JEEMECS (Journal of Electrical Engineering, Mechatronic and Computer Science ISSN 2614-4859
Vol. 6, No. 1, February 2023, pp. xx-xx

Designing a Facial Expression Detection System to Determine the Level of Customer Satisfaction with K-Nearest Neighbor Method

Diana Rahmawati [a,1,*], Koko Joni [a,2], Muhammad Iqbal Maulana [a,3]

[a] Electrical Engineering, Faculty of Engineering, Trunojoyo University of Madura, Bangkalan, Indonesia
[1] diana.rahmawati@trunojoyo.ac.id [*]; [2] kokojoni@trunojoyo.ac.id; [3] 170431100041@student.trunojoyo.ac.id
[*] corresponding author

Keywords
Conventional System
Facial Expression
Customer Satisfaction
Viola-Jones
K-Nearest Neighbor

ABSTRACT

Facial expressions are one of the ways humans communicate to convey one's emotions to their communication partner nonverbally. Therefore, human facial expressions can be used for various purposes, one of which is knowing customer satisfaction. So far, customers of Bank Rakyat Indonesia (BRI) provide feedback on service quality using only a polling system, namely by filling out a criticism and suggestion form and then entering it in a suggestion box which is distinguished between satisfied and dissatisfied. However, such a method is less effective because it can be easily manipulated by customers and customers are often indifferent to the feedback. So that the improvement of service quality tends to be less effective. This research will design a system that can recognize human facial expressions to determine the level of customer satisfaction with input data in the form of video data taken by a webcam camera with the viola-jones method to detect faces and determine facial patterns. Then the facial data will be classified using the K-Nearest Neighbor method to determine the type of facial expression. Determination of the value of k will determine the success rate of facial expression detection. The processed data will be displayed on a liquid crystal display (LCD) and then stored in a MySQL database. The results showed that the accuracy of facial expression detection was 80.77% from 52 facial expression data.

1. Introduction

In the world of banking, customer satisfaction is the company's priority. The level of customer satisfaction can be a benchmark for the credibility of a company, especially in the banking sector. The most important factor influencing the level of customer satisfaction is service quality. Dimensions of service quality that can be implemented properly is a key factor that has an influence on the success of a company in creating customer satisfaction [1].

Facial expressions are one way for humans to communicate in conveying one's emotions to their nonverbal communication partner. Micro-expressions are facial movements that can reveal the emotions someone is trying to hide. In 1969, Ekman analyzed video interviews of people with mental distress who had attempted suicide and found micro-expressions. Micro-expressions will appear in no time. The valid duration is 0.5 seconds. In addition, micro expression generally occurs at low intensity. Facial expressions are considered more than just emotional reactions. As is well known, facial expressions coordinate social interactions through information and motivational functions. According to Darwin's theory of natural selection (1872), emotions are universal because they have several forms of expression and function at the level of communication between humans, regardless of cultural prejudices [2]. Therefore, human facial expressions can be used for various purposes, one of which is knowing customer satisfaction.

In practice, so far, customers of Bank Rakyat Indonesia (BRI) provide feedback on service quality using only a polling system, namely by filling out a criticism and suggestion form and then entering it in a suggestion box which is distinguished between satisfied and dissatisfied. However, such a method is less effective because it can be easily manipulated by customers and customers are often

indifferent to the feedback. So that the improvement of service quality tends to be less effective. Therefore, the basis for making a system to determine the level of satisfaction by detecting facial expressions from customers.

From the background of the problems above, the author uses technology and science in digital images to detect a person's facial expression. The camera is used to detect and determine facial patterns using the Viola-Jones method, then facial data will be classified using the K-Nearest Neighbor method to determine the type of facial expression. The processed data will be displayed on the LCD and then stored in the MySQL database.

2. The Proposed Method

2.1. Facial Expression

Facial expression is one of the human methods of communicating to convey one's emotions to the other person's communication nonverbally. In general, people tend to recognize facial expressions that show emotions of anger, sadness, joy, and fear [3]. In terms of the impact it causes, emotions are divided into positive emotions and negative emotions. Positive emotions are emotions that are always coveted by everyone, such as happiness, pleasure, satisfaction and the like. Conversely, negative emotions are emotions that are not expected to occur in a person [4].

In general, there are seven micro-expressions, namely: disgust, anger, fear, sadness, happiness, surprise and contempt. Micro-expressions that occur generally have the following characteristics:

1. Sadness emotions, generally sad emotions are characterized by physical characteristics of the upper eyes drooping downwards, the eyes are not concentrated, and the corners of the lips are slightly lowered;
2. Anger emotions, generally angry emotions have physical characteristics of both eyebrows pressing together at the top of the nose, sharp eyes, and slightly narrowed lips;
3. Emotions of pleasure, generally happy emotions have physical characteristics of movement of the muscles that orbit around the eyes so that they have many wrinkles or hollows, the cheeks are pushed up, and the lips are wide;
4. Emotions of humiliation, generally the emotion of humiliation is characterized by the physical characteristics of the corners of the lips tightening and lifting only on one side of the face;
5. Emotions of fear, in general the emotion of fear is characterized by the physical characteristics of both eyebrows being raised simultaneously, the upper eyelids being raised, the lips being slightly stretched horizontally, pulling towards the ears;
6. Disgust emotion, generally disgust emotion is characterized by the physical characteristics of wrinkles at the top of the nose and around the eyes, the upper lip is lifted;
7. Emotions of surprise, generally the emotion of surprise or surprise has physical characteristics of eyebrows with a raised position, the upper eyelids are raised, and the jaw is slightly open [5].

2.2. Viola-Jones

The most commonly used method for detecting faces is the Viola-Jones. Viola-Jones has developed a face detection framework that can analyze images quickly and has a high level of detection accuracy. Many libraries are provided in the viola jones framework to perform the feature extraction process, the haar-like feature is the basic function to improve the selection process of a feature. Viola-Jones introduces a relatively new image representation known as the Integral Image feature. By using the integral image feature, you can quickly calculate haar-like features at any scale or location. The combination of Haar-Like Feature, Integral Image, Adaboost and Cascade classifier are the four main keys of the Viola Jones approach [6]. The face detection algorithm used in the viola jones method can be seen in the image below.

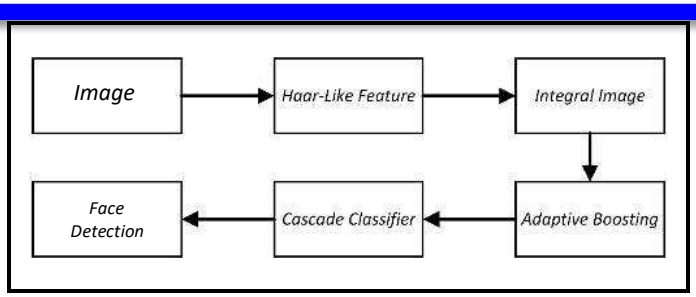


Fig 1. Face Detection Process Using the Viola-Jones Scheme

The following is the face detection process using the Viola-Jones method.

1. First read the image taken by the camera;
2. The reading of haar features is carried out on the obtained image by processing the input image into boxes to get the difference in the threshold value of the dark and light areas of the imag;.
3. Next, integral image to calculate pixel values on hundreds of features more efficiently;
4. Then a machine learning algorithm called adabost is used to select the right Haar feature to be used and to set a threshold value. AdaBoost creates a strong classifier by combining many weak classifiers;
5. Cascade classifier is the next stage. The weights that AdaBoost applies determine the order of the filters in the cascade. The filter with the highest weight is placed in the initial process with the aim of quickly removing non-faced image areas;
6. The last stage is the image detection results will be displayed whether there is a face or not [7].

2.3. K-Nearest Neighbor

K-NN (K-Nearest Neighbor) is a supervised learning algorithm. K-NN can classify data sets based on k number of nearest neighbors. Among the many classification algorithms, K-NN is the simplest and easiest to understand algorithm.

The idea underlying K-NN is to classify objects based on the "closeness" of the object to the "majority" of its neighbors. For example, if an object is near several squares, then K-NN will conclude that the object is a box [8]. In conducting the classification, K-NN has the following stages:

1. Determine the value of k, which is the nearest number of data;
2. Calculate the square of the Euclidean distance with training data;
3. Sorting data from the highest data to the lowest data;
4. Collecting parameters that will be used to classify K-NN using the value of k;
5. The results are obtained from the most majority values.

In K-NN there are several ways to calculate distance from training data, one of which is Euclidean Distance, Cityblock, Correlation, Hamming and Cosine Distance. Here is the mathematical form for calculating the square of the distance with Euclidean:

$$d_1 = \sqrt{\sum_{i=1}^{p}(x_{2i} - x_{1i})^2} \qquad (1)$$

In K-NN there is one general rule in choosing the value of k is to choose k is odd (not even). This is because, the probability of a classification failure is quite large when using k with an even value [18]. Based on suggestions from many machine learning experts, we recommend choosing a value of k that follows the following formula:

$$k \approx \sqrt{n/2} \qquad (2)$$

K-NN is non-parametric statistical, meaning that there is no type of statistical distribution that really fits every data set. In other words, the best value of k really depends on the condition of each data set. K-NN is generally good enough to be applied to very large data sets. On the other hand, a very large number of data sets will cause the distance calculation process (d) to be more complicated so that the computation time increases sharply [9].

2.4. Gray Level Co-Occurance Matrix

GLCM is a co-occurrence matrix that is formed from an image by examining paired pixels with the same gray level intensity. The use of this approach is based on the assumption that in iteration the configuration or pair of grays will appear in the texture or fiber. There are four computational angles in GLCM, namely 0°, 45°, 90°, and 135° [10].

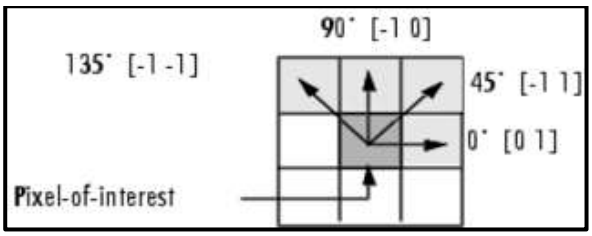


Fig 2. GLCM Computing Angle

The following are the steps in calculating GLCM:

1. The initial GLCM matrix is formed from pairs of two pixels aligned in the 0°, 45°, 90° or 135° directions;
2. By multiplying the initial GLCM matrix with the transpose value, you will get a symmetrical matrix;
3. By dividing each matrix element by the number of pixel pairs to normalize the GLCM matrix;
4. Then the feature extraction process, including:
   a. ASM (Energy)
      The Angular Second Moment (ASM) feature or commonly referred to as energy is used to measure image homogeneity based on the sum of the squares of image elements in the co-occurrence matrix.
      $Energy = \sum_{i,j} p(i,j)^2$ (3)
      Based on the above equation, where i = row, j = column, N = number of pixels.
   b. Entropy
      The more non-uniform the pixel value, the smaller the entropy will be. Conversely, when the entropy value is greater, the image is more uniform.
      $Entropy = -\sum_i \sum_j p(i,j) \cdot \log p(i,j)$ (4)
      Based on the above equation, where i = row, j = column, N = number of pixels.
   c. Contrast
      The difference in intensity between a single pixel and an adjacent pixel is known as contrast. For constant images, the value will be zero.
      $Contrast = \sum_{i,j} |i-j|^2 p(i,j)$ (5)
      Based on the above equation, where i = row, j = column, N = number of pixels.
   d. Homogeneity
      The homogeneity value describes the homogeneity of an image that has a similar degree of gray. Homogeneous image has a large homogeneity.
      $Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|}$ (6)
      Based on the above equation, where i = row, j = column, N = number of pixels.
   e. Correlation
      Correlation is a linear relationship between pixels with respect to other pixels at a certain position. In the same gray-level region higher values can be obtained.
      $Correlation = \sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma i \sigma j}$ (7)
      Based on the above equation, where i = row, j = column, N = number of pixels.

3. Method

3.1. Hardware Design

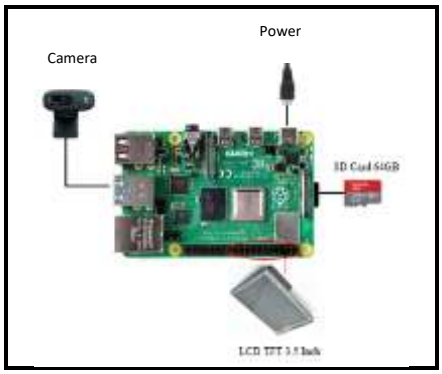


Fig 3. Hardware Design

The hardware used in detecting facial expressions includes a 3.5-inch TFT LCD screen, web camera, Raspberry Pi, switching adapter (5V 2.5A), and Micro SD. The micro SD used in this design has a capacity of 32GB. The LCD screen used in this design is a TFT type measuring 3.5 inches and a touch screen. Here is a single line for the hardware design used:

3.2. Software Design

On Figure.4 below is a flowchart of a program describing software that aims to detect a person's facial expression that has been designed in this study.

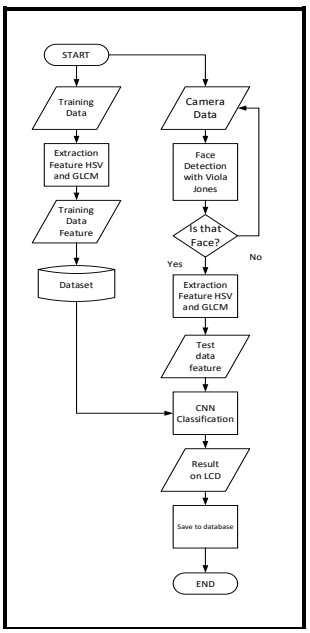


Fig 4. Flowchart System

The system begins with the input of training data that has been provided. Then the training data were extracted using the HSV and GLCM methods. After obtaining the features from the data, the features or characteristics of the data will be stored as a dataset for the classification process. Then after the dataset is available, the program will turn on the camera and take the data in the form of a video image. The data is then identified using the viola jones method to find out whether there is facial data or not. If there is none, the process will be restarted with input data, then if there is facial data, the process will continue with the feature extraction process using the HSV and GLCM methods. After the features are obtained, the next process is the classification process using the K-Nearest Neighbor (KNN) method based on the proximity of the test data to the training data based on the Euclidean distance. After the classification results are obtained, the results will be displayed on the LCD and stored in the MySQL database.

3.3. Design System

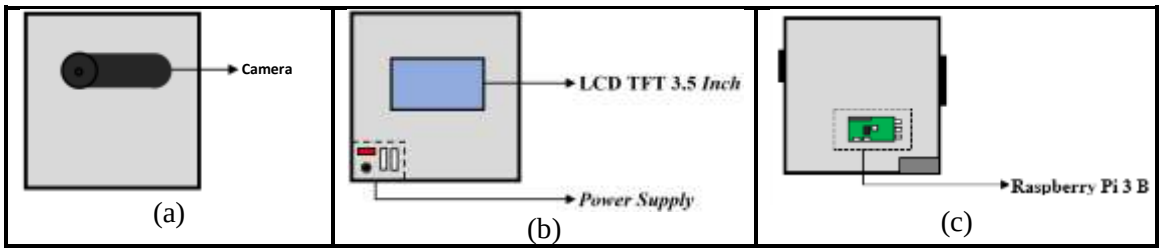


Fig 5. a) Front View Design, b) Back View Design, c) Side View Design

The tool frame design is made of plywood with a thickness of 9 mm and has a black color. The framework of the tool has a height of 20 cm and a width of 20 cm assuming a table height of 100 cm for customer service, because this tool will be placed on the customer service desk.

4. Results and Discussion

4.1. Face Detection Test

1. Testing Based on Face Position

   In the face detection test based on the position of the face, it is carried out in four positions, namely up, down, right side, and left side. The following is the result of the face detection test based on the position of the face.

Table 1. Face Detection Test Result Data Based on Face Position

Position	Detection Result	Description
Up		Detected
Down		Detected
Right		Not Detected
Left		Not Detected

From the test data above, it is found that the system can only detect faces only in the front position. This is intended so that the facial expression detection process can work optimally so that the face detection process can only be carried out in the front position to get the overall facial expression detection. Then in the testing process, the level of lighting is very influential on the

results detected by the system. Bright enough lighting can produce accurate detection results. And vice versa when lighting is lacking, the detected results tend to be less accurate (errors).

2. Testing Using Accessories

The face detection test using facial accessories is carried out under two conditions, namely using transparent glasses and using black glasses. The following is the data from the face detection test using facial accessories**.**

Table 2. Face Test Results Data Using Facial Accessories

Type of Accessories	Detection Result	Description
Face Shield		Detected
Glasses		Detected

In this test, when using facial accessories in the form of a face shield, there is a discrepancy in the results because there is light that reflects on the surface of the face shield and causes a glare effect so that the detection results are not optimal and less accurate. However, when using facial accessories in the form of transparent glasses, the face can be detected as a whole so that the detection process is much more optimal and accurate.

3. Testing Based on Effectiveness Distance

In the face detection test based on the effectiveness distance, it is used to determine the effective distance from the face detection system so that the optimal distance from the face detection system is obtained. The following is the result of the face detection test based on the effectiveness distance.

Table 3. Test Result Data Based on Effectiveness Distance

Distance (cm)	Description
32	Not Detected
33	Detected
100	Detected
150	Detected
187	Detected
188	Not Detected

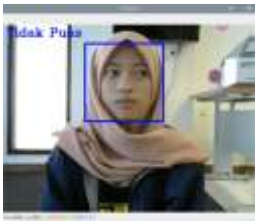
From the test data above, namely testing the effectiveness distance of face detection, the results show that the minimum detection distance on this tool is 33 cm and the maximum detection distance is 187 cm. If the face object is less than 33 cm or more than 187 cm, then the tool cannot detect the face object and greatly affects the detection results. Then the optimal distance from this tool is between 60 cm to 130 cm. If the face object is less than or more than the optimal distance, the tool can still detect. However, the accuracy of the detection results will be reduced because it is too close or too far away.

4.2. Facial Expression Detection Test

In this facial expression detection test, there were 26 respondents to determine the expression of each respondent related to the quality of service at the Trunojoyo University library, Madura. From the expressions of respondents who have been detected, an expression that shows the level of satisfaction is shown in the following Table. 4.

Table 4. Facial Expression Detection Test Results Data

Detection Result	Expert Testimony	Description
	Neutral	Correct
	Satisfied	Correct
	Satisfied	Correct
	Not Satisfied	Correct
	Neutral	Incorrect
	Not Satisfied	Correct
	Not Satisfied	Correct
	Satisfied	Correct

Detection Result	Expert Testimony	Description
	Satisfied	Correct
	Not Satisfied	Correct
	Satisfied	Correct
	Not Satisfied	Correct
	Not Satisfied	Correct
	Satisfied	Correct
	Satisfied	Correct

From the test data above, namely the facial expression detection test, the percentage of success was 80.77% and the percentage of failure was 19.23% from 52 facial expression data detected from 26 respondents. Facial expression data were obtained from interactions made by respondents when using library facilities at Trunojoyo University Madura. After the detection process is carried out and the expression is successfully detected, then the detected expression will be checked by experts who

have the capacity to read facial expressions and the results will be compared with the previously detected expressions to determine the suitability of the system when reading expressions from respondents.

5. Conclusion

The conclusions from this conducted research are described in several points below:

1. From the test results, it is found that the determination of the value of k in the classification process greatly determines the accuracy of the classification results obtained;
2. In the facial expression detection test, the percentage of success was 80.77% of 52 facial expression data detected from 26 respondents using the value of k = 13 as the number of neighbors in the KNN classification process;
3. The test results show that the minimum and maximum distances for facial expression detection are between 33 cm to 187 cm with the optimal distance between 60 cm to 130 c;
4. In this study, there is a delay constraint when there is a fast object movement. This is due to the algorithm used to provide a heavy computing load so that there is a delay;
5. From the results of testing the entire system, the results show that good lighting quality is very influential with the results obtained.

References

[1] S. Normasari, S. Kumadji, and A. Kusumawati, "Pengaruh Kualitas Pelayanan Terhadap Kepuasan Pelanggan, Citra Perusahaan dan Loyalitas Pelanggan," vol. 6, no. 2, pp. 1–9, 2013.

[2] N. Aminarto, Y. A. Sari, and R. C. Wihandika, "Pengenalan Emosi Berdasarkan Ekspresi Mikro Menggunakan Metode Local Binary Pattern," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 10, pp. 3230–3238, 2018.

[3] J. E. Prawitasllri, "Mengenal Emosi Melalui Komunikasi Nonverbal," *Bul. Psikol.*, vol. 3, no. 1, pp. 27–43, 2016, doi: 10.22146/bpsi.13384.

[4] Y. F. Nadhiroh, "PENGENDALIAN EMOSI (Kajian Religio-Psikologis tentang Psikologi Manusia)," *J. SAINTIFIKA Islam.*, vol. 2, no. 1, pp. 53–63, 2015.

[5] E. Setyati, Y. K. Suprapto, and M. H. Purnomo, "Facial emotional expressions recognition based on active shape model and radial basis function network," *CIMSA 2012 - 2012 IEEE Int. Conf. Comput. Intell. Meas. Syst. Appl. Proc.*, pp. 41–46, 2012, doi: 10.1109/CIMSA.2012.6269607.

[6] P. Viola, M. Jones, and M. Energy, "Robust Real-Time Face Detection Intro to Face Detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.

[7] M. H. Purnomo and A. Muntasa, *Konsep Pengolahan Citra Digital dan Ekstraksi Fitur*. Yogyakarta: Graha Ilmu, 2010.

[8] D. Rahmawati, M. P. P. I, M. Ulum, and M. Ulum, "Identification and Classification of Pathogenic Bacteria Using the K-Nearest Neighbor Method Identifikasi Dan Klasifikasi Bakteri Patogen Dengan," vol. 5, no. 1, 2021.

[9] D. Purnama, *Pengantar Machine Learning Konsep dan Praktikum Dengan Contoh Latihan Berbasis R dan PYTHON*. Bandung: INFORMATIKA, 2019.

[10] I. G. Rai *et al.*, "Ekstraksi Fitur Warna , Tekstur dan Bentuk untuk Clustered-Based Retrieval of Images (CLUE)," vol. 16, no. April, 2017.