Acne Vulgaris Detection and Classification: A Dual Integrated Deep CNN Model

Md Baharul Islam[1,2,*], Masum Shah Junayed[2,3], Arezoo Sadeghzadeh[2], Nipa Anjum[4], Afsana Ahsan Jeny[2,3], A. F. M. Shahen Shah[5]
[1]Department of CSE, Florida Gulf Coast University, Fort Myers, FL, USA
[2]Department of Computer Engineering, Bahcesehir University, Istanbul, Turkey
[3]Department of CSE, University of Connecticut, Storrs, CT, USA
[4]Department of CSE, Khulna University of Engineering & Technology, Khulna, Bangladesh
[5]Department of Electronics and Communication Engineering, Yildiz Technical University, Turkey
E-mail: bislam.eng@gmail.com
*Corresponding Author

Recognizing acne disease and evaluating its type is vital for the efficacy of the medical treatment. This report collects a dataset of 420 images and then labels them into seven different classes by a well-experienced dermatologist. After a pre-processing step, including local and global contrast enhancement and noise removal by a smoothing filter, the dataset size is enhanced using augmentation. The images of the dataset and the augmented ones are all fed into a novel integrated dual deep convolutional neural network (CNN) model to recognize acne disease and its type by classifying it into seven groups. 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 proposed architecture's optimizer, loss function, and activation functions are all tuned so that both CNN units are trained with minimum kernel size and fewer training parameters. Thus, the computational cost is minimized. Compared with three machine learning-based classifiers and five pre-trained models, our model achieves competitive state-of-the-art performance with an accuracy of 97.53% on the developed dataset.

Povzetek: Razviti sistem uporablja CNN model za prepoznavanje in klasifikacijo aken (acne vulgaris) z natančnostjo 97.53% na naboru 420 slik.

1 Introduction

Acne is an unwanted skin disease occurring in the pilosebaceous unit. It is most commonly seen on the face, forehead, chest, upper back, and shoulders. Acne is usually related to hormonal fluctuations during adolescence, although some adults continue to experience acne into their 40-50s, too (caused by oil and dead skin cells) [1]. According to the surveys conducted in [2, 3], around 80% of adolescents and young adults are affected by acne, and approximately 40-50 million Americans have acne problems. Acne causes pain, redness, bleeding, and many other physical problems. Its psychological and emotional effects on patients can be far worse than physical issues. The changes in the beauty of the skin's appearance result in several psychological problems such as anger, depression, anxiety, fear, shame, embarrassment, low self-esteem, poor self-image, etc. Acne also has negative impacts on the social life of the patients, e.g., lack of confidence, limited employment chances, social withdrawal, and suicidal tendencies at worst [7].

Dermatologists diagnose acne by a simple visual inspection based on comedones, pustules, nodules, cysts, etc. It is a subjective diagnosis depending on the experts' experience and ability. There is no particular test for acne, and in only special and critical cases, the X-ray, CT scan, or MRI tests are suggested by dermatologists [40]. Some professionals occasionally employ dermoscopic images for clinical diagnosis [9]. However, these images are acquired by a non-invasive method which is time-consuming. Additionally, there are several skin analysis systems, e.g., VISIA from Canfield and ANTERA 3D from Miravex, which are also expensive and cannot always detect acne accurately. These types of equipment also require to be operated and analyzed by well-experienced experts. Due to the lack of dermatologists, especially in under-developing countries, people do not receive timely treatments for acne. Even in developed countries, it is estimated that for an appointment with a dermatologist, patients are asked to wait for an average of 32 days[1], which delays the treatment procedure. On the other hand, if some types of acne, such as cysts (caused by severe infection), are not cured on time, they are likely to turn into permanent scars, or they need to be surgically removed by

---

[1]https://www.firstderm.com/appointment-wait-time-see-dermatologist/

dermatologists.

Hence, it is vital to provide an automated system to recognize acne and identify its type. In this case, patients can receive timely treatment without expensive equipment or expert help at the initial stage. Generally, providing such a system is challenging due to several main reasons: (a) high diversity of acne lesions and human skin tones, (b) significant variations in the size, shape, and position of acne, and (c) dependency on the age, gender, and skin types.

Skin disease detection, especially acne detection based on deep neural networks and machine learning techniques, has attracted much attention among researchers. Several approaches have been reported in the literature in the last decade [12, 16, 14, 34, 29]. Even some of them are for acne severity grading [27, 35, 39, 10]. However, limited research has been carried out in the acne classification [5, 32, 21] while it is a vital issue in getting appropriate treatment in the early stages. The main challenge of acne recognition systems is their inability to classify different types of acne vulgaris. On the other hand, acne classification approaches are also required to improve their performance.

Another main challenge in acne classification is the lack of an appropriate and publicly available dataset. One of the earliest datasets was used in [5] with 35 images in 5 different types (i.e., nail, comedones, papules, pustule, and nodule). However, these small numbers of images are not enough, especially for deep learning-based systems. Later, another dataset was developed in [32] with 3000 skin images in 7 classes (including normal skin, papule, cyst, blackhead, pustule, whitehead, and nodule).

Although this dataset is rich in the number of images and classes and sufficient for deep learning models, it is not publicly available. We have even tried to contact the corresponding author through email to access their dataset, but they are not reachable. Additionally, the dataset does not include two more acne types, i.e., excoriated and keloidalis acne. Another dataset was proposed in [21] (collected from `http://dermnet.com`) with 300 images in 5 classes (i.e. closed comedo (whitehead), cystic, keloidalis, open comedo (blackhead), pustular). Although it includes keloidalis acne and has a reasonable number of images, it still suffers from a lack of excoriated acne and has limited classes.

In [29], another dataset was utilized with total 871 images (in 4 classes). This dataset also has a limited number of classes and is not publicly available. Recently, a publicly available dataset has been developed in [35] called "ACNE04," which is suitable only for grading acne in 4 classes (i.e., mild, moderate, severe, and very severe), not for acne type classification.

A new acne dataset with seven classes has been developed to address the abovementioned challenges. Then, an acne recognition and classification system are proposed based on an integrated dual deep CNN model. Overall, the main contributions of this work can be summarized as follows:

- Due to the unavailability of the public acne classification dataset, an acne dataset is created, including 420 acne images in 7 classes. Contrary to the datasets proposed in [32] and [21], our developed dataset contains two different types of acne, i.e., Keloids and Excoriated. So, suppose the dataset of [32] is publicly available. In that case, its combination with ours results in a dataset with nine classes which can be significantly valuable for further research and analysis on acne classification problems.

- Providing an acne disease-free skin (DF) class (including normal skin and the skin with other diseases such as Eczema, Skin Cancer, Fungal Infection, etc.) enables our method not only to classify acne into six different types but also to distinguish it from the other skin diseases.

- The images of the dataset are captured using a smartphone camera. As smartphones are extensively available and used by almost everyone, the proposed method can be considered a remote screening system that ordinary people can efficiently utilize without needing expensive equipment and expert help.

- An integrated dual CNN-based automatic acne recognition and classification system is proposed. In this model, the extracted feature maps from two CNN models (without fully connected layers) are concatenated and aggregated, resulting in comprehensive rich information from input images and so high acne classification accuracy. The kernel size, the training parameters, and the number of layers is minimized by adjusting the optimizer, loss function, and activation functions, which reduces computational cost while the accuracy is still high.

- The performance of the proposed method is compared with the results of three machine learning-based classifiers, i.e., KNN, SVM, and MLP, and five pretrained deep learning-based models, i.e., GoogleNet, MobileNet, VGG-19, ResNet-50, and AlexNet. Our model receives competitive performance in the recognition and classification of acne.

The rest of the paper is organized as follows: Section 2 reviews the related works reported in the literature for acne detection, classification, and grading. The developed dataset, its setup, pre-processing, and augmentation steps are all explained in Section 3. Our proposed model is presented in detail in Section 4. Experimental results and the related discussions and comparisons are demonstrated in Section 5. Finally, conclusions are drawn in Section 6 as well as the future directions.

2 Related works

Many approaches have been reported in the literature for acne detection, classification, and grading in the last

decade. Generally, these approaches can be categorized into two main groups: conventional computer vision- and machine learning-based and deep learning-based methods. This section summarizes some of the primary studies in these two groups.

2.1 Conventional approaches

Conventional automatic acne detection/recognition methods are primarily based on feature extraction for classification. Malik et al. [27] proposed an acne grading system based on feature extraction and support vector machine (SVM). They classified the acne severity into four classes, i.e., mild, moderate, severe, and very severe. Khongsuwan et al. [24] proposed a method for counting the number of points for acne vulgaris. Their system achieved a prediction accuracy of 83.75% on the cropped part of the skin. The Ultra-Violet (UV) fluorescence light was applied to capture images before converting them to Gray-Scale and RGB. In this method, the quality of the images was improved using adaptive histogram equalization, and the number of points (acne) was counted based on the extended maxima transform. This image processing technique can quickly analyze acne images, but it would become tricky when the number of images is high.

Later, Alamdari et al [5] developed a mobile application for acne detection, classification and segmentation. They collected 35 images in five classes from various dermatology resources. They used Fuzzy C-Min (FCM) to classify images with no acne (normal skin) from the skin images with acne disease (with an accuracy of 100%). Moreover, they performed another classification task based on SVM (with the linear kernel) and fuzzy c-means techniques to distinguish acne scarring from inflammatory acne. This classification task achieved an average accuracy of 80% and 66.6% for FCM and linear SVM methods, respectively. As the images of their dataset were captured by a cellphone, they suffered from difficulty visualizing the small lesion and ascertaining the depth of involvement. Additionally, they used a limited number of images for segmentation and classification, i.e., 35 images, while more images from more subjects are required to provide an accurate evaluation.

Kittigul et al. [25] detected acne based on robust applied features and then classified it using five designed features. They achieved an average accuracy of 68%, which is insufficient for clinical purposes. Hameed et al. [13] presented a hybrid technique using Naive Bayes Classifier (NBC) and image processing to detect and classify acne into three different types. Using 40 images in each of the three classes, they achieved an accuracy of 93.42%. Acne patterns were segmented using adaptively regularized fuzzy c-means (ARFCM) clustering technique and Morphological opening, creating the mask for all training images. Fourteen Haralick features were extracted from all patterns of the masks, which were later fed into NBC to perform the classification.

Although numerous approaches have been proposed based on traditional image processing techniques, they still suffer from noise and low accuracy due to the variations in characteristics of acne vulgaris, such as color variations and color complexity.

2.2 Deep learning-based methods

Deep neural networks (DNNs) such as CNNs have been extensively used for image classification. According to their high recognition and classification ability, sometimes they can perform even better than human beings in specific tasks, such as traffic sign recognition, face recognition, and handwriting digit recognition [30, 15]. Contrary to conventional computer vision methods, CNN-based models extract more and deeper features which enhance the classification accuracy and enable the system to deal with more classification types [11]. Hence, they are the focus of interest among the researchers not only for acne detection and classification but also in every field of medical image analysis [23, 8, 26, 22, 38, 28, 17, 33, 18, 19, 20].

Shen et al. [32] proposed a new automatic CNN-based diagnosis method for facial acne vulgaris to classify different types of acne vulgaris. This method extracted image features based on CNNs, classified by a classifier. The skin area was detected by applying a binary classifier for skin and non-skin classes. Then, the type of acne was determined using a seven-class classifier. Zhao et al. [39] proposed a grading system to assess the severity of facial acne vulgaris using 4,700 selfie images in 5 groups from "clear" to "severe." Based on the transfer learning approach, the features of the images were extracted using a pre-trained model (ResNet 152). Then, the target severity level was learned from the labeled images by adding and training a fully connected layer. The irrelevant background was minimized using OpenCV models to find facial landmarks. Key skin patches were extracted from the selfie images based on these landmarks. They trained their model after rolling each skin patch to improve testing results. The Root Mean Squared Error (RMSE) was 0.482 when applying the skin patch rolling data augmentation.

Junayed et al. [21] utilized Deep Residual Neural Network to build a model called "AcneNet" in which 1800 acne images (original images plus the augmented ones) in five classes were used. Training, validation, and testing accuracy were 86.28%, 86.11%, and 95.89%, respectively. Two pre-trained models, Inception-V3 and MobileNet, were also implemented on the same dataset and compared with their proposed method. Although this method was slightly underfitting, its performance was competitive.

Alom et al. [6] worked on skin cancer segmentation and classification using dermoscopic images. They proposed NABLA-N Net based on the R2U-Net model, which was composed of three different architectures: NABLA-N Net (A), NABLA-N Net (B), and NABLA-N Net (AB). An Inception Recurrent Residual Convolution Network (RRCNN) was used for recognizing skin cancer from der-

moscopic images. They used the transfer learning technique with NABLA-2 Net (AB) and got a testing accuracy of 96.36% without applying any augmentations. This accuracy was increased to 96.03% by employing augmentation along with transfer learning. They classified images into seven classes and got a testing accuracy of 81.12% without using data augmentation. This accuracy was also increased to 87.09% after applying data augmentation. Although this model was not proposed to detect acne, it performed well in skin cancer segmentation and classification.

Another automatic diagnosis system for skin disease was proposed by Shanthi et al. in [31] based on AlexNet architecture. They used the DermNet dataset in which 105 images were used for training (its 10% was taken as validation) and 69 images for testing. The dataset had four classes: Acne, Urticaria, Eczema Herpeticum, and Keratosis. They obtained 96.32% training accuracy and 62.1% validation accuracy. Testing accuracy for each type of Acne, Keratosis, Eczema Herpeticum, and Urticaria was achieved as 85.7%, 92.3%, 93.3%, and 92.8%, respectively.

Rashatapruksa et al. [29] tried to overcome the weak performance of traditional image processing techniques in acne detection and classification. They compared the performance of the Faster Region-based Convolution Neural Network (Faster R-CNN) and Region-based Fully Convolutional Network (R-FCN) on a dataset with 871 images (four classes of acne). Achieving a mean average precision of 28.3% for R-FCN, they proved that R-FCN performed comparatively better than Faster R-CNN. Although this method was more accurate and faster than traditional image-processing methods, its accuracy was still low for real-life clinical applications.

3 Dataset development

Identifying the type of acne is a crucial factor for having a successful treatment. One of the main challenges in automatic acne classification is providing a proper dataset. It is essential to have a dataset with a sufficient number of images and classes, especially for deep learning-based systems. A dataset of 420 pictures in 7 different types is proposed to overcome this challenge. These seven classes include Acne of Closed Comedo (ACC with 68 photos), Acne of Cystic (AC with 50 images), Acne of Excoriated (AE with 56 images), Acne of Keloidalis (AK with 71 images), Acne of Open Comedo (AOC with 53 photos), Acne of Pustular (AP with 62 images), and acne disease-free skin (DF with 60 images) (including the images of normal skin and the skin images with other diseases rather than acne).

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 cam-

era. Some samples of these 77 images are illustrated in Fig. 2. An example for each of these acne types with the related descriptions is illustrated in Fig. 1. The rest of the images have been collected from public platforms of Baumann Cosmetic Dermatology (`http://www.derm.net/`) and New Zealand Dermatologists (`https://dermnetnz.org/`). All of the images in our dataset have been labeled by a well-experienced dermatologist in 7 classes. Due to different image sizes, all images are resized to 224×224 images.

4 Proposed CNN model

The main flowchart of the proposed method is illustrated in Fig. 3. This figure shows that the input images are fed into two CNN-based models after passing the pre-processing step and augmented. The extracted features from these dual CNN models are concatenated, aggregated using fully connected layers, and passed to the softmax classifier for classification. The details related to each step are presented in the following subsections.

4.1 Preprocessing and augmentation

Contrast is essential in medical imaging to better represent the images, especially for acne recognition and classification. Unlike the reported approaches in the literature in which only one contrast enhancement technique, i.e., local contrast or global contrast, was applied, a novel preprocessing technique is proposed in this paper through which local and global contrasts are incorporated. The primary goal of combining the local and global contrasts is to create an informational image that clearly shows the acne's location while simultaneously improving the image quality. A novel statistical function is generated to enhance the local contrast of the acne images in the region of $Q(m, n)$ by using the local mean of (E) and local standard deviation of (σ). The following equations (1 and 2) calculate the E and σ, respectively:

$$E(Q(m,n)) = \frac{1}{(2u+1)^2} \sum_{m=0}^{h} \sum_{n=0}^{w} (Q(m,n)) \quad (1)$$

$$\sigma = \sqrt{\frac{1}{(2u+1)^2} \sum_{m=0}^{h} \sum_{n=0}^{w} [Q(m,n) - E(Q(m,n))]} \quad (2)$$

where $(2u+1)^2$ and $E(Q(m,n))$ represent the local contrast and the mean of the original input image, respectively. Here, $Q(m,n)$ denotes a region with the height of m and the width of n in 3 channels of RGB, i.e. $(m \times n \times 3)$, in which $(m, n) \in \mathbb{R}$. A statistical function that utilizes these parameters is described as:

$$Q_L(m,n) = E(Q(m,n)) + \phi[Q(m,n) - E(Q(m,n))]$$
$$(3)$$

Type/Photo	Description	Type/Photo	Description
ACC	It seems as though tiny pimples have risen to the surface of the skin. Unlike common pimples, these lesions are not inflammatory and do not pain. When a clogged comedo forms, it is because a block of skin cells and oil got stuck within the hair follicle.	AK	Acne Keloidalis is an inflammatory disease that lasts a long time. It can cause blisters and lesion-s on the back of the neck, which can progress to scars.
AC	Acne Cystic is the most severe kind of acne. These are formed when cysts grow deep beneath the skin. Bacteria, oil, and dry skin cells may gather in pores, creating an environment that encourages acne.	AOC	Whitehead is essentially the same as folliculitis, except that material within the follicle gets oxidized, and the follicles become blocked. Its sebaceous glands get blocked by excessive sebum. Pores filled with blackheads are known as open comedones.
AE	A kind of acne caused by an atypical circumstance is excoriated acne. To peel off the skin is the meaning of the term excoriate. Blotchy, picky, and touchy acne is a characteristic of overwhelming urges to pinch or squeeze every pore on the skin, even minor blemishes.	AP	Papules are almost identical to pustules. However, pustules have a white or yellow substance on the surface that looks like sebum.

Figure 1: Six different types of acne, including Acne of Closed Comedo (ACC), Acne of Cystic (AC), Acne of Excoriated (AE), Acne of Keloidalis (AK), Acne of Open Comedo (AOC), and Acne of Pustular (AP), and their characteristics.
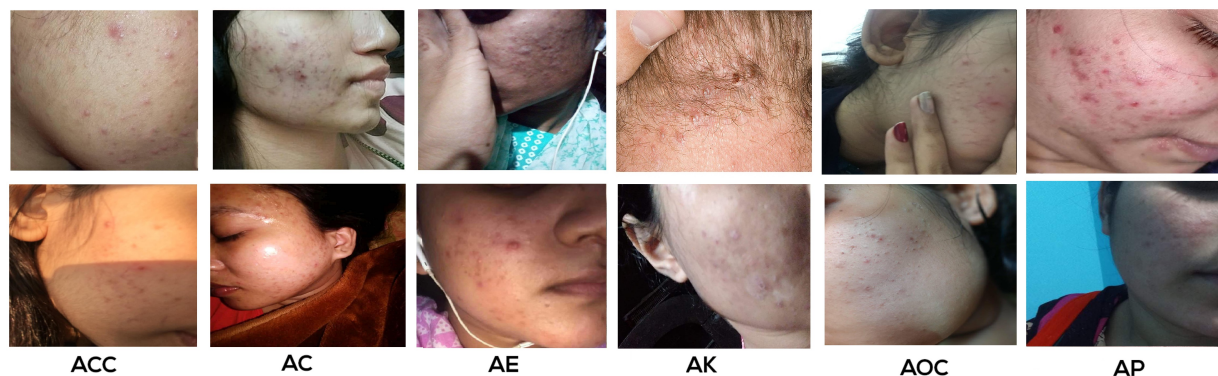

Figure 2: A sample of collected acne dataset

Here, the local contrast-enhanced image and the contrast gains (range greater than 1) are represented by $Q_L(m, n)$ and ϕ, respectively. Afterward, a top-hat maximization technique is used to enhance the global contrast. The top-hat filter operation is accomplished and stated as follows:

$$Q_{top}(m, n) = Q(m, n) - Q(m, n)^{\circ} s_e \qquad (4)$$

In this step, $\circ$ and s_e denote the opening operation and the structural element, respectively. The opening operation is employed here to boost the global contrast, and the procedure is simple enough to be completed in a minimal amount of time. The output of $Q_{top}(m, n)$ is fed into the maximizing function, which creates the improved picture. The final enhanced image is created by combining the outputs of $Q_g(m, n)$ and $Q_{con}(m, n)$ calculated by equations 5 and 6, respectively:

$$Q_g(m, n) = \max_{z \in \Delta(m)} \left(\max_{\alpha \in ((m,n))} (Q_{top}(m, n)) \right) \qquad (5)$$

$$Q_{con}(m, n) = \sum (Q_g(m, n), Q_L(m, n)) - Q(m, n) \qquad (6)$$

It is worth mentioning that, in addition to the contrast enhancement, all images pass a smoothing filter through which their probable noises are removed.

Providing a considerable amount of training data is critical in Deep Learning (DL)-based models. If the dataset number is enormous, the overall model has a very flexible function with many tunable parameters for training. Additionally, increasing the number of training data in CNN reduces the probability of overfitting, generalizes the model to different input patterns, and so makes it robust [37]. Thus, to take advantage of the essential training data, seven other augmentation techniques, i.e., scaling, flipping horizontally, rotating 30° randomly to the right or left, shading, padding, affine transformation, and translation, are employed. These augmentations enhance the number of images in the dataset by producing additional images equal to seven times the original set.
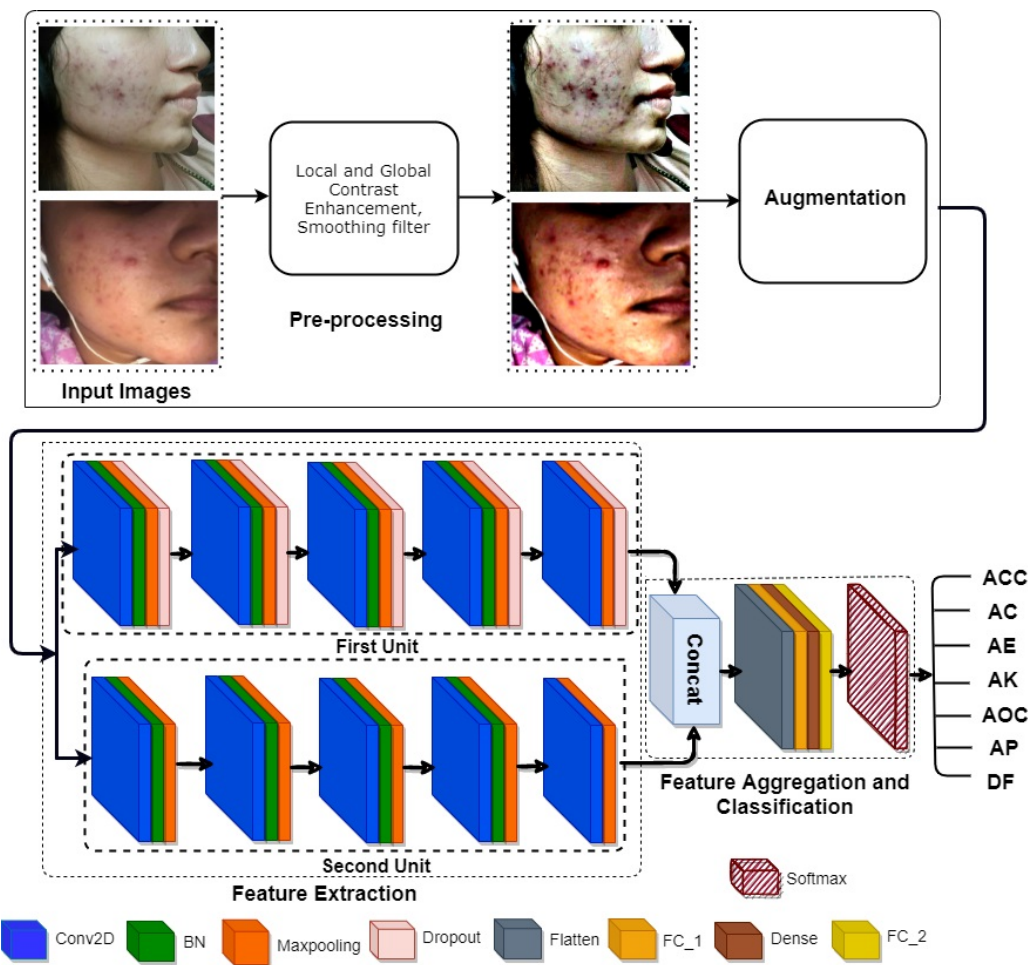
Figure 3: An overview of the proposed deep CNN-based acne classification system for identifying and categorizing acne vulgaris. In this system, the input images are fed into an integrated dual-CNN model after applying pre-processing and augmentation.

4.2   Dual CNN-based feature extractor

With the emergence of high GPUs and the considerable number of data, deep CNN-based models have been the focus of interest and extensively applied for detecting and classifying disease images in the last decade[36]. Deep neural networks with many layers obtain high accuracy for feature extraction and classification. However, increasing the number of layers raises the requirement for many training images, parameters, and high computational time. An integrated deep CNN-based feature extractor is proposed for acne recognition and classification to solve these limitations and get highly informative feature maps with fewer layers. In our proposed method, two separated CNN models (with no fully connected and classification layers), i.e., first and second units in Fig. 3, are designed and trained parallel for feature extraction.

The architecture details of these two CNN models, including the number of layers and the filter, kernel, and output sizes, are all summarized in Table 1. This table shows that both CNN units take the input images of size $224 \times 224 \times 3$ (RGB images with three channels). The

first unit comprises five convolution blocks, containing a convolutional layer, a Max Pooling (MP) layer (for reducing the space size for data representation), and two regularization layers, i.e., a batch normalization layer (BN) and a dropout layer. Regularization is a strategy to improve the model by changing the learning algorithm. It also improves the model's performance on invisible information, reduces overfitting, and enhances generalization with improved convergence. In this unit, the filters in 5 blocks are 16, 32, 64, 96, and 128, respectively. The kernel size (KS) in all convolutional and MP layers are 3×3 and 2×2, respectively. The padding for the first three convolutional layers is applied as "same" and for the rest two convolutional layers as "valid." ReLu activation function is used in all of the convolutional layers as follows:

$$RELU(x) = MAX(0, X) \tag{7}$$

The negative values of the matrix are considered 0, and the positive values are kept unchanged. Five dropout layers in 5 blocks are set as 0.25, 0.25, 0.4, 0.4, and 0.25, which means 25%, 25%, 40%, 40%, 40%, and 25% of neurons in

Table 1: The summary of the proposed two CNN feature extractors, including layers, their configurations, and output shape.

First Unit					Second Unit				
Layers	**Filter**	**Configuration**	**Stride**	**Output Shape**	**Layers**	**Filter**	**Configuration**	**Stride**	**Output Shape**
Conv2D	16	KS: 3×3; padding: same; ReLU	2	$224 \times 224 \times 16$	Conv2D	32	KS: 7×7; padding: same; ReLU	1	$224 \times 224 \times 32$
BN	-	-	-	$224 \times 224 \times 16$	BN	-		-	$224 \times 224 \times 32$
MP	-	KS: 2×2	2	$112 \times 112 \times 16$	MP	-	KS: 2×2	2	$112 \times 112 \times 32$
Dropout	-	0.25	-	$112 \times 112 \times 16$	Conv2D	32	KS: 5×5; padding: same; ReLU	1	$112 \times 112 \times 32$
Conv2D	32	KS: 3×3; padding: same; ReLU	2	$112 \times 112 \times 32$	BN	-		-	$112 \times 112 \times 32$
BN	-		-	$112 \times 112 \times 32$	MP	-	KS: 2×2	2	$56 \times 56 \times 32$
MP	-	KS: 2×2	2	$56 \times 56 \times 32$	Conv2D	48	KS: 3×3; padding: same; ReLU	1	$56 \times 56 \times 48$
Dropout	-	0.25	-	$56 \times 56 \times 32$	BN	-		-	$56 \times 56 \times 48$
Conv2D	64	KS: 3×3; padding: same; ReLU	1	$56 \times 56 \times 64$	MP	-	KS: 2×2	2	$28 \times 28 \times 48$
BN	-	-	-	$56 \times 56 \times 64$	Conv2D	64	KS: 3×3; padding: same; ReLU	2	$28 \times 28 \times 64$
MP	-	KS: 2×2	2	$28 \times 28 \times 64$	BN	-	-	-	$28 \times 28 \times 64$
Dropout	-	0.4	-	$28 \times 28 \times 64$	MP	-	KS: 2×2	2	$14 \times 14 \times 64$
Conv2D	96	KS: 3×3; padding: valid; AF: ReLU	1	$28 \times 28 \times 96$	Conv2D	128	KS: 3×3; padding: same; ReLU	2	$14 \times 14 \times 128$
BN	-	-	-	$28 \times 28 \times 96$	MP	-	KS: 2×2	2	$7 \times 7 \times 128$
MP	-	KS: 2×2	2	$14 \times 14 \times 96$					
Dropout	-	0.4	-	$14 \times 14 \times 96$					
Conv2D	128	KS: 3×3; padding: valid; ReLU	1	$14 \times 14 \times 128$					
MP	-	-	-	$7 \times 7 \times 128$					
Dropout	-	0.25	-	$7 \times 7 \times 128$					

hidden layers. These are set to 0 at each training phase update to prevent the model from overfitting while improving the accuracy. The output of the first CNN unit is a set of feature maps produced from the last dropout layer and has a high level of detail on acne disorders useful for acne type classification.

Similarly, the second unit contains five convolution blocks, each composed of a convolutional layer, a BN layer, and an MP layer, but no dropout layers. Another difference with the first unit is its filters 32, 32, 48, 64, and 128 in five blocks, respectively. The kernel sizes in the first and second convolutional layers are 7×7 and 5×5, respectively. In the rest three convolutional layers, the kernel size is 3×3. Padding in all convolutional layers is applied as "same." Other characteristics of the second unit, such as the kernel size in the MP layer and the activation function, are the same as the first unit. The second unit provides different feature maps from the first one.

4.3 Feature aggregation and classification

The acquired sets of feature maps from two CNN-based feature extractors have specific information about the input images. Consequently, their combination forms a comprehensive feature map, resulting in rich, robust, deep information from the inputs and high classification accuracy. As illustrated in Fig. 3, these two sets of feature maps are first assembled in the concatenation layer to obtain powerful feature aggregation and high-dimension feature representation with fewer semantic correlations. More discriminative shape information is provided by aggregating all features using a flattened layer, two fully connected layers (FC_1 and FC_2), and a dense layer. The number of neurons employed in FC_1, dense, and FC_2 layers are 1048, 128, and 512, respectively. The final aggregated features are fed into a softmax classifier to recognize and classify the acne disease.

To better predict and classify, categorical cross-entropy is employed as the loss function, and the model is trained using the ADAM optimizer. The multi-class cross-entropy loss function is defined as follows:

$$Loss = -\sum_{i=1}^{N} y_i log(\hat{y}_i) \qquad (8)$$

where $y_i = \begin{cases} 1, & \text{if the element is in class } i \\ 0, & \text{otherwise} \end{cases}$ and $\hat{y}_i$ is the probability that the element is in class i. The minus sign shows that the loss value gets smaller as the distributions become closer. Adam Optimizer helps the CNN model minimize errors, making it more reliable and efficient. In our proposed model, we use the automatic learning rate reduction technique. The initial learning rate is 0.0003.

5 Experimental results and discussion

In this section, the experimental setup and results are presented. Additionally, the performance of our proposed method is compared with three conventional machine learning classifiers and five pre-trained models, i.e., GoogleNet, MobileNet, VGG-19, ResNet-50, and AlexNet, on our developed dataset.

5.1 Experimental setup

All experiments presented in this paper are carried out using an intel core i9 PC with a 3.60 GHz CPU, 64 GB RAM, and Nvidia Geforce Rtx 2080 super GPU with 8 GB video RAM. All training and testing are conducted in an Anaconda python environment with a visual code editor using Keras and TensorFlow frameworks. The number of epochs is defined as 60 with a batch size of 32.

Evaluation metrics

The performance of the proposed method is evaluated in terms of accuracy ($\frac{(TP+TN)}{(TP+TN+FP+FN)}$), Precision ($\frac{(TP)}{(TP+FP)}$), Recall/Sensitivity ($\frac{(TP)}{(FN+TP)}$), F1-

score $(\frac{2*(Precision*Recall)}{(Precision*Recall)})$, Specificity $(\frac{(TN)}{(FP+TN)})$, and Matthews Correlation Coefficient (MCC= $\frac{(TP*TN)(FP*FN)}{\sqrt{((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN))}}$) Score. In these evaluation metrics, TP, TN, FP, and FN stand for True Positive, True Negative, False Positive, and False Negative, respectively. True Positive refers to the correct prediction done by the classifier when the actual class of the data and the predicted class are both 1 (True). On the other hand, when the actual class is 0 (False) and the predicted class is also 0 (False), it is considered True Negative. In False Positive, the actual class of the data is 0 (False), while the classifier predicts it as 1 (True). It is named False because the model mispredicted the class and Positive due to the predicted class being 1 (True). Conversely, a False Negative happens when the actual class is 1 (True) and the predicted one is 0 (False). Similarly, False shows the misclassification, and Negative refers to the predicted class as 0 (False).

Accuracy, as the classification rate, is defined as the number of correct predictions divided by the total number of predictions. Recall (Sensitivity) is the true positive rate that observes the actual positive values correctly identified. The precision determines the number of positive class predictions which belong to the positive class, while the F1-score is the consonant mean of Precision and Recall, which measures testing accuracy. MCC measures the accuracy of the classifier by comparing observed and expected results.

Table 2: Performance comparison of our method with and without applying data augmentation in terms of sensitivity, specificity, and accuracy.

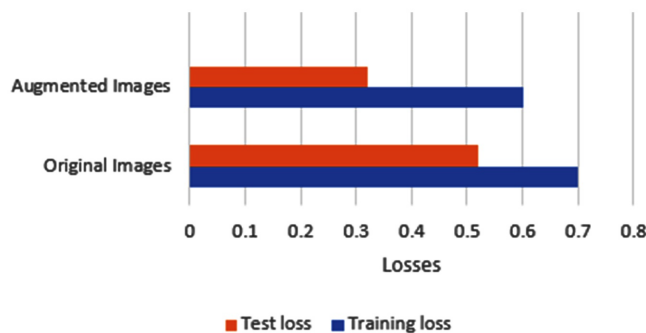| Dataset	Sensitivity (%)	Specificity (%)	Accuracy (%)
Without Augmentation	70.11	82.76	82.18
With Augmentation	91.42	98.56	97.53

Figure 4: Performance comparison of our method with and without applying data augmentation in training and test loss.

5.2 Performance assessment

The performance of the proposed method for the automatic acne classification is evaluated on the developed dataset in

terms of sensitivity, specificity, and accuracy. The generated dataset's images are resized to 224×224 RGB images as input. Then, a 10-fold cross-validation strategy is employed for two dataset scenarios: without and with data augmentation. Hence, the whole dataset without augmentation (420 images) is randomly divided into ten equal-size subsamples. Among them, a single subsample is selected as a testing set (i.e., 10% of the whole dataset, which is 42 images), and the remaining nine subsamples (i.e., 90% of the dataset) are used as the training set. This process is repeated ten times, while each of the ten subsamples is used exactly once as the testing set during the whole validation process. Similarly, a 10-fold cross-validation strategy is also applied for the dataset with augmentation. Still, this time the number of images is increased to 3360 (i.e., 420 original images plus 2940 augmented images based on seven augmentation techniques).
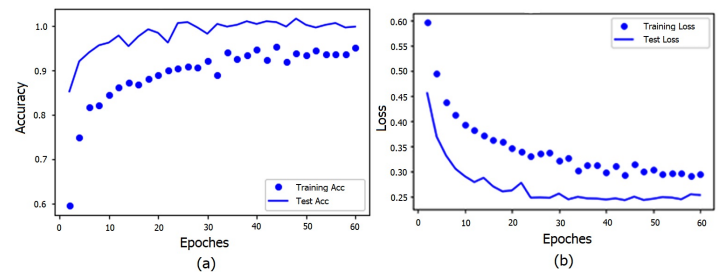

Figure 5: Performance of the proposed system on our developed dataset (with augmentation) in terms of (a) training and validation (test) accuracy, and (b) training and validation (test) loss.

The average results of 10-fold cross-validation for both scenarios are summarized in Table 2. As presented in the table, sensitivity, specificity, and accuracy are increased by 21.31% (from 70.11 to 91.42%), 15.8% (from 82.76% to 98.56%), and 15.35% (from 82.18% to 97.53%), respectively, for augmented images. Additionally, the impact of the augmentation on the performance of the proposed method is investigated in terms of training and test loss in Fig. 4. As illustrated in this figure, after augmentation, the training loss and test loss are decreased by 0.1 (from 0.7 to 0.6) and 0.2 (from 0.52 to 0.32), respectively. These results conclude that the augmentation prevents the model from overfitting and improves the system's performance. Hence, in all experiments, the system's performance is evaluated on the augmented dataset.

Moreover, the effects of the convolution block numbers in both CNN units are further investigated to select the optimum number. Twenty-five different combinations of convolution blocks are implemented with two optimizers (i.e., stochastic gradient descent (SGD) and ADAM) as presented in Table 3. According to the feature maps of each block (which are not shown here for brevity), the first convolution blocks usually detect and extract the edges of the images. As the number of blocks increases and the network becomes more profound, the feature maps look more like an

Table 3: The number of convolution blocks and the type of optimizers applied on the model.

No of Conv. Blocks in the First Unit	Filters of the First Unit	NO of Conv. Blocks in the Second Unit	Filters of the Second Unit	Optimizer	Accuracy
3	32, 64, 96	3	32, 48, 64	SGD	82.31
				ADAM	83.63
		4	32, 32, 48, 64	SGD	79.78
				ADAM	76.92
		5	32, 32, 48, 64, 128	SGD	81.66
				ADAM	82.01
		6	16, 32, 32, 48, 64, 128	SGD	83.33
				ADAM	83.98
		7	16, 32, 32, 48, 64, 128, 256	SGD	81.50
				ADAM	82.23
4	16, 32, 64, 128	3	32, 48, 64	SGD	78.81
				ADAM	83.19
		4	32, 32, 48, 64	SGD	82.47
				ADAM	81.62
		5	32, 32, 48, 64, 128	SGD	87.02
				ADAM	85.28
		6	16, 32, 32, 48, 64, 128	SGD	83.26
				ADAM	84.55
		7	16, 32, 32, 48, 64, 128, 256	SGD	87.29
				ADAM	89.61
5	16, 32, 64, 96, 128	3	32, 48, 64	SGD	91.56
				ADAM	93.83
		4	32, 32, 48, 64	SGD	94.34
				ADAM	95.17
		5	32, 32, 48, 64, 128	SGD	96.71
				ADAM	**97.53**
		6	16, 32, 32, 48, 64, 128	SGD	95.09
				ADAM	94.18
		7	16, 32, 32, 48, 64, 128, 256	SGD	93.41
				ADAM	95.62
6	16, 32, 64, 64, 96, 128	3	32, 48, 64	SGD	95.01
				ADAM	94.98
		4	32, 32, 48, 64	SGD	95.27
				ADAM	95.71
		5	32, 32, 48, 64, 128	SGD	95.55
				ADAM	96.48
		6	16, 32, 32, 48, 64, 128	SGD	94.98
				ADAM	95.24
		7	16, 32, 32, 48, 64, 128, 256	SGD	95.38
				ADAM	95.77
7	16, 32, 64, 64, 96, 128, 256	3	32, 48, 64	SGD	95.35
				ADAM	96.50
		4	32, 32, 48, 64	SGD	95.11
				ADAM	94.96
		5	32, 32, 48, 64, 128	SGD	95.75
				ADAM	96.10
		6	16, 32, 32, 48, 64, 128	SGD	95.83
				ADAM	95.75
		7	16, 32, 32, 48, 64, 128, 256	SGD	96.00
				ADAM	95.89

abstract representation than the original image. The simple patterns, such as edges and shapes, are detected based on lower-level feature maps, while the high-level concepts are encoded using deeper feature maps. In our integrated dual-CNN feature extractor, the required features of acne and its types are extracted with five convolution blocks for both CNN units. As summarized in Table 3, using less than five convolution blocks provides fewer features insufficient for getting high accuracy. Although more features are extracted by increasing the number of convolution blocks, it does not necessarily always increase the accuracy. And instead, it leads to overfitting and false positives. Going deeper results in sparser feature maps (the filters detect fewer features). Consequently, the deeper feature maps provide more information about the class of the image than the image itself, which is helpful but less visually interpretable. In the first convolution blocks, simple shapes (available in every image) are detected, while the deeper networks seek more complex features that don't appear in every image. This phenomenon happens in our system when the number of convolution blocks in both CNN units is more than 5. Hence, the highest accuracy (96.71%

with SGD and 97.53% with ADAM) is achieved using five convolution blocks in both CNN-based feature extractors. Comparing the performance of our model based on two optimizers of SGD and ADAM, it is observed that ADAM achieves better results as it is an extension of SGD.

Hence, the performance of the final proposed system is presented in Figs. 5 (a) and (b) in terms of accuracy and loss, respectively, on the augmented dataset. The dotted line is related to the training set in both graphs, and the solid line is related to validation/test data. The X and Y axes in this figure demonstrate the number of epochs and accuracy/loss, respectively. It is observed that the model's performance is almost stable after 30 epochs and the training and test accuracy reach 94.81% (with a loss of 0.38) and 97.53% (with a loss of 0.29), respectively, after 60 epochs. Overall, the test set is well performed than the training set.

5.3 Comparison with classifiers

In this study, acne classification is carried out with conventional machine learning classifiers to demonstrate the capability of our proposed dual CNN-based acne classifi-

Table 4: Comparison of the proposed model performance with different machine learning-based classifiers.

Classifiers	Precision	F1-Score	Sensitivity	Specificity	Accuracy
KNN	72.35%	74.41%	75.89%	90.25%	90.10%
MLP	75.79%	76.52%	76.92%	93.45%	92.86%
SVM	78.14%	82.48%	80.82%	94.45%	94.06%
Proposed Softmax	91.37%	91.36%	91.42%	98.56%	97.53%

cation system. The softmax classifier is replaced with machine learning-based classifiers. The extracted integrated feature maps from the dual-CNN feature extractor are fed into three classifiers: SVM, MLP, and KNN. Their tuned hyperparameters, i.e., initial learning rate, minimum batch size, learning algorithm, maximum epochs, and learning factor of f_c, are 0.0003, 32, ADAM, 60, and 10, respectively.

The performance of these classifiers is evaluated in terms of precision, F1-score, sensitivity, specificity, and accuracy and compared with our proposed model based on the softmax classifier in Table 4. As presented in this table, the accuracy of KNN, MLP, and SVM classifiers and our model is 90.10%, 92.86%, 94.06%, and 97.53%, respectively. Comparing the results, our proposed CNN model based on a softmax classifier achieves at least 3.5% more accuracy than the other classifiers. Additionally, its processing time is less than the different conventional classifiers.

5.4 Comparison with pre-trained models

Our proposed method can determine the type of acne and distinguish between acne and other skin diseases such as eczema and cancer. To further highlight our proposed acne classification capability, it is compared with five pre-trained models: GoogleNet, MobileNet, VGG-19, ResNet-50, and AlexNet. Table 5 displays the performance of these models as well as ours in terms of accuracy, precision, F1-score, sensitivity, specificity, and MCC for each of the acne types (i.e., ACC, AC, AE, AK, AOC, AP, DF) and their average results. As presented in this table, the average accuracy of GoogleNet, MobileNet, VGG19, ResNet50, AlexNet, and our proposed model is 94.13%, 94.90%, 95.58%, 96.24%, 95.89%, and 97.53%, respectively, among which ours is the highest. Not only in terms of accuracy but also in terms of other evaluation metrics, i.e., precision (91.37%), F1-score (91.36%), sensitivity (91.42%), specificity (98.56%), and MCC (89.94%), our proposed method outperforms the others. Additionally, the accuracy of each class in our proposed method is higher than that of other methods. The highest accuracy belongs to the DF class, which refers to the acne-free skin images. As our developed dataset has the benefit of having a type containing normal skin and other skin disease images (except acne), our proposed method has successfully trained for recognizing acne disease. If the probe image does not contain acne, it is classified as DF with high accuracy of 99.40%.

Fig. 6 presents the confusion matrices of our proposed model and the other five pre-trained deep learning-based models. As demonstrated in this figure, the number of true positives in all seven classes is higher in our proposed model, which proves its competitive performance compared to GoogleNet, MobileNet, VGG-19, ResNet-50, and AlexNet.

As another evaluation tool, our proposed method's Receiver Operating Characteristic (ROC) Curve and five pre-trained models are illustrated in Fig. 7. It presents the performance of the models at different thresholds, in which the x-axis is the false positive rate and the y-axis is the true positive rate. In this probability curve, the Area Under the ROC Curve (AUC) indicates the classification capability of the corresponding model. As illustrated in this figure, the AUC scores of GoogleNet, MobileNet, VGG-19, ResNet-50, AlexNet, and our proposed CNN are 91.32, 92.91, 91.76, 91.88, 93.24, and 94.67, respectively. These models are all implemented with the ADAM optimizer. As the higher AUC shows a better performance, our proposed method obtains the best performance having the highest AUC of 94.67.

5.5 Comparison with the state-of-the-arts

To have a fair comparison between our proposed method and the state-of-the-art approaches, we must implement them on the proposed dataset. No available source codes are found for the related works to implement them. Consequently, our model is only compared with one state-of-the-art acne classification approach proposed in [21] by implementing it on the same dataset accessible in `http://dermnet.com`. The results are presented in Table 6. As it is noted in this table, our proposed dual CNN-based acne classification system achieves higher performance (96.74%) than the state-of-the-art approach of [21] (with a reported accuracy of 95.89%) on the same dataset of 1800 acne images in 5 different classes.

Additionally, as the proposed model has competitive performance in acne classification, it also inspired us to evaluate it for acne grading. Hence, it is implemented on an acne grading dataset called "ACNE04" [35], which is publicly available. The grading system is implemented on the same dataset of 1457 acne images in 4 classes, i.e., mild, moderate, severe, and very severe. The results are also summarized in the same table (Table 6). Comparing the results, our proposed acne classification system can be successfully performed for acne grading by achieving higher accuracy of 86.36% which is 2.25% higher than that of the state-of-the-art approach in [35].

5.6 Failure cases

Although our proposed acne classification system can successfully recognize and classify acne into six different types, there are a few misclassification cases, as shown in Fig. 8 with its actual class and the predicted label. How-

Table 5: Comparison between the performance of the proposed model and the other five pre-trained models.

Methods	Disease types	Accuracy (%)	Precision (%)	F1-Score (%)	Sensitivity (%)	Specificity (%)	MCC (%)
GoogleNet	ACC	93.45	75.00	76.60	78.26	95.86	72.81
	AC	94.05	77.08	78.72	80.43	96.21	75.29
	AE	94.94	81.25	82.11	82.98	96.89	79.17
	AK	93.75	77.08	77.89	78.72	96.19	74.26
	AOC	94.35	79.17	80.00	80.85	96.54	76.71
	AP	90.48	72.92	68.63	64.81	95.39	63.19
	DF	97.92	93.75	92.78	91.84	98.95	91.57
Total Average		94.13	79.46	79.53	79.70	96.57	76.14
MobileNet	ACC	94.64	81.25	81.25	81.25	96.88	78.13
	AC	95.54	79.17	83.52	88.37	96.59	81.11
	AE	93.75	81.25	78.79	76.47	96.84	75.17
	AK	94.94	83.33	82.47	81.63	97.21	79.52
	AOC	94.94	79.17	81.72	84.44	96.56	78.84
	AP	92.26	77.08	74.00	71.15	96.13	69.54
	DF	98.21	93.75	93.75	93.75	98.96	92.71
Total Average		94.90	82.14	82.21	82.44	97.02	79.29
VGG-19	ACC	95.24	83.33	83.33	83.33	97.22	80.56
	AC	95.54	85.42	84.54	83.67	97.56	81.93
	AE	94.94	83.33	82.47	81.63	97.21	79.52
	AK	96.43	87.50	87.50	87.50	97.92	85.42
	AOC	94.35	77.08	79.57	82.22	96.22	76.35
	AP	93.75	79.17	78.35	77.55	96.52	74.70
	DF	98.81	95.83	95.83	95.83	99.31	95.14
Total Average		95.58	84.52	84.51	84.53	97.42	81.95
ResNet-50	ACC	96.43	87.50	87.50	87.50	97.92	85.42
	AC	94.64	79.17	80.85	82.61	96.55	77.76
	AE	95.24	85.42	83.67	82.00	97.55	80.91
	AK	96.54	81.25	83.87	86.67	96.91	81.34
	AOC	94.94	79.17	81.72	84.44	96.56	78.84
	AP	93.75	83.33	79.21	75.47	97.17	75.67
	DF	99.11	97.92	96.91	95.92	99.65	96.39
Total Average		96.24	84.81	84.82	84.94	97.47	82.33
AlexNet	ACC	95.54	81.25	83.67	86.67	96.91	81.34
	AC	95.24	83.33	83.33	83.33	97.22	80.56
	AE	96.73	89.58	88.66	87.76	98.26	86.75
	AK	97.02	93.75	90.00	86.54	98.94	88.35
	AOC	96.43	85.42	87.23	89.13	97.59	85.19
	AP	93.75	77.08	77.89	78.72	96.19	74.26
	DF	96.54	81.36	88.89	97.96	96.31	87.39
Total Average		95.89	84.54	85.67	87.16	97.35	80.41
Proposed model	ACC	97.92	89.58	92.47	95.56	98.28	91.33
	AC	96.13	87.50	86.60	85.71	97.91	84.34
	AE	96.73	89.58	88.66	87.76	98.26	86.75
	AK	97.92	93.75	92.78	91.84	98.95	91.57
	AOC	98.21	95.83	93.88	92.00	99.30	92.86
	AP	96.43	85.42	87.23	89.13	97.59	85.19
	DF	99.40	97.92	97.92	97.92	99.65	97.57
Total Average		**97.53**	**91.37**	**91.36**	**91.42**	**98.56**	**89.94**

Table 6: The performance comparison of the AcneNet and ACNE04 datasets in the model.

Datasets	No. of classes	Dataset Sizes	Accuracy
AcneNet [21] (Classification)	5	1800	95.89 [21] **96.74 (Our)**
ACNE04 [35] (Grading)	4	1457	84.11 [35] **86.36 (Our)**

ever, among these eight instances, two misclassifications (Figs. 8 (a) and (b) has the low confidence score close to 0.5. Deeply analyzing the acne images of these misclassification cases, we draw the inferences that the proposed model mostly has difficulties classifying the tiny acne, as illustrated in Figs. 8 (d) and (h). Misclassifications with high confidence scores have occurred in Figs. 8 (c), (e), (f), and (g). Several reasons can be included, such as low image quality, a unique acne case, etc.

6   Conclusions and future works

This paper introduced a deep learning-based lightweight system to recognize and classify different acne vulgaris using a novel acne dataset. Firstly, an acne dataset with 420 images in 7 classes was developed. Then, these images were modified by applying a pre-processing system. The number of shots was increased to 3360 by using seven different augmentation methods. An integrated dual CNN-based model was proposed to recognize acne and classify it into seven groups. The whole feature extractor was composed of two CNN models with different conv2d, BN, and max-pooling layers. The extracted feature maps from these two models were first concatenated and then aggregated using fully connected layers. The final comprehensive feature maps were fed into a softmax layer for classification. The performance was investigated for convolution blocks in two feature extractor units and two different SGD and ADAM optimizers.

In addition, its performance was evaluated for both original images without augmentation and the extended dataset with augmented images to analyze the influence of augmentation. It was compared with three conventional machine learning-based classifiers and five pre-trained deep learning-based models and received competitive performance. The proposed method's feasibility has been confirmed by conducting experiments and achieving a state-of-the-art accuracy of 97.53% in acne classification. It was
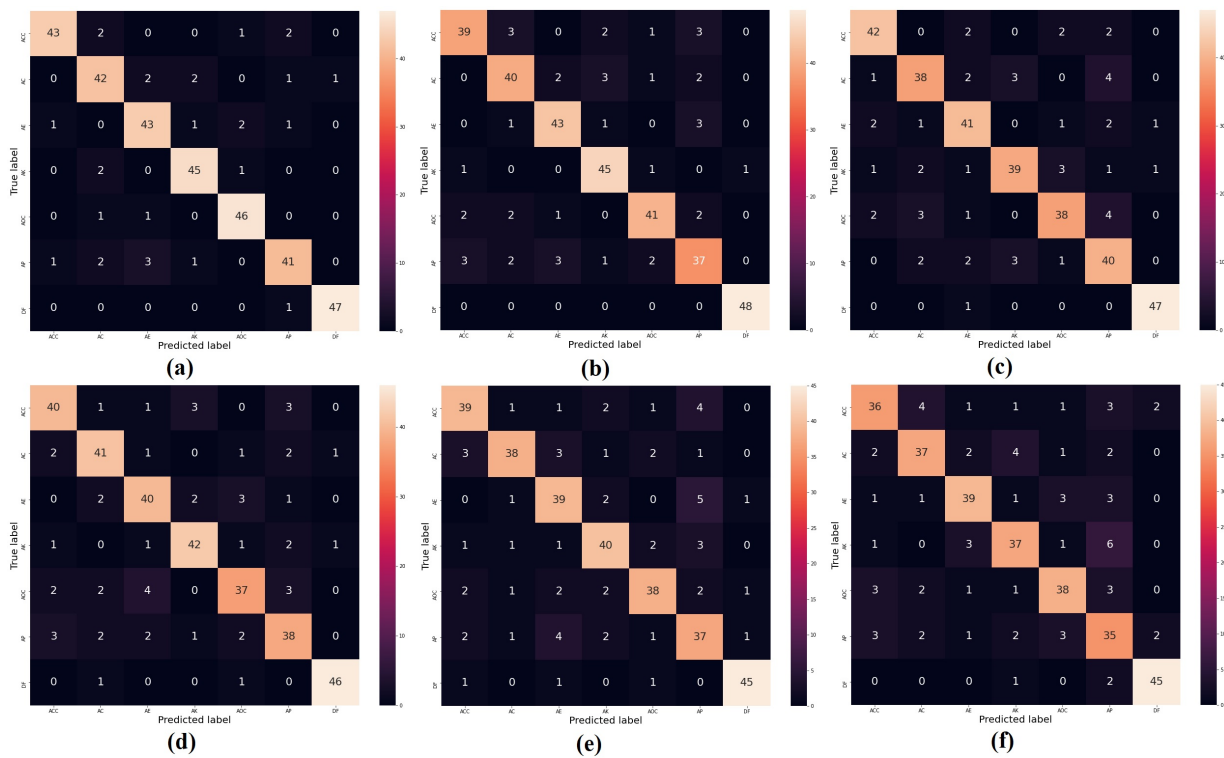
Figure 6: Confusion matrix for acne classification. Here, (a), (b), (c), (d), (e), and (f) represent the confusion matrix of the proposed model, AlexNet, ResNet-50, VGG-19, MobileNet and GoogleNet
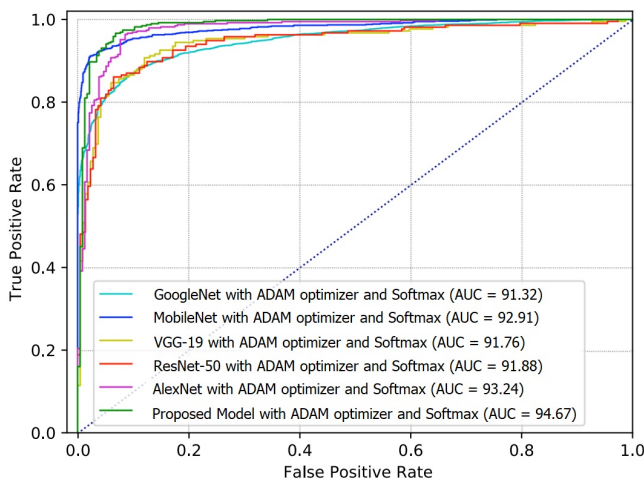

Figure 7: Comparison between the AUC of the ROC curves belonging to our proposed deep dual-CNN model and five pre-trained models.

also implemented on an acne grading dataset and achieved good performance with an accuracy of 86.36%. In the future, we want to expand our study with an updated and extended form of our proposed acne dataset with additional images and acne classes. As our proposed model is highly accurate for acne classification and computationally efficient, it will also be adjusted and applied to other similar dermatological disease classification and identification tasks in our future research direction.

Experimental code availability

For further investigation, comparison, and analysis of this study by the research community, the experimental code and model are accessible upon request to the corresponding author through email.

Compliance with ethical standards

This article does not contain any studies with human participants and/or animals performed by any authors.

Conflict of interest

We (authors) certify that this article has no actual or potential conflict of interest.

Authors' contributions

All authors contributed to this paper. **Afsana Ahsan Jeny, Masum Shah Junayed, Nipa Anjum:** Methodology, Experiment, Writing- Original draft preparation. **Md Baharul Islam:** Investigation, Conceptualization, Supervision, Writing- Reviewing and Editing, **Arezoo Sadeghzadeh, A. F. M. Shahen Shah: Writing- Reviewing and Editing**.
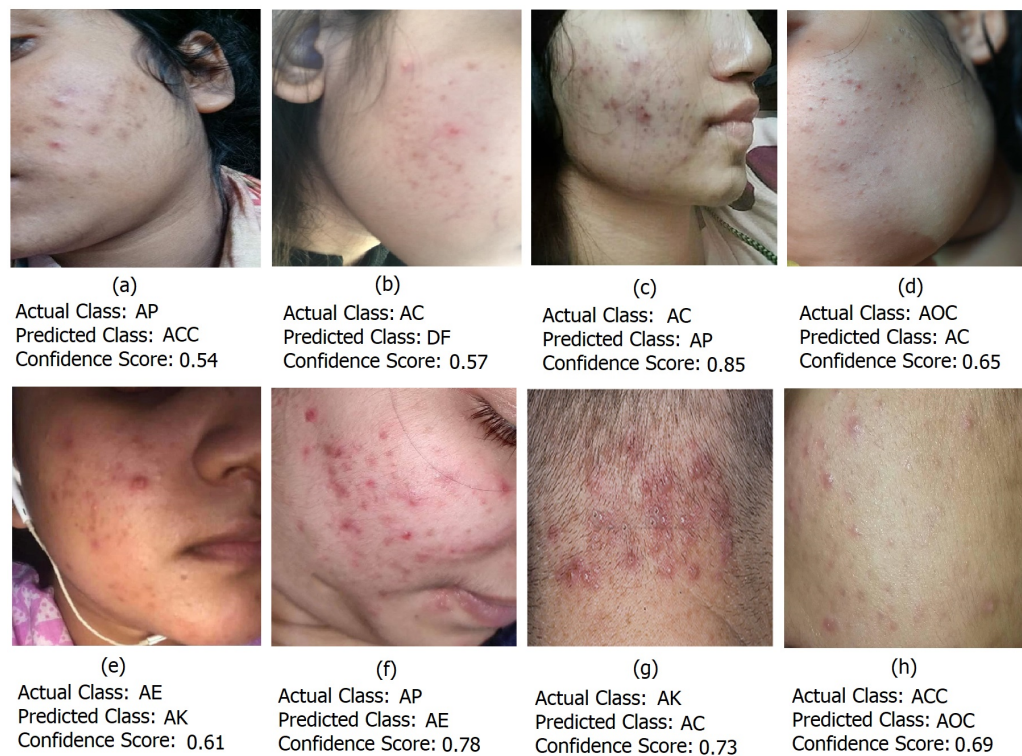
(a)
Actual Class: AP
Predicted Class: ACC
Confidence Score: 0.54

(b)
Actual Class: AC
Predicted Class: DF
Confidence Score: 0.57

(c)
Actual Class: AC
Predicted Class: AP
Confidence Score: 0.85

(d)
Actual Class: AOC
Predicted Class: AC
Confidence Score: 0.65

(e)
Actual Class: AE
Predicted Class: AK
Confidence Score: 0.61

(f)
Actual Class: AP
Predicted Class: AE
Confidence Score: 0.78

(g)
Actual Class: AK
Predicted Class: AC
Confidence Score: 0.73

(h)
Actual Class: ACC
Predicted Class: AOC
Confidence Score: 0.69

Figure 8: Misclassification cases of the proposed model.

References

[1] Goulden V, Clark SM, Cunliffe WJ. Post☐adolescent acne: a review of clinical features. British journal of dermatology. 1997 Jan;136(1):66-70. `https://doi.org/10.1046/j.1365-2133.1997.d01-1144.x`

[2] Dharshana S, Singh A, Sharma S, Mohan S, Joshi A. Depression, mood change and self-esteem among adolescents aged 12-25 years with acne vulgaris in India. Annals of Tropical Medicine and Public Health. 2016;9(1). `https://doi.org/10.4103/1755-6783.168712`

[3] Picardo M, Eichenfield LF, Tan J. Acne and rosacea. Dermatology and therapy. 2017 Jan;7(1):43-52. `https://doi.org/10.1007/s13555-016-0168-8`

[4] Zaenglein AL. Acne vulgaris. New England Journal of Medicine. 2018 Oct 4;379(14):1343-52. `https://doi.org/10.1056/NEJMcp1702493`

[5] Alamdari N, Tavakolian K, Alhashim M, Fazel-Rezai R. Detection and classification of acne lesions in acne patients: A mobile application. In IEEE international conference on electro information technology (EIT) 2016 May 19 (pp. 0739-0743). IEEE. `https://doi.org/10.1109/eit.2016.7535331`

[6] Alom MZ, Aspiras T, Taha TM, Asari VK. Skin cancer segmentation and classification with NABLA-N and inception recurrent residual convolutional networks. arXiv preprint arXiv:1904.11126. 2019 Apr 25. `https://doi.org/10.48550/arXiv.1904.11126`

[7] Ayer J, Burrows N. Acne: more than skin deep. Postgraduate medical journal. 2006 Aug 1;82(970):500-6. `https://doi.org/10.1136/pgmj.2006.045377`

[8] Abayomi-Alli OO, Damasevicius R, Misra S, Maskeliunas R, Abayomi-Alli A. Malignant skin melanoma detection using image augmentation by oversampling in nonlinear lower-dimensional embedding manifold. Turkish Journal of Electrical Engineering and Computer Sciences. 2021;29(8):2600-14. `https://doi.org/10.3906/elk-2101-133`

[9] Barata C, Celebi ME, Marques JS, Rozeira J. Clinically inspired analysis of dermoscopy images using a generative model. Computer Vision and Image Understanding. 2016 Oct 1;151:124-37. `https://doi.org/10.1016/j.cviu.2015.09.011`

[10] Chen L, Li Y, Han L, Yuan L, Sun Y, Tang X. Classification and Treatment System for Facial Acne Vulgaris Based on Image Recognition. InElderly Health Services and Remote Health Monitoring 2020 (pp. 65-71). Springer, Singapore. `https://doi.org/10.1007/978-981-15-7154-1_6`

[11] Cireşan DC, Meier U, Masci J, Gambardella LM, Schmidhuber J. High-performance neu-

ral networks for visual object classification. arXiv preprint arXiv:1102.0183. 2011 Feb 1. https://doi.org/10.48550/arXiv.1102.0183

[12] Gu Y, Ge Z, Bonnington CP, Zhou J. Progressive transfer learning and adversarial domain adaptation for cross-domain skin disease classification. IEEE journal of biomedical and health informatics. 2019 Sep 23;24(5):1379-93. https://doi.org/10.1109/jbhi.2019.2942429

[13] Hameed AZ, Awad WK, Irsan NA, Abdulbaqi AS. Hybrid technique for skin pimples image detection and classification. International Journal of Horticultural Science and Technology. 2020;29:4102-9. http://sersc.org/journals/index.php/IJAST/article/view/5164

[14] Hameed N, Shabut AM, Ghosh MK, Hossain MA. Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques. Expert Systems with Applications. 2020 Mar 1;141:112961. https://doi.org/10.1016/j.eswa.2019.112961

[15] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. InProceedings of the IEEE international conference on computer vision 2015 (pp. 1026-1034). https://doi.org/10.1109/iccv.2015.123

[16] Jeny AA, Sakib AN, Junayed MS, Lima KA, Ahmed I, Islam MB. SkNet: a convolutional neural networks-based classification approach for skin cancer classes. In 23rd International Conference on Computer and Information Technology (ICCIT) 2020 Dec 19 (pp. 1-6). IEEE. https://doi.org/10.1109/iccit51783.2020.9392716

[17] Jiang S, Li H, Jin Z. A visually interpretable deep learning framework for histopathological image-based skin cancer diagnosis. IEEE Journal of Biomedical and Health Informatics. 2021 Jan 15;25(5):1483-94. https://doi.org/10.1109/jbhi.2021.3052044

[18] Junayed MS, Islam MB, Sadeghzadeh A, Rahman S. CataractNet: An automated cataract detection system using deep learning for fundus images. IEEE Access. 2021 Sep 15;9:128799-808. https://doi.org/10.1109/access.2021.3112938

[19] Almoosawi NM, Khudeyer RS. ResNet-34/DR: a residual convolutional neural network for the diagnosis of diabetic retinopathy. Informatica. 2021 Dec 17;45(7). https://doi.org/10.31449/inf.v45i7.3774

[20] Ullah Z, Odeh A, Khattak I, Hasan MA. Enhancement of Pre-Trained Deep Learning Models to Improve Brain Tumor Classification. Informatica. 2023. 47; 165–172. https://doi.org/10.31449/inf.v47i6.4645

[21] Junayed MS, Jeny AA, Atik ST, Neehal N, Karim A, Azam S, Shanmugam B. AcneNet-a deep CNN-based classification approach for acne classes. In 12th International Conference on Information & Communication Technology and System (ICTS) 2019 Jul 18 (pp. 203-208). IEEE. https://doi.org/10.1109/icts.2019.8850935

[22] Junayed MS, Sakib AN, Anjum N, Islam MB, Jeny AA. Eczemanet: A deep cnn-based eczema diseases classification. In2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS) 2020 Dec 9 (pp. 174-179). IEEE. https://doi.org/10.1109/ipas50080.2020.9334929

[23] Kaymak R, Kaymak C, Ucar A. Skin lesion segmentation using fully convolutional networks: A comparative experimental study. Expert Systems with Applications. 2020 Dec 15;161:113742. https://doi.org/10.1016/j.eswa.2020.113742

[24] Khongsuwan M, Kiattisin S, Wongseree W, Leelasantitham A. Counting number of points for acne vulgaris using UV fluorescence and image processing. Biomedical Engineering International Conference 2012 Jan 29 (pp. 142-146). IEEE. https://doi.org/10.1109/bmeicon.2012.6172038

[25] Kittigul N, Uyyanonvara B. Acne detection using speeded up robust features and quantification using K-Nearest neighbors algorithm. InProceedings of the 6th International Conference on Bioinformatics and Biomedical Science 2017 Jun 22 (pp. 168-171). https://doi.org/10.1145/3121138.3121168

[26] Loram I, Siddique A, Sánchez MB, Harding P, Silverdale M, Kobylecki C, Cunningham R. Objective analysis of neck muscle boundaries for cervical dystonia using ultrasound imaging and deep learning. IEEE journal of biomedical and health informatics. 2020 Jan 9;24(4):1016-27. https://doi.org/10.1109/jbhi.2020.2964098

[27] Malik AS, Ramli R, Hani AF, Salih Y, Yap FB, Nisar H. Digital assessment of facial acne vulgaris. In 2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings 2014 May 12 (pp. 546-550). IEEE. https://doi.org/10.1109/i2mtc.2014.6860804

[28] Pour MP, Seker H. Transform domain representation-driven convolutional neural networks for skin lesion segmentation. Expert Systems with Applications. 2020 Apr 15;144:113129. https://doi.org/10.1016/j.eswa.2019.113129

[29] Rashataprucksa K, Chuangchaichatchavarn C, Triukose S, Nitinawarat S, Pongprutthipan M, Piromsopa K. Acne detection with deep neural networks. In2020 2nd International Conference on Image Processing and Machine Vision 2020 Aug 5 (pp. 53-56). `https://doi.org/10.1145/3421558.3421566`

[30] Sermanet P, Chintala S, LeCun Y. Convolutional neural networks applied to house numbers digit classification. InProceedings of the 21st international conference on pattern recognition (ICPR2012) 2012 Nov 11 (pp. 3288-3291). IEEE. `https://doi.org/10.48550/arXiv.1204.3968`

[31] Shanthi T, Sabeenian RS, Anand R. Automatic diagnosis of skin diseases using convolution neural network. Microprocessors and Microsystems. 2020 Jul 1;76:103074. `https://doi.org/10.1016/j.micpro.2020.103074`

[32] Shen X, Zhang J, Yan C, Zhou H. An automatic diagnosis method of facial acne vulgaris based on convolutional neural network. Scientific reports. 2018 Apr 11;8(1):1-0. `https://doi.org/10.1038/s41598-018-24204-6`

[33] Son HM, Jeon W, Kim J, Heo CY, Yoon HJ, Park JU, Chung TM. AI-based localization and classification of skin disease with erythema. Scientific Reports. 2021 Mar 5;11(1):1-4. `https://doi.org/10.1038/s41598-021-84593-z`

[34] Tang P, Liang Q, Yan X, Xiang S, Zhang D. GP-CNN-DTEL: Global-part CNN model with data-transformed ensemble learning for skin lesion classification. IEEE journal of biomedical and health informatics. 2020 Feb 28;24(10):2870-82. `https://doi.org/10.1109/jbhi.2020.2977013`

[35] Wu X, Wen N, Liang J, Lai YK, She D, Cheng MM, Yang J. Joint acne image grading and counting via label distribution learning. InProceedings of the IEEE/CVF International Conference on Computer Vision 2019 (pp. 10642-10651). `https://doi.org/10.1109/iccv.2019.01074`

[36] Yadav SS, Jadhav SM. Deep convolutional neural network-based medical image classification for disease diagnosis. Journal of Big Data. 2019 Dec;6(1):1-8. `https://doi.org/10.1186/s40537-019-0276-2`

[37] Zhang C, Tavanapong W, Wong J, Groen PC, Oh J. Real data augmentation for medical image classification. In Intravascular imaging and computer-assisted stenting, and large-scale annotation of biomedical data and expert label synthesis 2017 Sep 14 (pp. 67-76). Springer, Cham. `https://doi.org/10.1007/978-3-319-67534-3_8`

[38] Zhang L, Wang X, Yang D, Sanford T, Harmon S, Turkbey B, Wood BJ, Roth H, Myronenko A, Xu D, Xu Z. Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation. IEEE transactions on medical imaging. 2020 Feb 12;39(7):2531-40. `https://doi.org/10.1109/tmi.2020.2973595`

[39] Zhao T, Zhang H, Spoelstra J. A computer vision application for assessing facial acne severity from selfie images. arXiv preprint arXiv:1907.07901. 2019 Jul 18. `https://doi.org/10.48550/arXiv.1907.07901`

[40] Zhao X, Hu X, Liao Y, He T, Zhang T, Zou X, Tian J. Accurate MR image super-resolution via lightweight lateral inhibition network. Computer Vision and Image Understanding. 2020 Dec 1;201:103075. `https://doi.org/10.1016/j.cviu.2020.103075`