Advances in Social Science, Education and Humanities Research, volume 528
Proceedings of the 7th International Conference on Research, Implementation, and Education of Mathematics and Sciences (ICRIEMS 2020)

Exploitation of Nasolabial Folds for Happy Smile Recognition on an Image Using ANN

Sri Andayani[1,*] Nur H. Waryanto[2] Bambang S. H. Marwoto[2] Monica M. Hapsari[2]

[1,2] *Mathematics Study Program, Universitas Negeri Yogyakarta*
**Corresponding author. Email: andayani@uny.ac.id*

ABSTRACT
Nasolabial folds are the wrinkle around the mouth that can be used to mark facial expressions. In this study, nasolabial folds were used to identify happy smile expressions on facial images. The recognition consists of 2 main stages. The first stage for facial area detection is carried out using the viola-jones algorithm, then the second stage for identification of happy smile expressions is performed using artificial neural network (ANN), i.e. the backpropagation algorithm. The processing of feature extraction exploits the calculation of the area of the nasolabial folds in the segmented image. The data of feature extraction is then used as input for the process of training and testing of the backpropagation neural network. The accuracy rate of the system for identifying happy smile expressions on 150 facial images is 92% for training dataset and 82% for testing dataset.

Keywords: *Nasolabial folds, happy smile recognition, backpropagation.*

1. INTRODUCTION

Facial expression analysis deals with visual recognition of facial movements and changes in facial features. The basic components of a facial expression analysis system are face detection, facial feature extraction, facial expression recognition. A face detection system is in charge of determining the presence or absence of facial parts and retrieving facial areas from the detected image. While the facial feature extraction system performs the task of extracting the facial data that has been obtained which will then be used for the basis of the facial expression recognition system.

Facial expression recognition is one of the human brain's ability to identify or recognize something. Facial expressions accounted for 55% of message delivery, while language and voice contributed 7% and 38%, respectively [1]. Facial expressions are important for a psychologist in treating patients because they contain important information related to emotions. Ekman [2] defines emotions as including sadness, anger, surprise, fear, disgust, contempt, and enjoyment.

A smile is defined as a facial expression in expressing one's feelings, friendship, approval, and respect for others [3]. Smiling expression is a type of expression due to facial muscle contraction which is a form of body response due to emotional stimulation received by the brain [4]. Smile expressions have a variety of emotions such as happiness, disgust, insult and various other emotions. The expression of a happy smile is a positive emotion that provides reciprocity for someone in controlling emotions, namely that a smile is not only influenced by emotions but can also affect the emotions themselves.

Image processing methods that are often used in the process of identifying facial expressions are the Viola-Jones method for face detection and the Backpropagation algorithm for facial expression recognition. The Viola-Jones method is a method for detecting faces in an image quickly by producing a high degree of accuracy.

Backpropagation is a systematic method for neural network multiplayer training. This method could strike a balance between the ability of the network to recognize patterns during training and the ability of the network to respond correctly to input patterns that are similar to the patterns used during training [5]. The backpropagation method neural network consists of three layers, namely the input layer, the hidden layer, and the output report.

The indicators that are often used in identifying facial expressions generally use features such as eyes, eyebrows, nose, or mouth [6]. One of the studies to determine real-time facial expressions was carried out using features by calculating the distance between features, namely two locations at the center of the eye, two locations at the tip of the deep eyebrows and two locations at the end of the mouth [7]. Another feature is to use the center point of the eyes and the center of the mouth as a feature and calculate the distance between features in determining the pattern of happy expressions [8]. Another feature extraction is in the form of facial characteristics, which is based on the calculation of facial coordinates and the difference in the distance between one feature and another [9].

Nasolabial folds are the wrinkle around the mouth that can be used to mark facial expressions. The expression of a happy smile of a facial image can be indicated by the presence of nasolabial folds that form around the mouth. Therefore, in contrast to previous studies, this study uses a new way to determine the expression of a happy smile by calculating the number of pixels with a value of 1 in the image around the mouth where there are nasolabial folds.

2. BASIC CONCEPTS

2.1. Expressions and Smiles

Emotion is a complex pattern of change that includes components of physiological awakening, subjective feelings, cognitive processes, and behavioural reactions. The process of physiological awakening is triggered by the presence of subjective stimuli experienced by a person and is usually accompanied by behavioural reactions in the form of facial expressions, attitudes, or other reactions.

2.2. Nasolabial fold

The nasolabial angle (NLA) is defined as the angle between a line drawn through the midpoint of the nostrils and a line drawn perpendicular to the Frankfurt horizontal while cutting the sub nasal. The 90 to 120 degree random range for nasolabial angles is usually stated in the literature [10] Four common ways to measure the nasolabial angle found in the rhinoplasty literature used in surveys include: the angle between the columella and the line intersecting the sub nasal and superius labral; the angle between the columella and the tangent to the cutaneous upper lip; the angle between the long axis of the nostril and the perpendicular to Frankfort's horizontal; and the angle between the long axis of the nostril and the line that intersects the glabella and pogonion [11].

2.3. Backpropagation

Artificial Neural Networks (ANN) are mathematical constructs made up of a large number of elements called neurons, each of which can make simple mathematical calculations. By being connected in a network, together the neurons can handle complex problems and questions, and produce accurate problem solutions. A simple neural network consists of three layers of neurons that process input and produce output [12].

The backpropagation algorithm is commonly used to train neural networks. This algorithm works by refining the weighting of the neural network based on the error rate obtained in the previous iteration. When a neural network is initialized, weights are assigned to its neurons. The inputs are loaded, they pass through the neuron network, and the network provides an output for each neuron, with an initial weight. Backpropagation helps to adjust the weight of the neurons so that the results are closer to what is known. Precise weight adjustment ensures a lower error rate, making the model reliable by increasing its generalizability [13].

3. METHOD

The data used are 150 RGB images size 260 x 360 pixels, which show two types of facial expressions, namely a happy smile expression and a sad/neutral expression. Data taken from Centro Universitario da FEI Face Database (https://fei.edu.br). Figure 1 shows some examples of the data.

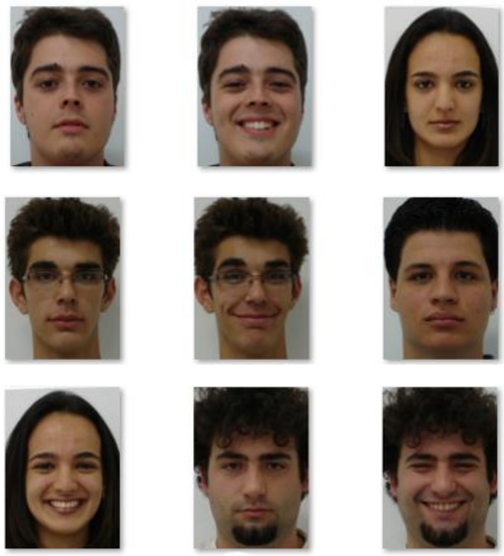


Figure 1 Some data images that show happy smile and sad/neutral expression.

The steps used are divided into two stages, namely, image processing to improve image quality and pattern recognition stages. The identification/classification process uses the backpropagation algorithm for training and network testing. Image quality is improved through contrast stretching, face detection, RGB to grayscale conversion, and resize. In the pattern recognition process, nasolabial folds are used as indicators/characteristics in determining the expression of a happy smile so that in the feature extraction process, the calculation process is carried out by calculating the area in the image of the nasolabial fold area. Feature extraction data are used as input data in the training process and testing of backpropagation neural networks.

The training and testing output data are then compared with the target data to obtain a level of accuracy that can be calculated with the formula:

$$Accuration = \frac{\sum right\ output}{\sum data} \times 100\% \quad (1)$$

4. RESULTS AND DISCUSSION

4.1. Results

Nasolabial folds formed smile line is used as the main feature in recognizing the image of a happy face. Therefore, part of the image used in feature extraction is around the mouth which contains all the strokes around the mouth. Figure 2 shows the part around the mouth where there is a nasolabial fold in the image used in the feature extraction process.

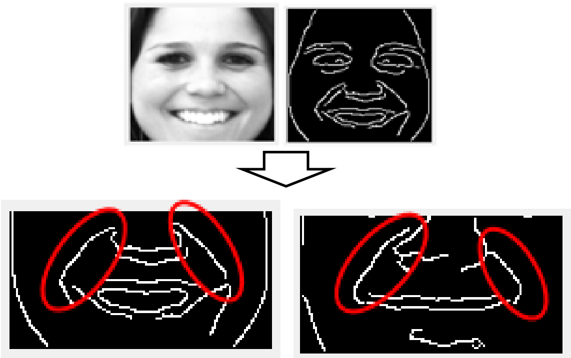
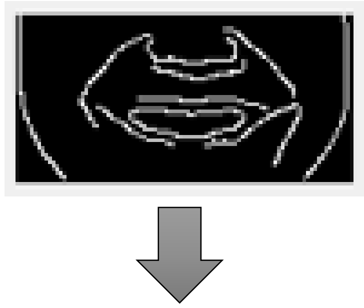


Figure 2 Nasolabial folds around the mouth

The matrix representation of the image fragments containing the nasolabial folds shows the values of 0 and 1, where the number 1 represents a white dot, and 0 represents a black dot. A matrix that represents image fragments presented in the Figure 3.



	1	2	3	4	5	6	7	8	9	10	...	120	121	122
1	0	1	0	0	0	0	0	0	0	0	..	0	0	0
2	0	1	0	0	0	0	0	0	0	0	..	0	0	0
3	0	1	0	0	0	0	0	0	0	0	..	0	0	0
4	0	1	0	0	0	0	0	0	0	0	..	1	0	0
5	0	1	0	0	0	0	0	0	0	0	..	1	0	0
6	0	1	0	0	0	0	0	0	0	0	..	1	0	0
7	0	1	0	0	0	0	0	0	0	0	..	1	0	0
8	0	1	0	0	0	0	0	0	0	0	..	1	0	0
	..	..	..	..	..	..	..	..	..	..	...	..	..	..
59	0	0	0	0	0	0	0	0	0	0	..	0	0	0
60	0	0	0	0	0	0	0	0	0	0	..	0	0	0
61	0	0	0	0	0	0	0	0	0	0	..	0	0	0

Figure 3 Matrix representation of nasolabial folds around the mouth

Therefore, the feature extraction process is carried out by finding the area of the nasolabial fold by counting the sum of pixels that have value of 1, using the formula (2).

$$S = \sum X_{rc} \quad (2)$$

where X_{rc} is a matrix element that have value of 1.

Table 1. Experimental results for optimal accuracy

Number Hidden layer	error Target	epoch	(%)
	1×10^{-6}		50
1	1×10^{-3}	100	50
	1×10^{-1}		50
	1×10^{-6}		92
2	1×10^{-3}	100	92
	1×10^{-1}		89
	1×10^{-6}		92
3	1×10^{-3}	100	92
	1×10^{-1}		91
	1×10^{-6}		91
4	1×10^{-3}	100	91
	1×10^{-1}		92

At the training stage, the artificial neural network is trained to analyze the types of facial expressions. The training data used is the result of calculations in feature extraction of 100 facial images data. With several attempts to get the optimal training accuracy which can be seen in table 1, the results obtained an accuracy of the training process is 92%.

Based on the test results in table 1, the best network architecture used is 1 unit in the input layer, 2 layers in the hidden layer and 2 classifications at the output layer. Regarding the two types of output, i.e., sad/neutral, and happy smile expression, the network uses the binary sigmoid (*logsig*) as activation function in the hidden layer and the output layer, and the target error is 0.000001 and the maximum epoch is 100.

Table 2. Several outputs of testing process

Images (Happy)	Nasolabial fold	area	Identification results
		**2451**	sad/ neutral
		462	sad/ neutral
		541	happy
		708	happy
		601	happy

The testing process was carried out on 50 test data for happy smile expression images. The test results of several images of recognizable and unrecognizable happy smile expressions are described in Table 2.

4.2. *Discussion*

Overall, the testing process provides the accuracy of 84%, meaning that the network was able to recognize 42 images as a happy smile expression from the 50 inputted images. From the test results, it can be known that the image of a happy smile expression that cannot be recognized by the network has the largest area and is 483, while the image of a happy smile expression that can be recognized by the network has the smallest area of 521 and the largest is 899. This means that the width of the nasolabial folds whose value is below 500 will provide incompatible results.

There are several possible factors that influence the success rate of the network in recognizing images of happy smile expressions, including the selection process at the image processing stage and in the network architecture creation process. Selection of optional image processing can affect the final image in the calculation for the identification/classification process. Determination of parameters in the network architecture, such as the number of hidden layers, error and epoch or other parameters, can also affect the network process in recognizing the input image.

5. CONCLUSION

Based on the description above, it can be concluded that nasolabial folds can be used as indicators/characteristics in determining the expression of a happy smile on facial images. The feature extraction process is carried out by calculating the area in the image of the nasolabial fold area. Feature extraction data are used as input data in the training process and testing of backpropagation neural networks. The results of the program provide an accuracy rate of 92% in the training process and in the testing process of 82%.

REFERENCES

[1] Z. Abidin, Pengembangan Sistem Pengenalan Ekspresi Wajah Menggunakan Jaringan Syaraf Tiruan Backpropagation (Studi Kasus Pada Database Mug). *Epsilon Jurnal Matematika Murni dan Terapan* Vol 5 no 1 : 2011, pp. 21-30.DOI: https://doi.org/10.20527/epsilon.v5i1.67

[2] P. Ekman, *Emotion Revealed*. New York: Henry Holt and Company, LLC, 2003.

[3] R.J. Mackley, An Evaluation of Smile Before and After Orthodontic Treatment. Angle Orthod, vol 63, 1993, pp. 183-190. DOI: 10.1043/0003-3219(1993)063<0183:AEOSBA>2.0.CO;2

[4] Abdurahman, Anatomi Senyum: Kajian Kinesiologi. Airlangga University Press. Surabaya, 2018.

[5] S. Kusmaryanto, Jaringan Saraf Tiruan Backpropagation untuk Pengenalan Wajah Metode Ekstraksi Fitur Berbasis Histogram. Jurnal EECCIS, [S.l.], v. 8, n. 2, p. pp.193-198, aug. 2014. ISSN 2460-8122. Available at: <https://jurnaleeccis.ub.ac.id/index.php/eeccis/article/view/283/246>.

[6] X. Zhao, S. Zhang, A review on facial expression recognition: feature extraction and classification. IETE Tech. Rev. 33, 2016, 505–517. DOI: https://doi.org/10.1080/02564602.2015.1117403.

[7] R Sigit, D. Pramadihanto, R. *Sulaiman* Sistem Pengenalan Ekspresi Wajah Secara Real Time. SNATI, ISBN: 979-756-061-6, 2005, pp. 27-30. Available at https://journal.uii.ac.id/Snati/article/view/1363

[8] P.R. Kalbuadi Analisis Pola Ekspresi Wajah Manusia. *Skripsi*. Universitas Negeri Yogyakarta. 2018

[9] J. Supardi, B. Anindita Pengembangan Arsitektur Jaringan Syaraf Tiruan untuk Pengenalan Ekspresi Wajah Manusia. Jurnal Geniric 10. 2015, http://repository.unsri.ac.id/id/eprint/9362

[10] P.A.U. Dewantari, M.A. Mutiara Dewi, G. Wirata, Aspek Anatomi Facial Exercise pada Nasolabial Fold, Bali Anatomy Journal (BAJ) 2019, Volume 2, Number 2: 45-50. DOI: https://doi.org/10.36675/baj.v2i2.29

[11] NN. Do facial exercises combat signs of aging? [Internet]. American Society for Aesthetic Plastic Surgery; 2019. Accessed on: December 10th 2019. http://www.surgery.org/consumers/plastic-surgery-news-briefs

[12] A.I. Galushkin, Neural Networks Theory Springer Berlin Heidelberg ISBN 978-3-540-48124-9. 2007.

[13] R.A Priyanka, C. Ashwitha., R. Arun Chakravarthi, R. Prakash, Face Recognition Model Using Back Propagation, Int. J. of Eng & Techn, 7 (3.34) (2018) 237-240. DOI: 10.14419/ijet.v7i3.34.18973