Acne Vulgaris Severity Analysis Application

B Nethravathi (nethravathi.sai@gmail.com)  
JSS Academy of Technical Education, Bangalore

C Aradita  
JSS medical college, JSS Academy of Higher Education and Research

S Veeranna  
JSS medical college, JSS Academy of Higher Education and Research

Vikram Patil  
JSS medical college, JSS Academy of Higher Education and Research

Sushruth Nagaraj  
JSS Academy of Technical Education, Bangalore

Suvanta A Kulkami  
JSS Academy of Technical Education, Bangalore

Research Article

Keywords: Acne vulgaris, Deep learning, Automated assessment, Severity analysis, Mobile app

Posted Date: June 29th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3109197/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Abstract

The most common skin problem, acne vulgaris, may have serious financial and psychological implications for individuals who have it, thus it's crucial to have an accurate grading system for effective treatment. Artificial intelligence (AI)-based skin image analysis has gained a lot of relevance in recent years, particularly for analyzing and assessing the skin images captured by mobile phones. The difficulty in accurately assessing the severity of acne lies in the similarity of lesion appearances and the challenge of counting lesions. The study suggested a mobile app that can identify different forms of acne to solve this problem by using photos of facial skin. This study employed the You Only Look Once (YOLO) deep learning algorithm to find and identify acne. Comedone, papule, pustule, and nodule are the four forms of acne vulgaris taken into consideration. The dataset used to train and test the model is taken from the ACNE04 dataset and a private dataset from the dermatology OPD of JSS Medical Hospital, Mysuru, Karnataka, India. The app showed positive outcomes in severity analysis, showing dermatologist-level diagnosis. This application could be a valuable tool for clinicians with a smart phone to assess acne severity quickly and conveniently, anywhere and at any time.

1. Introduction

Acne vulgaris is a prevalent skin ailment that affects millions of individuals worldwide. It is distinguished by the appearance of various types of lesions on the skin. These lesions can cause physical discomfort as well as psychological distress for affected individuals. To effectively manage and treat acne, an accurate grading system is necessary. The Global Acne Grading System (GAGS) is a widely used grading methodology that considers the type, number, and distribution of lesions to assign a severity score. However, the subjective nature of lesion counting and the similarity in appearance of different lesion types can make grading challenging for clinicians.

Recent studies have explored the use of computer-based image analysis techniques, such as deep learning models, to automate acne grading. In this study, the researchers present an automated system for acne severity assessment that combines lesion classification with the GAGS metric. A convolutional neural network (CNN) model is trained on facial skin images to accurately identify and classify different acne lesions. The severity score is then calculated based on the GAGS metric, taking into account lesion type, number, and distribution. To make acne severity assessment easy and efficient for clinicians, the researchers developed a mobile phone application. This application allows clinicians to obtain dermatologist-level diagnosis anywhere and anytime. The proposed method significantly improves upon existing state-of-the-art models in lesion classification accuracy, demonstrating the potential of using mobile applications in clinical settings.

In summary, acne vulgaris is a common skin disorder that requires an accurate grading system for effective treatment planning and management. The use of computer-based image analysis techniques, such as deep learning models, can automate acne grading and improve accuracy. The development of a
A smartphone application for assessing the severity of acne has the potential to make diagnosis and treatment more accessible and efficient for clinicians.

2. Related Work

In (Alzahrani, Saeed et.al, 2022) an attention mechanism guided deep regression model for grading acne severity was presented. The proposed model was based on a ResNet-50 architecture with an attention mechanism and was trained on a dataset of 600 facial images. The model achieved high accuracy in predicting the severity grades of acne lesions. The merit of this study is that the attention mechanism helped to identify the most important features for grading acne severity. However, the study was limited by the small size of the dataset. In (Quattrini A et.al 2022) a deep learning-based facial acne classification system using a convolutional neural network (CNN) was developed. The model was trained on a dataset of 4,000 facial images and achieved high accuracy in classifying different types of acne lesions. The merit of this study is that the proposed system is efficient and accurate for classifying acne lesions. However, the study was limited by the lack of diversity in the dataset.

In (Zhang et.al, 2022) an acne detection method based on ensemble neural networks was proposed. The method was based on a CNN architecture and an ensemble learning approach that combined multiple CNNs. The model was trained on a dataset of 2,000 facial images and achieved high accuracy in detecting acne lesions. The merit of this study is that the ensemble learning approach improved the accuracy of acne detection. However, the study was limited by the small size of the dataset.

In (Hasanah et.al, 2022) a method for identifying different types of acne vulgaris in facial images using gray-level co-occurrence matrix (GLCM) feature extraction and an extreme learning machine (ELM) algorithm was presented. The model was trained on a dataset of 200 facial images and achieved high accuracy in identifying different types of acne vulgaris. The merit of this study is that the proposed method is efficient and accurate for identifying different types of acne vulgaris. However, the study was limited by the small size of the dataset.

In (Wen, Hao et al, 2022) an interpretable convolutional neural network (CNN) model for acne detection and severity evaluation was proposed. The model was based on a ResNet-50 architecture with an attention mechanism and was trained on a dataset of 6,000 facial images. The model achieved high accuracy in detecting and grading acne severity. The merit of this study is that the attention mechanism helped to identify the most important features for acne detection and severity grading. However, the study was limited by the lack of diversity in the dataset.

In (Wang, J et.al, 2022) a cell phone app for facial acne severity assessment was developed. They utilized a deep learning model and a smartphone camera to capture and analyze facial images for acne grading. Their method showed high accuracy in grading acne severity, and the app can be used for self-assessment and monitoring by patients.
In (Yi Lin et.al, 2022) a unified acne grading framework on face images, which combines feature extraction and deep learning techniques was presented. They used a large-scale dataset and achieved high accuracy in acne grading. However, their approach requires a significant amount of computational resources.

In (Huynh et.al, 2022) an automatic acne object detection and acne severity grading system using smartphone images and artificial intelligence was developed. They utilized convolutional neural networks and achieved high accuracy in acne grading. Their method has the potential to be used as a screening tool in clinics.

In (Duc Tri Phan et.al, 2021) a smart LED therapy device with an automatic facial acne vulgaris diagnosis based on deep learning and internet of things application was developed. Their system includes a camera and a deep learning model for acne detection and diagnosis. They also included an LED therapy component for acne treatment. However, the system's effectiveness in acne treatment needs further validation.

In (Neha Yadav et.al, 2021) a method for facial acne detection using the HSV model-based segmentation driven deep learning approach was proposed. Their method showed promising results in acne detection, but their dataset was relatively small.

In (Amandip Sangha et.al, 2021) a deep learning object detection-based method to detect acne from facial images was presented. The authors used the YOLOv5 object detection model and their own dataset containing 208 facial images. The proposed method achieved an accuracy of 95.98% in detecting acne lesions. However, the small size of the dataset may limit the generalization of the proposed method.

In (E. Malgina et.al, 2021) a mobile application for assessing facial acne severity from photos was developed. The application employs a convolutional neural network (CNN) model to classify acne severity into four levels. The authors used the Labeled Faces in the Wild (LFW) dataset to train the CNN model. The proposed application provides a user-friendly and convenient way to assess acne severity, but the accuracy and generalization of the model were not reported.

In (Nur Ashiqin et.al, 2021) an acne type recognition system for a mobile-based application using YOLO object detection model was proposed. They used a dataset of 447 facial images, which was manually annotated. The proposed system achieved an accuracy of 97.54% in classifying different types of acne. However, the size of the dataset is relatively small, which may affect the generalization of the proposed system.

In (M. S. Junayed et.al, 2022) a novel deep CNN model called ScarNet for acne scar classification was presented. They collected a new dataset consisting of 1,980 images of different types of acne scars, which were annotated by dermatologists. The proposed model achieved a classification accuracy of 94.09%, outperforming several state-of-the-art models. The authors also provided an online demo of their
model, which can be accessed by dermatologists and other healthcare professionals for accurate
diagnosis of acne scars.

In (A. Dragomir et.al, 2020) a smart mirror for skin type identification was developed. The authors used a
Raspberry Pi, a camera, and a display to create the mirror. The system employs a CNN model to analyze
facial images and identify the user's skin type. The authors used the Fitzpatrick skin type classification
system and a dataset of 150 facial images to train the CNN model. The proposed system provides a
convenient and personalized way for users to determine their skin type, but the accuracy and
generalization of the model were not reported.

In (X. Wu et al, 2019) a joint acne image grading and counting approach that combines deep learning and
label distribution learning was presented. The proposed method uses a ResNet-50 model to extract
features from input images and a label distribution learning framework to perform both acne grading and
counting tasks. The proposed method achieved state-of-the-art performance on the public dataset used in
the experiment.

In (K Karunanayake et.al, 2020) a skin disease detection system called CURETO that uses image
processing and convolutional neural networks (CNNs) was developed. The system employs a pre-trained
ResNet-50 model as a feature extractor, and then fine-tunes the model for skin disease detection. The
authors evaluated the system on a dataset containing various skin diseases, including acne, and
achieved high accuracy.

In (Kuladech Rashataprucksa et.al, 2020) a deep neural network-based approach for acne detection that
uses a customized version of the VGG-16 model was presented. The proposed method involves training
the model on a large dataset of facial images labeled with acne and non-acne categories. The proposed
method achieved high accuracy on the test set.

In (Shen, X et.al, 2018) an automatic diagnosis method for facial acne vulgaris using a convolutional
neural network (CNN) was proposed. The proposed method uses a deep CNN model to classify input
images into different acne severity levels. The authors evaluated the proposed method on a dataset
containing 3,102 facial images and achieved high accuracy.

In (N Alamdari et.al, 2016) a mobile application for detecting and classifying acne lesions in acne
patients was developed. The system uses a support vector machine (SVM) classifier to classify input
images into different acne lesion types. The authors evaluated the system on a dataset containing 400
acne images and achieved high classification accuracy. However, the dataset used in the experiment is
relatively small, which may limit the generalizability of the proposed approach.

In (Zhao et.al, 2019) a computer vision application for acne severity assessment from selfie images
using image segmentation and feature extraction was developed. The method utilizes color-based
segmentation, morphological operations, and histogram of oriented gradients (HOG) feature extraction.
The dataset consists of 90 self-taken selfies. The merit of this approach is that it provides a low-cost and
convenient way for acne severity assessment. The limitation is that the dataset used is relatively small and may not be representative of the general population.

In (Lim et.al, 2020) a deep learning approach with convolutional neural networks to automatically grade acne vulgaris was proposed. They used a dataset of 13,000 images of acne vulgaris with corresponding grades assigned by dermatologists. The model achieved a high level of accuracy in grading acne vulgaris, and the method potentially reduces human error and subjectivity. However, the limitation of the deep learning approach is the lack of interpretability.

In (M. S. Junayed et.al, 2019) a deep convolutional neural network-based approach for classifying acne classes was presented. The dataset used for training and testing the model consists of 6,534 images of different types of acne lesions. The method achieved high accuracy in classifying acne lesions. The limitation of this approach is that the model may not generalize well to new populations and datasets.

In (Natchapol Kittigul et.al, 2017) an acne detection method using Speeded Up Robust Features and K-Nearest Neighbors algorithm was presented. The dataset used consists of 200 images of acne vulgaris. The method achieved high accuracy in detecting acne lesions. However, the limitation is that the dataset is relatively small and may not be representative of the general population.

In (N. Kittigul et.al, 2016) an automatic acne detection system for medical treatment progress report using the Speeded Up Robust Features algorithm was developed. The dataset consists of 100 images of acne vulgaris. The method achieved high accuracy in detecting acne lesions, and the system may potentially reduce human error and subjectivity in medical treatment progress reports. However, the limitation is the small dataset used.

In (Chantharaphaichit,et.al, 2015) a Bayesian classifier for automatic acne detection for medical treatment was used. The method used clinical images and achieved high accuracy in detecting acne lesions. The merit of this approach is that it is a low-cost and convenient method for acne detection. However, the limitation is the relatively small dataset used.

In (T. Chantharaphaichi et.al, 2015) an automatic acne detection method for medical treatment using image processing techniques and decision-making rules was developed. The dataset consists of clinical images of acne vulgaris. The method achieved high accuracy in detecting acne lesions, and the system may potentially reduce human error and subjectivity in medical treatment progress reports. The limitation is that the generalizability of this method to different populations is unknown.

In (C. Chang and H. Liao, 2013) an automatic facial spot and acne detection system based on a support vector machine (SVM) classifier was proposed. The system uses digital image processing techniques to extract features from facial images, and the SVM classifier is trained on a dataset of images labeled by dermatologists. The system achieved an accuracy of 85% in detecting facial spots and 75% in detecting acne lesions.
In (Thang, Do et.al, 2022) an application called Acne Scan that uses machine learning algorithms to assess the severity of acne and recommend treatments was developed. The application uses a dataset of facial images labeled with acne severity scores by dermatologists. The machine learning models used in the application include convolutional neural networks (CNNs) and support vector regression (SVR). The application was found to be effective in accurately assessing acne severity and recommending appropriate treatments. In (Firas Gerges et.al, 2022) an automated diagnosis system for acne and rosacea using convolutional neural networks (CNNs) was proposed. The system uses transfer learning to fine-tune a pre-trained CNN on a dataset of facial images labeled with acne or rosacea diagnoses by dermatologists. The system achieved an accuracy of 90% in diagnosing acne and 83% in diagnosing rosacea. In (Y. Lin et.al, 2021) an acne grading framework for facial images based on skin attention and SFNet (a type of neural network) was presented. The system uses attention mechanisms to highlight regions of the image that are most important for acne grading, and then uses the SFNet to classify the severity of acne in those regions. The system achieved high accuracy on a dataset of facial images labeled with acne severity scores by dermatologists. In (Nguyen et.al, 2021) a system for assessing the severity of facial acne using machine learning algorithms was proposed. The system uses a dataset of facial images labeled with acne severity scores by dermatologists and extracts features using a combination of histogram of oriented gradients (HOG) and local binary patterns (LBP). The system uses a support vector regression (SVR) model to predict acne severity scores based on the extracted features. The system achieved good performance in predicting acne severity scores.

In (Mohammad Amini et.al, 2018) an automated facial acne assessment system based on smartphone images was presented. The system uses a convolutional neural network (CNN) to classify facial images as either having or not having acne, and to predict the severity of acne if present. The system achieved high accuracy on a dataset of smartphone images labeled with acne severity scores by dermatologists. In (A. Hanifa Setianingrum et.al, 2020) a system for acne type identification based on gray-level co-occurrence matrix (GLCM) and support vector machine (SVM was proposed. The system uses a dataset of facial images labeled with acne types by dermatologists and extracts features using GLCM. The SVM classifier is then trained on the extracted features to identify the type of acne in the image. The system achieved good performance in identifying acne types.

3. Methodology

The following section outlines the approach employed in this study, which encompasses five primary stages, namely data collection, grading methodology, YOLO-based model creation, model evaluation, and application development. The proposed methodology for this study is depicted in the accompanying figure, and each phase will be discussed in detail below.

3.1 Data collection and preprocessing

The facial images for this study were obtained from three sources. The publicly available ACNE04 and DermNet dataset and a private dataset from the dermatology OPD of JSS Medical Hospital, Mysuru of
facial acne patients were considered. The total number of images used is 2000. All participants provided informed consent, and all procedures were carried out according to approved guidelines and regulations. All the images were first randomly split into a training set, a validation set and a test set with a percentage of 80, 15, 5 respectively. Data augmentation was used during the training process, and all images were resized to 416x416. The Acne lesion types considered are comedones, papules, pustules and nodules.

3.2 Developing ML model using YOLOv5

Object detection is a vital area of computer vision that involves identifying specific objects in digital images. Recently, deep learning-based models have emerged as the dominant technique for object detection. These models are deep convolutional neural networks which have been trained on extensively annotated datasets to anticipate rectangular bounding boxes enclosing the objects of interest depicted in the images.

The You Look Only Once (YOLO) is a deep learning algorithm that offers high detection and recognition accuracy. YOLO allows for end-to-end learning and has been applied in various applications, including autonomous vehicles, intelligent systems, and virtual reality. Using features from the entire image, YOLO simultaneously predicts bounding boxes. The input image is partitioned into a grid of S x S cells, with each cell producing bounding boxes and their corresponding confidence scores.

For this study, the YOLOv5 model is utilized, a convolutional neural network with 773 layers and 141.8 million parameters. YOLOv5 is an example of a single-stage detector, where a single neural network predicts bounding boxes and class probabilities from full images directly in one evaluation. This is in contrast to two-stage detectors, such as the Fast R-CNN family of models, which use separate neural networks for proposing regions of interest and predicting bounding boxes within the proposed regions. YOLOv5 P615 model was pre-trained on the COCO9 dataset for object detection, making it a reliable and accurate option for this study.

3.3 Grading and Severity Level Assessment

The Global Acne Grading System (GAGS) is a quantitative method for evaluating the severity of acne. It involves calculating a total severity score by combining five sub-scores that correspond to different regions of the face. To obtain each regional score, the highest weighted lesion in the area is multiplied by a factor that varies depending on the location (2 for each cheek and forehead, 1 for the nose, and 1 for the chin). The weights for the different types of lesions (1 for ≥ one comedone, 2 for ≥ one papule, 3 for ≥ one pustule, and 4 for ≥ one nodule) were determined based on considerations of the size and density of pilosebaceous units in the region.
### Table 1
Grading and Severity Level Table

<table>
<thead>
<tr>
<th>Location</th>
<th>Factor X Grade (0–4) = local score</th>
<th>[Global score = 0 = None, 1–18 = Mild, 19–30 = Moderate, 31–38 = severe]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forehead</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Right Cheek</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Left Cheek</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Nose</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Chin</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

#### 3.4 Performance Evaluation Metrics

To assess the deep learning model's prediction performance, the error between the real classes and the expected outcomes is calculated using the following metrics:

1. Intersection over Union (IoU): \( \text{IoU} = \frac{\text{Area of overlap}}{\text{Area of union}} \)
2. Precision: \( \text{Precision} = \frac{tp}{(tp + fp)} \)
3. Recall: \( \text{Recall} = \frac{tp}{(tp + fn)} \)
4. F1-Score (F1): \( F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \)
5. Accuracy (Acc): \( \text{Acc} = \frac{(tp + tn)}{(tp + tn + fp + fn)} \)
6. Mean Average Precision (mAP): \( \text{mAP} = \frac{1}{N} \sum_{i=1}^{N} \text{AP}_i \)

Where, \( tp = \text{True Positive}, \) \( tn = \text{True Negative}, \) \( fp = \text{False Positive}, \) \( fn = \text{False Negative}, \) \( \text{AP} = \text{Average Precision}, \) \( N = \text{Number of Classes in the Dataset} \)

#### 3.5 App Development

With the rise of the mobile Internet and the Internet of Things, there is a growing need for deep learning algorithms to be directly deployed on smart devices. The deployment of intelligent acne diagnosis on smart devices not only provides convenience to acne patients, but also reduces the workload of dermatologists. Recent advancements in mobile phone hardware technology have made it possible for apps to be used for auxiliary diagnoses. The use of a proposed acne detection and severity assessment has resulted in the development of a mobile app that can automatically analyze facial acne.

The mobile app developed has the login or sign-up screen which on registering redirect the users to the interface which prompts the user to take 3 images of their face (front, left and right) from a distance of 15 cm from the device with proper lighting. These images of the user are then fed to the model for acne detection. Based on the detection, the grading of the acne is done using the Global Acne Grading System.
4. Results

In The YOLO algorithm was trained on a combination of both private and publicly available dataset with a train, test, and validation split of 80, 15 and 5 percent respectively. Data augmentation was used during the training process, and all images were resized to 416x416. The confidence threshold was set at 0.5 for the calculation of precision and recall. Accuracy was not chosen as the evaluation metric due to the problem of class imbalance. The mean Average Precision value is a more reliable metric for such a scenario. Overall, the model achieved an mAP value of 85%.

The mobile app achieved high accuracy in classifying acne severity levels compared to manual assessments by dermatologists. The app's performance was also robust to changes in lighting and facial expressions. About 85% accuracy is achieved.

The graphs for precision, recall and mAP are as follows:

5. Discussions

Model Strengths:

There are several advantages of the acne detection model. It is proficiently trained multi-class detection of varying severity levels, allowing it to identify different types of acne, such as comedones, papules, pustules, and nodules. Additionally, the model is also reliable for different skin tones. The model performs satisfactorily on blurry and low lighting condition images.

Model Weakness:

The regions such as the folds around the lips and nostrils, and occasionally hair follicles, hinders the model's ability to differentiate the regions of the body. The similarity in shape and color between these body parts and inflamed acne lesions is likely the cause of the model's confusion. Additionally, the model may be confused by hyperpigmented lesions, as post-inflammatory hyperpigmentation often accompanies acne. To address these issues, exposing the model to more photos without acne is proposed so that it can learn to recognize the subtle differences. This approach would provide more complex negative examples for the model to learn from.

6. Conclusion and Future Work

In this paper, the proposed automated system for acne severity assessment combines lesion classification with the GAGS metric. A CNN model is trained using the YOLOv5 architecture on a combination of private and publicly available dataset to accurately identify and classify different acne lesions. The proposed method significantly improves upon existing state-of-the-art models in lesion classification accuracy, achieving an accuracy of 85% in the experiments. Additionally, a mobile
application that can be used by clinicians for easy and efficient acne severity assessment was developed, providing dermatologist-level diagnosis anywhere and anytime.

There is still room for improvement in the proposed method. One possible avenue for future work is to investigate the use of different CNN architectures and training techniques to further improve the accuracy of lesion classification. Additionally, expanding the dataset to include a more diverse range of skin types and demographics may also improve the model's generalization capabilities. Finally, implementing active learning can improve the model's performance.

One potential improvement for the mobile application is to incorporate a feature that allows patients to track their acne progress over time, enabling them to monitor the effectiveness of their treatment plan. Additionally, integrating a telemedicine platform that allows patients to consult with dermatologists remotely could greatly improve access to care, particularly for individuals living in rural or remote areas. Finally, incorporating machine learning algorithms to provide personalized treatment recommendations based on patient-specific data could further improve the efficacy of acne treatment.

**Declarations**

**Author contributions:** All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by NB, AC, VS, SN and SK. The first draft of the manuscript was written by SN and SK and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Data availability:** The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Declarations Conflict of interest:** The authors do not have any financial or non-financial interests directly or indirectly related to the work submitted for publication

**References**


11. Amandip Sangha, Mohammad Rizvi medRxiv "Detection of acne by deep learning object detection" 2021.12.05.21267310; doi: https://doi.org/10.1101/2021.12.05.21267310


29. Thang, Do & Tran, Duy & Tran, Tham & Pham, Nguyen. (2022). Acne Scan - Application In Assessment Of Acne And Supports Treatment Road. 10.29007/bg75.


**Figures**

![Diagram](image.png)

**Figure 1**

Methodology
Figure 2
Acne vulgaris types

Figure 3
Working of YOLO
Figure 4

Screenshots of the proposed mobile app
Figure 5

Flowchart of the application
Figure 6

Precision, Recall and mAP graphs