Acne Detection with Deep Neural Networks

Kuladech Rashataprucksa*
Research Group on Applied Computer Engineering Technology for Medicine and Healthcare, Chulalongkorn University, Bangkok, Thailand

Chavalit Chuangchaichatchavarn
Pan Rajdhevee Suphanahong Foundation, Pathumthani, Thailand

Sipat Triukose
Research Group on Applied Computer Engineering Technology for Medicine and Healthcare, Chulalongkorn University, Bangkok, Thailand

Sirin Nitinawarat
Research Group on Applied Computer Engineering Technology for Medicine and Healthcare, Chulalongkorn University, Bangkok, Thailand

Marisa Pongprutthipan
Division of Dermatology, Department of Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand

Krerk Piromsopa
Research Group on Applied Computer Engineering Technology for Medicine and Healthcare, Chulalongkorn University, Bangkok, Thailand

ABSTRACT

In this paper, a baseline for automated acne detection with deep neural networks is presented. The goal is to overcome the poor performance of traditional acne detection methods. Acne vulgaris is a disease of sebaceous gland usually presented on the face. Nowadays, dermatologists diagnose acne by counting the number of pimples, which is time consuming. We directly compared the effectiveness of Faster-RCNN and R-FCN models. We used mean average precision (mAP) to evaluate the performance. The results confirm that acne detection with deep learning is indeed promising. The data set was taken from Pan Rajdhevee Group Public Co., Ltd. Our proposed model achieved the mean average precision of 28.3%. This is not only more accurate, but also faster than that of traditional image-processing methods.

CCS CONCEPTS

• **Computing methodologies**; • **Object detection**;

KEYWORDS

Additional Key Words and Phrases: Acne vulgaris, Faster R-CNN, R-FCN

ACM Reference Format:

Kuladech Rashataprucksa, Chavalit Chuangchaichatchavarn, Sipat Triukose, Sirin Nitinawarat, Marisa Pongprutthipan, and Krerk Piromsopa. 2020. Acne Detection with Deep Neural Networks. In *2020 2nd International Conference on Image Processing and Machine Vision (IPMV 2020), August 05–07, 2020, Bangkok, Thailand.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3421558.3421566

*Corresponding Author E-mail:.6070434921@student.chula.ac.th

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IPMV 2020, August 05–07, 2020, Bangkok, Thailand

© 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-8841-2/20/08…$15.00
https://doi.org/10.1145/3421558.3421566

1 INTRODUCTION

Acne is a chronic skin condition which is a disease of the sebaceous gland that mostly exists on the face, forehead, neck, back, chest and shoulders [1]. Acnes have several types depending on the severity of the conditions. There are two main types: non-inflammatory and inflammatory lesions. Whitehead and blackhead are non-inflammatory lesions. Papule, nodule, pustule, and cyst are of the inflammatory type. Sebum production, follicular hyperkeratosis, seborrhea, and inflammatory reactions are the most important factors leading to acne. Acne treatment varies depending on how severe it is. Thus, the best way of treatment and prevention of acne scars is an individual treatment assessment [2][3]. The reference method for assessing acne severity is acne lesions counting. Currently, dermatologists have to manually count the number of acne lesions. This method is not well suited to actual clinical practice as it is time consuming and inaccurate. Furthermore, the assessment results are highly depended on physicians' experience and knowledge. Therefore, an automated localization and classification of acne is necessary to improve the diagnosis accuracy.

The traditional methods of automated acne detection use image processing and machine learning technique. The characteristics of acne vulgaris is needed for acne detection and classification. Because of color variation and color complexity of acne vulgaris, these methods still have noises and cannot achieve good results.

Deep learning has become the prominent position not just this work but, in every field, especially in medical image analysis. Convolutional neural network (CNN) has been widely used for image classification. To provide a better result and performance, we directly compared Faster-RCNN and R-FCN methods with each other. Our investigation reports that R-FCN has achieved better performance with the acne data set.

The remainder of this paper is structured as follows. In section II, related works are presented. The proposed methodology is explained in section III. The experimental setup is explained in section IV. The results are explained in section V. The discussion is described in section VI. Lastly, the conclusions are presented in section VII.

2 RELATED WORKS

There are two groups of related works in acne detection. The former one is based on image processing, and the latter on deep learning.

Chantharaphaichi et al. [4] proposed a facial acne area detection by using image processing technique. The model was based on HSV and gray scale color space. A binary threshold on frontal facial images was applied to classify acne spots. However, detection results this method still have noises because of color variation and lighting condition. Maroni et al. [5] developed an imaging method for automated acne detection and diagnosis. The discrimination between frontal face and body was performed using the Haar Cascade Classifier algorithm which was proposed by Paul Viola [6]. They obscured the mouth to avoid being classified as acne vulgaris because of its color. Then to segment the skin, an ensemble of random forest models, trained on FSD data set, was performed with color, texture, shape, and unsupervised descriptor as features. The a* channel of the CIELab model was performed to discriminate between skin and non-skin area. Acne extraction was applied by using adaptive thresholding in the previous step. Finally, the blob detection based on the Laplacian Of Gaussian was used to mark acne spots.

Chiun-Li Chin et al. [7] proposed a facial pore detection which is based on CNNs. The model contained 3 convolution layers, 3 pooling layers and 4 fully connected layers. They achieved nearly 90% of accuracy rate. Xiaolei Shen et al. [8] presented a method for facial acne vulgaris which is based on CNNs. This model can classify six types of acnes and healthy skin. They used sliding window method to crop the input facial image into small area. The skin and non-skin of each small area were detected by using a binary-classifier with CNN. The acne classification was performed by a seven-classifier with CNN. Sophie Seite et al. [9] developed an artificial intelligent algorithm (AIA). The proposed model can classify the input images into five classes (GEA 0,1,2,3,4+) by using the Global Acne Severity Scale (GEA scale) [10]. The model can differentiate types of acne lesions (comedonal, inflammatory) and postinflammatory hyperpigmentation (PIH).

3 PROPOSED METHODOLOGY

3.1 Architectures

The objective of this work is to provide a good baseline for acne detection. Thus, we investigated two models, Faster R-CNN, and R-FCN, for our experiments

3.1.1 Faster R-CNN.. Faster Region-based Convolution Neural Network (Faster R-CNN) [11] is one of the region-based family, evolved from R-CNN [12] and Fast R-CNN [13]. Faster R-CNN is a two-stage detection algorithm. The first stage generates the region proposals in an image that might contain an object. The second stage classifies and refines each region into object classes.

3.1.2 R-FCN.. Region-based Fully Convoluftional Networks (R-FCN) [14] uses region-based fully convolutional network instead of fully connected network, which is more accurate and faster than Faster R-CNN. Faster R-CNN applies a costly per-region subnetwork hundreds of times. In contrast, the region-based detectors of R-FCN are fully convolutional with most computation shared among the entire images.

Table 1: Total number of acnes per type

Acne types	Number of acnes
Type I	9413
Type III	2769
PIE	2599
PIH	1136
Total acnes	15917

4 EXPERIMENTAL SETUP

In this section, we present the information about the data set, system environment and evaluation metrics.

4.1 Data Set and Preprocessing

We have a total of 871 annotated images. The data set was separated into training data and testing data. We used 783 images for training and 88 images for testing. The details about the total number of acnes per type are shown in table 1. The data were supplied by Pan Rajdhevee Group Public Co.,Ltd. The images were labeled by dermatologists. There are 4 types of acne vulgaris in this data set.

Type I: A noninflamed skin-colored papule with whitehead or blackhead, or the so-called comedones.

Type III: An inflamed erythematous papule to pustules and a deeply inflamed erythematous nodules, nodulocystic or cystic.

Postinflammatory erythema (PIE): A damaged or dilated capillaries (blood vessels) caused by inflammation and trauma which ranges from pink to red.

Postinflammatory hyperpigmentation (PIH): An overproduction of melanin which ranges from brown to black.

4.2 Environment Systems

The classifier was built using TensorFlow (v.1.13.1). Our experiments were conducted on Intel(R) Xeon(R) CPU E5-1650 v3 @ 3.50GHz with Nvidia GeForce RTX 2080 Ti GPU.

To train the object detection model, we attempted to optimize the model performance by searching for the greatest hyperparameters. The final parameters are shown in table 2

4.3 Evaluation Metric

We measured the performance of the models by using precision, recall and mean average precision (mAP) metric. For object detection, we needed classification and localization. Thus, to measure the experimental results, the intersection over union (IoU) was used. IoU is the area of the intersection divided by the area of the union of the data set as shown in equation 1. A regional proposal is considered positive if the IoU exceeds a threshold. If the IoU score is above the threshold and the predicted value matches the actual value, it is observed as true positive (TP). Otherwise, it is a false positive (FP). If the group of truth objects has no associated predicted objects, it is considered as a false negative (FN).

$$IoU = \frac{|A1 \cap A2|}{|A1 \cup A2|} \tag{1}$$

Where

- A1: the area of the ground truth

Table 2: Our final values of hyperparameters used in the experiments

Parameters	Parameters search	Final parameters
Backbone architecture	ResNet-50, ResNet-101	ResNet-101
Image size (width, height)	(512,512), (900,900)	(900,900)
Batch size	1, 2, 4 (maximum batch size)	2
Anchor size	(32,32), (64,64)	(64,64)
Anchor scales	0.25,0.75,0.5,1,1.5, 1.75, 2	0.25,0.5,1,1.5,2
Anchor aspect ratios	0.5,0.75,1.0,1.25,1.5,2	0.5,1.0,1.5,2
Anchor stride	(4,4), (8,8)	(8,8)

Table 3: The confusion matrix using Faster R-CNN

	I	III	PIE	PIH	Reject
I	487	33	6	12	410
III	39	171	9	19	54
PIE	1	27	34	31	25
PIH	16	13	23	121	151

- A2: the area of the prediction

Precision describes the reliability of an object detection model. Calculated by the function shown in equation 2.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Recall effectively describes the ability of an object detection model. Calculated by the function shown in equation 3.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Accuracy describes how close a result of a model is to the actual value. Calculated by the function shown in equation 4.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

Mean average precision (mAP) is widely used for evaluating object detection. AP is the area under the precision-recall curve. We can compute mAP by computing the average of the AP for each class.

5 EXPERIMENTAL RESULTS

In this section, we show the experiment results by comparing the performance of each object detection model. Note that data cleansing and augmentation were applied before performing the experiments.

We evaluated object detection models for acne detection using four acne types and compared the results between Faster R-CNN and R-FCN. An example of the test results is shown in figure 1 and 2. The summary of confusion matrix for each classification are shown in table 3 and 4. Each row in the table represents a true label and each column represents a corresponding predicted label. The precision recall and AP for each model are shown in table 5. The mAP for each model is shown in table 6

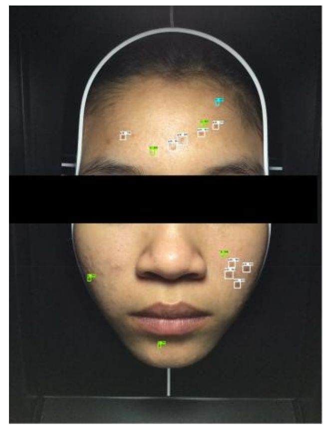

Figure 1: The result of Faster R-CNN

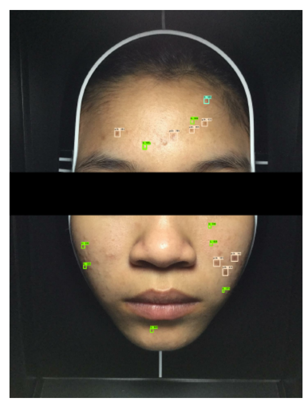

Figure 2: The result of R-FCN

Table 4: The confusion matrix using R-FCN

	I	III	PIE	PIH	Reject
I	384	33	3	10	518
III	30	156	12	17	77
PIE	4	25	34	20	35
PIH	10	15	14	125	160

Table 5: The results of the experiments on acne detection

Type	Model	Precision	Recall	Accuracy	AP
1	Faster R-CNN	0.521	0.513	0.591	0.189
	R-FCN	0.609	0.405	0.594	0.203
3	Faster R-CNN	0.593	0.585	0.892	0.320
	R-FCN	0.593	0.534	0.878	0.375
PIE	Faster R-CNN	0.354	0.288	0.934	0.203
	R-FCN	0.465	0.288	0.938	0.262
PIH	Faster R-CNN	0.458	0.373	0.844	0.241
	R-FCN	0.518	0.385	0.842	0.287

Table 6: mAP(@.5)

	mAP
Faster R-CNN	0.233
R-FCN	0.283

6 DISCUSSION

We have trained an acne detection model using two models: Faster R-CNN and R-FCN. R-FCN performed reasonably well with the mAP of up to 28.3 %. The result suggests that the performance of R-FCN model is better than that of the Faster R-CNN with the acne data set. As shown in table 5, type I has low recall and high precision, type III has precision and recall over 50%. PIE and PIH both have low recall. Hence, our conclusion that this model can effortlessly detect acne type III, because the distinguished appearance of this acne type, the redness, and its large size. Since acne type I is smaller with the color being typically paler or yellowish, it may cause the model to miss more acne spots. Because the characteristics of PIE and PIH vary depending on skin color, the model may falsely predict them. Compared to the image processing methods based on color spaces [4, 7–9], our method provides better result and performance.

7 CONCLUSIONS AND FUTURE WORK

In this work, we have established the baseline for acne detection with deep neural network. We investigated two models: Faster R-CNN and R-FCN. Faster R-CNN consists of two subnetworks. The first is the region proposals, which is used to generate object proposals and the second is used to predict the object classes of the object. R-FCN uses region-based fully convolution network with all computation shared on the entire image. The model that we used can classify acne into 4 types from facial pictures. From the experiment, Faster-RCNN gives mAP of 23.3 % and R-FCN gives mAP of 28.3%. So, we concluded that the best model is R-FCN. The model performs well on acne type III. Due to the color features and sizes of acne type I, PIH and PIH, they are harder to be precisely classified.

In the future, we will use a more standardized data set by collecting more facial images including as many types of acne as possible to build a more effective detection system. Besides, one of the challenges for applying object detection in medical imaging is the lacked labeled data. We plan to train a model by using semi-supervised learning which combined a labeled data and unlabeled data during training.

ACKNOWLEDGMENTS

The authors would like to acknowledge Pan Rajdhevee Group Public Co.,Ltd for providing the PAN data set and for the helpful comments. We also would like to thank the following people for their advice and supports throughout this project: Ms.Manita, Ms.Yada, Ms.Poonnapa and Mr.Napat. This work was also supported in part by the scholarship from Department of computer engineering, Chulalongkorn university.

REFERENCES

[1] Abhineetha Hosthota, Swapna Bondade, and Vinay Basavaraja. 2016. Impact of acne vulgaris on quality of life and self-esteem. Cutis 98, 2 (2016), 121–124.

[2] Diane M Thiboutot, Brigitte Dréno, Abdullah Abanmi, Andrew F Alexis, Elena Araviiskaia, Maria Isabel Barona Cabal, Vincenzo Bettoli, Flordeliz Casintahan, Steven Chow, Adilson da Costa, *et al.* 2018. Practical management of acne for clinicians: An international consensus from the Global Alliance to Improve Outcomes in Acne. Journal of the American Academy of Dermatology 78, 2 (2018), S1–S23.

[3] Andrea L Zaenglein, Arun L Pathy, Bethanee J Schlosser, Ali Alikhan, Hilary E Baldwin, Diane S Berson, Whitney P Bowe, Emmy M Graber, Julie C Harper, Sewon Kang, *et al.* 2016. Guidelines of care for the management of acne vulgaris. Journal of the American Academy of Dermatology 74, 5 (2016), 945–973.

[4] Thanapha Chantharaphaichi, Bunyarit Uyyanonvara, Chanjira Sinthanayothin, and Akinori Nishihara. 2015. Automatic acne detection for medical treatment. In 2015 6th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES). IEEE, 1–6.

[5] Gabriele Maroni, Michele Ermidoro, Fabio Previdi, and Glauco Bigini. 2017. Automated detection, extraction and counting of acne lesions for automatic evaluation and tracking of acne severity. In 2017 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 1–6.

[6] Paul Viola and Michael Jones. 2001. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001, Vol. 1. IEEE, I–I.

[7] Chiun-Li Chin, Zih-Yi Yang, Rui-Cih Su, and Cheng-Shiun Yang. 2018. A Facial Pore Aided Detection System Using CNN Deep Learning Algorithm. In 2018 9th International Conference on Awareness Science and Technology (iCAST). IEEE, 90–94.

[8] Xiaolei Shen, Jiachi Zhang, Chenjun Yan, and Hong Zhou. 2018. An Automatic Diagnosis Method of Facial Acne Vulgaris Based on Convolutional Neural Network. Scientific reports 8, 1 (2018), 5839.

[9] Sophie Seité, Amir Khammari, Michael Benzaquen, Dominique Moyal, and Brigitte Dréno. 2019. Development and accuracy of an artificial intelligence algorithm for acne grading from smartphone photographs. Experimental dermatology (2019).

[10] B Dreno, F Poli, H Pawin, C Beylot, M Faure, M Chivot, N Auffret, D Moyse, F Ballanger, and J Revuz. 2011. Development and evaluation of a Global Acne Severity scale (GEA scale) suitable for France and Europe. Journal of the European Academy of Dermatology and Venereology 25, 1 (2011), 43–48.

[11] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems. 91–99.

[12] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition. 580–587.

[13] Ross Girshick. 2015. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision. 1440–1448.

[14] Jifeng Dai, Yi Li, Kaiming He, and Jian Sun. 2016. R-fcn: Object detection via region-based fully convolutional networks. In Advances in neural information processing systems. 379–387.