How to cite this article:

Liliana, D. Y., Basaruddin, T., Widyanto, M. R., & Oriza, I. I. D. (2020). High-Level fuzzy linguistic features of facial component for human emotion recognition. *Journal of Information and Communication Technology, 19(1)*, 103-129. https://doi.org/10.32890/jict2020.19.1.5

# HIGH-LEVEL FUZZY LINGUISTIC FEATURES OF FACIAL COMPONENTS IN HUMAN EMOTION RECOGNITION

**[1]Dewi Yanti Liliana, [1]Tarzan Basaruddin, [1]Muhammad Rahmat Widyanto & [2]Imelda Ika Dian Oriza**

*[1]Faculty of Computer Science, Universitas Indonesia, Indonesia*
*[2]Faculty of Psychology, Universitas Indonesia, Indonesia*

*dewiyanti.liliana@tik.pnj.ac.id; chan@cs.ui.ac.id; widyanto@cs.ui.ac.id; imelda.ika@ui.ac.id*

## ABSTRACT

Emotion is an important element in an interaction since it conveys human perception and response of an event. Unlike verbal words that can be manipulated, emotion is brief, spontaneous and provides more honest information. There are several classes of basic primary human emotions that differ from one another. These classes are happy, sad, fearful, surprised, disgusted, and angry. Meanwhile, a psychologist has developed a set of rules to recognize emotions based on facial expressions. This research aims to develop an artificial intelligent model based on psychological knowledge to recognize emotions by analyzing facial expressions. Moreover, the proposed model has defined high-level fuzzy linguistic features of facial components which distinguish it from existing methods that commonly use low-level image features (e.g. color, intensity, histogram, texture). High-level linguistic features (e.g. opened eyes, wrinkled nose) are better at representing human minds than low-level features which are only understood by machines. The model functions by detecting facial points first to locate important facial components; then extracting geometric facial components

**Received:** 4/12/2018 **Revised:** 22/5/2019 **Accepted:** 18/09/2019 **Published:** 23/12/2019

features which are then applied to a fuzzy facial components inference system resulting in high-level linguistic facial features. In the last step, the high-level linguistic features are applied to a fuzzy emotion inference system which classifies the input image into its respective emotion class based on psychological rules. Experiments conducted using facial expression dataset gave a high accuracy rate of 98.26% for fuzzy facial components linguistic identification. The proposed model also outperformed other classifiers (Fuzzy C-Means, Fuzzy Inference System, and Support Vector Machine). This intelligent model can contribute in various fields, including psychology, health, and education, especially in helping people with emotional disorders (e.g. Alexithymia, Asperger syndrome, and Autism) to recognize emotions.

**Keywords:** Basic emotion, emotion recognition, facial expression, facial components, fuzzy system, high-level linguistic features.

## INTRODUCTION

Automatic emotion recognition is an active research focus in the affective computing field. The objective is to design an intelligent agent to recognize human emotions. This topic intersects computer science with psychology disciplines. Many researches have been proposed to address automatic emotion recognition problems using different approaches based on artificial intelligence and machine learning (Kumari, Rajesh, & Pooja, 2015).

In communication, emotion is a means to convey messages through nonverbal signals such as facial expression, prosody, gesture and bodily expression (Pantic et al., 2011). Among these signals, facial expression becomes the central feature of emotion (Ekman, 1992). The problem in emotion recognition is the variability in facial expressions. This is a challenging task since there are various subjective ways for humans to express emotions. A psychologist, Paul Ekman has defined basic emotion as a separate discrete emotion that differs from one another (Ekman, 1992). Basic emotion is a universal and primary emotion. The six categories of basic emotions are happy, sad, angry, disgusted, fearful, and surprised.

Facial expression is the movement of facial muscles in response to a given stimulus (Ekman, Friesen, & Hager, 2002). According to Ekman (2003), there are more than 10000 combinations of facial muscle movements (Ekman, 2003). For example, when surprised, people tend to raise their outer eyebrows

with their mouths wide open or when angry, people tend to tighten their lips with their inner eyebrows coming close together. The task of emotion recognition is selecting the most prominent emotions from facial components combination. Existing methods such as Support Vector Machine, Neural Network and Deep Learning function by training large numbers of images and learning from extracted patterns to classify images into corresponding classes (Chen et al., 2012; Tang, Guo, Shen, & Du, 2018). However, this is not representative of what psychologists do in recognizing emotions. Psychologists have defined knowledge about human emotions and this definition is derived from years of studies and observations on various facial expressions across cultures (Ekman & Friesen, 1975). Ekman and Friesen (1975) have developed a set of rules to interpret emotions from facial expressions. This knowledge has been utilized by people who work on emotion recognition such as psychotherapists, interrogators, nurses, lawyers, managers, salesmen, and actors.

In this paper, we propose a model of emotion recognition based on psychological knowledge. Our proposed model is different from existing models which commonly use low-level feature representation (e.g. color intensity, histogram, or texture values). Instead, we define high-level fuzzy linguistic features of facial components (e.g. opened mouth, closed eyes, wrinkled nose) as image features. This novel feature is understood by the human mind, since it uses human natural language representation. Moreover, the nature of fuzzy system in dealing with vague problems, as well as its ability to work with ambiguous data is suitable to be applied in emotion recognition tasks which has varying degrees and different intensities of facial expressions as input.

The proposed model works as follows. First, we elicit psychological knowledge of emotions from facial component analysis and denote it as a set of emotion rules. For each input image, the model detects facial points and locates the coordinates of facial components (eyebrows, eyes, nose, and lips). Next, the model extracts geometric face feature parameters and applies these parameters in the fuzzy facial component inference system, resulting in high-level linguistic features of facial components. For the last step, we feed the respected fuzzy linguistic parameters into a fuzzy emotion recognition system which classifies the facial image into basic emotion classes based on psychological rules.

Our contribution in the proposed model is adopting the natural way to recognize emotions automatically using psychological knowledge which focuses on facial component rules. Moreover, we enhance the feature extraction process by using high-level linguistic features of facial components which is simple to compute yet powerful as facial expression descriptors.

Our model is useful in various fields, such as psychology, health, education, robotics, and entertainment. The next section explains related studies on emotion recognition and analysis. We will discuss our proposed model in the methodology section and elaborate our findings from experiments in the results and discussion section. In the last section we provide our conclusion and suggest future work.

## RELATED WORK

Research on facial component analysis is based on the assumption that by decomposing the human face into facial components it enables the exploration of local facial features, instead of processing the whole face directly or using global features (Li, Lian, & Lu, 2012). There are three types of feature processing techniques which are based on template model, mathematical model, and deep learning methods. The first technique uses template matching to increase facial points detection accuracy via Active Appearance Model (AAM) framework (Pratiwi, Widyanto, Basaruddin, & Liliana, 2017; Wang, Li, & Wang, 2014). The second technique is applied by calculating feature values using a mathematical model (Loconsole, Miranda, Augusto, Frisoli, & Orvalho, 2014). A recent trend is using deep learning methods for feature processing (Das & Chakrabarty, 2016; Pitaloka, Wulandari, Basaruddin, & Liliana, 2017), but it requires a large amount of data for the training process; and since deep learning is a black-box method, we cannot observe the process inside it.

A large number of approaches on facial expression analysis are based on appearance and geometric features. Appearance feature is related to facial texture variations (e.g. furrows, wrinkles). This feature can be gained using filtering techniques such as Gabor filter which has high accuracy performance under different illuminations, poses, and expressions (Sudhakar & Nithyanandam, 2017); Principle Component Analysis (PCA) to reduce feature dimensionality (Chakrabarti & Dutta, 2013); and Local Binary Pattern (LBP) to extract various types of textures on face images (Lekdioui, Messoussi, Ruichek, Chaabi, & Touahni, 2017). Another appearance-based feature extracts skin color for facial component classification which works fast on a pixel level (Mayer, Wimmer, & Radig, 2010). Generally, obtaining appearance feature descriptors require more computational cost, in terms of memory and time.

Geometric features describe the geometric shape of facial components. It also represents facial component movements. Geometric features are

known for its robustness in handling expression variability (Sadeghi, Raie, & Mohammadi, 2013). Geometric features from facial components are extracted from other meaningful descriptors by using a mathematical model (Loconsole et al., 2014) or through deep learning methods (Tang, Guo, Shen, & Du, 2018). Loconsole et al. (2014) used geometric features extraction based on pixel coordinates and transformed it into a linear feature representation in a simple and fast way. Their work has inspired this proposed model; hence we explored more aspects on geometric features to enhance facial components analysis performance. Geometric based features extraction employed by other researchers include Nicolai and Choi (2015) who used threshold values for facial components and Chaturvedi and Tripathi (2014) used the Euclidean distance between facial components as a geometric features descriptor. Other approaches combined geometric and appearance features and gained advantage from both sides but suffered from memory and time computation as a consequence (Chen et al., 2012; Sadeghi et al., 2013).

In geometric-based features extraction, facial landmarks become a crucial starting point. Therefore, facial points detection must be performed with high accuracy results. Active Appearance Model (AAM) is a robust framework for facial points detection which works on face images with varying positions and scales (Cootes, Edwards, & Taylor, 2001). AAM is a template-based model which combines shape and texture features to locate facial fiducial points (Wang et al., 2014). We applied AAM as an intermediate system to detect facial points (Liliana, Widyanto, & Basaruddin, 2017). We also utilized AAM framework as an initial step for our fuzzy facial components analysis.

In contrast to existing fuzzy approaches for emotion recognition, our proposed model uses a high-level fuzzy linguistic feature of facial components as input for determining emotion class as output. Hence, we do not force mapping of six basic emotions in a single dimensional output unlike other studies (Chaturvedi & Tripathi, 2014; Halder, Bhattacharjee, Nasipuri, Basu, & Kundu, 2010; Sujono & Gunawan, 2015). Instead we used each emotion class as a separate output. In addition to our proposed model, we employed simple geometric features extraction methods which worked on a pixel basis to enhance the features extraction process. These geometric features served as input for the fuzzy facial component analysis. We defined a separate fuzzy rule-based model for each facial component which takes geometric features values as input and performs high-level features extraction to produce high-level linguistic features of facial components as output.

## FUZZY FACIAL COMPONENTS ANALYSIS FOR EMOTION RECOGNITION

Generally, there are three main steps in emotion recognition (Kumari et al., 2015). The first step is detecting face among other objects and locating important facial fiducial points. The second step is extracting facial features that become facial expression descriptors using specific methods. The last step is performing classification by using facial features as input to recognize emotions. For the first step, Ekman (2003) has created a face atlas of facial components which consists of three parts of the face: upper part (eyebrows and forehead); center part (eyes and nose); and lower part (cheek, mouth, and chin). These components are involved in the change in facial expressions. For the second step, we utilized geometric facial features to produce high-level fuzzy linguistic features of facial components. In the last step, we performed fuzzy rule-based emotion classification.

We have developed a model of emotion recognition which consists of several processing steps. Initially, the input is a static face image under conditions of a single subject and frontal view without occlusion. Each subsequent image is processed in the following manner. First, the face region is detected, resulting in facial points. The next geometric facial components features extraction step is transforming facial points into facial components features using a mathematical model and resulting in geometric facial features parameters. After that, fuzzy facial components inference system is performed to generate a set of linguistic conditions of facial components as well as parameter values as input for the emotion recognition subsystem. Lastly, the fuzzy emotion inference system determines the emotion result. The proposed emotion recognition is illustrated in Figure 1.

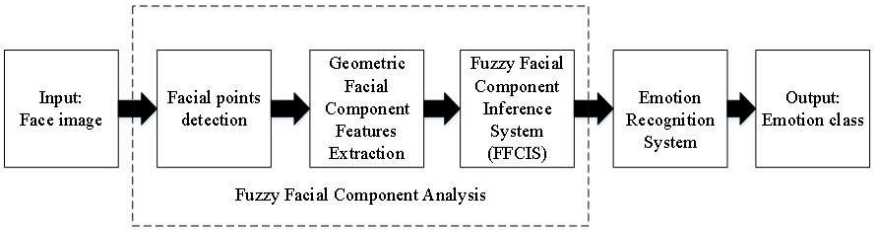


**Figure 1**. Diagram of emotion recognition model.

Figure 1 shows the diagram of our proposed emotion recognition model. Four subsystems are involved in the model: facial points detection; geometric facial components features extraction; Fuzzy Facial Components

Inference System (FFCIS); and Fuzzy Emotion Inference System (FEIS). The rectangles inside the dashed-lines represent the features extraction phase. The process converts low-level image features (geometric features) into high-level fuzzy linguistic features using FFCIS. We divided our proposed model into five subsections. Subsection one discusses facial components analysis. Subsection two talks about facial point detection techniques. Subsection three discusses the geometric facial components features extraction process. Subsection four talks about the fuzzy facial components inference system. The last subsection discusses the emotion recognition system.

**Facial Components Analysis**

Initially, the face area is detected by performing Active Appearance Model (AAM) which works rapidly in locating the face region and marking facial points (FPs) (Cootes et al., 2001). FPs which indicate the location of facial landmarks are extracted using our proposed geometric features extraction methods. Only FPs located on facial components are used and this makes our method work fast. As a result, a set of geometric features parameters is extracted from 10 facial components. Table 1 describes the geometric facial features for the corresponding facial components (FC).

Table 1

*Geometric Features of Facial Components*

| FC Number | Facial Component | Feature Description | Feature Label | FC Output |
|---|---|---|---|---|
| 1 | Left eyebrow ($FC_1$) | eccentricity<br>distance ratio | $FC_1GF_1$<br>$FC_1GF_2$ | $yFC_1$ |
| 2 | Right eyebrow ($FC_2$) | eccentricity<br>distance ratio | $FC2GF_1$<br>$FC_2GF_2$ | $yFC_2$ |
| 3 | Inner eyebrow ($FC_3$) | distance ratio<br>height ratio | $FC_3GF_1$<br>$FC_3GF_2$ | $yFC_3$ |
| 4 | Left eye ($FC_4$) | eccentricity<br>opening ratio | $FC_4GF_1$<br>$FC_4GF_2$ | $yFC_4$ |
| 5 | Right eye ($FC_5$) | eccentricity<br>opening ratio | $FC_5GF_1$<br>$FC_5GF_2$ | $yFC_5$ |
| 6 | Nose ($FC_6$) | wrinkle ratio<br>height ratio | $FC_6GF_1$<br>$FC_6GF_2$ | $yFC_6$ |

(continued)

| FC Number | Facial Component | Feature Description | Feature Label | FC Output |
|---|---|---|---|---|
| 7 | Upper lip ($FC_7$) | eccentricity<br>thickness ratio | $FC_7GF_1$<br>$FC_7GF_2$ | $yFC_7$ |
| 8 | Lower lip ($FC_8$) | eccentricity<br>thickness ratio | $FC_8GF_1$<br>$FC_8GF_2$ | $yFC_8$ |
| 9 | Inner mouth ($FC_9$) | eccentricity<br>opening ratio | $FC_9GF_1$<br>$FC_9GF_2$ | $yFC_9$ |
| 10 | Outer mouth ($FC_{10}$) | eccentricity<br>opening ratio | $FC_{10}GF_1$<br>$FC_{10}GF_2$ | $yFC_{10}$ |

Geometric facial components features parameters are fed into the fuzzy facial components inference system (FFCIS). FFCIS consists of 10 fuzzy rule-based engines, each engine for each facial component. The result of FFCIS is a set of fuzzy linguistic features of facial components. These fuzzy linguistic features become parameters for the fuzzy emotion recognition, resulting in emotion classification. Each facial component has a different number of feature descriptors, depending on the shape of the facial component. We utilized 10 facial components and labeled them as *FCi* as in Table 1, where *i* refers to the facial component number. We defined geometric features for each component based on geometric shape and characteristics of facial components. A total of 20 geometric features (GF) of facial components were designed and denoted as *FCiGFj*, where *i* refers to the facial component number and *j* refers to the order of the facial component geometric feature. Basically, there are two types of GF which are distance ratio and ellipse eccentricity. We will explain more about these geometric features extraction methods in the next subsection. The last column of Table 1 is the output of fuzzy facial components analysis, denoted as *yFCi*, where *i* represents the corresponding FC number.

**Facial Points Detection**

AAM is utilized to perform facial points detection. AAM is effective and flexible for object tracking, especially for facial landmarks detection (Wang et al., 2014). We applied fast AAM framework by (Tzimiropoulos & Pantic, 2013) for facial points detection. Figure 2(a) shows the screenshot image of the AAM process in locating facial points, while Figure 2(b) shows the FPs detection results.

The AAM generated 68 facial points; each of them had information about the coordinates of its position. The next step after FPs detection was geometric features extraction. We used FPs pixel coordinates as input and

processed it through a simple calculation method which utilized x-axis and y-axis values of FPs and transformed them into fuzzy facial components values.

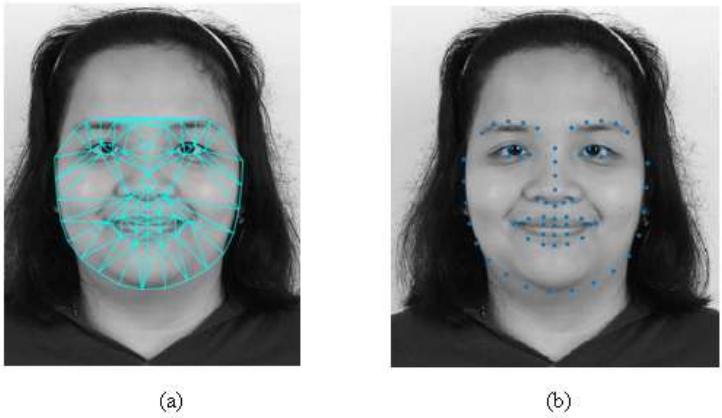


**Figure 2**. AAM facial points detection (a) process (b) results.

**Geometric facial components features extraction**

We registered all FPs coordinates of their positions but excluded some FPs which were located on the jaw area in the features extraction process because these FPs did not produce rapid signals for geometric features. We processed each facial component separately using the respective FPs as explained in the previous subsection. Unlike other studies, we did not use any reference template for features extraction. Generally, we have proposed two types of geometric features: ellipse eccentricity and distance ratio. The first feature comes from the idea that several facial components have an elliptic or half-elliptic shape (e.g. eyes, eyebrows, lips, and mouth). We utilized the eccentricity parameter of ellipse as the first type of geometric features. Eccentricity represents the elliptical curvature level. This value ranges from zero to one, where zero value indicates a circle. The larger the value, the more elliptical is the object. We related this to the facial component case; wide-opened eyes or wide-opened mouth implies a smaller eccentricity value, and vice versa. The second type of geometric feature is distance ratio. We used ratio since it is independent of the measurement unit and it indicates the distance in comparison with the height or width of facial components.

$FC_1$ and $FC_2$ have the same geometric features as mentioned in Table 1. Thus, feature formulation is the same for both FCs. Figure 3(a) shows $FC_1$ and $FC_2$, and Figure 3(b) shows $FC_3$ areas with their corresponding FPs.

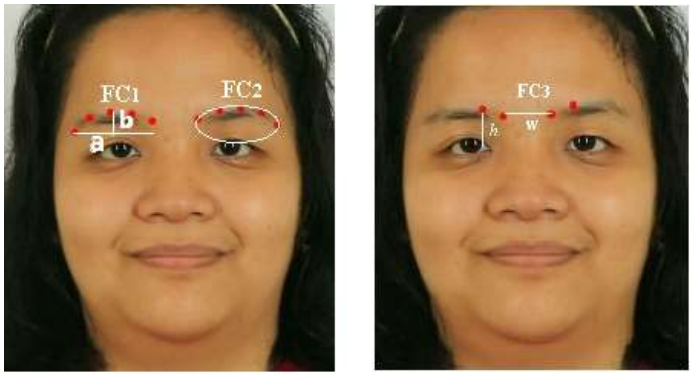


**Figure 3**. Facial component of eyebrows showing (a) left eyebrow FC1 and right eyebrow FC2 (b) inner eyebrow FC3.

The first feature of $FC_1$ and $FC_2$ is eyebrows eccentricity. Based on this study, we assume that $FC_1$ and $FC_2$ are half-ellipse shape objects (Loconsole et al., 2014), thus we used ellipse eccentricity property to extract $FC_1$ and $FC_2$ geometric features. $FC_1$ has three parameters a, b, and c, and so has $FC_2$; where b is the height of $FC_1$ obtained by subtracting $FC_1$ y-axis maximum value with $FC_1$ y-axis minimum value; $a$ is the half width of $FC_1$, obtained by subtracting the x-axis value of $FC_1$ outermost FP to the x-axis value of $FC_1$ innermost FP. We calculated the eccentricity parameter using Equation 1.

$$FC_1GF_1 = \frac{\sqrt{a^2 - b^2}}{a}, \tag{1}$$

where $FC_1GF_1$ is the first geometric feature of $FC_1$, which refers to the eccentricity feature. We also applied Equation 1 for FCs which used eccentricity features, as listed in Table 1. The same calculation process was applied to $FC_2GF_1$ using Equation 1 for the right eyebrow.

The second feature of $FC_1GF_2$ and $FC_2GF_2$ is the height ratio which is defined as the ratio of eyebrow height to face height which is obtained using Equation 2, where $FC_1GF_2$ is the second geometric feature of $FC_1$, which refers to the height ratio between, eyebrow height to face height denoted by $F_{height}$ using Equation 2.

$$FC_1GF_2 = \frac{a}{F_{height}}. \tag{2}$$

$FC_3$ geometric features are distance ratio and height ratio. Distance ratio of inner eyebrow ($FC_3$) in Figure 3(b) is formulated in Equation 3.

$$FC_3GF_1 = \frac{w}{F_{width}}, \tag{3}$$

$FC_3GF_1$ is the distance ratio between $FC_3$ with face width denoted by $F_{width}$, where $w$ is the distance between the left and right inner eyebrow. Meanwhile, the height ratio of $FC_3$ is calculated using Equation 4.

$$FC_3GF_2 = \frac{h}{F_{height}}, \tag{4}$$

where $FC_3GF_2$ is the height ratio of $FC_3$ and $h$ is the distance between the inner FP to the center of the eye.

Figure 4(a) shows the left eye ($FC_4$) and right eye ($FC_5$) and their FPs.

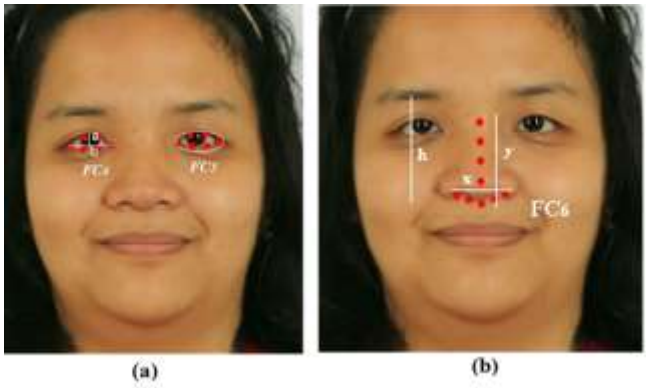


**Figure 4**. Facial components showing (a) left eye FC1 and right eye FC2 (b) nose FC6 .

Eyes have an elliptic shape, so we assume that geometric feature for eyes is ellipse eccentricity. Two geometric features for $FC_4$ and $FC_5$ are eyes eccentricity and eyes opening ratio. Equation 1 is applied to gain $FC_4GF_1$ and $FC_5GF_1$, where $a$ is the half width of $FC_4$ or $FC_5$ and $b$ is the half height of $FC_4$ or $FC_5$. For the second feature, we applied Equation 5 to obtain the opening ratio $FC_4GF_2$ and $FC_5GF_2$.

$$FC_4GF_2 = \frac{b}{a}. \tag{5}$$

Figure 4(b) shows nose facial component ($FC_6$) with two geometric features: wrinkle ratio and height ratio. Wrinkle ratio $FC_6GF_1$ indicates how shrinking the nose is measured using Equation 6.

$$FC_6GF_1 = \frac{x}{y}, \tag{6}$$

where $x$ is $FC_6$ width and $y$ is $FC_6$ height. The larger the $FC_6GF_1$ value, the more wrinkled the nose, and vice versa. Height ratio $FC_6GF_2$ is obtained using Equation 7.

$$FC_6GF_2 = \frac{y}{h}, \tag{7}$$

where $h$ is the distance between the top FP of the eyebrow and the lowest FP of the nose.

Figure 5 shows $FC_7$, $FC_8$, $FC_9$, and $FC_{10}$. The four FCs use eccentricity as the first geometric feature and labeled as $FC_iGF_1$, where i= 7,…,10 refer to the FC numbers.

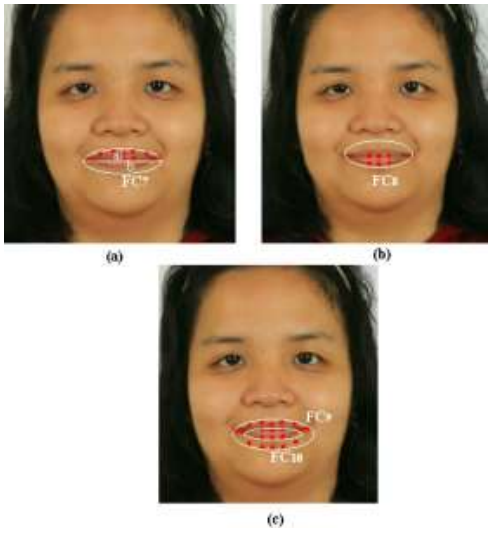


**Figure 5**. (a) upper lip FC7 (b) lower lip FC8 (c) inner mouth FC9 and outer mouth FC10.

We applied Equation 1 to obtain these feature values. The second geometric feature for $FC_7$ and $FC_8$ is the thickness ratio, while for $FC_9$ and $FC_{10}$ is the opening ratio. These values are obtained using Equation 5. We obtained geometric feature values for all facial components after the features extraction process. The next step is fuzzifying these facial components values and entering them into the Fuzzy Facial Components Inference System (FFCIS).

**Fuzzy Facial Components Inference System (FFCIS)**

We have produced the relevant geometric feature parameters for fuzzy facial component analysis in the previous step. Using resultant parameters as input, we designed FFCIS for each facial component and utilized the Fuzzy Mamdani Inference method (Mamdani, 1974). Thus, from this phase we obtained 10 results of fuzzy facial components parameters *yFCi*, where $i = 1,\ldots,10$ is the number of facial components. The design of an FFCIS is shown in Figure 6. Two inputs for each FFCIS: $FC_iGF_1$ and $FC_iGF_2$ refer to the first geometric feature of *i*-th FC and the second geometric feature of *i*-th FC, respectively. A set of rules are mapping input into output using the Fuzzy Mamdani inference method. The output is the facial components parameter $yFC_i$.

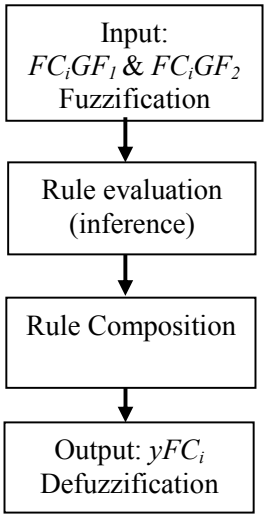


**Figure 6.** Fuzzy facial components inference system flowchart.

In this study, psychological knowledge is involved in determining fuzzy parameters. Three factors that need to be considered in fuzzy parameters definition are the: type of membership function, number of membership

functions in each fuzzy variable, and interval value for each linguistic variable. The first factor is determined based on the characteristics of the problem related to fuzzy variables. Triangular membership function is widely used because it is simple and yields a good result (Chaturvedi & Tripathi, 2014). The second factor is the number of membership function for each fuzzy variable which is determined based on psychological knowledge, regarding various states of facial components linguistic conditions. While the last factor, or the interval value of linguistic variables is obtained through experiments on various facial expression images.

The flowchart in Figure 6 describes the FFCIS process. Four main steps are input fuzzification, rule evaluation, rule composition, and defuzzification. In the first step, each input linguistic variable is defined as low, medium, and high. Geometric feature values are fuzzified into fuzzy values using triangular membership functions. Fuzzy triangular membership function is a function which maps a crisp value in the x-axis into a fuzzy value in the y-axis. It uses three parameters α1, α2, α3 of a triangular curve for fuzzification of any input *x* using Equation 8.

$$
\begin{aligned}
f(x) &= 0, x \le \alpha_1 \text{ or } x \ge \alpha_3 \\
f(x) &= \frac{x-\alpha_1}{\alpha_2-\alpha_1}, \alpha_1 x \le \alpha_2 \\
f(x) &= \frac{\alpha_3-x}{\alpha_3-\alpha_2}, \alpha_2 x \le \alpha_3
\end{aligned}
\tag{8}
$$

where f(x) is the degree of membership or fuzzy value; $\alpha_1$, $\alpha_2$, $\alpha_3$ are left, middle, and right parameters of fuzzy triangular membership function. For output, the membership function is a triangular curve, while the linguistic variable depends on the type of facial components. Table 2 lists the output linguistic variables.

Table 2

*FFCIS Output Linguistic Var*

| No. | FFCIS | Output | Fuzzy Linguistic Variable | | |
|---|---|---|---|---|---|
| 1 | Left eyebrow | $yFC_1$ | lower | normal | raised |
| 2 | Right eyebrow | $yFC_2$ | lower | normal | raised |
| 3 | Inner eyebrow | $yFC_3$ | closer | normal | - |
| 4 | Left eye | $yFC_4$ | narrow | normal | wide |

(continued)

| No. | FFCIS | Output | Fuzzy Linguistic Variable | | |
|---|---|---|---|---|---|
| 5 | Right eye | $yFC_5$ | narrow | normal | wide |
| 6 | Nose | $yFC_6$ | normal | wrinkled | - |
| 7 | Upper lip | $yFC_7$ | thin | normal | thick |
| 8 | Lower lip | $yFC_8$ | thin | normal | thick |
| 9 | Inner mouth | $yFC_9$ | tight | normal | open |
| 10 | Outer mouth | $yFC_{10}$ | tight | normal | open |

As we can see in Table 2, our fuzzy facial components model has 10 separated FFCIS and 10 output parameters. Each output has different fuzzy linguistic variables. For example, the eyes have three linguistic variables: narrow, normal, and wide; while the mouth has three different linguistic variables: tight, normal, and wide. The output linguistic variables depend on facial components traits and psychological knowledge. The interval value is from zero to one.

The next step is rule evaluation. Each FFCIS owns separate rules. A set of rules has been stored in each FFCIS' inference engine. We enumerated all possible input linguistic conditions at the antecedent part and related it with output at the consequence part based on psychological knowledge. As an example, the left eyebrow FFCIS has nine rules. One of its rules is:

IF eccentricity is *low* AND distance ratio is *low* THEN $yFC_1$ is *lower*.

Rule composition uses fuzzified input to evaluate FFCIS rules. This step yields an area of rule aggregation which matches with the input conditions. The last step, defuzzification is applied to the output area using centroid of gravity method to achieve the result of $yFC_i$ value. We implemented our FFCIS model using MATLAB 2014 software. The results were a set of fuzzy facial components (FFC) parameters, $yFC_i$. Next, these output FFC parameters became input and proceeded to the emotion recognition inference system. Figure 7 shows the example of FFCIS membership functions for $FC_1$ with (a) eccentricity input; (b) distance ratio input; and (c) $yFC_1$ output.

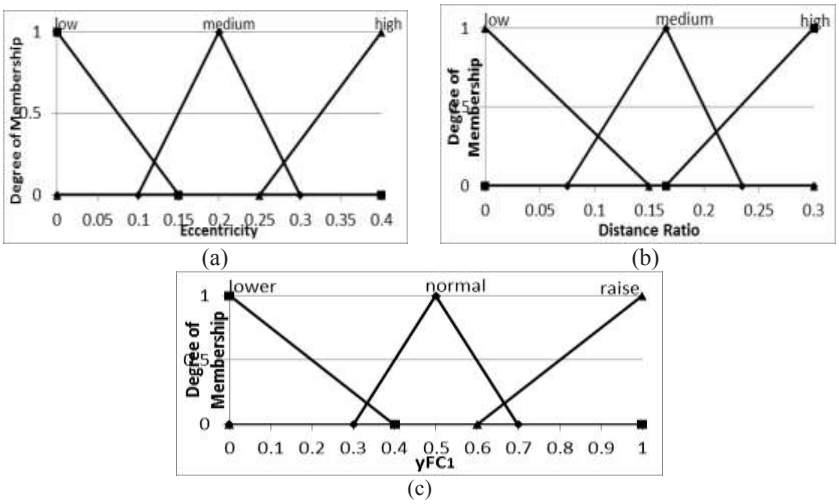


**Figure 7**. Fuzzy membership function for FFCIS.

**Emotion recognition**

Emotion recognition is the last process in our proposed model. The objective is to classify facial expression into six basic emotion classes (happy, sad, angry, disgusted, fearful, and surprised). We designed an inference engine for each basic emotion; six Fuzzy Inference Systems (FIS) using Sugeno method for six basic emotion classes. The output from each FIS is a value between 0 (low intensity) and 1 (high intensity). This value represents the degree of emotion displayed on a face image. We took the highest emotion FIS output value as the result of emotion classification.

The Fuzzy emotion process is the same as the FFCIS process. A vector of facial components parameters *yFCi* becomes the input of Fuzzy emotion. Each Fuzzy emotion has its own rules which are stored in an inference engine. The IF-THEN rules are constructed from the facial components linguistic condition as the antecedent, and the emotion as the consequent. An example of emotion rule for surprised is:

> IF $yFC_1$ is *raise* AND $yFC_2$ is *raise* AND $yFC_4$ is *wide* AND $yFC_5$ is *wide* AND $yFC_7$ is *thick* AND $yFC_8$ is *normal* AND $yFC_{10}$ THEN *Surprised Emotion* is 1 (*high*).

The rule evaluation is the process of scanning relevant rules for some input linguistic conditions. Rules which are triggered by the input are

composed to build the output area. The last step is the defuzzification process using weighted average sum method which determines the degree of emotion value as in Equation 9.

$$Z = \frac{\sum_{i=1}^{n} \alpha_i * z_i}{\sum_{i=1}^{n} \alpha_i}, \tag{9}$$

where Z is the defuzzification result or the intensity of fuzzy emotion, α is the membership degree of input, and z is a real input value on x-axis. The highest *Z* value determines the emotion classification result.

## EXPERIMENTAL RESULTS

We carried out an experiment to investigate the performance of our proposed model. We monitored two important aspects: facial component linguistic analysis and emotion recognition performances. The first aspect was gained by testing our proposed fuzzy facial components inference system accuracy using CK+ facial expression dataset (Lucey et al., 2010). The second aspect was obtained by testing the proposed emotion recognition system using four different datasets: CK++, JAFFE, DISFA, and our own facial expressions dataset. We also compared our model to other recognition methods to observe performance using Fuzzy C-Means (FCM) clustering, Fuzzy Inference System (FIS), and Support Vector Machine (SVM).

In the first experiment, we tested the performance of our proposed fuzzy facial components inference system by observing the output; a set of linguistic conditions related to an input image which describes states of the facial components. The objective was to measure the correct facial components linguistic conditions. An example of the correct FFCIS output is given in Figure 1 as input: Left eyebrow is normal; right eyebrow is normal; inner eyebrow is normal; left eye is normal; right eye is normal; nose is normal; upper lip is thin; lower lip is thin; inner mouth is tight; outer mouth is normal.

The first experimental result is summarized in Table 3. The first column is emotion classes, while the rest of the columns are the normalized value of correct facial components identification. The last row shows the percentage of correct identification of facial components.

Table 3

*Facial Components Identification Resul*

| No. | Emotion Class | Facial Component | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $FC_1$ | $FC_2$ | $FC_3$ | $FC_4$ | $FC_5$ | $FC_6$ | $FC_7$ | $FC_8$ | $FC_9$ | $FC_{10}$ |
| 1 | Angry | 0.98 | 0.98 | 0.98 | 0.93 | 0.93 | 0.98 | 0.98 | 0.98 | 0.98 | 1.00 |
| 2 | Disgusted | 1.00 | 1.00 | 1.00 | 0.96 | 0.96 | 0.96 | 1.00 | 1.00 | 1.00 | 0.98 |
| 3 | Fearful | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.92 | 1.00 | 1.00 | 1.00 | 1.00 |
| 4 | Happy | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 1.00 | 0.98 | 0.98 | 0.98 |
| 5 | Sad | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.96 | 0.91 | 0.91 | 0.91 | 0.96 |
| 6 | Surprised | 0.98 | 0.98 | 1.00 | 1.00 | 1.00 | 0.98 | 1.00 | 1.00 | 0.93 | 0.98 |
| | Accuracy (%) | 99.26 | 99.26 | 99.63 | 98.22 | 98.22 | 95.96 | 98.18 | 97.85 | 96.74 | 98.24 |

We analyzed the output by comparing facial component linguistic features output with the input image states. We ran the experiment on CK+ dataset consisting of 238 annotated images of six basic emotions: 45 angry, 50 disgusted, 25 fearful, 50 happy, 23 sad and 45 surprised. From Table 3 we can see that the highest identification results is FC3 or inner eyebrow (99.63%) and the lowest identification result is FC6 or nose (95.96%). The average accuracy rate and standard deviation of facial components analysis are 98.15% and 1.07, respectively.

Table 4. shows the example of facial components linguistic features identification results. Six examples for each basic emotion input image describes the high-level linguistic features of facial components in human natural language generated by the system. The blue dots are the AAM facial points detection results. We validated these linguistic conditions based on psychological expert judgment.

The second experiment observed the performance of our proposed model using several facial expression datasets. The objective was to measure the recognition accuracy of our proposed model on different datasets. Figure 8 shows images and labels of six basic emotions recognition results using different facial expression datasets.

Our proposed model selected the highest intensity value amongst six emotion intensity values as a result of fuzzy emotion inference system and classified the input images using the six basic emotions. We used four datasets: CK+ (238 images), JAFFE (183 images), DISFA (149 images), and our own facial expression dataset, Indonesian Mixed Emotion Dataset or IMED (270 images) to test emotion recognition performance. The dataset size is 640x490

pixels for CK+; 256x256 pixels for JAFFE; 1280x720 pixels for DISFA; and 720x480 pixels for IMED. Each image contains a single face with frontal orientation.

Table 4

*Examples of FFCIS output: Facial components linguistic features*

| Input Image | FFCIS Output: Linguistic Feature | Input Image | FFCIS Output: Linguistic Feature |
|---|---|---|---|
|  Angry | left eyebrow is normal<br>right eyebrow is normal<br>inner eyebrow is normal<br>left eye is narrow<br>right eye is narrow<br>nose is wrinkled<br>upper lip is thin<br>lower lip is thin<br>inner mouth is tight<br>outer mouth is tight |  Happy | left eyebrow is normal<br>right eyebrow is normal<br>inner eyebrow is normal<br>left eye is normal<br>right eye is normal<br>nose is normal<br>upper lip is thin<br>lower lip is normal<br>inner mouth is normal<br>outer mouth is open |
|  Disgusted | left eyebrow is normal<br>right eyebrow is normal<br>inner eyebrow is closer<br>left eye is narrow<br>right eye is narrow<br>nose is wrinkled<br>upper lip is thin<br>lower lip is normal<br>inner mouth is normal<br>outer mouth is normal |  Sad | left eyebrow is raised<br>right eyebrow is raised<br>inner eyebrow is closer<br>left eye is normal<br>right eye is normal<br>nose is normal<br>upper lip is thin<br>lower lip is thin<br>inner mouth is tight<br>outer mouth is tight |
|  Fearful | left eyebrow is raised<br>right eyebrow is raised<br>inner eyebrow is normal<br>left eye is normal<br>right eye is normal<br>nose is normal<br>upper lip is thin<br>lower lip is thin<br>inner mouth is tight<br>outer mouth is normal |  Surprised | left eyebrow is raised<br>right eyebrow is raised<br>inner eyebrow is normal<br>left eye is wide open<br>right eye is wide open<br>nose is wrinkled<br>upper lip is normal<br>lower lip is normal<br>inner mouth is open<br>outer mouth is open |

Furthermore, we displayed the confusion matrix which represents the emotion recognition results of our proposed model on facial expression datasets: CK+, JAFFE, DISFA, and IMED. Confusion matrix represents the performance of the classification model. Rows display the number of prediction results, while columns indicate the actual classes. Table 4(a) represents the confusion matrix for CK+ dataset; Table 4(b) is the confusion matrix for JAFFE dataset; Table 4(c) is the confusion matrix for DISFA dataset; and Table 4(d) is the confusion matrix for IMED dataset.

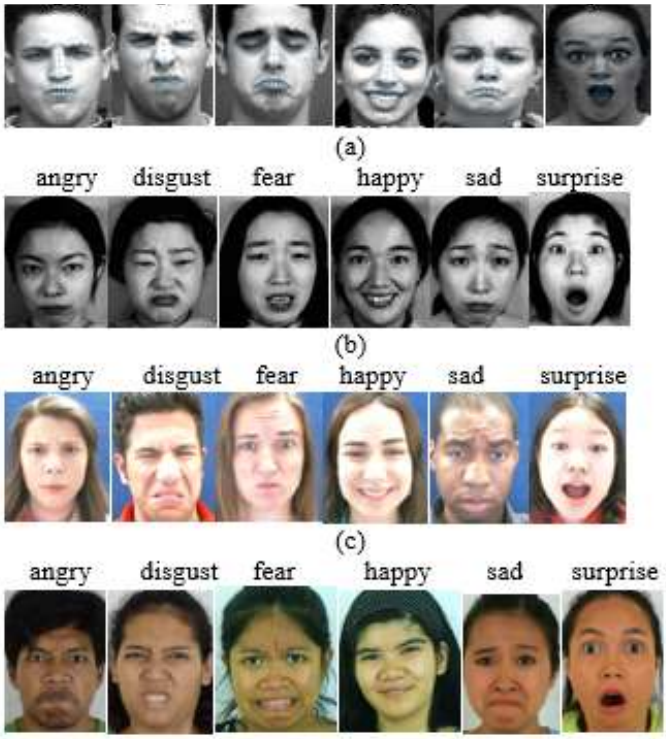


**Figure 8**. Emotion recognition using different datasets (a) CK+; (b) JAFFE; (c) DISFA; (d) IMED.

Table 4

*Confusion matrices using different datasets*

| **(a) CK+** | | | | | | | | **(b) JAFFE** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Emo. | An | dis | fe | ha | sa | su | Prec. | Emo. | An | dis | fe | ha | sa | su | Prec. |
| an | 43 | 1 | 1 | 0 | 0 | 0 | 0.96 | an | 25 | 0 | 0 | 2 | 1 | 2 | 0.83 |
| di | 0 | 48 | 1 | 0 | 1 | 0 | 0.96 | di | 2 | 26 | 0 | 1 | 0 | 1 | 0.87 |
| fe | 0 | 0 | 23 | 0 | 2 | 0 | 0.92 | fe | 2 | 0 | 26 | 0 | 0 | 3 | 0.84 |
| ha | 0 | 0 | 0 | 49 | 0 | 1 | 0.98 | ha | 4 | 0 | 1 | 26 | 0 | 0 | 0.84 |
| sa | 1 | 1 | 0 | 0 | 21 | 0 | 0.91 | sa | 2 | 0 | 2 | 0 | 25 | 2 | 0.81 |
| su | 0 | 0 | 1 | 0 | 0 | 44 | 0.97 | su | 1 | 0 | 0 | 0 | 1 | 28 | 0.93 |
| Rec. | 0.98 | 0.96 | 0.88 | 1 | 0.87 | 0.98 | | Rec. | 0.69 | 1 | 0.90 | 0.90 | 0.93 | 0.78 | |

(continued)

| (c) DISFA | | | | | | | | (d) IMED | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Emo. | an | dis | fe | ha | sa | su | Prec. | Emo. | An | dis | fe | ha | sa | su | Prec. |
| an | 22 | 1 | 0 | 2 | 1 | 0 | 0.85 | an | 39 | 0 | 2 | 2 | 2 | 0 | 0.87 |
| di | 1 | 28 | 1 | 0 | 0 | 0 | 0.93 | di | 2 | 40 | 1 | 2 | 0 | 0 | 0.89 |
| fe | 0 | 0 | 22 | 0 | 1 | 1 | 0.92 | fe | 3 | 1 | 39 | 2 | 0 | 0 | 0.87 |
| ha | 2 | 0 | 0 | 21 | 0 | 0 | 0.91 | ha | 0 | 1 | 3 | 40 | 1 | 0 | 0.89 |
| sa | 1 | 0 | 0 | 2 | 18 | 0 | 0.86 | sa | 3 | 0 | 0 | 2 | 40 | 0 | 0.89 |
| su | 1 | 0 | 0 | 1 | 0 | 23 | 0.92 | su | 1 | 0 | 0 | 2 | 1 | 41 | 0.91 |
| Rec. | 0.81 | 0.97 | 0.96 | 0.81 | 0.90 | 0.96 | | | 0.81 | 0.95 | 0.87 | 0.80 | 0.91 | 1 | |

*an = angry; dis = disgusted; fe = fearful; ha = happy; sa = sad; su = surprised; prec = precision; rec = recall

We can see from Table 4 precision and recall scores for each emotion class. Precision indicates the prediction correctness, while recall indicates how correct the recognition result is to the actual data. Table 4(a) is the CK+ confusion matrix and the highest precision score is 0.98 for happy. This means that 49 images from the happy class were correctly classified as happy by the system, with only one image misclassified as surprised. The lowest precision score is 0.91 for sad, with two misclassified images: angry and disgusted. The highest recall score in Table 4 is 1 for happy. This implies that none of the images were misclassified as happy. While the lowest recall score is sad, with one image of disgusted and two images of fearful misclassified as sad. Table 4(b) is the confusion matrix for the JAFFE dataset. It shows that the highest precision score is 0.93 for surprised and the lowest precision score is 0.81 for sad. Meanwhile, the highest recall score is 1 for disgusted and the lowest recall score is 0.69 for angry. Recall score is low for angry in JAFFE because some images from the other classes besides happy are misclassified as angry. Table 4(c) is the confusion matrix for DISFA dataset. It shows that the highest precision score is 0.93 for disgusted and the lowest precision score is 0.85 for angry. Meanwhile, the highest recall score is 0.97 for disgusted and the lowest recall score is 0.81 for angry and happy. Table 4(d) is the confusion matrix for IMED dataset. It shows that the highest precision score is 0.91 for surprised and the lowest precision score is 0.87 for angry and fearful. The highest recall score is 1 for surprised and the lowest recall score is 0.80 for happy.

We summarized confusion matrices in Table 5 and displayed the results using classification measurement tools: average precision, average recall, average accuracy, and F1-score of multiclass to measure recognition performance. The result is displayed in Table 5.

Table 5

*Recognition results from different datasets*

| Dataset | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| CK+ | 0.958 | 0.951 | 0.946 | 0.948 |
| JAFFE | 0.853 | 0.853 | 0.865 | 0.859 |
| DISFA | 0.899 | 0.898 | 0.901 | 0.899 |
| IMED | 0.885 | 0.885 | 0.890 | 0.887 |

We can see from each row that the performance of our proposed model is consistent and high, indicated by the uniform values of average precision, recall, accuracy and F1 score in each class.

The last experiment compared the performance of the proposed model with other classifiers: Fuzzy C-Means (FCM), Fuzzy Inference System (FIS), and Support Vector Machines (SVM). The experiment used CK+ dataset (238 images) and classified six classes of basic emotions. A comparison of the results is shown in Table 6.

Table 6

*Recognition results using different methods*

| Dataset | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| FCM | 0.833 | 0.830 | 0.830 | 0.830 |
| FIS | 0.724 | 0.720 | 0.730 | 0.725 |
| SVM | 0.925 | 0.930 | 0.930 | 0.930 |
| Proposed model | 0.958 | 0.950 | 0.950 | 0.950 |

Table 6 shows that our proposed model obtained the highest accuracy, precision, recall, and F1-scores; followed by SVM, FCM, and lastly, FIS.

Figure 9 (a) is a graphic representation of the results in Table 5, similarly, the graph in Figure 9(b) is a graphic representation of the results in Table 6. The bar graph represents different measurement tools (accuracy, precision, recall, and F1-score) in Figure 9(a). The bars have uniform height in each dataset (CK+, JAFFE, DISFA, and IMED). This implies that the proposed model showed good performance in different datasets. Meanwhile, in Figure 9(b), our proposed model (the yellow bar) outperformed other classifiers in terms of average accuracy, precision, recall, and F1-scores.

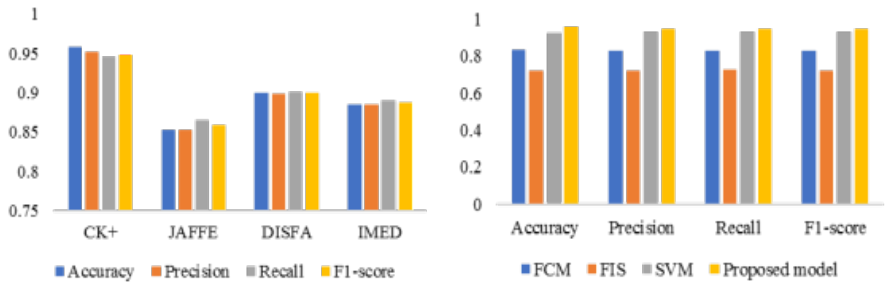


**Figure 9**. Fuzzy facial components inference system flowchart.

## DISCUSSION

From the results of the experiment we can observe two aspects: the performance of facial components linguistic analysis and the performance of emotion recognition using our proposed fuzzy linguistic facial features and fuzzy emotion inference system. The first aspect has been reviewed in Table 3, given that the average accuracy and standard deviation of facial components identification is 98.15%±1.07. This high value is reached because the geometric features extraction has given the best feature descriptors to the FFCIS subsystem. FFCIS rule-based is also powerful in analyzing facial components linguistic conditions. Here we demonstrated that the knowledge developed from psychological rules is the determining factor, resulting in the correct identification of facial components linguistic conditions. The second aspect is the emotion recognition performance. Our proposed model gained a satisfactory recognition result with an accuracy rate of 0.958 surpassing other classifiers. The strength of our proposed model in comparison to other methods is that we applied a fuzzy facial components inference and fuzzy emotion inference, where knowledge is stored in the system. Thus, we did not require any training data for the classification process, unlike FCM and SVM. The difference between the proposed model and the other three classifiers (FCM, FIS, and SVM) is in the image features. We used high-level linguistic facial components features while other classifiers processed whole face images as input. In our model, we processed only important facial points and avoided processing whole images which did not contribute emotion signals in certain parts. Thus, we accelerated the features extraction process by using geometric facial components features.

## CONCLUSION

In this paper, we have adopted a novel approach in emotion recognition employed by psychology experts using high-level linguistic features of facial components. We designed an emotion recognition model which comprises four main methods: facial point detection by AAM; geometric features extraction; fuzzy facial inference system; and emotion recognition using fuzzy inference system. The objective is to classify six basic classes of emotion: happy, sad, angry, disgusted, fearful, and surprised. The performance of the proposed model was analyzed by investigating the facial components linguistic analysis and the emotion classification results. The proposed model showed high recognition in both facial components linguistic analysis and emotion recognition tested on different facial expression datasets and compared with other classifiers. The strength of the model is in the adaptive fuzzy rules which apply psychological knowledge. The enhancement in knowledge and refinement of fuzzy parameters are crucial to secure more accurate recognition results in future. Fuzzy facial components analysis is potentially applicable in various fields which consider facial components linguistic features in observations such as pain detection, stress detection, lie detection, and animation reconstruction. Moreover, fuzzy emotion recognition can help people with emotional disorders (e.g. Alexithymia, Asperger syndrome, and Autism) to recognize emotions.

## ACKNOWLEDGMENT

The authors would like to express their appreciation to colleagues for contributing their expertise in this research. The first author would also like to thank the Indonesia Endowment Fund for Education for funding this doctoral study.

## REFERENCES

Chakrabarti, D., & Dutta, D. (2013). Facial Expression Recognition Using Eigenspaces. *Procedia Technology*, *10*, 755–761. doi: 10.1016/j.protcy.2013.12.419

Chaturvedi, A., & Tripathi, A. (2014). Emotion Recognition using Fuzzy Rule-based System. *International Journal of Computer Applications*, *93*(11), 8887. Retrieved from http://ijcaonline.com/archives/volume93/number11/16260-5920

Chen, J., Chen, D., Gong, Y., Yu, M., Zhang, K., & Wang, L. (2012). Facial expression recognition using geometric and appearance features. *Proceedings of the 4th International Conference on Internet Multimedia Computing and Service - ICIMCS '12*, 29. doi: 10.1145/2382336.2382345

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(6), 681–685. doi: 10.1109/34.927467

Das, D., & Chakrabarty, A. (2016). Emotion recognition from face dataset using deep neural nets. *Proceedings of the 2016 International Symposium on INnovations in Intelligent SysTems and Applications, INISTA 2016*. doi: 10.1109/INISTA.2016.7571861

Ekman, P., & Friesen, W. V. (1975). Unmasking the face. Englewood Cliffs, NJ: Prentice-Hall.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3–4), 169–200. doi: 10.1080/02699939208411068

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial action coding system. manual and investigator's guide, Salt Lake City, UT: Research Nexus.

Ekman, P. (2003). *Emotions revealed: Understanding faces and feelings*. Times Books, US: New York. Orion. doi: 10.1002/bsl

Halder, S., Bhattacharjee, D., Nasipuri, M., Basu, D. K., & Kundu, M. (2010). Fuzzy Classification of Facial Component Parameters. *CoRR*, *abs/1006.5*. Retrieved from http://arxiv.org/abs/1006.5945

Kumari, J., Rajesh, R., & Pooja, K. M. (2015). Facial Expression Recognition: A Survey. *Procedia Computer Science*, *58*, 486–491. doi: 10.1016/j.procs.2015.08.011

Lekdioui, K., Messoussi, R., Ruichek, Y., Chaabi, Y., & Touahni, R. (2017). Facial decomposition for expression recognition using texture/shape descriptors and SVM classifier. *Signal Processing: Image Communication*, *58*(July), 300–312. doi: 10.1016/j.image.2017.08.001

Li, B., Lian, X. C., & Lu, B. L. (2012). Gender classification by combining clothing, hair and facial component classifiers. *Neurocomputing*, *76*(1), 18–27. doi: 10.1016/j.neucom.2011.01.028

Liliana, D. Y., Widyanto, M. R., & Basaruddin, T. (2017). Human emotion recognition based on active appearance model and semi-supervised fuzzy C-means. In *2016 International Conference on Advanced Computer Science and Information Systems, ICACSIS 2016* (pp. 439–445). doi: 10.1109/ICACSIS.2016.7872744

Loconsole, C., Miranda, C. R., Augusto, G., Frisoli, A., & Orvalho, V. (2014). Real-time Emotion Recognition - Novel Method for Geometrical Facial Features Extraction. In *Proceedings of the 9th International*

*Conference on Computer Vision Theory and Applications*, *1,* 378–385. doi: 10.5220/0004738903780385

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, CVPRW 2010*, (May), 94–101. http://doi.org/10.1109/CVPRW.2010.5543262

Mamdani, E. H. (1974). Application of fuzzy algorithms for control of simple dynamic plant. *Proceedings of the Institution of Electrical Engineers*, *121*(12), 1585–1588. doi: 10.1049/piee.1974.0328

Mayer, C., Wimmer, M., & Radig, B. (2010). Adjusted pixel features for robust facial component classification. *Image and Vision Computing*, *28*(5), 762–771. doi: 10.1016/j.imavis.2009.07.012

Nicolai, A., & Choi, A. (2015). Facial Emotion Recognition Using Fuzzy Systems. *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 2216–2221. doi: 10.1109/SMC.2015.387

Pantic, M., Cowie, R., D'Errico, F., Heylen, D., Mehu, M., Pelachaud, C., … Vinciarelli, A. (2011). Social Signal Processing: The Research Agenda. *Visual Analysis of Humans*, 511–538. doi: 10.1007/978-0-85729-997-0_26

Pitaloka, D. A., Wulandari, A., Basaruddin, T., & Liliana, D. Y. (2017). Enhancing CNN with Preprocessing Stage in Automatic Emotion Recognition. *Procedia Computer Science*, *116*, 523–529. doi: 10.1016/j.procs.2017.10.038

Pratiwi, N., Widyanto, M. R., Basaruddin, T., & Liliana, D. Y. (2017). Nonlinear Fuzzy Robust PCA on Shape Modelling of Active Appearance Model for Facial Expression Recognition. In *Proceedings of the International Conference on Video and Image Processing - ICVIP 2017* (pp. 68–72). New York, NY, USA: ACM. doi: 10.1145/3177404.3177444

Sadeghi, H., Raie, A. A., & Mohammadi, M. R. (2013). Facial expression recognition using geometric normalization and appearance representation. In *Iranian Conference on Machine Vision and Image Processing, MVIP* (pp. 159–163). IEEE. doi: 10.1109/IranianMVIP.2013.6779970

Sudhakar, K., & Nithyanandam, P. (2017). An accurate facial component detection using gabor filter. *Bulletin of Electrical Engineering and Informatics*, *6*(3), 287–294. doi: 10.11591/eei.v6i3.683

Sujono, & Gunawan, A. A. S. (2015). Face Expression Detection on Kinect Using Active Appearance Model and Fuzzy Logic. *Procedia Computer Science*, *59*(Iccsci), 268–274. doi: 10.1016/j.procs.2015.07.558

Tang, X., Guo, F., Shen, J., & Du, T. (2018). Facial landmark detection by semi-supervised deep learning. *Neurocomputing*, *297*, 22–32. doi: 10.1016/j.neucom.2018.01.080

Tzimiropoulos, G., & Pantic, M. (2013). Optimization Problems for Fast AAM Fitting in-the-Wild. In *2013 IEEE International Conference on Computer Vision* (pp. 593–600). doi: 10.1109/ICCV.2013.79

Wang, L., Li, R., & Wang, K. (2014). A Novel Automatic Facial Expression Recognition Method Based on AAM. *Journal of Computers*, *9*(3), 608–617. doi: 10.4304/jcp.9.3.608-617