Multiple Smile Detection Using Histogram of Oriented Gradient and Support Vector Machine Methods

Multiple Smile Detection Using Histogram of Oriented Gradient and Support Vector Machine Methods

R Syaputra[1*] D Syamsuar[1] and E S Negara[1]
[1]Computer Science, Universitas Bina Darma Palembang, Indonesia

E-mail: rezkisyaputraharahap@gmail.com

Abstract. The face or the front of the head consists of the eyes, nose and mouth. Each face has its uniqueness from the many human faces. The face is used to show happy and sad expressions and feelings. Smiling also includes self-expression from others. The analysis of facial expressions plays a key role in analyzing human emotions and behavior. Smile detection is a specialized task in facial expression analysis with a variety of potential applications such as photo selection, user experience analysis, smile payments, and patient monitoring. Conventional approaches often extract low-level face descriptors and smile detection based on strong binary classifiers. In this paper, we propose an effective Histogram of Oriented method supporting vector engines for smile detection. The experimental results show that our proposed network outperforms the current state of the method. The test results using the Histogram of Oriented Gradient and Support Vector Machine method for smile detection were 87% for a precision value of 88% and a recall value of 83% and accuracy. In the future, we want to exploit some of the latest effective designs. we will try to update our mouth model so we can support a bigger head turn and face size scale.

1. Introduction

Common facial expressions in everyday life are smiles that express emotional feelings and underlie feelings such as joy, happiness, sadness, boredom, anger, satisfaction, and dissatisfaction. The need for smile monitoring is of importance and challenges the development and advancement in the real world to assist humans in various potential applications, such as photo selection, user analysis, payment using a smile, applications in the healthcare industry, application of artificial intelligence, and patient feeling conditions. Smile detection has already attracted the interest and attention of researchers, professionals in the development of this new science.

The efficient approach and development method used is the difference in data intensity between pixels in the image database using a simple grayscale feature. Then a figure database was adopted to select and combine the devices using a classification method based on pixels to form a robust classification for smile detection.

This study, representing a face using three artificial features, 1) Local Binary Pattern (LBP), 2) Local Phase Quantization (LPQ), 3) Histogram of Oriented Gradient (HOG). Use the gradient orientation histogram method to represent image data (faces) and to classify image data (faces) using the support vector machine method.

In recent years, researchers have witnessed many great technological achievements in computer science. The number of smart technologies that have been successfully created by researchers and professionals and combine computer science and other sciences that provide concrete examples such

as IBM Watson, Health robots, Google AI driverless cars, and delivery systems using drones on Amazon developed in advances in computer technology. In the advancement of computer technology, it is necessary to have in-depth learning about computer technology that can help humans, especially using the Convolutional Neural Network (CNN) method, this method contributes to the development of this technology. Convolutional neural networks are considered to be the deepest and most effective neural network architectures for data classification such as images. Researchers have started implementing a convolutional neural network to detect smiles. a convolutional neural network that exists to extract higher-level features from an image database (faces). Uses a classification support vector machine to detect smiles for classification tasks. This study aims to develop and apply artificial intelligence-based methods to recognize the object of a smile. The development of this research is very useful for detecting the object of a smile that can be used in various ways.

2. Literature Review

2.1. Smile

Smile is a facial expression to show someone's feeling's whether someone is sad, happy and so forth. There are two types of smile, one social smile is a voluntary facial expression that is intentionally made like when getting acquainted with others, making a passport photo, two spontaneous smiles are facial expressions that are accidentally made, natural and caused by emotional overflow. Smiles in each person are different where the shape of each person's smile is not same but the smile has the same pattern, where the pattern of the right and left edge lip is higer than the middle lib. There are some extraction features for detecting smiles such as haar-like feature extraction over gray colors and standardized positions such as hear done by pinky and manihs, histogram equalization refers more to the color of each block done by eduard and ivanna, features gabor extraction is more directed toward frequency and the edge detections extraction feature which is more directed towards the outline of a shape is done by eduard in the same year.

The use of haar-like features, histogram equalization and edge detection is susceptible to light interference even though it has a pretty good shape recognition result, which is for haar-like ti get 85% accuracy which is better than histogram equalization which has an accuracy of 50% and gabor 81,6%, while the use frequencies allows misleading to identify features. While the hog feature has a better shape recognition ability which is shown by aryuanto therefore this study uses hog ectraction feature which aims to recognize the shape of a smile.

2.2. Support Vector Machine

Support Vector Machine is looking for the best hyperplane that functions as a separator of two data classes to form a model that will be used in the classification process or Support Vector Machine (SVM) is a binary classification method that utilizes a hyperplane equation that divides the two classes like (Figure 2). The two label classes or y in SVM are $\in \{-1,1\}$. Data xd is a feature vector with d = 1,2, ..., Nd dimensions. For d = 2 the hyperplane is a line, while for d = 3 it is a plane. Hyperplane equation is a classification equation shown by equation [12] with the classification parameter is the value w is the weight value and b is biased.

$$f_{svm}(x) = w.x + b$$

To find the values of w and b, equations (13) and (14) are used

$$b = -\frac{1}{2} w.x^{+} + w.x^{-}$$

$$w = \sum_{i=1}^{N} a_i \, y_i \, x_i$$

N: the amount of data
α: coefficient that functions or is used to find weight values
y: a label for each lip
x: is the feature vector of the HOG.

Support Vector Machine is also used to manage datasets by utilizing kernel tricks to find hyperplanes, one of which is the Linear Kernel shown by equation in [13]-[15].

$$K(x_i, x_j) = x_i, x_j$$

Place the Linear Kernel Support Vector Machine Visualization image From the kernel results obtained, Support Vector Machine makes a classification equation that is adjusted to the kernel used, such as equation (16) with class classification values based on equation [16]-[17].

$$f_{svm}(x) = \sum \alpha_i \, y_i \, K(x_i, x) + b$$

$$class = \begin{cases} 1, f_{svm}(\mathrm{X}) \geq 0 \\ -1, selanjutnya \end{cases}$$

The determination of the value of α_i is done by solving the LD (Lagrange Multiplier Duality) optimization problem, which is shown by equation (18) and by the formula in the equation.

$$Ld = \max_{\alpha i \, \epsilon \, R} \sum\nolimits_{i \epsilon N} \alpha_i \sum\nolimits_{i \epsilon N} \alpha_i \, \alpha_j y_i \, y_j$$

Which has two conditions 0 ≤ α_i ≤C and ∑_ (i∈N) $\sum_{i \epsilon N} \alpha_i \; y_i = 0$where C is the constant value with, Ld (x_i), the kernel function used.

2.3. Histogram of Oriented Gradient (HOG)

Histogram of Oriented Gradient (HOG) is the extraction of features used in computer vision and image processing by calculating the gradient value in an image to get the results that will be used to detect the objects. Histogram of Oriented Gradient is a form of local object and a value that is used from the gradient intensity. Histogram of Oriented Gradient (HOG) is the extraction of features used in image processing computers by calculating the gradient value in an image to get the results that will be used to recognize the characteristics of the object.

The histogram of Oriented Gradient is the shape of the object's locale and the value used for the intensity of the gradient. The process in using Histogram of Oriented is to look for the gradient orientation and vertical gradient values and then look for the magnitude and binning orientation values of the original image size then divide the image into several blocks that have a size of 2x2 which later in the block there are several cells with a size of 8x8 which have a gradient orientation 9 bin so that the feature vector is obtained, each cell block has a direction gradient value or an edge orientation value for each image pixel. In improving the performance of the resulting gradient values carry out a binning orientation by normalizing in contrast, then this value is used to describe each block of the results of the normalized value. The first stage of the Histogram of Oriented process is to find the value of the gradient calculation using equation [1][2].

$$g_x(r,c) = g(r,c+1) - g(r,c-1)$$
$$g_y(r,c) = g(r+1,c) - g(r-1,c)$$

Or use the Central Difference Gradient Operator equation:

$$g_x = [-1\ 0\ 1]$$
$$g_y \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Where g = gradient r = row c = column

After the x and y values have been obtained, the next process is to find the magnitude μ value to find out how big the gradient is and look for the value bin binning orientation using equations [3] and [4].

$$\mu = \sqrt{g_x^2 + g_y^2}$$

$$\theta = \frac{180}{\pi} (tan_2 {}^{-1} (ly, lx)\ mod\ \pi$$

After getting the θ value of each cell in the block, then look for the histogram for each cell for bin j, Cid a where the value of B is determined from the number of bins, namely [9] bin using equations [5] and [6].

$$bin\ j = \left[\frac{\theta}{w} - \frac{1}{2}\right] mod\ B$$
$$w = \frac{180}{B}$$

Then the value obtained does vote bin which is divided into two types namely vote bin Vj indicated by equation [7] and vote bin (+ 1) indicated by equation [8] and the boundary of the bin is shown in equation.

$$Vj = \mu \frac{c_{j+1^{-\theta}}}{w}$$
$$V_{j+1} = \mu \frac{\theta^{w} - c_j}{w}$$

With Wi-Fi Limits, W (i + 1)

$$C_i = W\left(i + \frac{1}{2}\right)$$

Later, the temple is looking for block distribution. If a vote has been obtained from each bin, then normalize the block which is used to get the b value of each block to reduce the change in contrast between blocks of objects to find block normalization using equation [10].

$$b \leftarrow \frac{-b}{\sqrt{||b||2_{-\varepsilon}}}$$

After getting the block value, the next step is to look for the Histogram of Oriented feature where the normalized blocks are combined into one Histogram of Oriented feature vector using equation [11].

$$h \leftarrow \frac{-h}{\sqrt{||h||2_{-\varepsilon}}}$$

The process in applying hog is to divide the image into blocks that later in the block there are image. In increasing the performance of the resulting gradient value, the binning orientation is done by normalizing in contrast, then this value is used to describe each block from the normalized value. Following the implematation of Histogram of Oriented can be seen in Fig 1.

Figure 1. . Implementation of Histogram of Oriented

In fig 1, the steps of Histogram of Oriented implementation are starting from gradient calculation, binning orientation, then block descriptor by Dalal, Trig and Derry, while the next stage in the Histogram of Oriented process is finding the gradient calculation value first to get value of gradient calculation using equation [1][2].

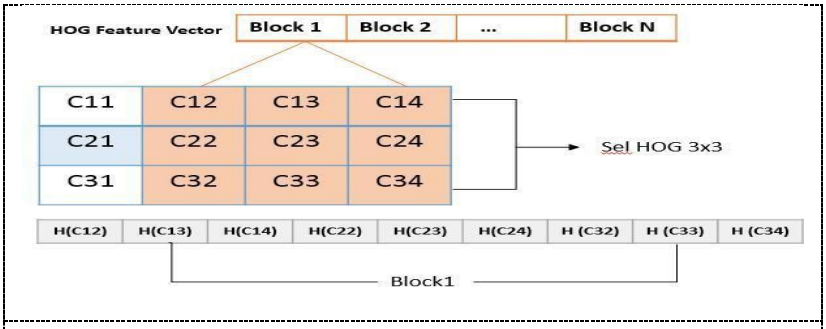

Figure 2. Division of Block for Each Cell

Fig 2 is a block division scheme for each cell, where this value will be used as a block val the probability of H (prior), P (X | H) is the probability of X in the hypothesis H, P (X) is the probability of X.

3. Results and Discussion

3.1. Framework

At the initial stage, the face image performs a grayscale process to speed up the duration of the program process and then continues at the image extraction stage Histogram of Oriented uses the existing functions in python and the results of the image extraction enter the Support Vector Machine training stage, this model is to load Support Vector Machine structs and groups. After the Support Vector Machine model training process is complete, the Support Vector Machine model data has been obtained and can be used to test data. Here are some steps are taken to identify multiple smiles using the Histogram of Oriented (HOG) Method and Support Vector Machine (SVM):

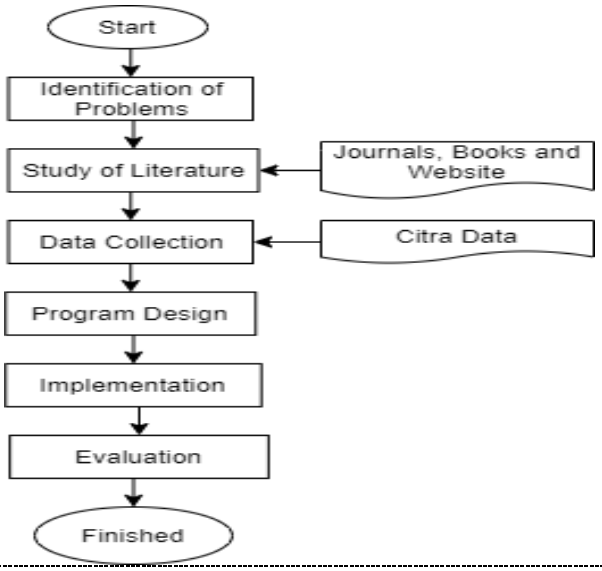

Figure 3. Framework

Given a set of pictures of a person into the system, it compares these images and determines which one has the best smile. Images can be from the same person or from several different individuals. To identify the subject's face in the photo using and the same algorithm is used to locate the subject's mouth. Furthermore, the Shi-Tomasi angle detection algorithm is carried out through the mouth area; it locates the edges and features of the mouth (smile wrinkles, teeth, mouth shape). It uses the points obtained from the detection angle; the most suitable second-degree polynomial line is drawn. By taking the derivative of the most suitable line it determines the concave point, and from there one can detect whether the subject is smiling or not.

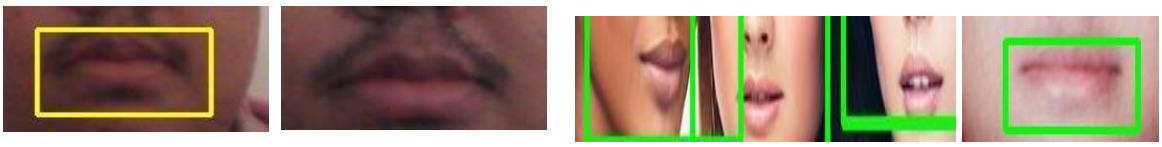

Figure 4. Smile Detection results

Our methods and algorithms are implemented using Python. The researcher tested the algorithm using facial images from the database. The database contains color images of the front faces with different lighting, expressions and backgrounds. From the testing process, the researcher found that our proposed method shows the best detection rate, both lip detection rate and smile detection rate. This is because some of the lip areas obtained with the mouth detector are too small (they do not cover allof the mouth/lips). So that a smile cannot be detected properly.

4. Conclusion

This study uses a gradient-oriented histogram method and a supporting vector machine to assist in handling smiles. Previous research used convolutional neural networks. Several previous studies used shallow convolutional neural networks with few hidden layers. Apart from that, this research also

shows inefficiency and data argumentation techniques (figure) in overcoming lack of data to avoid overfitting.

By detecting the characteristics and the angle of detection, we can detect whether the person is smiling or not in a photo with Face Detection. This Automatic Smile Identification has many potential applications, including enhanced camera functionality for advanced functionality. The test results using the Histogram of Oriented Gradient and Support Vector Machine method for smile detection were 87% for a precision value of 88% and a recall value of 83% and accuracy. In the future researchers want to develop some of the model more effective. we will try to update our mouth model so we can support a bigger head turn and face size scale.

References

[1] L. An, S. Yang, and B. Bhanu. Effecient smile detection by extreme learning machine. *Neurocomputingi*, 149:354-363, 2015.

[2] V. Jain and J. L. Crowley. Smile detection using multiscale gaussian derivatives. In 12th WSEAS *International Conference on Signal Processing, Robotics and Automation*, 2013.

[3] C. Junkai, O. Qihao, C. Zheru, and F. Hong. Smile detection in the wild with deep convolutional neural network. *Machine Vision and Applications*, 28(1):173–183, 2017.

[4] M. Liu, S. Li, S. Shan, and X. Chen. Enhancing recognition in the wild with unlabeled referencedata. *In Asian Conference on Computer Vision*, pages 577–588. Springer, 2012.

[5] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *J. Edik Inform.*, vol. 2, no. 2, p. 7, 2017.

[6] C. Shan. Smile detection by boosting pixel differences. *IEEE transactions on image processing*, 21(1):431–436, 2012

[7] Imagenet large scale visual recognition challenge 2014 (ilsvrc2014).*http://www.imagenet.org/challenges / LSVRC/2014/*. Accessed: 2017-06-15.

[8] M. B. Pranoto dan K. N. Ramadhani, "Face Detection System Menggunakan Metode Histogram of Oriented Gradients (HOG) dan Support Vector Machine (SVM) Face Dtection System using Histogram of Oriented Gradients (HOG) Method amd Support Vector Machine (SVM)," *e-Proceeding Eng.*, vol. 4, no. 3, hal. 5038–5045, 2017.

[9] D. Amputri, S. Nadra, G. Gasim, dan M. E. Al Rivan, "Perbandingan jarak potret dan resolusi kamera pada tingkat akurasi pengenalan angka kwh metermenggunakan svm," *J. Inform. Glob.*, vol. 8, no. 1, hal. 7–12, 2017.

[10] L. Cunhe and W. Chenggang, "A new semi- supervised support vector machine learning algorithm based on active learning," in *2010 2nd International Conference on Future Computer and Communication*, Wuhan, China, 2010, pp. V3-638-V3-641, doi: 10.1109/ICFCC.2010.5497471.

[11] M. Hofmann, "Support Vector Machines — Kernels and the Kernel Trick," *Notes*, vol. 26, no. 3, p. 16, Jun. 2006.

[12] C. Manning, P. Raghavan, and H. Schuetze, "*Introduction to Information Retrieval*," p. 581, Apr. 2009.

[13] T. Sutabri, A. Suryatno, D. Setiadi, and E. S. Negara, "Improving Naïve Bayes in Sentiment Analysis For Hotel Industry in Indonesia," in *2018 Third International Conference on Informatics and Computing (ICIC)*, Palembang, Indonesia, Oct. 2018, pp. 1–6, doi: 10.1109/IAC.2018.8780444.

[14] M. Raza, F. K. Hussain, O. K. Hussain, M. Zhao, and Z. ur Rehman, "A comparative analysis of machine learning models for quality pillar assessment of SaaS services by multi-class text classification of users' reviews," *Future Gener. Comput. Syst.*, vol. 101,

pp. 341–371 Dec.2019, doi: 10.1016/j.future.2019.06.022.

[15] C. Darujati and A. B. Gumelar, "Pemanfaatan Teknik Supervised Untuk Klasifikasi Teks Bahasa Indonesia," *J. Link*, vol. 16, no. 1, p. 8, Feb. 2012.

[16] K. Zhang, Y. Huang, H. Wu, and L. Wang. Facial smile detection based on deep learning features. *In Pattern Recognition (ACPR), 2015 3rd IAPR Asian Conference* on, pages 534–538. IEEE, 2015.

[17] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–9, 2015.