing studies primarily employ fuzzy logic to classify disgust emotion based on alternative participant data sources. Table 4 presents this comparison, where Column 1 lists the method, Column 2 provides the bibliographic reference, and Column 3 reports the best performance achieved by each study.

Table 4. Comparison with related studies.

Method	Author	Displeasure %
Speech emotion recognition using a fuzzy approach	(Ton-Tha, 2019)	62%
Fuzzy system for facial emotion recognition	(Gupta et al, 2020)	95%
Modeling adaptive E-learning environment using facial expressions and fuzzy logic	(Megahed, 2020)	63%
A fuzzy logic approach to reliable real-time recognition of facial emotions-multimedia tools and applications	(Bahreini, 2019)	68.2%

Emotion recognition based on interval type-2 fuzzy logic from facial expression.	(Dirik, 2020)	85%
Fuzzy Inference System based Analysis. of Facial Expressions for Emotion Recognition	(Das, 2021)	88%
Facial expression recognition using fuzzified Pseudo Zernike Moments and structural features	(Ahmady, 2022)	91.30%
A fuzzy ensemble-based deep learning model for EEG-based emotion recognition	(Dhara, 2024)	98%
Proposed method	Fu. Sys (RP3) Fu. Sys (RP4)	96.7275% 99.9975%

Table 4 presents a comparison of the proposed algorithm with previous studies in literature. Most of these works employ facial expressions or speech-based features to classify the disgust emotion using fuzzy systems.

When comparing performance, the proposed method clearly outperforms the previous approaches. For instance, Gupta et al. (2020) report a 95% classification rate, Ahmady (2022) achieves 91.30%, and Dhara (2024) reaches 98%. In contrast, the proposed fuzzy system attains a 96.73% accuracy using the third-degree polynomial (RP3) and 99.9975% using the fourth-degree polynomial (RP4), averaged across all normalization methods and F1 scores.

These results demonstrate that the proposed approach surpasses the classification rates of the compared studies, highlighting the effectiveness of using head-movement-derived signals in combination with fuzzy inference and polynomial regression for disgust emotion recognition.

4.2 Discussion and Limitations

The results demonstrate that the combination of Hu moments, polynomial regression, signal normalization, and fuzzy logic classification enables highly accurate detection of the disgust emotion based on macro- and micro-head movement signals.

The use of third-degree (RP3) and fourth degree (RP4) polynomial regressions yield strong and consistent performance, achieving near-perfect classification metrics across all evaluated normalization strategies, as reported in Tables 2 and 3. In particular, RP4 exhibits greater stability and higher average performance, indicating its effectiveness in capturing subtle signal variations. To support the reliability of the proposed approach, confusion matrices were analyzed for representative test configurations (Figure 16). These matrices reflect a conservative classification behavior, where misclassifications are limited and primarily associated with borderline signal patterns rather than systematic errors.

The effectiveness of the proposed method is further confirmed through comparison with existing studies (Table 4). While most previous approaches rely on facial expressions or speech-based features, the proposed head-movement-based method achieves superior classification accuracy—up to 99.9975% with RP4—highlighting the added value of motion-derived signals combined with fuzzy inference. Normalization strategies performance an important role in system performance, particularly for RP3, where slight variations across methods are observed. This indicates that normalization influences the sensitivity of the fuzzy system to signal amplitude and frequency characteristics.

Despite these promising results, several limitations should be considered:

1. Dataset size: The evaluation was conducted using data from 24 participants. Although consistent performance was observed across multiple tests and normalization methods, the current analysis is limited to the available dataset. Additional experiments with a larger number of participants would allow further assessment of performance stability under broader experimental conditions.
2. Stimulus dependency: The algorithm analyzes audiovisual stimuli specifically designed to elicit disgust. Variations in stimulus intensity or type may influence the magnitude and duration of participants' head-movement responses.
3. Environmental conditions: Factors such as ambient lighting, camera angle, and participant posture may affect landmark detection accuracy using Mediapipe, potentially impacting signal quality and subsequent feature extraction.

4. Risk of overfitting: Although multiple polynomial degrees and normalization strategies were evaluated, the extremely high performance achieved by RP4 suggests sensitivity to the extracted feature representations. While confusion matrix analysis and metric consistency indicate controlled behavior, future work should explore additional validation strategies, such as cross-dataset evaluation, to further assess robustness.
5. Signal variability: Head-movement responses vary across participants; some exhibit subtle motion patterns, while others produce more pronounced or prolonged movements. This variability can introduce differences in signal disturbances at similar temporal points. The choice of polynomial degree helps mitigate this effect, with RP4 providing improved sensitivity to subtle variations.

Overall, the proposed method demonstrates strong potential for emotion recognition based on head-movement signals. By integrating polynomial regression, signal normalization, and fuzzy inference, the approach provides a reliable alternative to facial- or voice-based systems while maintaining stable performance and mitigating overfitting risks across preprocessing strategies.

5 Conclusions

This work demonstrates that head-movement signals can be effectively used to recognize the disgust emotion through a hybrid framework that integrates Hu moments, polynomial regression, signal normalization, and fuzzy logic classification. By transforming macro- and micro-head movements into numerical signals, the proposed approach provides a robust alternative to traditional emotion recognition methods based on facial expressions or speech.

The experimental results confirm that the proposed preprocessing strategy—combining polynomial regression for signal standardization and normalization for numerical stability—significantly enhances classification performance. In particular, the fuzzy system trained with fourth-degree polynomial regression consistently achieves superior results compared to the third-degree polynomial, yielding near-perfect accuracy, precision, sensitivity, and F1 scores across multiple normalization methods. These findings indicate that higher-degree regression more effectively captures subtle variations in head-movement signals while maintaining stable classification behavior.

Performance evaluation using confusion matrices and standard classification metrics further supports the reliability of the proposed method. The results show limited misclassification and consistent behavior across different preprocessing configurations, highlighting the robustness of the fuzzy inference system when applied to motion-derived emotional signals. Compared to existing studies reported in the literature, which evaluate performance in terms of recognition rate or overall classification accuracy, the proposed method achieves a higher percentage of correctly classified instances for the disgust emotion, demonstrating the added value of head-movement-based signal analysis.

In future work we will explore alternative signal standardization techniques and additional classification models, such as neural networks, k-nearest neighbors, and support vector machines, to further assess the general applicability of the proposed framework. Moreover, the inclusion of additional emotional categories and the creation of a larger multimodal database incorporating signals from multiple body parts will be investigated, with the objective of achieving reliable emotion recognition using a minimal and efficient set of features

Acknowledgements
The research presented in this paper was partially supported by Secihti.

References

Ahmady, M. M. (2022). Facial expression recognition using fuzzified pseudo-Zernike moments and structural features. *Fuzzy Sets and Systems, 438*, 155–172. https://doi.org/10.1016/j.fss.2022.03.003

Antonio, M., & Alessandro, A. (2019). Spontaneous blink rate as an index of attention and emotion during film clips viewing. *Physiology & Behavior, 204*, 256–263. https://doi.org/10.1016/j.physbeh.2019.02.031

Assiri, B., & Hossain, M. A. (2023). Face emotion recognition based on infrared thermal imagery by applying machine learning and parallelism. *Mathematical Biosciences and Engineering, 20*(1), 913–929. https://doi.org/10.3934/mbe.2023042

Bahreini, K., van der Vegt, W., & Westera, W. (2019). A fuzzy logic approach to reliable real-time recognition of facial emotions. *Multimedia Tools and Applications, 78*, 18943–18966. https://doi.org/10.1007/s11042-018-6734-2

Cabada, R. Z., & Estrada, M. L. (2019). Multimodal recognition of learning-oriented emotions. *Research in Computing Science, 148*, 153–165.

Crenn, A. M. (2020). Generic body expression recognition based on synthesis of realistic neutral motion. *IEEE Access, 8*, 207758–207767. https://doi.org/10.1109/ACCESS.2020.3037683

Das, A., & Mohanty, S. (2021). Fuzzy inference system-based analysis of facial expressions for emotion recognition. *Turkish Journal of Computer and Mathematics Education, 12*(2), 274–280.

Demiral, Ş. B., Liu, C. K., Benveniste, H., Tomasi, D., & Volkow, N. D. (2023). Activation of brain arousal networks coincident with eye blinks during resting state. *Cerebral Cortex, 33*(11), 6792–6802. https://doi.org/10.1093/cercor/bhad116

Dhara, T. S. (2024). A fuzzy ensemble-based deep learning model for EEG-based emotion recognition. *Cognitive Computation, 16*, 1364–1378. https://doi.org/10.1007/s12559-023-10189-7

Dirik, M. C. (2020). Emotion recognition based on interval type-2 fuzzy logic from facial expression. *Journal of Soft Computing and Artificial Intelligence, 1*(2), 1–17.

Golondrino, G. E. (2020). Application of polynomial regression for the characterization of the COVID-19 curve using machine learning techniques. *Investigación e Innovación en Ingenierías, 8*(2), 87–105. https://doi.org/10.17081/invinno.8.2.3953

Golondrino, G. E. (2022). Fuzzy logic-based system for the estimation of the usability level in user tests. *International Journal of Computers Communications & Control, 17*(4), 48–54. https://doi.org/10.15837/ijccc.2022.4.4675

Gupta, K., Gupta, M., Christopher, J., & Arunachalam, V. (2020). Fuzzy system for facial emotion recognition. In *Proceedings of the International Conference on Intelligent Systems Design and Applications* (pp. 536–552). Springer. https://doi.org/10.1007/978-3-030-16660-1_53

Hasan, R. T. (2021). Face detection and recognition using OpenCV. *Journal of Soft Computing and Data Mining, 3*(1), 86–97.

Henderi, H., Wahyuningsih, T., & Rahwanto, E. (2021). Comparison of min-max normalization and z-score normalization in the k-nearest neighbor algorithm to test the accuracy of types of breast cancer. *International Journal of Informatics and Information Systems, 4*(1), 13–20. https://doi.org/10.47738/ijiis.v4i1.73

Hodson, T. O. (2022). Root mean square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geoscientific Model Development, 15*(14), 5481–5491. https://doi.org/10.5194/gmd-15-5481-2022

Izonin, I. I. (2022). Towards data normalization task for the efficient mining of medical data. In *2022 12th International Conference on Advanced Computer Information Technologies (ACIT)* (pp. 480–484). IEEE. https://doi.org/10.1109/ACIT54803.2022.9913177

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., … Grundmann, M. (2019). *MediaPipe: A framework for building perception pipelines* (arXiv:1906.08172). arXiv. https://arxiv.org/abs/1906.08172

Mahanama, B., Jayawardana, Y., Rengarajan, S., Jayawardena, G., Chukoskie, L., Snider, J., & Jayarathna, S. (2022). Eye movement and pupil measures: A review. *Frontiers in Computer Science, 3*, 733531. https://doi.org/10.3389/fcomp.2021.733531

Matthiopoulou, O. G. (2024). Towards the automated analysis of expressive gesture qualities in full-body movement: The perceived origin of movement. *Human-Centric Computing and Information Sciences, 14*, 106–118. https://doi.org/10.22967/hcis.2024.14.1.106

Megahed, M. (2020). Modeling adaptive e-learning environment using facial expressions and fuzzy logic. *Expert Systems with Applications, 159*, 113221. https://doi.org/10.1016/j.eswa.2020.113221

Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). *SN Applied Sciences, 2*, 446. https://doi.org/10.1007/s42452-020-2233-1

Michelini, Y., Acuña, I., Guzmán, J. I., & Godoy, J. C. (2019). LATEMO-E: A film database to elicit discrete emotions and evaluate emotional dimensions in Latin-Americans. *Trends in Psychology, 27*, 473–490. https://doi.org/10.9788/TP2019.2-07

Mozaffari, L. (2023). Facial expression recognition using deep neural networks. In *Proceedings of the International Conference on Applied Artificial Intelligence* (pp. 1–9).

Paul, B., Bera, S., Dey, T., & Phadikar, S. (2024). Machine learning approach of speech emotion recognition using feature fusion technique. *Multimedia Tools and Applications, 83*, 8663–8688. https://doi.org/10.1007/s11042-023-16587-6

Sharma, S., Choudhary, S., & Sharma, A. (2022). Image watermarking in frequency domain using Hu's invariant moments and firefly algorithm. *Multimedia Tools and Applications, 81*, 1–15.

Shashidhar, R. S. (2022). Mouse cursor control using facial movements—An HCI application. In *International Conference on Sustainable Computing and Data Communication Systems* (pp. 367–371). IEEE. https://doi.org/10.1109/ICSCDS53736.2022.9760908

Tan, C., Ceballos, G., Kasabov, N., & Puthanmadam Subramaniyam, N. (2020). FusionSense: Emotion classification using feature fusion of multimodal data and deep learning in a brain-inspired spiking neural network. *Sensors, 20*(18), 5328. https://doi.org/10.3390/s20185328

Ton-That, A. H., & Cao, N. T. (2019). Speech emotion recognition using a fuzzy approach. *Journal of Intelligent & Fuzzy Systems, 36*(2), 1587–1597. https://doi.org/10.3233/JIFS-169908

Wang, M., & Deng, W. (2021). Deep face recognition: A survey. *Neurocomputing, 429*, 215–244. https://doi.org/10.1016/j.neucom.2020.10.081

Xu, S. F. (2022). Emotion recognition from gait analysis: Current research and future directions. *IEEE Transactions on Computational Social Systems, 9*(2), 363–377. https://doi.org/10.1109/TCSS.2021.3128807

Yan, J. L. (2024). Multimodal emotion recognition based on facial expressions, speech, and body gestures. *Electronics, 13*(7), 1334. https://doi.org/10.3390/electronics13071334

Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion, 59*, 103–126. https://doi.org/10.1016/j.inffus.2020.01.011

Zhang, Z., Lai, C., Liu, H., & Li, Y.-F. (2020). Infrared facial expression recognition via Gaussian-based label distribution learning in dark illumination environment for human emotion detection. *Neurocomputing, 409*, 341–350. https://doi.org/10.1016/j.neucom.2020.05.092