Acne Classification with Gaussian Mixture Model based on Texture Features

Alfa Nadhya Maimanah, Wahyono*, Faizal Makhrus
Department of Computer Science and Electronics, Universitas Gadjah Mada, Yogyakarta, Indonesia

***Abstract*—This paper presents an acne detection method on face images using a Gaussian Mixture Model (GMM). First, the skin area in the face image is segmented based on color information using the GMM. Second, the candidates of the acne region are then extracted using a Laplacian of Gaussian-based blob detection strategy. Then, texture features are extracted from acne candidates using either a Gabor Filter or Gray Level Co-occurrence Matrix (GLCM). Lastly, these features are then utilized as input in the GMM for verifying whether these regions are acne or not. In our experiment, the proposed method was evaluated using face images from ACNE04 dataset. Based on the experiment, it is found that the best classification results were obtained when GLCM features in the Cr-YCbCr channel are applied. In addition, the proposed method has competitive performance compared to K-Nearest Neighbor (KNN).**

Keywords—Acne; GLCM; Gabor filter; Gaussian mixture model

I. Introduction

In recent years, the beauty technology industry has experienced significant growth along with increasing public enthusiasm and awareness of skin health. One major concern of skin problems is *acne vulgaris*. Acne occurs when sebum is trapped in the hair follicles and attacked by *Propionibacterium acnes* [1]. Adolescents to adults may experience this problem. It could leave scars that influence the level of confidence of some people. A common method to determine the acne severity is through manual counting by a dermatologist. The process is susceptible to subjectivity factors, both inter-observers—the same patient has different assessments by different dermatologists—and intra-observer—the same patient has different assessments by the same dermatologist on different days [1]. In addition, the technique is not effective in terms of time and effort spent by a dermatologist [2]. Therefore, technology is needed to assist the process of acne severity assessment through face images.

There have been several developed technologies which are employed to assess facial skin. In addition to helping determine the appropriate and accurate type of treatment, the presence of technology can also attract the attention of customers. The types of technology developed are quite diverse, ranging from instruments equipped with skin analysis systems such as VISIA, Internet of Things (IoT) such as LG U+ LTE Magic Mirror, mobile-based applications such as TroveSkin, and Skin Genius by L'Oréal Paris, as well as Software as a Service (SaaS) such as Haut.AI. Facial skin analysis models are developed based on Artificial Intelligence (AI) which is trained using the data, the company had collected.

This research proposes a study of comparison of texture-based feature extraction methods for assessing the acne severity on human face images. To the best of our knowledge, the research related to acne images generally aims to segment and/or classify acne types. However, research that aims to determine the severity of acne suffered by patients is still limited. Research [3] contributed to determining the severity of acne, but it was still based on the area of acne on the right and left cheeks without using appropriate standards. Since there are still few studies on acne images with standardized acne severity, this research used the criteria formulated by Hayashi [4], which was also utilized in the dataset developed by Wu et al [5]. This criterion estimates the acne severity based on the number of papules and pustules detected on a face image captured with an angle of 70° from the front side.

Furthermore, this research utilizes texture-based features, namely GLCM and Gabor Filter. GLCM was chosen because it was inspired by research related to the acne types classification conducted by Ramadhani [6] which achieved an accuracy value of 72%. In addition, research conducted by Chang and Lio [7] successfully detected acne on face images using GLCM features with accuracy of 99.40%. It was proved that GLCM has the ability to extract the features needed for acne detection. On the other hand, Gabor Filter was chosen because research conducted by Jeon and Cheoi [8] could successfully detect abnormal areas on skin images, including small acne and regions with low contrast levels, which open the possibility of implementing this method for acne detection. To classify acne and non-acne areas, this research conducted an experiment to use the Gaussian Mixture Model (GMM) method. This method is usually used as a density estimator so it is suitable for clustering problems. In addition, the GMM has also been proved as a good approach for classification which was implemented by Dey [9] for skin classification as well as Wan et al. [10] for classifying 10 kinds of datasets from the UCI Machine Learning Repository.

The rest of this paper is organized as follows. Section II discusses the research related to this topic. Section III explains the details of the proposed method. Section IV demonstrates the experimental result. Conclusions are given in Section V.

II. Literature Review

Based on the feature extraction method, the research related to acne images is generally divided into two categories: hand-crafted and deep learning-based feature extractions. The features from the hand-crafted strategies can be obtained based on color, shape, and texture [11]. A comprehensive experiment and observation are required to

*Corresponding Author.

determine the best features as the basis for acne detection, unlike deep learning where the model can extract its own features automatically.

Deep learning is the latest research trend regarding images with acne objects. Zhao et al. [12] used a regression model with transfer learning on ResNet-152 to determine the acne severity but it did not achieve a good performance since the data were imbalanced, which is dominated by one class. Arifianto and Muhimmah [13] used transfer learning on ResNet-50 to detect acne and obtained an accuracy of 63.2%. Both studies faced the problem of limited data with good quality which resulted in low accuracy. Junayed et al. [14] also utilized deep learning for classifying five classes of acne with accuracy over 94%, but it required an expensive computation time.

On the other hand, the hand-crafted feature extraction based on either color or texture was also implemented. Acne color features tend to be inconsistent with variations in lighting and skin color on the same type of acne as in [15] and [16]. With the input of the cropped acne area from the face image, the author [6] utilizes feature textures for classification of acne types. Authors in [7] extracted Gray Level Co-occurrence Matrix (GLCM) statistical features on a whole face image which was then divided into blocks. However, the image acquisition process required a special device with standard camera settings, lighting, and shooting distance parameters that had been determined beforehand. It was less flexible to be implemented as a mobile application only has a smartphone camera available. Nevertheless, the hand-crafted approach could achieve a high accuracy of 99.40% with less computation time compared to deep learning approach. Therefore, this research would be conducted to study more on the utilization of hand-crafted texture-based features to detect acne on face images.

The input images used in previous studies related to acne recognition were quite diverse. Some studies used whole human face images with various levels of acne severity, as in the research [12] and [13], while other studies used cropped images on the acne object only, such as [14] which used a dataset from Dermnet consisting of five types of acne.

Furthermore, research related to acne images can be grouped into three objectives: acne segmentation, acne classification, and acne severity assessment. Research conducted by Maroni et al. [2] detected and counted the number of acnes with a sequence of processes starting from body part detection, skin segmentation, heatmap creation, acne extraction, and blob detection. The best skin segmentation was generated by the Random Forest method based on the 15 most informative features explored. Acne extraction was carried out using the adaptive thresholding of heatmap images. After that, acne was detected and counted using the Laplacian of Gaussian (LoG) filter. Research conducted by Jeon and Cheoi [8] aimed to cluster abnormal areas on human skin using the density-based spatial clustering of applications with noise (DBSCAN) algorithm based on the features obtained from the Gabor Filter. In their study, the skin segmentation was not carried out first even though there was an image input in the form of a face. The proposed method had better performance than the [17] method because it could detect small acne and regions with low contrast levels.

Lastly, some researchers classified acne based on texture features. Ramadhani [6] utilized GLCM to obtain texture characteristics including contrast, energy, entropy, correlations, and dissimilarity. Among these characteristics, entropy was the most influential statistical feature since it represents texture irregularities. The overall accuracy value obtained was 72%. Some studies classified acne based on color features, like a research by Darmawan et al. [15] where RGB color intensity was used. Although it had been able to detect types of acne, there were still limitations due to lighting factors during image acquisition resulting in color values discrepancy and accuracy. Gunawan et al. [16] conducted segmentation using Region Growing and classified the acne types using Self Organizing Map (SOM). RGB histogram feature was used as SOM input. The classification accuracy was still not ideal due to acne color variations influenced by diversity of skin color and lighting conditions in each image. Several studies used the segmentation results to classify the types of acne, like in Arora and Sarvani [17] who explored the methods of acne segmentation and machine learning models for acne classification. Compared to color and texture segmentation, the 2-level K-means clustering had the best accuracy at 70%. While the classification of acne and acne scars had an average accuracy of 80% using the Fuzzy C-Means (FCM).

III. METHODS

A. Classification Model Development

A total of 40 images from the ACNE04 dataset—10 images from each Hayashi Criteria—are used for evaluation. These images have non-uniform dimensions therefore they are resized to 320×320. The bounding box coordinates in the Extensible Markup Language (XML) annotation document are also adjusted to the same dimension which is then used for image cropping to obtain acne blocks. On the other hand, for non-acne blocks, a whole image is divided into a uniform block size of 20×20 which is then curated manually to remove the ones with acne. In total, there are 908 acne blocks and 870 non-acne blocks.

The first texture features are extracted from GLCM. The distance and angles chosen are one pixel and (0°, 45°, 90°, and 135°), respectively. For comparison purposes, this feature extraction is carried out on four different channels of color space, i.e. Grayscale, Hue-HSV, Red-RGB, and Cr-YCrCb. There are six features calculated from the GLCM including contrast, dissimilarity, correlation, energy, homogeneity, and ASM, each of which was the average of the four neighboring angles. These features are then normalized to the range of [0,1]. Finally, there will be six GLCM features obtained from each block. The second texture features are extracted from Gabor filtered images. The two-dimensional Gabor filter is a Gaussian kernel function modulated by a sinusoidal wave [18]. The Gabor filter bank is created with the size of 3×3 by adjusting the five variables in formula 1. The first three variables followed the research by [8]: $\Psi=\pi/2$; $\lambda=0.8$; $\gamma=0$. For comparison purposes, the last two variables are set on two different configurations: $\theta=0°$, 30°, 45°, 60°, 90°, 120°, 135°, 150°; σ=1 and θ= 0°, 45°, 90°, 135°; σ=1,2. The visualization of the Gabor filters used is displayed in Fig. 1. After the filtering process, two statistical features, namely, mean and variance are calculated. Finally, there will be 16 Gabor features obtained from each block. After both features are gathered, they are trained for acne and non-acne classification using GMM.

$$g(x, y; \Psi, \lambda, \gamma, \theta, \sigma) = exp\left(\frac{-x'^2+\gamma^2 y'^2}{2\sigma^2}\right) cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (1)$$

The GMM classifier performs acne and non-acne classification using two GMMs, one model for acne features and the other for non-acne features. After training those models, a new testing feature is classified as acne if the log-likelihood value in acne GMM is higher than in non-acne GMM. To determine the optimal number of Gaussian components in the two GMMs, experiments are carried out on the number of Gaussian components in the range of 2-20 so that the lowest Bayesian Information Criterion (BIC) value could be obtained. Below is the pseudocode to create a GMM Classifier.

Input: GLCM/Gabor Features, K_{max} = 20, and C = 2 (acne and non-acne)
Output: K_{opt} for every class c
for c=0:C do
    for k=2: K_{max} do
        Apply GMM−EM with k number of Gaussian components;
        Calculate BIC based on maximum parameters obtained;
    end for
    K_{opt} = arg min k (BIC)
end for

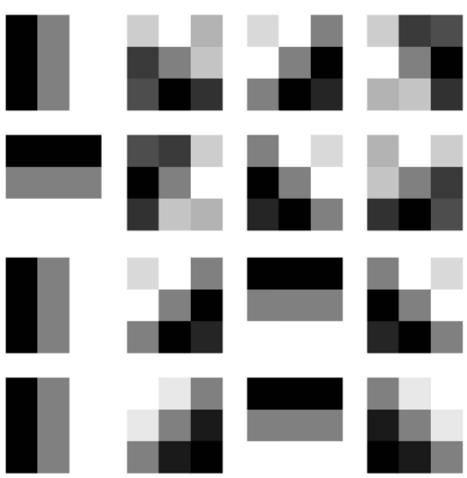

Fig. 1. Gabor Filter Bank: 8 Angles-1 Standard Deviation (Row 1 and 2), 4 Angles-2 Standard Deviations (Row 3 and 4)

B. Skin Segmentation

To evaluate the classification model, skin segmentation steps are implemented as shown in Fig. 2. A total of 20 images from the ACNE04 dataset—five images from each Hayashi Criteria—are used for testing. These images are also resized to 320×320 dimensions. Then an enhancement process is carried out on the a* CIELab channel. The reason behind this is that the a* channel represents the level of pixel redness that is independent of lighting, making it stronger for detecting acne that tends to be redder than the surrounding skin. The enhancement started with the unsharp mask to sharpen the image after going through the resizing step. Then, the Difference of Gaussian (DoG) process is applied by subtracting the unsharp masked image from the Gaussian Blurred image following the research conducted by [19]. The resulting image highlights the pixels that tended to be red. The Gaussian Blur filter parameters use kernel size of 19×19 and σ=13. The next step is skin segmentation. For comparison purposes, two skin segmentation methods are used. The first is GMM segmentation based on BGR skin and nonskin pixel values from the Skin Segmentation Dataset, UCI Machine Learning Repository. The second method is Otsu Thresholding on the median blurred Cr image. This channel is selected since it is a red chromatic channel that could help the process of blurring pixels whose intensity did not resemble the skin—in which the color tends to be reddish. The size of the median blur filter is 21×21. The result of Otsu Thresholding is then used as the mask on the resulting image from a* CIELab enhancement step. Face parts such as the mouth, right eye, left eye, right eyebrow, and left eyebrow could be susceptible to being misdetected as acne, especially those with similar features in terms of color intensity, such as the mouth. Therefore masking is done on these parts. Face parts detection was done with the help of the DLIB library. The resulting image is then being used as a mask. The next step is to do acne candidate thresholding on the already masked a* CIELab enhanced image with the threshold value of 128. To remove the noise, a morphological opening operation is performed with a kernel size of 3×3.

C. Acne Detection

To calculate the number of acne candidates, the blob detection method is used with the Laplacian of Gaussian (LoG) kernel. The parameters used include the minimum standard deviation of the Gaussian kernel = 1, the maximum standard deviation of the Gaussian kernel = 5, the number of intermediate values, standard deviation = 15, threshold = 0.2, and overlap = 0.1. Each detected blob stored the coordinates of the blob's center (x,y) and its radius. They are used to determine the bounding box coordinates of the acne candidates which are then used for the cropping process. The cropping process is carried out by taking the pixels in the bounding box's coordinates range. The cropped pixel blocks are stored for texture features calculation. The features obtained from all blocks of acne candidates are used to predict the acne classification using the GMM Classifier model that had been trained previously. To reduce the number of overlapping bounding boxes while recognizing the same object, the Non-Maximum Suppression (NMS) process is carried out. The output image is the original image with a bounding box on each of the detected acne. In addition, the text of the number of detected acne and the severity are also displayed.

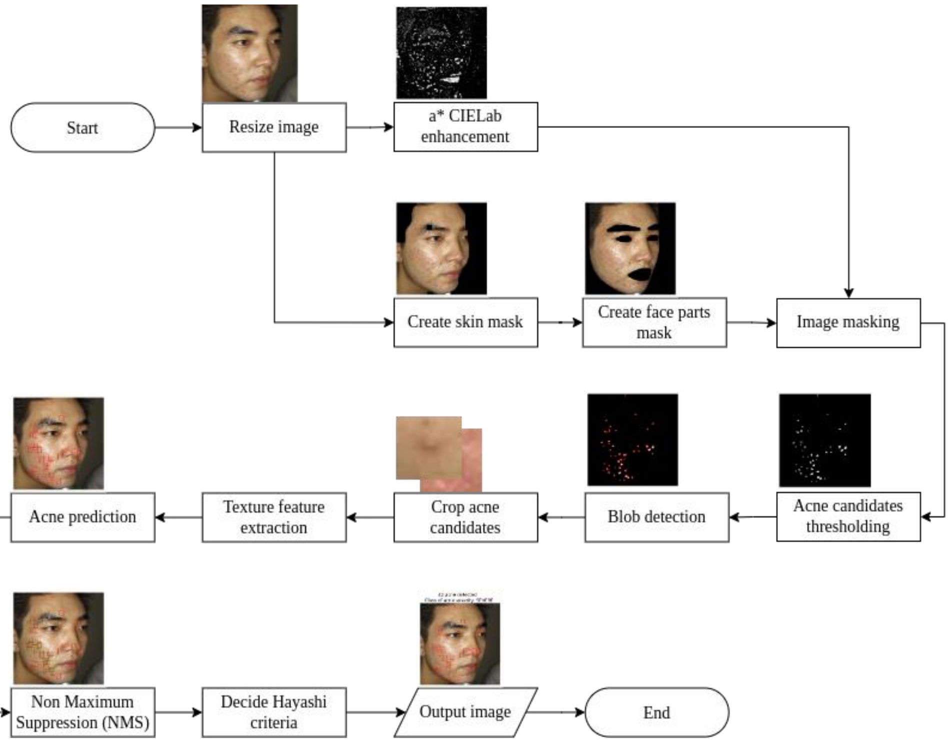

Fig. 2. Flowchart of the Proposed System.

IV. Experimental Results and Discussions

A. Skin Segmentation Result

Skin segmentation using the GMM method can separate facial skin areas in some images although hair, background, and clothing areas are still visible. Based on observations, these non-facial areas occur because of the morphological operation process carried out to remove noise and holes from the initial segmentation results. Several kernel sizes (3×3,5×5,7×7), the number of iterations (1,2,5,7), and the order of operations morphology are tried, however the obtained results still include 20-30% of non-facial regions. Fig. 3 on the left is the result of skin segmentation using the GMM method by performing closing and dilation operations using a kernel size of 5×5 and iterating 7 and 2 times respectively. Fig. 3 on the right is the result of skin segmentation using Otsu Thresholding on Median Blurred Cr Image which is better than the GMM method based on the minimum visible hair area, almost no visible background area, and minimum visible clothing area factors. The computation time of this second method is faster with an average computation time of 0.005 seconds, compared to the GMM method with an average computation time of 0.419 seconds. Therefore, based on the quality of the segmentation results and computational time, the experiments use Otsu Thresholding on Median Blurred Cr Image.

B. Classification Result with GLCM Features

Acne classification using the GMM method with the input of GLCM features is shown in Table I. In addition, Table II presents the details of Table I. Viewed from the prediction accuracy, all channels have accuracies below 55%. The highest total performance is obtained by Red-RGB channel based on the validation and test accuracies and the second place is Grayscale channel. Meanwhile, viewed from the precision value, all channels have relatively low precision in the range of 0.45-0.52, which means that there are still quite a lot of false positives. On the other hand, based on the recall value, Cr-YCrCb channel has the highest value at 0.78 that it becomes the best among the other three channels in correctly predicting acne close to the ground truth. Cr-YCrCb also has the best F1-Score value of 0.56.

Fig. 3. Example of Skin Segmentation Result with GMM Segmentation (Left) and Otsu Thresholding on Median Blurred Cr Image (Right)

TABLE I. Comparison of Testing and Prediction Accuracy with GLCM Features

Channel	K of non-acne	K of acne	Val Acc	Test Acc
Grayscale	10	7	95.22%	48.89%
Red-RGB	6	5	94.38%	53.20%
Cr-YCrCb	11	12	87.92%	45.14%
Hue-HSV	11	15	71.34%	41.75%

TABLE II. MODEL EVALUATION WITH GLCM FEATURES

Channel	Acc	Prec	Rec	F1	Time	IoU	Hayashi Correct
Grayscale	48.89%	0.52	0.58	0.50	1.228	0.466	10
Red-RGB	53.20%	0.53	0.53	0.49	1.209	0.442	10
Cr-YCrCb	45.14%	0.50	0.78	0.56	1.260	0.463	11
Hue-HSV	41.75%	0.45	0.70	0.52	1.236	0.452	12

Based on the number of images where the Hayashi class [4] is determined correctly, Hue-HSV is the best with 12 images. Having the recall value of 0.70, Hue-HSV channel performs below the Cr-YCrCb because more acne was still predicted as non-acne (it has more false negatives). In terms of computation time, the four channels do not differ much in the range of 1.2 seconds since the computational load tends to be the same as seen from the number of calculated features and steps. Based on the Intersection over Union (IoU) values [20], all four channels are above 0.4 which mean that they are quite good at detecting the location of acne. Nonetheless, since the size of the ground truth bounding boxes varies while the size of the prediction bounding boxes is uniform, the IoU value is difficult to approach 1. From all the experiment results, Cr-YCrCb is chosen as the best performance of GMM classification since it has the highest recall value of 0.78 although its accuracy value is low—in third place. An example of the output images can be seen in Fig. 4. The recall metric is chosen since in this problem, it is more important to correctly identify positive acne (the fewer false negatives are the better). Cr-YCrCb still has many errors in detecting non-acne, as indicated by the high number of false positives (low accuracy). Although in this case, this error is not life-threatening, it is much better to reduce it.

C. Classification Result with Gabor Features

Table III shows the comparison of validation accuracy and testing accuracy of GMM classification using Gabor features. The detailed evaluation results are also shown in Table IV. Viewed from the prediction accuracy, all filters do not achieve high accuracy. The highest was obtained by 4 degrees-2 standard deviations (4deg2sd) filter variation at 55.43%. Meanwhile, viewed from the precision value, both still have low values around 0.5 which mean that there are still many false positives. Compared to the recall these two filters do not differ too much around 0.5. Therefore, they are not good at predicting the correct acne. In addition, since both precision and recall from those two filters have small differences then the F1-Score values differ by only 0.02.

These two filters have the same number of images where the Hayashi class is determined correctly. It means that even though the number of detected acne is close to ground truth, the false negative predictions are still a lot. In terms of the computational time, they are both in the range of 0.4 seconds. The reason for the computational load which tends to be similar is because of the similarity in the number of features and steps. Based on the IoU value, the results show that both are above 0.4 which means that they are quite good at detecting the location of acne. Nevertheless, since the size of the ground truth bounding boxes varies while the size of the prediction bounding boxes is uniform; hence the IoU value is difficult to approach 1. Based on these two experiments, the 8 degrees-1 standard deviation (8deg1sd) filter parameter is chosen as the best filter since it has the highest recall value of 0.54.

TABLE III. COMPARISON OF TESTING AND PREDICTION ACCURACY WITH GABOR FEATURES

Filter	K of non-acne	K of acne	Val Acc	Test Acc
4deg2sd	3	3	67.69%	55.43%
8deg1sd	3	3	67.82%	54.27%

TABLE IV. MODEL EVALUATION WITH GABOR FEATURES

Filter	Acc	Prec	Rec	F1	Time	IoU	Hayashi Correct
4deg2sd	55.43%	0.52	0.51	0.48	0.449	0.446	10
8deg1sd	54.27%	0.54	0.54	0.50	0.472	0.411	10

Fig. 4. Example of the Output Images from the Proposed Method (Detected Acnes Represented by Red Bounding Boxes).

D. GLCM Features vs Gabor Filter Features

An important factor that affects the performance of the model is the quality and ability of the features to represent acne and non-acne characteristics. Viewed from the number of features, Gabor Filter [8] has 16 features while the GLCM has 6 features (initially there are 24, but then only the average value of all neighboring directions of each feature is used [7]). Having fewer features, the GLCM has a generally better recall value than the Gabor Filter although they tend to have lower accuracy. Therefore it is concluded that the GLCM features tend to be better at representing acne and non-acne objects even though in terms of computation time it is longer than the Gabor features. Models with GLCM features require computation time around 1.2 seconds while Gabor features only 0.4 seconds. Another factor that affects the model performance is the classification class which is limited to only two classes, namely acne and non-acne. Based on observations, the model detected several blocks as positive with characteristics close to acne but they are not, such as acne scars, spots, moles, and image lighting that make the skin tend to look red. This causes the high number of false positives in the prediction.

E. Color vs Texture

To determine the performance of the texture features in acne detection, some experiments are conducted. One reason to conduct this research is that color features are not good enough to detect acne since it tends to be inconsistent with variations in lighting and skin color during the image acquisition process even for the same type of acne as stated in [15] and [16]. Therefore the use of the texture features is proposed in order to improve the detection results. However, from the results presented in Table V, it shows out that the color features are still better than the texture features. The color features have the highest recall, F1-Score, and IoU values. With a faster computation time due to less computational load, the accuracy values of color features are generally near to the accuracy values of texture features. Therefore, it is concluded that the use of texture features for acne detection is not better than color features.

TABLE V. TEXTURE FEATURES COMPARED WITH COLOR FEATURES

Feature	Acc	Prec	Rec	F1	Time	IoU	Hayashi Correct
Color	44.15%	0.52	0.79	0.58	0.383	0.473	12
GLCM Grayscale	48.89%	0.52	0.58	0.50	1.228	0.466	10
GLCM Red-RGB	53.20%	0.53	0.53	0.49	1.209	0.442	10
GLCM Cr-YCrCb	45.14%	0.50	0.78	0.56	1.260	0.463	11
GLCM Hue-HSV	41.75%	0.45	0.70	0.52	1.236	0.452	12
Gabor 4deg2sd	55.43%	0.52	0.51	0.48	0.449	0.446	10
Gabor 8deg1sd	54.27%	0.54	0.54	0.50	0.472	0.411	10

F. GMM vs KNN (K-Nearest Neighbor)

To determine the performance of GMM as a classification model, a comparison is made with KNN as one of the supervised machine learning models. The inputs of KNN are GLCM features since they produce better performance than the Gabor Filter features in GMM classification. Based on the number of images where the Hayashi class is correctly determined by using KNN classifier, Hue-HSV is the best channel with 14 correct images. KNN with the Hue-HSV channel is good at predicting acne blocks that are quite close to the ground truth given the recall value of 0.69. However, Cr-YCrCb achieved the highest recall value at 0.77. Despite that, both of those channels have equal F1-Score at 0.54. The computational time of four channels is approximately 1.4 seconds for the reason that the computational loads are similar. Based on the IoU values which are approximately 0.4, they are quite good at detecting the location of acne.

Viewed from the overall accuracy values, the KNN model is better than the GMM. However, the overall recall and F1-Score values are still below GMM except using Cr-YCrCb channel which has a high recall at 0.77. The overall computation time and IoU between KNN and GMM do not differ much. GMM Classifier is better in general at predicting Hayashi class correctly. It can be concluded that GMM as a classification model has competitive performance compared to KNN based on the evaluation parameters. The performances of GMM and KNN with the GLCM features can be seen in Table VI.

TABLE VI. GMM CLASSIFIER AND KNN CLASSIFIER EVALUATION WITH GLCM FEATURES

Channel	Acc	Prec	Rec	F1	Time	IoU	Hayashi Correct
GMM Grayscale	48.89%	0.52	0.58	0.50	1.228	0.466	10
GMM Red-RGB	53.20%	0.53	0.53	0.49	1.209	0.442	10
GMM Cr-YCrCb	45.14%	0.50	0.78	0.56	1.260	0.463	11
GMM Hue-HSV	41.75%	0.45	0.70	0.52	1.236	0.452	12
KNN Grayscale	60.12%	0.57	0.39	0.43	1.322	0.358	7
KNN Red-RGB	59.15%	0.58	0.32	0.39	1.434	0.374	7
KNN Cr-YCrCb	41.76%	0.47	0.77	0.54	1.484	0.473	12
KNN Hue-HSV	49.90%	0.52	0.69	0.54	1.422	0.444	14

All experiments with GMM Classifier have low accuracies in the range of 40-50% since there are still many false positives—or in the other word there are still many parts of the skin that are misdetected as acne. Therefore, the number of detected acne is often bigger than the ground truth. The reason behind this is the limited annotation of the dataset—only acne is labeled—causing the model to fail in recognizing non-acne objects such as acne scars, spots, and moles to acne objects. Adding professional annotations by dermatologists for those objects may improve the classification performance to let the model learn better.

V. Conclusion

The best acne classification result based on recall value is achieved by using the GMM classifier with the GLCM features in Cr-YCrCb channel as the input. The recall is 0.78. It also has a faster computation time, which is about 0.3 seconds compared to the worst method using texture features which is about 1.2 seconds. This classification method is also compared to a standard classification method which is KNN and shows that it outperforms all the evaluation criteria (see Table VI). Some suggestions for further research include the using of a face frame during image acquisition to keep the distance and the captured face size to be uniform, trying other classification algorithms to improve the performance, and increasing the number of training images as well as complementing them with a wider variety of colors and skins.

References

[1] R. Ramli, A. Malik, A. Hani, and A. Jamil, "Acne analysis, grading and computational assessment methods: an overview", Skin Research and Technology, vol. 18, no. 1, pp. 1-14, 2012.

[2] G. Maroni, M. Ermidoro, F. Previdi, and G. Bigini, "Automated detection, extraction and counting of acne lesions for automatic evaluation and tracking of acne severity", IEEE Symposium Series on Computational Intelligence 2017, pp. 1-6, 2018.

[3] C. Hsia, T. Lin, J. Lin, H. Prasetyo, S. Chen, and H. Tseng, "System for recommending facial skincare products", Sensors and Materials, vol. 32, no. 10, pp. 3235-3242, 2020.

[4] N. Hayashi, H. Akamatsu, and M. Kawashima, "Establishment of grading criteria for acne severity", The Journal of Dermatology, vol. 35, pp. 255-260, 2008.

[5] X. Wu, W. Ni, L. Jie, Y.K. Lai, S. Cheng, Ming-Ming, and J. Yang, "Joint acne image grading and counting via label distribution learning", IEEE International Conference on Computer Vision, 2019.

[6] M. Ramadhani, "Classification of acne types based on texture using the GLCM method", "Klasifikasi jenis jerawat berdasarkan tekstur dengan menggunakan metode GLCM", e-Proceeding of Engineering, vol. 5, no. 1, pp. 870-876, 2018.

[7] C. Chang and H. Liao, "Automatic facial spots and acnes detection system", Journal of Cosmetics, Dermatological Sciences and Applications, vol. 3, pp. 28-35, 2013.

[8] M. Jeon and K. Cheoi, "Detection of abnormal region of skin using gabor filter and density-based spatial clustering of applications with noise", Journal of Korea Multimedia Society, vol. 21, no. 2, pp. 117-129, 2018.

[9] S. Dey, Python Image Processing Cookbook. Birmingham: Packt Publishing, 2020.

[10] Hu. Wan, H. Wang, B. Scotney, and J. Liu, "A novel Gaussian mixture model for classification", IEEE International Conference on Systems, Man and Cybernetics (SMC), pp. 3298-3303, 2019.

[11] M. Dhanashree, S. Kalel, M. Pooja, M. Pisal, M. Ramdas, P. Bagawade, and B. Scholar, "Color, Shape and Texture feature extraction for Content Based Image Retrieval System: A Study", International Journal of Advanced Research in Computer and Communication Engineering, vol. 5, no. 4.

[12] T. Zhao, H. Zhang, and J. Spoelstra. "A computer vision application for assessing facial acne severity from selfie images", ArXiv, vol. 1907.07901, 2019.

[13] J. Arifianto and I. Muhimmah, "Face acne detection web application using deep learning algorithm with TensorFlow", "Aplikasi web pendeteksi jerawat pada wajah menggunakan algoritma deep learning dengan TensorFlow", Prosiding Automata UII, vol. 2, no. 2, 2021.

[14] M. Junayed, A. Jeny, A. Atik, N. Neehal, A. Karim, S. Azam, and B. Shanmugam, "AcneNet - A deep CNN based classification approach for acne classes", 12th International Conference on Information Communication Technology and System, pp. 203-208, 2019.

[15] A. Darmawan, A. Rositasari, and I. Muhimmah, "The identification system of acne type on Indonesian people's face image", IOP Conference Series: Materials Science and Engineering, vol. 803, no. 1, 2020.

[16] A. Gunawan, R. Adipranata, and G. Budhi, "Making acne segmentation and classification applications using region growing and self-organizing map methods", "Pembuatan aplikasi segmentasi dan klasifikasi jerawat dengan metode region growing dan self organizing map", Jurnal Infra, vol. 5, no. 1, pp. 1-6, 2017.

[17] N. Alamdari, K. Tavakolian, M. Alhashim, and R. Fazel-Rezai, "Detection and classification of acne lesions in acne patients: A mobile application", IEEE International Conference on Electro Information Technology, pp. 739-743, 2016.

[18] N. Arora and G. Sarvani, "A review paper on Gabor filter algorithm & its applications", International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), vol. 6, no. 9, 2017.

[19] F. Vasefi, W. Kemp, N. MacKinnon, M. Amini, M. Valdebran, K. Huang, and H. Zhang, "Automated facial acne assessment from smartphone images", Proceeding SPIE Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XVI, vol. 10497, 2018.

[20] R. Padilla, S. Netto, and E. Silva, "A Survey on Performance Metrics for Object-Detection Algorithms", Proceedings of International Conference on Systems Signals and Image Processing, pp. 237-242, 2020.