System Analysis of Acne Diagnostics Using Artificial Intelligence

Shavkat Fazilov[1[0000-0002-4217-6659]] and Shukhrat Mamadjanov[1,2[0009-0001-3036-6607]]

[1] Digital Technologies and Artificial Intelligence Development Research Institute, Tashkent, Uzbekistan
[2] Tashkent International University, Tashkent, Uzbekistan
shukhratjonmamadjanov@gmail.com

Abstract. Acne, considered one of the dermatological disorders, is a widespread skin disease that often occurs during crucial developmental stages in adolescents and can lead to long-term psychosocial consequences. Traditional methods, such as assessing overall severity and counting rash elements, are limited by subjectivity and time constraints. This article focuses on a systematic review of the latest advancements in diagnosing acne, detecting and counting rash elements, and assessing severity using artificial intelligence (AI), highlighting the potential of AI-based methods to improve objectivity, reproducibility, and clinical effectiveness. A comprehensive literature search was conducted in PubMed, Scopus, arXiv, Embase, and Web of Science databases for studies published from 2017 to February 2026. The search strategy included terms related to "acne" and various AI methodologies (e.g., "neural network," "deep learning," "convolutional neural network"). During the research process, 391 articles were reviewed, of which 39 articles met the final criteria. Data were extracted on study design, dataset characteristics (including internal and open-access resources such as ACNE04 and AcneSCU), AI architectures (primarily CNN-based models), and performance indicators. Although AI-based models have demonstrated high accuracy under controlled conditions, the scarcity of large public datasets, the predominance of data from specific ethnic groups, and the lack of comprehensive external validation reveal significant barriers to implementation in clinical practice. The research findings indicate that while AI has the potential to standardize acne assessment, reduce observer variability, and enable self-monitoring through mobile platforms, there are considerable challenges in achieving reliable real-world application. Future research should prioritize the creation of large, diverse, and openly accessible datasets and conduct prospective clinical trials to ensure fair and effective dermatological care.

Keywords: Acne, artificial intelligence, acne position, dermatology, image classification, computer vision.

1 Introduction

Facial acne is one of the most common skin problems, especially prevalent among adolescent boys and girls, and is a condition that affects a large portion of the younger generation [1]. This ailment typically emerges during youth, at a crucial stage of a person's social, psychological, and physical development, potentially leading to dissatisfaction with one's appearance and negatively impacting social relationships and self-confidence. Furthermore, in many cases, this condition persists beyond puberty, resulting in a decrease in the quality of life [2].

Accurate measurement and assessment of acne is crucial for effective treatment and the reproducibility of scientific research [2]. However, due to the spontaneous variability of the acne and the uneven distribution of the rash, this process can be complicated. In another study, a new hybrid model was presented, combining DeepLabV3 and InceptionV3 for classification, which offers an improved solution for detecting acne. The DeepLabV3 model separates acne lesions and creates distinct segmentation masks, while InceptionV3 effectively classifies various types of acne and increases overall diagnostic accuracy [3].

Traditional methods used to determine the degree of acne are divided into two main areas: general assessment and counting of inflammatory elements. However, both of these methods are limited and time-consuming, as they are based on personal considerations[4]. Subjective assessment of severity on a general scale is carried out by comparing the clinical picture with a standard picture or written description. Currently, the most commonly used scale for assessing acne is the Investigator Global Assessment (IGA), which ranges from 0 to 4 (0, pure; 1, almost transparent; 2, light; 3, average; 4, severe) [4] and the Global Acne Severity Scale (GEA) (0, no lesions; 1, almost no injuries; 2, light; 3, average; 4, heavy; 5, very heavy) [4].

Another possible approach is the method of counting lesions, shown in the example of the Hayashi scale. This method assesses the severity of the acne by counting inflammatory lesions (papules and pustules) on one side of the face and divides it into four categories: mild (0-5 lesions), moderate (6-20), severe (21-50) and very severe (>50) [4]. Assessment of the overall severity is usually simpler and faster than counting lesions, but this method depends more on the evaluator and is less sensitive to changes in the disease over time. On the contrary, counting lesions provides a more accurate and reproducible method, but it is often impractical, time-consuming, and does not take into account such parameters as the nature of the lesion's spread, size, or the presence of erythema[4]. The criteria for counting lesions and assessing the severity of acne can be further divided into two types: criteria that take into account only the number of lesions, and criteria that take into account both the number and type of lesions (e.g., open and closed comedones, papules, pustules, and nodules) [4].

Recently, thanks to the rapid development of artificial intelligence (AI) and computer vision technologies, new opportunities have emerged for automatic and objective assessment and detection of acne. This process is carried out, in particular, by segmenting the affected areas and calculating their quantity, as well as by assessing the overall situation. For the purpose of high-precision detection and grouping of acne symptoms, models based on artificial intelligence, prepared using vast amounts of data and advanced computational methods, allow for the assessment of this skin problem and the effectiveness of treatments performed in relation to it using standardized and continuously repeatable methods [5]. The use of automated methods for assessing acne based

on the principles of artificial intelligence allows for faster results compared to conventional assessment methods, ensures high accuracy, increases recovery, and normalizes the assessment process. This serves to minimize differences (variations) in assessment between different observers and within one observer [6]. AI algorithms can also lead to self-monitoring through mobile applications, evaluating treatment responses, processing large volumes of epidemiological data[7], and simplifying comparisons between studies in meta-analysis[8].
The purpose of this systematic review is to identify acne, quantify their prevalence, and comprehensively cover the latest developments and applications of artificial intelligence in assessing the severity of the disease.

2 Materials and methods

A systematic review of the literature included a thorough examination of the analyzed articles, paying particular attention to the structure of the computational models used (e.g., CNNs, deep residual transformations, and integrated systems), the nature of the data sources, the methods used for marking/tagging, and the procedures for evaluating the results. The comparative aspect focuses on key performance indicators, including accuracy, retention, true negative score, F1 score, and AUC score. All of this process strictly adhered to the guidelines set out in the PRISMA 2020 documents to guarantee accuracy in methodology, reproducibility, and thoroughness in selected studies. A comprehensive search was carried out in the databases PubMed, Scopus, Web of Science, IEEE Xplore, arXiv, Embase. A methodological search plan was compiled for searching relevant literature, which included designated headings and general linguistic terms for the subject. The resulting survey combined acne-related terms with artificial intelligence and visual data evaluation phrases, using logical connectors such as AND and OR. In the process of selecting preliminary results, comprehensive reviews or studies combined with meta-analysis were deliberately skipped, which ensured the emphasis on basic, live, or laboratory research in the final selection. The study included works published in the period from January 2017 to February 2026. We included in this analysis any study covering patients of all ages with a confirmed presence of acne and using artificial intelligence or computer vision algorithms to identify the damage or assess its severity. Studies that yielded good results had to compare their models with clinical data based on "correct" standards evaluated by doctors (dermatologists) or approved by specialists. At the same time, they had to show at least one numerical performance metric (e.g., accuracy level, Dice coefficient, sensitivity, or specificity). The study examined observational studies published in English. We excluded conference summaries, non-English-language articles, descriptive reviews and meta-analyses, as well as any research that reused data sets already used elsewhere to avoid duplication.

3 Results

3.1 Sorting research and transparency

As a result of the systematic search, 391 articles were identified, of which 115 studies underwent a full-text review. As a result, 39 studies were selected. The analysis was focused on the methods of artificial intelligence used in the diagnosis and assessment of the severity of acne, the data sets used, and the results achieved. Of these, 17 (43.6%) were based solely on internal datasets, while 22 (56.4%) used publicly available datasets (Face Datasets [3], Dermnet [9], ACNE04 [11], AcneSCU [12], MedDialog [13]). More specifically, four studies (10.2%) used only the ACNE04 dataset, two studies (5%) used the AcneSCU dataset, nine studies (23%) combined ACNE04 with internal data, one study (2.5%) used the Face Datasets dataset, two studies (5%) used the Dermnet dataset, two studies (5%) used the MedDialog dataset, and three studies (7.6%) combined ACNE04, AcneSCU, and internal data.

In addition, in one study, a synthetic data set created by artificial intelligence trained using a Generative Adversarial Network (GAN) was used in the ACNE04 dataset [14]. In another study[14], an open data set of faces (2037 images from CelebAMask-HQ and Flickr-Faces-HQ) was used, a small portion of which suffered from acne. It should be noted that in most studies, images from outpatient clinics were used, but in seven studies (28%) images taken by the patients themselves using smartphones were used. Publicly available software code was available in ten studies (25.6%).

3.2 Demographic characteristics of participants and diversity of the dataset

The age of the participants was recorded in only five studies (12.8%). Data on the ethnicity of the participants were available for 16 studies (41%), including those who used only the open ACNE04 database. Of these, 15 (38.4%) mainly included persons of East Asian or Chinese nationality. The only exception was a study by Seitè et al., which analyzed 5,972 images of 1,072 acne patients representing various ethnic groups, such as Europeans, Africans, Asians, Latin Americans, and Indians, in countries such as France, South Africa, China, and India. Images were obtained using both iOS and Android smartphones, which ensured a wide range of skin colors and types.

3.3 Algorithms

In 13 (33.3%) of the analyzed 39 studies, multiple deep learning (DL) algorithms were used. CNN and deep convolutional neural networks (DCNN) were the main architectures, accounting for 72% (28/39) of all studied models (Table 1). The main focus of the study was the segmentation and counting of acne (17 studies, 43.6%) and determining the severity (20 studies, 51.2%), while only two studies (5.2%) studied the binary classification (presence and absence of acne). The training datasets varied significantly, ranging from 276 to over 212,374 images. Key methodological approaches include CNN-based architectures, semi-controlled teaching methods, and generative competitive networks (GANs) such as StyleGAN2-ADA. These were used to create synthetic datasets, thereby eliminating data shortages and increasing the generalizability of models. The training data was typically divided into training and validation sub-sets, with 80-90% directed to training and the remainder to validation.

Table 1. Description of the works included in the study.

Authors	Application Category	Methods	Dataset	Dataset Source	Publicly Available Code	Number of Photos
Sharmin et al. [3] (2025)	A Hybrid CNN Framework DLI-Net for Acne Detection with XAI	Introduces a novel hybrid model that combines DeepLabV3 for precise image segmentation with InceptionV3 for classification, offering an enhanced solution for acne detection.	Face Datasets consists of 1725 images of inflammatory acne, 594 images of non-inflammatory acne, and 8460 images of clear skin.	To address the class imbalance, oversampling techniques such as data augmentation (rotation and flipping) were applied, resulting in a balanced dataset with 5051 images of inflammatory acne, 1962 images of non-inflammatory acne, and 5325 images of clear skin.	Datasets available at: https://www.kaggle.com/datasets/xtvgie/face-datasets	10779
Shen et al. [16] (2017)	Automatic Acne Diagnosis	Binary classifier for skin detection and seven-class classifier for acne lesion classification, using CNNs with data augmentation techniques.	Binary classifier: 3000 cutaneous and 3000 non-cutaneous images. Seven-class classifier: 6000 images per class (blackheads, whiteheads, papules, pustules, cysts, nodules, normal skin).	Clinical images, augmented with rotation, shift, shear, scaling, and horizontal flip.	No	Not specified
Ayesha et al. [9] (2025)	Optimizing Acne Severity Detection	A transfer learning approach was implemented by extracting image features using a ResNet-152 pretrained model, then a fully connected layer was added and trained to learn the target severity level from labelled images.	DermNet NZ acne image sets used as the training dataset, and 100 test images were obtained to serve as a 'golden set' for model evaluation from Flicker Faces HQ Dataset .	The dataset acquired for the training process consisted of 1,148 selfie images of different people from different nationalities (European and South Asian), skin colors and features of mixed resolutions (high and low resolution) containing more than one acne lesion all over the face.	Datasets available at: https://www.kaggle.com/datasets/shubhamgoel27/dermnet	1148
Zhao et al. [17] (2019)	Acne Severity Assessment using Selfie Images	Deep learning model using transfer learning with ResNet 152 and	4700 selfie images, with 4470 used for training and 230 used for testing.	Selfie images collected by Nestlé SHIELD and labeled by 11 dermatologists.	Code available on GitHub:https://github.com/	4700

		a novel rolling image augmentation approach to improve CNN model generalization.			Microsoft/nestle-acne-assessment (accessed on 5 March 2025)	
Zhang et al. [13] (2024)	Advancements in acne detection	For the development of an advanced detection system, a multi-functional method for detecting acne using the CenterNet-based learning paradigm is proposed.	Through image usability testing and inclusion criteria, 153,183 acne lesions images were finally selected, from which a training set containing 150,219 images and a validation set containing 2,322 acne lesions images were constructed for subsequent modeling.	To train and validate our algorithm, we collected 212,374 original acne lesion images from dermatology clinical clinics in 15 hospitals in different provinces of China using smartphones from January 2020 to October 2022.	No	212374
Wu et al. [11] (2019)	Joint Acne Image Grading and Counting via Label Distribution Learning	Label Distribution Learning framework that integrates global grading and local lesion counting.	1457 facial images from the ACNE04 dataset with 18,983 annotated lesions.	Clinical images collected under the Hayashi grading criterion, taken at a 70-degree angle.	Code and dataset available on GitHub: https://github.com/xpwu95/ldl (accessed on 5 March 2025)	1457
Alzahrani et al. [19] (2022)	Attention Mechanism Guided Deep Regression Model for Acne Severity Grading.	The developed architecture incorporates dilated UNet dense regressor for density regression with the information of bounding boxes generated from Faster R-CNN network, producing a hybrid detection–regression framework.	To conduct the experiments in this research work, a publicly available dataset named ACNE04 is used.	The ACNE04 contains 1457 images of lesions with 18,983 bounding boxes.	Code and dataset available on GitHub: https://github.com/xpwu95/LDL	1457
Seitè et al. [6] (2019)	Acne Grading from Smartphone Photographs	AI Algorithm trained on a large and diverse dataset, using deep learning techniques for acne grading and lesion identification.	5972 images from 1072 patients with acne, including 2939 inflammatory lesions, 7603 non-inflammatory lesions, and 5702 PIHP instances.	Smartphone images (iOS and Android) across various racial groups.	No	5972
Phuc et al. [22] (2024)	ACNE8M - An acnes detection and differ-	Given the promising potential of YOLOv8, this model was selected for acne detection training.	Acnes detection dataset utilized for this model is sourced from Roboflow and authenticated by our dermatologists.	All images are resized to 800 x 800 pixels, with several augmentations applied: mosaic, horizontal and vertical flips, and	Datasets available at:	9440

	ential diagnosis system using AI technologies.			rotation augmentation within a range of -25 to 25 degrees.	https://universe.roboflow.com/testdg2/fold_02	
Lim et al. [18] (2019)	Automated Acne Grading using Deep Learning and CNNs	Deep learning models (DenseNet, Inception v4, ResNet18) trained on high-resolution images to automatically calculate IGA scores.	472 frontal facial images from 416 acne patients, with training set of 314 images and testing set of 98 images, augmented to 6248 images.	Clinical images captured at $\sim 2000 \times 3000$ pixel resolution.	No	472
Bernhard et al. [26] (2024)	Automatic Acne Severity Grading with a Small and Imbalanced Data Set of Low-Resolution Images.	A total of 1374 triplets of images (frontal and lateral views) from 391 different patients suffering from acne labeled with the IGA severity grade by an expert dermatologist were used to train and validate a deep learning model that predicts the IGA severity grade.	The study involved 391 Chinese patients	Pictures were taken with iPhone 11 with a ring light and a resolution of 3024×4032 pixels.	No	1374
Peris Fajarnés et al. [20] (2020)	Segmentation Methods for Acne Vulgaris Images using Fluorescence Imaging	K-means clustering algorithm implemented in MATLAB R2018b for automated segmentation and counting of acne lesions in fluorescence images.	54 processed fluorescence images of mild acne patients, captured using iPhone X with Wood's lamp and LED lamp.	Fluorescence images captured using iPhone X smartphone.	No	54
Aldi et al. [28] (2025)	Multi-Class Facial Acne Classification using the EfficientNetV2-S Deep Learning Model.	Employed a quantitative experimental approach using facial acne image data, which were classified into five categories through a Convolutional Neural Network (CNN) model based on the EfficientNetV2-S architecture.	This dataset was obtained from a public repository on Kaggle, originally sourced from DermNet, and was preprocessed to support research in multi-class acne classification using digital facial images.	The dataset is divided into three subsets: 2,834 images for training, 921 for validation, and 918 for testing.	Datasets available at: https://www.kaggle.com/datasets/shubhamgoel27/dermnet/data	4673

Rashatapruck sa et al. [21] (2021)	Acne Detection	Faster R-CNN and R-FCN models used for acne detection, focusing on various types of acne lesions.	871 annotated images with 15,917 lesions categorized into four types: Type I, Type III, PIE, and PIH.	Clinical images provided by Pan Rajdhevee Group Public Co., Ltd.; Bangkok, Thailand.	No	871
Yang et al. [23] (2021)	Acne Vulgaris Assessment	Inception-v3 architecture used for model construction, trained on clinical images categorized into four severity grades.	5871 clinical images, with 1565 images in the training set, 392 in the validation set, and 40 in the test set.	Clinical images captured with single-lens reflex cameras.	No	5871
Watanabe et al. [31] (2025)	Deep Learning-Based Acne Severity Classification Using Standardized Facial Images of Japanese Patients.	An EfficientNet-B2 model was trained using a two-stage curriculum learning strategy, class weighting, and extensive data augmentation techniques.	A dataset of 349 facial images was collected from Japanese acne patients.	The patient cohort ranged in age from six to 76 years, with a median age of 21 and a mean age of 25.7 years. The dataset comprised 226 images from female patients (61.4%) and 142 from male patients (38.6%).	No	349
Quattrini et al. [15] (2022)	Facial Acne Classification	Used a VGG-19 based CNN for classification and BiSeNet for semantic segmentation, trained on re-annotated images from the FFHQ dataset.	2307 re-annotated images from the Flickr-Faces-HQ (FFHQ) dataset, used for acne severity classification.	High-resolution images from the FFHQ dataset, re-annotated for acne severity.	Datasets available at: https://www.kaggle.com/datasets/arnaud58/flickrfaceshq-dataset-ffhq	2307
Min et al. [24] (2022)	Acne Detection using Mask-Aware Attention with Dynamic Context Enhancement	ACNet integrates Composite Feature Refinement, Dynamic Context Enhancement, and Mask-Aware Multi-Attention for enhanced feature representation and detection accuracy.	ACNE04 dataset with 1457 facial images and 18,983 annotated lesions.	Clinical photographs with varying resolutions, annotated by dermatologists.	Datasets available at: https://www.kaggle.com/datasets/manuelhettich/acne04	1457
Mascarenhas et al. [36] (2025)	Improving acne severity detection.	Traditional methods to address blurring fail to recover fine details, making them unsuitable for teledermatology. To resolve this	This study utilized photographic images from four datasets: 1) the Acne Recognition Dataset 2) the CUHK dataset 3) the AR dataset 4) the XM2GTS dataset.	228 people with skin problems rated 3-4 stage on the IGA scale took 3 selfies: frontal, two-thirds to the left, and two-thirds to the right.	Code and dataset available on https://github.com/Princiya1990/CATDeblurring	650

		issue, the study proposes a framework based on generative networks.				
Zhang et al. [25] (2022)	Acne Detection	Ensemble neural network composed of ResNet50-based classification and YOLOv5-based localization modules for simultaneous severity, count, and bounding-box prediction	ACNE04: 1457 facial images with expert-annotated lesion counts, severity grades, and bounding boxes	ACNE04 dataset	Datasets available at: https://www.kaggle.com/datasets/manuelhettich/acne04	
Wang et al. [29] (2022)	Cell Phone App for Facial Acne Severity Assessment	Acne-RegNet, a lightweight model for lesion detection and classification, combined with metadata for personalized severity assessment.	1455 images from ACNE04 and 1515 images from a private dataset at Xiangya Hospital, processed into 80×80 lesion patches and resized for the model.	Clinical photographs, including metadata, collected from a private dataset and the ACNE04 dataset.	Datasets available at: https://www.kaggle.com/datasets/manuelhettich/acne04	2970
Islam et al. [40] (2023)	Acne Vulgaris Detection and Classification	First, two CNN-based units are designed to extract deep feature maps, later combined in a feature aggregation module. The aggregated features provide rich input information and classify the acne by a softmax.	The authors have captured seventy-seven images of this dataset from the subjects who visited the Department of Dermatology at Bangabandhu Sheikh Mujib Medical University (BSMMU) and Dhaka Medical College (DMC).	The informed consent was obtained from all subjects before capturing the acne images by the 13-MP smartphone camera.	Datasets available at http://www.derm.net https://dermnetnz. org	420
Wang et al. [27] (2023)	Acne Detection and Severity Quantification	Two-stage deep learning scheme	1803 facial images (1374 for training/test, 429 for clinical validation)	Smartphone images captured with iPhone 5s at Xiangya Hospital, annotated by dermatologists	No	1803
Huynh et al. [30] (2022)	Automatic Acne Object Detection and Severity Grading Using	Faster R-CNN deep learning model for acne object detection and LightGBM machine learning	1572 facial images with 41,859 labeled lesions, captured by iOS and Android smartphones.	Smartphone images collected via the 'Skin Detective' mobile application.	No	1572

	162					
	Smartphone Images and AI	model for grading acne severity based on the IGA scale.				
Lin et al. [32] (2022)	Unified Acne Grading Framework on Face Images	KIEGLFN framework with Key Information Enhancement and Global Local Fusion Network, including a transfer fine-tuning strategy for new grading criteria.	1457 images from the ACNE04 dataset, 200 images from the PLSBRACNE01 dataset, annotated for lesion location, severity, and type.	Clinical images from the ACNE04 and PLSBRACNE01 datasets.	Code available on GitHub: https:// github.com/linyi0604/KIEGLFN (accessed on 5 March 2025)	1657
Liu et al. [33] (2022)	Ensemble Pruning of Deep Learning Base Models for Acne Grading	Ensemble pruning of deep learning models, fivefold cross-validation, and transfer learning applied to both acne and skin cancer datasets.	1457 images from the ACNE04 dataset, annotated for lesion location and severity grading; 3297 dermoscopic images in the skin cancer dataset, categorized as benign or malignant.	Clinical images from the ACNE04 and skin cancer datasets.	Datasets available at: https://github.com/xpwu95/ldl and https: //www.kaggle. com/datasets/ fanconic/skin-cancer-malignant-vs-benign (accessed on 5 March 2025)	4754
Wen et al. [34] (2022)	Acne Detection and Severity Evaluation	Interpretable CNN-based object detection models, with a focus on interpretability	1457 images from the ACNE04 dataset, with 1222 high-quality images used after data cleaning.	Clinical images with annotations for lesion location, severity, and bounding boxes.	Code available on GitHub: https:// github.com/wenh06/acne_detection/ and https://github.com/wenh06/yolov4_acne_torch	1457
Kim et al. [35] (2023)	Acne Lesion Detection and Counting for Severity Evaluation	CNN-based algorithm, with a focus on binary and five-class classification of acne lesions.	1213 images with 20,699 labeled lesions (closed comedones, open comedones, papules, nodules/cysts, pustules).	Clinical images acquired in standardized settings using Canon EOS 550D and NIKON D7100 cameras.	Detailed code and hyperparameters provided in the Electronic Supplementary Materials.	1213
Wei et al. [37] (2023)	Accurate Acne Detection via Decoupled Sequential Detection Head	Decoupled Sequential Detection Head mechanism applied to mainstream two-stage detectors.	ACNE-DET: 276 high-resolution images with 31,777 labeled lesions; ACNE04: 1457 images with 18,983 bounding box annotations.	Clinical images acquired with VISIA system and annotated by dermatologists.	Code and dataset are publicly available, repository not specified.	276

Chen et al. [44] (2025)	Feature Feedback-Based Pseudo-Label Learning for Multi-Standards in Clinical Acne Grading.	Feature Feedback-Based Pseudo-Label Learning (FF-PLL) (1) an acne feature feedback (AFF) (2) all-facial skin segmentation (AFSS) (3) the AcneAugment (AA) strategy	This study proposes two different standard acne grading datasets, namely ACNE04 and ACNE-ECKH, to validate the effectiveness of the proposed method.	ACNE-ECKH dataset comprises 8242 images for acne image classification, consisting of 1500 labeled images and 6742	Datasets available at: https://www.kaggle.com/datasets/manuelhettich/acne04	9699
Li et al. [38] (2023)	AI-Powered Acne Grading System with Lesion Identification	ResNet50 architecture, with 945 images used for training and 240 for testing, incorporating lesion identification and severity grading.	1501 standardized photos from 1501 acne patients, with detailed labeling of 15,922 lesions across 10 categories.	Clinical images obtained using the VISIA complexion analysis system.	No	1501
Lin et al. [39] (2023)	Acne Severity Grading via Diagnostic Evidence Distillation	Teacher-student structure using CNNs for both global estimation and lesion detection, integrating lesion counting and grading.	ACNE04: 1457 facial images with 13,633 lesions; PLSBRACNE01: 200 images categorized by severity.	Clinical images labeled according to Hayashi and Pillsbury criteria.	Code available on GitHub:https://github.com/linyi0604/DED (accessed on 5 March 2025)	1457
Li et al. [7] (2023)	AI Analysis of Dark Circles and Facial Skin Aging in Chinese Population	CNNs based system for grading facial features from selfies, validated against dermatologist assessments.	Over 1.9 million selfies from 1,939,586 participants, graded for facial aging indicators.	Selfies taken with smartphones, collected via a mobile app.	No	1,939,586
Li et al. [41] (2024)	Acne severity assessment	CNN-based automatic grading	1,156,703 women aged 18-85	Selfies taken with smartphones, collected via a mobile app. Annotated by dermatologists.	No	1,156,703
Kim et al. [42] (2024)	Facial Acne Segmentation	Semi-supervised learning using bidirectional copy-paste and deep learning with U-Net.	Acne images cropped from facial images	Clinical images (likely taken with specialized skin diagnostic equipment)	No	2000

Zhang et al. [43] (2024)	Acne Detection using CenterNet	CenterNet-based deep learning framework, utilizing deep hierarchical aggregation strategies for feature extraction, combined with anchor-free detection for improved accuracy in locating and identifying acne lesions.	153,183 images collected from 15 hospitals in different provinces of China using smartphones.	A mix of clinical and smartphone images	No	153,183
Prokhorov et al. [45] (2024)	Acne Image Grading using Label Distribution Smoothing	Gaussian label distribution generation and label smoothing techniques combined with a scale-adaptive approach, validated on the ACNE04 dataset using ResNet-50 architecture.	ACNE04 dataset with 1457 images and 18,983 bounding boxes of lesions.	Clinical images from the ACNE04 dataset, specifically annotated for acne grading.	Source code is publicly available at http://github.com/ openface-io/acnelds	1457
Zein et al. [14] (2024)	Acne Dataset Generation Using Generative Adversarial Networks (GANs)	GAN-based method with StyleGAN2 to generate synthetic acne images at different severity levels, used for training a CNN-based classification system.	1473 real images initially collected, with additional synthetic images generated across Mild, Moderate, and Severe severity levels.	Synthetic images generated using GANs based on clinical images.	Code and dataset available on GitHub: https: //github.com/hz0 1/AcneGAN/tree/ main	1473
Gao et al. [46] (2025)	Acne lesion detection and severity grading in online and offline scenarios	Deep learning model (AcneDGNet) that combines a vision transformer-based feature extractor with a CNN-based lesion detection module and a severity grading module	2157 facial acne images compiled from two public datasets (ACNE04, AcneSCU) and three self-built datasets (AcnePA1, AcnePA2, AcnePKUIH)	Mixed capture sources: digital cameras, VISIA system, and smartphone images	Datasets available at: https://www.kaggle.com/datasets/manuelhettich/acne04	2157

4 Conclusions

Artificial intelligence-based acne assessment systems are expected to improve dermatological care by providing standardized assessments, reducing variability among observers, and improving clinical effectiveness. However, data discrepancies, limited data set diversity, and difficulties in integrating these tools into existing clinical work processes continue to limit their widespread use.

The widespread use of the method for determining the degree of acne based on artificial intelligence can standardize dermatological studies, allow for further comparison of studies conducted in different regions and clinical conditions, and facilitate the development of treatment methods based on reliable and repeatable data. In addition, the ability of artificial intelligence to process large amounts of data from different skin types can expand our understanding of the spread and geographical distribution of acne. However, although some models have demonstrated high accuracy under controlled conditions [46], the implementation of these results into clinical practice remains a complex task: to date, none of the proposed algorithms have been clinically confirmed, and no research on real-life situations has been published. To confirm the effectiveness of these artificial intelligence systems in everyday dermatological practice, it is necessary to conduct prospective studies with testing in real conditions.

Unlike other dermatological diseases, which have a broad public database supporting the development of a standardized artificial intelligence algorithm, acne is rare in these sources [47]. Confidentiality issues associated with the mass sharing of images of faces affected by acne remain a major obstacle to the creation of large, open datasets. Consequently, most datasets used in studies related to acne are not publicly available [15]. Quattrini et al. [15] used a public data set of faces: Flickr-Faces-HQ (FFHQ), consisting of 70,000 high-quality PNG images of human faces measuring 1024×1024 and containing significant differences in terms of age, ethnicity, and image background. However, since most images did not show acne, the data set was very disproportionate: approximately 88% of images fell under the "No acne" category. This discrepancy limited the study to determining the presence of acne rather than assessing its severity.

In conclusion, none of the published algorithms to date demonstrated the generality, reproducibility, or external and prospective clinical validation necessary for integration into conventional dermatological practice. Future work on acne diagnostics and severity assessment based on artificial intelligence should address these gaps, in particular, the need for prospective real-world clinical trials to assess model effectiveness in real-world settings[46]; the creation of large, publicly accessible and demographically diverse imaging repositories[14]; the study of federated education, allowing for the training of interagency models to maintain confidentiality; and the introduction of intelligible artificial intelligence methods to enhance clinicians' ability to interpret and perceive. Despite current limitations, the results are encouraging: AI-based tools promise to fundamentally change acne research and clinical care, reduce variability between and within observers, and provide truly personalized, timely treatment. By improving model architectures, expanding and diversifying the existing data set, and firmly confirming effectiveness in real conditions, the industry can reveal the full potential of artificial intelligence in these conditions.

References

1. Chen, H., Zhang, T.C., Yin, X.L., Man, J.Y., Yang, X.R., Lu, M.: Magnitude and temporal trend of acne vulgaris burden in 204 countries and territories from 1990 to 2019: An analysis from the Global Burden of Disease Study 2019. *Br. J. Dermatol.* 186, 673–683 (2021).
2. Reynolds, R.V., Yeung, H., Cheng, C.E., Cook-Bolden, F., Desai, S.R., Druby, K.M., Freeman, E.E., Keri, J.E., Gold, L.F.S., Tan, J.K., et al.: Guidelines of care for the management of acne vulgaris. *J. Am. Acad. Dermatol.* 90, 1006.e1–1006.e30 (2024).
3. Sharmin, S., Farid, F.A., Jihad, M., Rahman, S., Uddin, J., Rafi, R.K., Hossan, R., Karim, H.A.: A Hybrid CNN Framework DLI-Net for Acne Detection with XAI. *J. Imaging* 11, 115 (2025).
4. Bae, I.H., Kwak, J.H., Na, C.H., Kim, M.S., Shin, B.S., Choi, H.: A Comprehensive Review of the Acne Grading Scale in 2023. *Ann. Dermatol.* 36, 65–73 (2024).
5. Nichols, J.A., Chan, H.W.H., Baker, M.A.B.: Machine learning: Applications of artificial intelligence to imaging and diagnosis. *Biophys. Rev.* 11, 111–118 (2018).
6. Seité, S., Khammari, A., Benzaquen, M., Moyal, D., Dréno, B.: Development and accuracy of an artificial intelligence algorithm for acne grading from smartphone photographs. *Exp. Dermatol.* 28, 1252–1257 (2019).
7. Li, J., Du, D., Zhang, J., Liu, W., Wang, J., Wei, X., Xue, L., Li, X., Diao, P., Zhang, L., et al.: Development and validation of an artificial intelligence-powered acne grading system incorporating lesion identification. *Front. Med.* 10, 1255704 (2023).
8. Jeong, H.K., Park, C., Henao, R., Kheterpal, M.: Deep Learning in Dermatology: A Systematic Review of Current Approaches, Outcomes, and Limitations. *JID Innov.* 3, 100150 (2022).
9. Singhapathiranage, A.P., Kelaart, D.D.K., Thathsarani, A.A.M., Liyanage, C.D.L., Lankasena, B.N.S.: Optimizing Acne Severity Detection: A Deep Learning Approach with Electronic Medical Record System Integration and Diverse Image Data. *Adv. Technol.* 5(1) (2025).
10. Van den Akker, O.R., Peters, G.Y., Bakker, C., Carlsson, R., Coles, N.A., Corker, K.S., Feldman, G., Moreau, D., Nordström, T., Pickering, J.S., et al.: Generalized Systematic Review Registration Form. Available at: https://osf.io/e8rhd (last accessed March 2025).
11. Wu, X., Wen, N., Liang, J., Lai, Y.-K., She, D., Cheng, M.-M., Yang, J.: Joint Acne Image Grading and Counting via Label Distribution Learning. In: *Proc. IEEE/CVF Int. Conf. Computer Vision (ICCV)*, Seoul, Republic of Korea, pp. 10641–10650 (2019).
12. Zhang, J., Zhang, L., Wang, J., Wei, X., Li, J., Jiang, X., Du, D.: SA-RPN: A Spatial Aware Region Proposal Network for Acne Detection. *IEEE J. Biomed. Health Inform.* 27, 5439–5448 (2023).
13. Zhang, D., Li, H., Shi, J., Shen, Y., Zhu, L., Chen, N., Wei, Z., Lv, J., Chen, Y., Hao, F.: Advancements in acne detection: application of the CenterNet network in smart dermatology. *Front. Med.* 11, 1344314 (2024).
14. Zein, H., Chantaf, S., Fournier, R., Nait-Ali, A.: Generative adversarial networks for anonymous acneic face dataset generation. *PLoS ONE* 19, e0297958 (2024).
15. Quattrini, A., Boër, C., Leidi, T., Paydar, R.: A Deep Learning-Based Facial Acne Classification System. *Clin. Cosmet. Investig. Dermatol.* 15, 851–857 (2022).

16.Shen, X., Zhang, J., Yan, C., Zhou, H.: An Automatic Diagnosis Method of Facial Acne Vulgaris Based on Convolutional Neural Network. *Sci. Rep.* 8, 5839 (2018).
17.Zhao, T., Zhang, H., Spoelstra, J.: A Computer Vision Application for Assessing Facial Acne Severity from Selfie Images. *arXiv preprint* arXiv:1907.07901 (2019).
18.Lim, Z.V., Akram, F., Ngo, C.P., Winarto, A.A., Lee, W.Q., Liang, K., Oon, H.H., Thng, S.T.G., Lee, H.K.: Automated grading of acne vulgaris by deep learning with convolutional neural networks. *Skin Res. Technol.* 26, 187–192 (2019).
19.Alzahrani, S., Al-Bander, B., Al-Nuaimy, W.: Attention Mechanism Guided Deep Regression Model for Acne Severity Grading. *Computers* 11, 31 (2022).
20.Fajarnés, G.P., Santonja, M.M., García, B.D., Lengua, I.L.: Segmentation methods for acne vulgaris images: Proposal of a new methodology applied to fluorescence images. *Skin Res. Technol.* 26, 734–739 (2020).
21.Rashataprucksa, K., Chuangchaichatchavarn, C., Triukose, S., Nitinawarat, S., Pongprutthipan, M., Piromsopa, K.: Acne Detection with Deep Neural Networks. In: *Proc. 2nd Int. Conf. Image Processing and Machine Vision*, Bangkok, Thailand (2020).
22.Nguyen, P.K., Le, T.D., Nguyen, B.A., Nguyen, P.A.: ACNE8M – An acne detection and differential diagnosis system using AI technologies. *VNUHCM J. Sci. Technol. Dev.* 27(3), 3550–3561 (2024).
23.Yang, Y., Guo, L., Wu, Q., Zhang, M., Zeng, R., Ding, H., Zheng, H., Xie, J., Li, Y., Ge, Y., et al.: Construction and evaluation of a deep learning model for assessing acne vulgaris using clinical images. *Dermatol. Ther.* 11, 1239–1248 (2021).
24.Min, S., Kong, H., Yoon, C., Kim, H.C., Suh, D.H.: Development and evaluation of an automatic acne lesion detection program using digital image processing. *Skin Res. Technol.* 19, e423–e432 (2012).
25.Zhang, H., Ma, T.: Acne Detection by Ensemble Neural Networks. *Sensors* 22, 6828 (2022).
26.Bernhard, R., Bletterer, A., Le Caro, M., et al.: Automatic Acne Severity Grading with a Small and Imbalanced Data Set of Low-Resolution Images. *Dermatol. Ther.* 14, 2953–2969 (2024).
27.Wang, J., Wang, C., Wang, Z., Hounye, A.H., Li, Z., Kong, M., Hou, M., Zhang, J., Qi, M.: A novel automatic acne detection and severity quantification scheme using deep learning. *Biomed. Signal Process. Control* 84, 104803 (2023).
28.Aldi, Y., Kusnawi.: Multi-Class Facial Acne Classification using the EfficientNetV2-S Deep Learning Model. *AVITEC* 7, 299–309 (2025).
29.Wang, J., Luo, Y., Wang, Z., Hounye, A.H., Cao, C., Hou, M., Zhang, J.: A cell phone app for facial acne severity assessment. *Appl. Intell.* 53, 7614–7633 (2022).
30.Huynh, Q.T., Nguyen, P.H., Le, H.X., Ngo, L.T., Trinh, N.-T., Tran, M.T.-T., Nguyen, H.T., Vu, N.T., Nguyen, A.T., Suda, K., et al.: Automatic Acne Object Detection and Acne Severity Grading Using Smartphone Images and Artificial Intelligence. *Diagnostics* 12, 1879 (2022).
31.Watanabe, K., Iinuma, K., Nakashima, C., et al.: Deep Learning-Based Acne Severity Classification Using Standardized Facial Images of Japanese Patients. *Cureus* 17(9), e91944 (2025).
32.Lin, Y., Jiang, J., Ma, Z., Chen, D., Guan, Y., You, H., Cheng, X., Liu, B., Luo, G.: KIEGLFN: A unified acne grading framework on face images. *Comput. Methods Programs Biomed.* 221, 106911 (2022).
33.Liu, S., Fan, Y., Duan, M., Wang, Y., Su, G., Ren, Y., Huang, L., Zhou, F.: AcneGrader: An ensemble pruning of the deep learning base models to grade acne. *Skin Res. Technol.* 28, 677–688 (2022).

34. Wen, H., Yu, W., Wu, Y., Zhao, J., Liu, X., Kuang, Z., Fan, R.: Acne detection and severity evaluation with interpretable convolutional neural network models. *Technol. Health Care* 30, 143–153 (2022).
35. Kim, D.H., Sun, S., Cho, S.I., Kong, H.-J., Lee, J.W., Lee, J.H., Suh, D.H.: Automated Facial Acne Lesion Detecting and Counting Algorithm for Acne Severity Evaluation and Its Utility in Assisting Dermatologists. *Am. J. Clin. Dermatol.* 24, 649–659 (2023).
36. Mascarenhas, P.P., Sannidhan, M.S., Pinto, A.J., Camero, D., Martis, J.E.: Improving acne severity detection: a GAN framework with contour accentuation for image deblurring. *Front. Bioinform.* 5, 1485797 (2025).
37. Wei, X., Zhang, L., Zhang, J., Wang, J., Liu, W., Li, J., Jiang, X.: Decoupled Sequential Detection Head for accurate acne detection. *Knowledge-Based Syst.* 284, 111305 (2023).
38. Li, T., Ma, X., Li, Z., Yu, N., Song, J., Ma, Z., Ying, H., Zhou, B., Huang, J., Wu, L., et al.: Artificial intelligence analysis of over a million Chinese men and women reveals level of dark circle in the facial skin aging process. *Skin Res. Technol.* 29, e13492 (2023).
39. Lin, Y., Jiang, J., Chen, D., Ma, Z., Guan, Y., Liu, X., You, H., Yang, J.: DED: Diagnostic Evidence Distillation for acne severity grading on face images. *Expert Syst. Appl.* 228, 120312 (2023).
40. Islam, M.B., Junayed, M.S., Sadeghzadeh, A., Anjum, N., Ahsan, A., Shah, A.F.M.S.: Acne Vulgaris Detection and Classification: A Dual Integrated Deep CNN Model. *Informatica* 47(4) (2023).
41. Li, T., Ma, X., Li, Z., Yu, N., Song, J., Ma, Z., Ying, H., Zhou, B., Huang, J., Wu, L., et al.: Facial adult female acne in China: An analysis based on artificial intelligence over one million. *Skin Res. Technol.* 30, e13693 (2024).
42. Kim, S., Yoon, H., Lee, J.: Semi-Supervised Facial Acne Segmentation Using Bidirectional Copy–Paste. *Diagnostics* 14, 1040 (2024).
43. Zhang, J., Zhang, L., Wang, J., Wei, X., Li, J., Jiang, X., Du, D.: Learning High-quality Proposals for Acne Detection. *arXiv preprint* arXiv:2207.03674 (2022).
44. Chen, Y.-Y., Chan, H.-T., Wang, H.-C., Wang, C.-S., Chen, H.-H., Chen, P.-H., Chen, Y.-J., Hsu, S.-H., Hsia, C.-H.: Feature Feedback-Based Pseudo-Label Learning for Multi-Standards in Clinical Acne Grading. *Bioengineering* 12, 342 (2025).
45. Prokhorov, K., Kalinin, A.A.: Improving Acne Image Grading with Label Distribution Smoothing. In: *Proc. IEEE Int. Symp. Biomedical Imaging (ISBI)*, Athens, Greece, pp. 1–5 (2024).
46. Gao, N., Wang, J., Zhao, Z., Chu, X., Lv, B., Han, G., Ni, Y., Xie, G.: Evaluation of an acne lesion detection and severity grading model for Chinese population in online and offline healthcare scenarios. *Sci. Rep.* 15, 1119 (2025).
47. Choy, S.P., Kim, B.J., Paolino, A., Tan, W.R., Lim, S.M.L., Seo, J., Tan, S.P., Francis, L., Tsakok, T., Simpson, M., et al.: Systematic review of deep learning image analyses for the diagnosis and monitoring of skin disease. *npj Digit. Med.* 6, 180 (2023).
48. Haggenmüller, S., Schmitt, M., Krieghoff-Henning, E., Hekler, A., Maron, R.C., Wies, C., Utikal, J.S., Meier, F., Hobelsberger, S., Gellrich, F.F., et al.: Federated Learning for Decentralized Artificial Intelligence in Melanoma Diagnostics. *JAMA Dermatol.* 160, 303–311 (2024).