Received: 16 April 2021 | Revised: 27 May 2021 | Accepted: 10 June 2021

DOI: 10.1111/exsy.12760

ORIGINAL ARTICLE

Expert Systems WILEY

HSV model-based segmentation driven facial acne detection using deep learning

Neha Yadav[1] | Sk Md Alfayeed[1] | Aditya Khamparia[2] | Babita Pandey[3] | Dang N. H. Thanh[4] | Sagar Pande[1]

[1]School of Computer Science and Engineering, Lovely Professional University, Jalandhar, India

[2]Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Satellite Centre, Tikarmafi, Amethi, India

[3]Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow, India

[4]Department of IT, University of Economics Ho Chi Minh City, Ho Chi Minh, Vietnam

Correspondence
Aditya Khamparia, Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Satellite Centre, Tikarmafi, Amethi, Tikarma.
Email: aditya.khamparia88@gmail.com

Abstract

Acne is a skin disease mainly caused by bacteria, the hair follicles exposed to oil, and dying skin cells. These sometimes trigger whiteheads, blackheads or pimples, usually on the neck, arms, arm and back of shoulders. Acne in adolescents is the most severe, even though it affects people of any generation. Doctors can easily detect acne by seeing a patient's skin, but automatic acne detection is not easy for machines. Deep learning (DL) approaches have been quite successful for various aspects like classification and detection of objects in real life. This paper proposes an enhanced DL CNN model with the Leaky ReLU activation function. DermNet NZ's facial acne images dataset is used for the experiments. Three different techniques- K-Means, Texture Analysis and HSV Model-Based Segmentation, are applied for image segmentation to extract the acne region from skin images. After applying all the above image segmentation methods five times for each method, output images from K-Means and HSV (5 + 5 images) are collected and combined with the dataset. Using that dataset, one SVM model using Scikit-learn and two CNN models- one with the ReLU activation function and another with the LeakyReLU activation function, is trained. Out of these three models, the proposed CNN (LeakyReLU) model achieved a 97.54% accuracy.

KEYWORDS

acne deep learning (DL), machine learning (ML), image processing (IP), CNN K-means

1 | INTRODUCTION

In the form of data training models, Machine Learning (ML) and Deep Learning (DL) uses data modelling to address various diagnostic tasks and problems, such as medical diagnosis, voice recognition, computer vision, fraud detection, and behaviour recognition, etc. The main topic of this paper is to propose a DL CNN model for facial acne detection.

Most of the men, women, even including children often suffer from different face skin-related diseases. Skin disorders do not only harm human skin but decrease people's confidence and become depressed. It is, therefore, critical to recognize the diseases early and diagnose seriously. It is also crucial to not being spread in other different parts of the body (Banerjee & Bandyopadhyay, 2018). The most important aspect of human anatomy is the facial skin. The indications and severity of skin disorders vary considerably. It may be transient or everlasting, painful or painless. Some may be genetic, and some may have situational reasons. Some diseases of the skin are mild, and others can pose a severe threat to life. It may induce many facial skin disorders, such as Acne, Vitiligo, Melanoma, Eczema, Rosacea, Moles, and so on. With the increasing advancement of medical cosmetology and improved technologies, people are increasingly worried about facial skin treatment.

Acne is an inflammatory or viral skin disease that affects the oil glands. Acne lesions can be categorized into various skin forms with or without comedones, pustule and reddish papules (Padilla-Medina et al., 2014).

Four main factors cause acne on the skin; bacteria, oil clogged hair follicles and dead skin cells, excess oil (sebum) production, and inflammation. Usually, acne occurs on the face, forehead, stomach, back and shoulders since the most oil (sebaceous) glands are located in these areas of skin. Hair follicles are connected to petroleum glands. The follicle wall may flourish and make a white head. Or the plug might be open and darkened to the surface, creating a blackhead. A blackhead might look like pores of dirt. The pore is congested with bacteria and oil, but when it is exposed to sunlight, it becomes brown. Pimples are elevated red spots with a white center, which are inflamed or infected with bacteria as blocked hair follicles (Mayo Clinic, 2020).

In some situations, though, it may be difficult to detect acne objectively from the colour image to properly evaluate acne lesions. At some point in life, it affects 85% of adults (Padilla-Medina et al., 2014). Traditional acne detection approaches such as standard flash photography and clear visual assessment were used by the physicians and dermatologists (Doukas et al., 2012). These techniques can be time-consuming from time to time. The accurate assessment of acne intensity is critical in the treatment of acne. So, such techniques are needed, which can easily detect the acne on the face using facial acne images and accurately classify them according to the characteristics of acne, and there should be a minimum error.

- This research is adopted because the most crucial part of the human body is the face and facial skin. With the increasing advancement of medical cosmetology and improved technologies, people are increasingly worried about facial skin treatment.
- Previously, many works have been done by various researchers on acne detection. However, there are significantly fewer papers available where DL approaches were used.
- Here, the idea is to build an enhanced acne detection CNN DL model to detect and classify the acne from the face.
- Three different methods, such as K-Means, Texture Analysis, and HSV Model-Based segmentation are applied for image segmentation.
- One SVM and two CNN models for image classification are trained for comparing the accuracy of each model.
- In this paper, the DL CNN approach is adapted because, in recent years, DL has earned considerable attention and outperformed itself in diverse image recognition and classification problems.
- This paper presents a DL CNN model using the Leaky ReLU activation function. This proposed model can classify facial acne skin from normal skin successfully with an accuracy of 97.54%.

The rest of the paper is organized as follows as Section 2 gives some ideas and a summary of previous related researches. Section 3 provides an explanation of the proposed methodology. Section 4 focuses on the dataset, preprocessing of the data, and architecture of the performed methods. Section 5 gives elaboration and analysis of the results with figures and tables. Finally, Section 6 delivers the conclusion and future scope.

2 | RELATED WORKS

In the past, many researchers have done many works on acne detection using various techniques. In 1999, (Kwon & da Vitoria Lobo, 1999) applied a few filters to separate wrinkles from face images in order to distinguish younger adults from older adults. Folds have been counted in many locations, for example, on the neck, next to the eye and on the cheekbones. The presence of irregularities in a field depends on the region's curves. Nonetheless, random initialization and a variety of rates generated implementation difficulties. The filter approach uses spatial orientation and allows the reaction to seam-like structures to be optimized.

In 2017, Alarifi et al. (2017) addressed some techniques of classification of skin that use conventional ML and convolutional neural networks (CNNs) to classify three types of facial skin problems: good normal skin with spots, patches with roughness, and wrinkles. They collected a total of 325 high-quality face images of different ethnicities for creating dataset. They performed many computer vision experiments using DL as well as ML. They used Sequential Minimal Optimization to train their SVM model and GoogLeNet (NAG) architecture for training the CNN model. The proposed SVM model achieved 81.5% accuracy, and the CNN model achieved 85% accuracy. Maroni et al. (2017) created a prototype framework for automated acne identification, lesion counting through mobile device remote image processing. The application pipeline is composed of the body part identification, skin segmentation, heat mapping, blob identification, and extraction of acne. Haar-Cascade classifier was used for body part detection. Random Forest was used for skin segmentation trained on the FSD dataset. At last Gaussian filtering and Adaptive Thresholding were used for acne extraction and blob detection.

Shen et al. (2018) presented a new automated diagnostic approach focused on CNNs for facial acne vulgaris. To address the limitations of prior approaches, which were unable to distinguish sufficient types of acne. The dataset collected by the authors consists of 3000 facial skin and 3000 non-skin images. Binary-classification was done for skin and non-skin classification, and seven-classifier classification was done to differentiate between acne and normal skin using 3000 facial skin images.

Jung et al. (2019) proposed a DL-based model to classify and detect four different types of facial acne (blemish, pustulosa, vulgaris, nodulocystic) using CNN sequential model with small loss rate (5.46%) and high accuracy (96.26%). The dataset was manually collected by the authors from various free legal websites. Python 3.6.3 with Keras, Tensorflow, and Scikitlearn was used for the experiments. Hameed et al. (2019) proposed a hybrid method for skin pimples detection combining Naive Bayes Classifier (NBC) and image processing techniques. They used an adaptive thresholding approach for pimples segmentation and ARFCM clustering technique for pimple pattern detection. They used NBC for feature extractions and image classification. They used 40 images for their experiments. In their classification, their method obtained 93.42% accuracy.

Lobo et al. (2020) proposed a mobile application for the detection of three types of acne - acne comedones, acne cyst, and acne pustules using the MobileNet version 2 model. They used a total of 3024 numbers of samples from various dermatologic sources as well as self-clicked images. They used ML Toolkit for face detection, facial landmark model for skin patches extraction. They considered MobileNet version 1 and added additional layers to propose the MobileNet version 2 model for image classification. They trained the model with 11,000 steps and achieved 91% accuracy. They stated that their proposed MobileNetV2 model performs better than the previous MobileNetV1 model. Chen et al. (2020) proposed Gaussian-adaptive bilateral filter (GABF) to resolve the image smoothing problem. The basic idea is to acquire a low-pass guidance for the range kernel by a Gaussian spatial kernel. Such low-pass guidance lead to a clean Gaussian range kernel for later bilateral composite. (Bharati et al., 2020) proposed hybrid DL model for lung cancer detection from X-ray images. (Chen et al., 2021) proposed two-pass (TP) BF, TP-based BF and an adaptive control scheme of range kernels for noise-invariant edge-preserving image smoothing.

Every previous work done by various researchers targeting acne detection is very appreciable. Many researchers have already proposed many DL models, especially CNN models, for acne detection. But most of them have used a very large amount of datasets or images to train their models because, generally, CNN needs a large amount of data for training to achieve very high accuracy. To overcome that large data requirement issue, in this paper, an enhanced CNN model based on the LeakyReLU activation function is proposed, which can detect acne with very high accuracy by using very less amount of dataset.

3 | PROPOSED METHODOLOGY

3.1 | K-means clustering

K-means clustering is a method of the unsupervised ML algorithm, where when we have unlabeled or mixed data. The aim is to separate a picture into regions that have a visual appearance rationally identical or the sections of items. The intensities of each pixel in an image involve the red, green, and blue intensities in a three-dimensional space, and each and every pixel of the image is known simply as the data point by the segmentation process (Narkhede & Adhiya, 2014) (Figure 1).

We will be using L^*a^*b colour space for our K-Means clustering method. That is why we need to do RGB to L^*a^*b Conversion defined as the following equation:

$$\begin{aligned} L^* &= 116f(G/G_n) - 16 \\ a^* &= 500[f(R/R_n) - f(B/Bn)] \\ b^* &= 200[f(R/R_n) - f(B/Bn)] \end{aligned} \tag{1}$$

where f is a calibration function.

$$\mathrm{f}(q) = \begin{cases} q^{1/3} & \text{if } q > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 q + \frac{4}{29} & \text{otherwise} \end{cases} \tag{2}$$

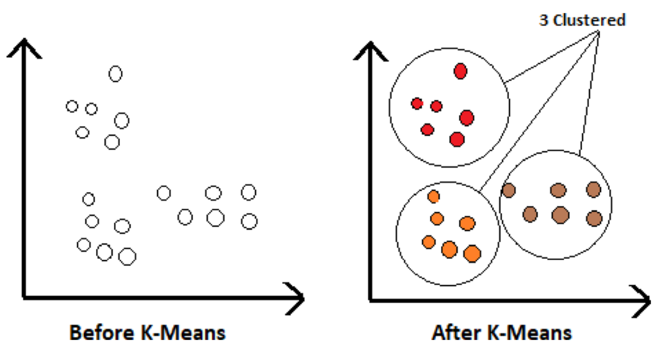

FIGURE 1 K-means clustering

K-means clustering algorithm partitions data into K-cluster (C1, C2, ……., Ck), represented by their centers or mean.

$$m = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^j - c_j \right\|^2 \tag{3}$$

where $\left\| x_i^j - c_j \right\|$ is chosen to be the distance measure between a x_i^j data point and the cluster center c_j, which is an indicator of the distance of n data points from their respective cluster center (Pixels, 2017).

3.2 | Texture analysis

The texture analysis is focused on the texture characteristics of an area and the sections of an image. The study of texture aims to quantify characteristics as smooth, rugged, bumpy or silky based on the spatial variation between the pixel size. In this program, the skin (background) is smooth, and acne (background) is rough.

3.3 | HSV model-based segmentation

The methods of splitting a digital image into several segments in digital image processing and computer vision are image segmentations. The aim of segmentation is to make the portrayal of a picture clearer and/or easier to analyse.

HSV is one of the most known and used models for cylindrical coordination. This layout rearranges the structure of the RGB to allow the images more natural and noticeable.

3.4 | SVM classification

Support vector machines (SVM) are usually known to be a classification method, which can be used both for classification problems and for regression. This can handle many categorical and continuous variables quickly. SVM constructs a hyperplane for various groups within multidimensional space. SVM produces an iterative and efficient hyperplane to minimize an error. The key concept for SVM is to consider a total max hyperplane, dividing the dataset into groups the best way possible (Figure 2).

Given training vectors $a_j \in \mathbb{R}^q$, $j = 1,\ldots, n$, in two classes, and a vector $y \in \{1, -1\}^n$, The following equation:

$$\begin{aligned} &\min_{c,b,\zeta} \frac{1}{2} c^M c + W \sum_{j=1}^{n} \zeta_j \\ &\quad \text{subject to } \; b_j \left(c^M \phi(a_j) + b \right) \geq 1 - \zeta_j, \\ &\quad \zeta_j \geq 0, \quad j = 1,\ldots,n \end{aligned} \tag{4}$$

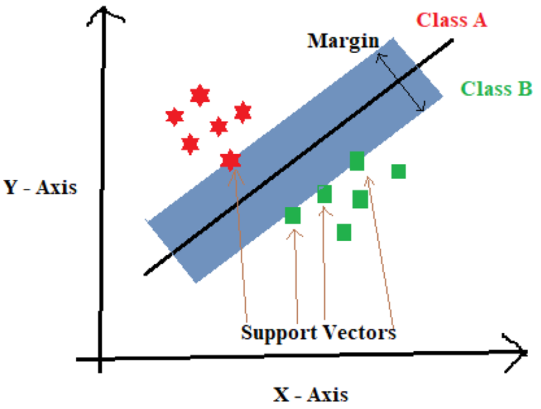

FIGURE 2 SVM classification

Its dual is

$$\min_{\alpha} \frac{1}{2}\alpha^T R\alpha - e^M\alpha \quad \text{subject to} \quad b^M\alpha = 0, \quad 0 \leq \alpha_j \leq W, \quad j = 1,\ldots,n \tag{5}$$

where e is the vector of all ones, $W > 0$ is the upper bound, R is an n by n positive semidefinite matrix, $R_{ji} \equiv b_j b_i L(a_j, a_i)$, where $L(a_j, a_i) = \phi(a_j)^M \phi(a_i)$ is the kernel. Here training vectors are implicitly mapped into a higher (maybe infinite) dimensional space by the function (Alamdari et al., 2016).

The decision function is:

$$\operatorname{sgn}\left(\sum_{j=1}^{n} b_j \alpha_j L(a_j, a) + \rho\right) \tag{6}$$

3.5 | CNN classification

CNN is a DL algorithm that can take any input image, allocate value (biases and learnable weights) to numerous objects/aspects in the picture, discriminate between them. The needed preprocessing in a ConvNet is substantially smaller than other classification algorithms. Although the filters are processed in rudimentary methods and are adequately trained, ConvNets can learn these characteristics/filters (Figure 3).

- Convolution is often represented mathematically with an asterisk * sign. If we have an input image represented as i and a filter represented with f, then the expression would be (Batool & Chellappa, 2014)

$$C = i * f \tag{7}$$

- Leaky Rectified Linear Unit (LeakyReLU) an activation function that also has an alpha α value like ReLU; the alpha value is preferred between 0.1 and 0.3. In this function, there is no 'dead ReLU' (or 'dying ReLU') problem. When the ReLU has values under 0, this completely blocks learning in the ReLU because of gradients of 0 in the negative part (Zheng et al., 2018).

$$A(C) = \begin{cases} C & \text{if } C > 0 \\ \alpha C & \text{if } C \leq 0 \end{cases} \tag{8}$$

- The output from the convolution layer was a 2D matrix. Ideally, we would want each row to represent a single input image. In fact, the fully connected layer can only work with 1D data. Hence, the values generated from the previous operation are first converted into a 1D format. Once the data is converted into a 1D array, it is sent to the fully connected layer. All of these individual values are treated as separate features that

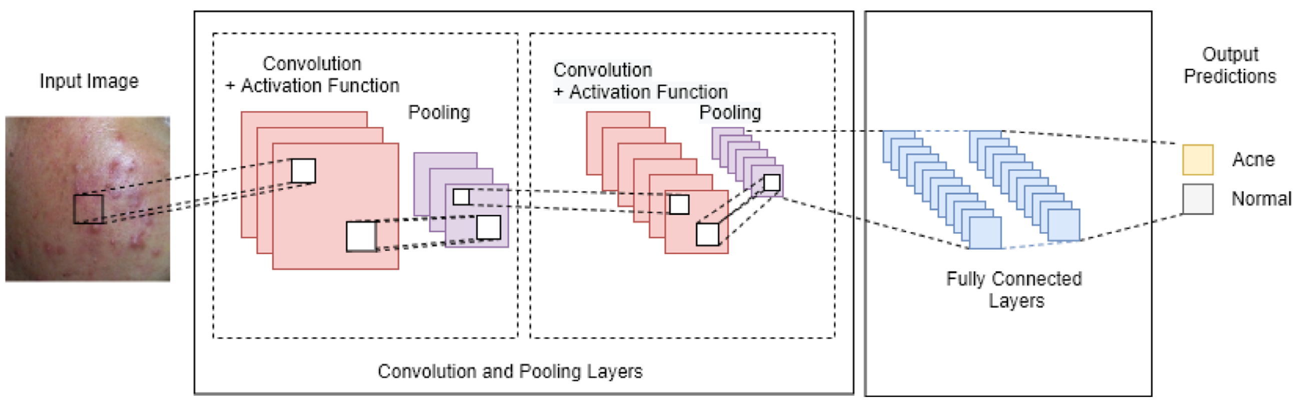

FIGURE 3 Basic CNN classification diagram

represent the image. The fully connected layer performs two operations on the incoming data – a linear transformation and a non-linear transformation.

$$F = W^T \cdot A + bias \tag{9}$$

4 | FACIAL ACNE CLASSIFICATION AND DETECTION

4.1 | Data description

Two types of the dataset were collected manually; one is the facial acne images dataset; another is the normal facial images dataset. The facial acne dataset has been obtained from (DermNet NZ, 2007), where many skin disease image datasets are available. Out of those various skin disease images, only facial acne images containing 200 images were collected. One hundred normal face images were obtained from various random websites that are available online for free. The normal face images were not directly usable; that is why each image was cropped manually. After the collection of both datasets, only 60 images each from both acne and normal datasets were selected for further process (Table 1).

4.2 | Image preprocessing and proposed workflow

The dataset should be cleaned and preprocessed before being processed by any algorithm. Here, the acne and normal datasets are not very much, so the dataset is manually cleaned, cropped, and selected the correct images that are to be processed.

After manual dataset handling, three different image segmentation techniques have been applied on five images from the acne image dataset as follows:

- First, two stages of k-means clustering have been applied (firstly with 2 clusters and secondly with 3 clusters, shown in Figure 4), which is the simplified ver. of the 'Color-based Clustering Algorithm'.
  - In 2 clusters clustering, the original image has been fed to the algorithm. Then the RGB image got converted to $L^*a^*b^*$ image shown in Figure 5(b). Then after applying the k-means algorithm, the image got classified into two different colour images, shown in Figure 5 (c) and (d).
  - In 3 clusters clustering, here also the original image has been fed to the algorithm. Then the RGB image got converted to $L^*a^*b^*$ image shown in Figure 6(b). Then after applying the k-means algorithm, unlike the 2 clusters clustering, here the image got classified into three different colour images, shown in Figure 6(c), (d), and (e).
- Second, the 3-D Gaussian method has been applied to smoothing the original picture. Then MATLAB Rangefilt function has been applied to calculate the local range of the picture. Then the image has been inverted by using the Complement function in MATLAB, shown in Figure 7.
- Lastly, HSV model-based segmentation has been applied using MATLAB; it is also called 'Color Blob Utility with Automatic Thresholding and Tolerance Calculations' method to segment any image.
  - In this method, the image has been loaded into the program; then the program gives an option to select the area of interest of the image. Then the area has been selected manually, shown in Figure 8(a). Then, the program finds the colour value for applying the mask on the image. Then the program generates the selected colour area, shown in Figure 5(c). Then the selected colour area got converted to an HSV colour space image. The initial filter has been applied over the original image using the HSV image. Then the program post-processes the filtered image in order to remove noises. In the end, the image gets filled with the colour, and the program generates output. The flow diagram is shown in Figure 9.

Now, as shown in Figure 10, five output images of K-Means and five output images of the HSV method have been selected, cropped, and resized manually. Here, the output images of Texture Analysis has been ignored. Then, a total of 10 output images of both have been replaced

TABLE 1 Data description table

Sl no.	Class	Collected images	Used images	Source
0	Acne	200	60	DermNet NZ
1	Normal	100	60	Google Images

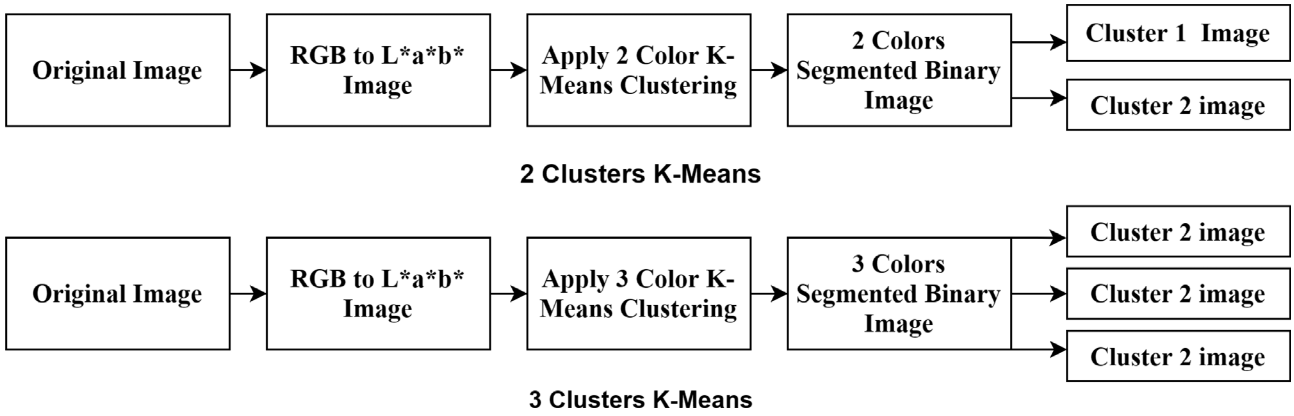

FIGURE 4 Workflow diagram of 2 stage K-means clustering

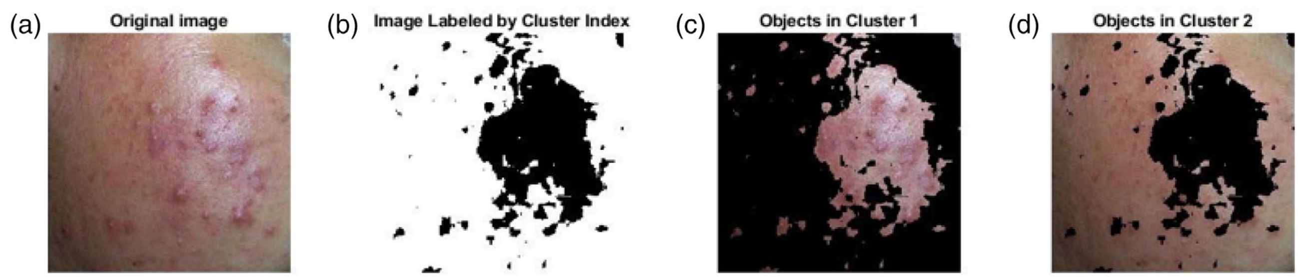

FIGURE 5 (a) Original (b) 2 Colours in '*a*b**' space (c) detected acne (d) skin

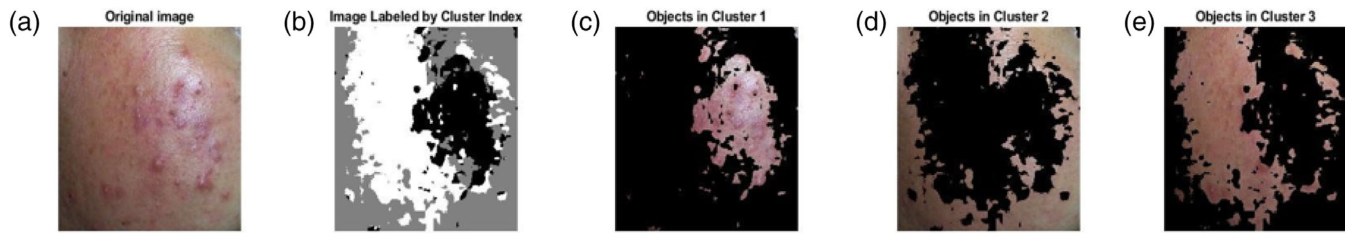

FIGURE 6 (a) Original (b) 3 Colours in '*a*b**' space (c) detected big acnes (d) skin (e) detected small acnes

FIGURE 7 Workflow diagram of texture analysis

with source input images (used for K-Means and HSV Model-Based Segmentation) inside the acne images dataset. After all of that image replacing, the newly mixed 60 acne images and 60 normal has been used for SVM and CNN classifications, shown in Figure 10.

- In this SVM classification, all the images have been fed to the program. Then the program automatically combines both acne+normal images and splits into 80% train data and 20% test data. Then the model has been trained with the SVM Parameter Optimization technique using python library Scikit-Learn with the module 'sklearn.svm.SVC'; the diagram is shown in Figure 11.

After SVM classification, two different CNN models have configured using different activation functions; First, with the "ReLU" activation function and secondly, with the 'LeakyReLU' activation function. Both of the CNN models consist of one dense layer with the 'ReLU/LeakyReLU' activation function, one dense layer with the 'Sigmoid' function, three convolution layers ($64 \times 3 \times 3$), two max-pooling layers ($64 \times 2 \times 2$). In

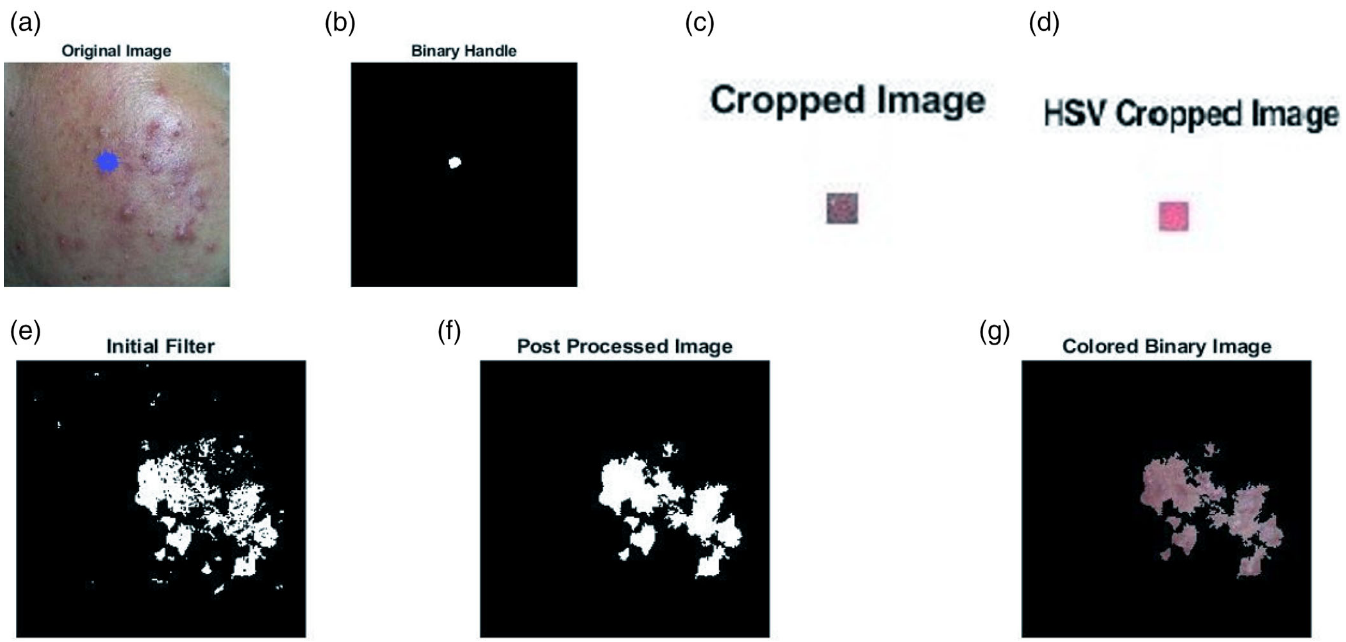

FIGURE 8 (a) Original image (b) binary representation of image (c) cropped image (d) HSV of cropped image (e) initial filter image (f) post processing image (g) desired output

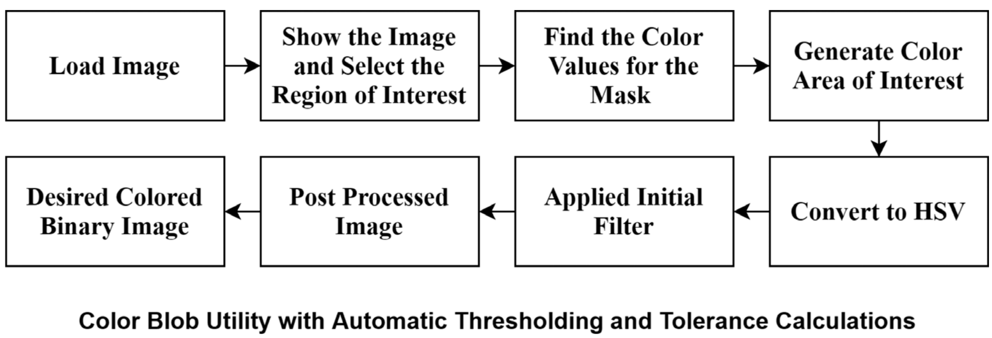

FIGURE 9 Workflow diagram of HSV model-based segmentation

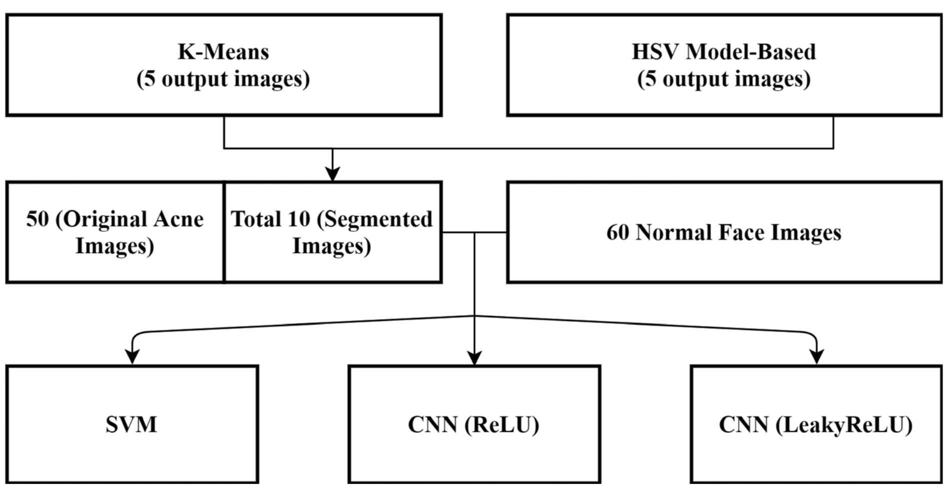

FIGURE 10 Workflow diagram after applying K-means and HSV model-based segmentation

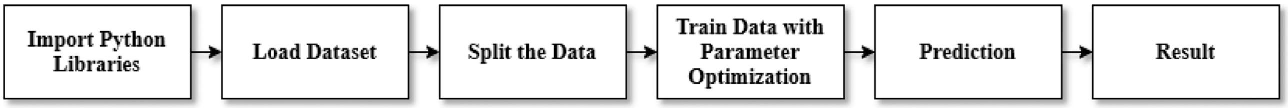

FIGURE 11 Workflow diagram of SVM using Scikit-learn

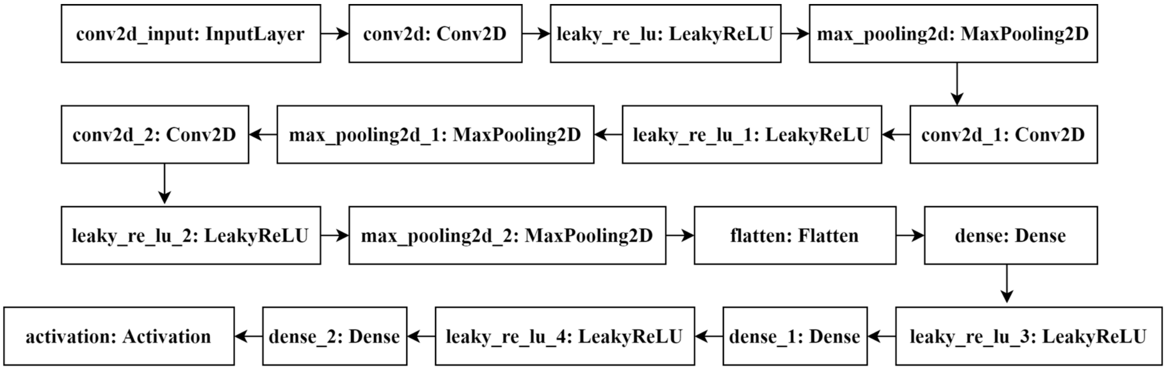

FIGURE 12 Workflow diagram of proposed CNN layers using LeakyReLU

the first CNN model, 5 ReLU activation function has been used. In the second CNN model, 5 LeakyReLU activation function has been used, shown in Figure 12.

- Like SVM, here also all the images that have been fed to the program. Then the program automatically combines both acne+normal images and splits into 80% train data and 20% test data. Then all the images get converted into greyscale images, resized into the same size, then converted to NumPy array. The two different models have been trained with a batch size of 1, 20 epochs, and 0.001 learning rate.

4.3 | Hardware and software used

In order to implement the proposed work, Acer Notebook is with specifications (Ryzen 5 Quad Core CPU with Radeon Vega 8 Integrated GPU, 240 GB SSD + 1 TB HDD and 8 GB of RAM) and Jupyter Notebook having Python 3.8 with (Scikit-Learn 0.23.1, Tensorflow 2.3.0 and Keras 2.4.3 libraries) have been used.

5 | RESULTS AND DISCUSSIONS

In the proposed works, For segmentation of the images, initially, three different methods have been applied: (i) K-Means, (ii) Texture Analysis (iii) HSV Model-Based Segmentation. After applying all the methods, visual results are shown below.

- After applying two stages of K-Means clustering, as shown in Figure 5(c), the acne regions are segmented from normal skin, and in Figure 5(d), the skin is segmented from acne regions. And by using 3 Clustered K-Means, as shown in Figure 6(c), the big acne regions are segmented from the skin and small acne regions. In Figure 6(d), the skin is segmented from big and small acne regions together, and in Figure 6(e), the small acne regions are segmented from normal skin and big acne regions.
- After applying Texture Analysis, as shown in Figure 13(e), there are some grey dot spots in the image, so by that it can be determined that the grey dot spots are the acne.
- After applying the HSV Model-Based method, as shown in Figure 8(g), the only red acne regions are segmented from the normal skin.

After applying all the image segmentation approaches, as per the visual assessment, it is determined that the Texture Analysis method cannot differentiate the acne from normal skin quite accurately. But the K-Means method accurately detects 70% of the acne from the normal skin region, and HSV Model-Based Method detects 50% of the acne from the normal skin region.

Now, K-Means and HSV Model-Based method's output images have been utilized for training SVM and CNN models.

- In the SVM classification using Scikit-learn, the model has been trained with a 2-class classification. Then the trained model has been tested, and an accuracy of 0.88 is achieved; detailed results are shown in Table 2. It means 88% of the test images are detected and classified into its correct class.
  - Class: Here, 0 is acne skin images, and 1 is normal skin images.
  - Precision: It analyzes how reliable a model is when positive labels are expected. Precision raises the question of how much a pattern has been positive, out of the number of times it has been positive? Precision reflects the appropriate percentage of the performance (Erika, 2019). The precision formula is as follows:

$$precision = \frac{TP}{TP+FN} \tag{10}$$

- Recall: It calculates the percentage of actual positives a model correctly identified (True Positive). We should use recall when the false negative cost is higher (Nitin, 2014). The formula for the recall is below:

$$recall = \frac{TP}{TP+FN} \tag{11}$$

- F1 Score: The f1-score gives you the harmonic mean of precision and recall. The scores corresponding to every class will tell you the accuracy of the classifier in classifying the data points in that particular class compared to all other classes (Aishwaria, 2020). The formula for F1 is below:

$$F = 2.\frac{precision.recall}{precision+recall} \tag{12}$$

- Support: It is the number of samples of the true response that lie in that class.
- Accuracy: It addresses the question, what percent of the predictions for models are correct?. The formula for accuracy is below:

$$accuracy = \frac{TN+TP}{TP+FP+TN+FN} \tag{13}$$

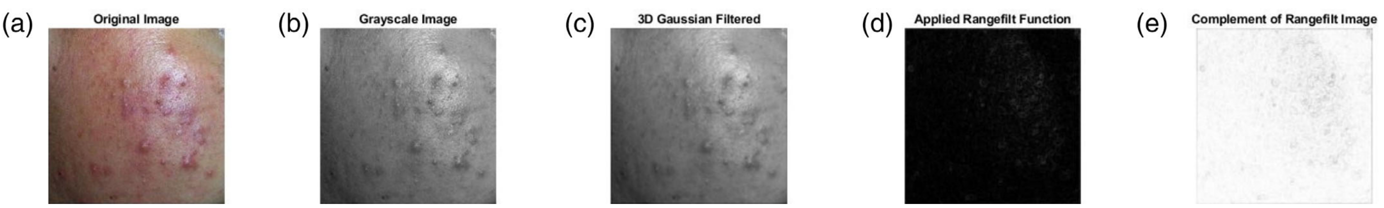

FIGURE 13 (a) Original image (b) grayscale image (c) 3D Gaussian filtered image (d) Rangefilt function image (e) desire textured image

TABLE 2 SVM using Scikit-learn classification report

Class	Precision	Recall	F10-score	Support
0	0.85	0.92	0.88	12
1	0.91	0.83	0.87	12
Avg./Total	0.88	0.88	0.88	24

At last, two CNN DL models have been trained; one with the 'ReLU' activation function and another with the 'LeakyReLU' activation function. After training both of the models, two different graphs of each model are obtained with loss and accuracy.

- In CNN with the ReLU model, as shown in Figure 14(a), the training loss started from 0.7 and then kept decreasing till the 18th epoch. After the 18th epoch, the loss became 0.1 and stayed below 0.1 till the last epoch. In the accuracy graph, as shown in Figure 14(b), the accuracy started from 0.64 and kept increasing till the 18th epoch. After the 18th epoch, the accuracy became 1.0 till the last epoch. In this model, an accuracy of 95.58% is achieved with a loss rate of 10.08%.
- In CNN with the LeakyReLU model, as shown in Figure 15(a), the training loss started from 0.69 and then kept decreasing till the 18th epoch. After the 18th epoch, the loss became 0.1 and stayed below 0.1 till the last epoch as like CNN with ReLU. But, in the accuracy graph, as shown in Figure 15(b), the accuracy started from 0.57 and kept increasing till the 17th epoch. After the 17th epoch, the accuracy became 1.0 till the last epoch. In this model, an accuracy of 97.54% is achieved with a low loss rate of 6.36%.

After training all the models, the results are shown in Table 3. Here, CNN with ReLU DL algorithm performs best out of rest two models with an accuracy of 97.54%. The CNN with ReLU model does perform well, but the accuracy of this is 95.58%, a little less than the LeakyReLU model. The SVM model performs poorly, and the accuracy of this model is 88%, which is not a good accuracy.

As per the prediction results shown in Table 3, out of five predictions, the SVM model has only detected four images correctly, but one incorrectly. But both CNN models have predicted five of the images correctly.

In most of the past proposed methods, authors have used a huge amount of datasets to achieve good accuracy. But the accuracy of the proposed CNN (LeakyReLU) model in this paper is better than many past proposed methods (shown in Table 4), even though only 120 images have been used to train the model.

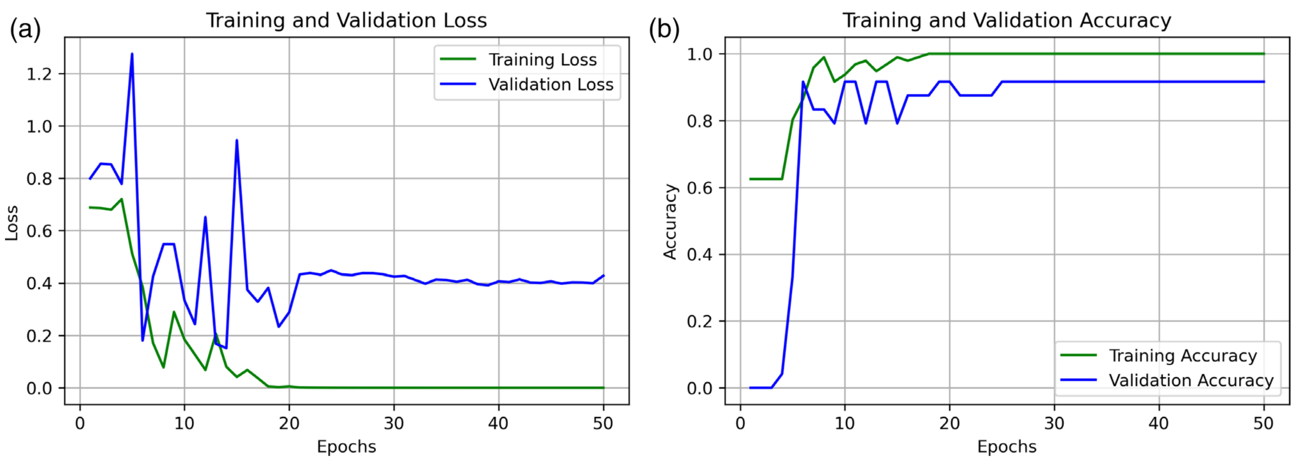

FIGURE 14 (a) Training and validation loss (ReLU) (b) training and validation accuracy (ReLU)

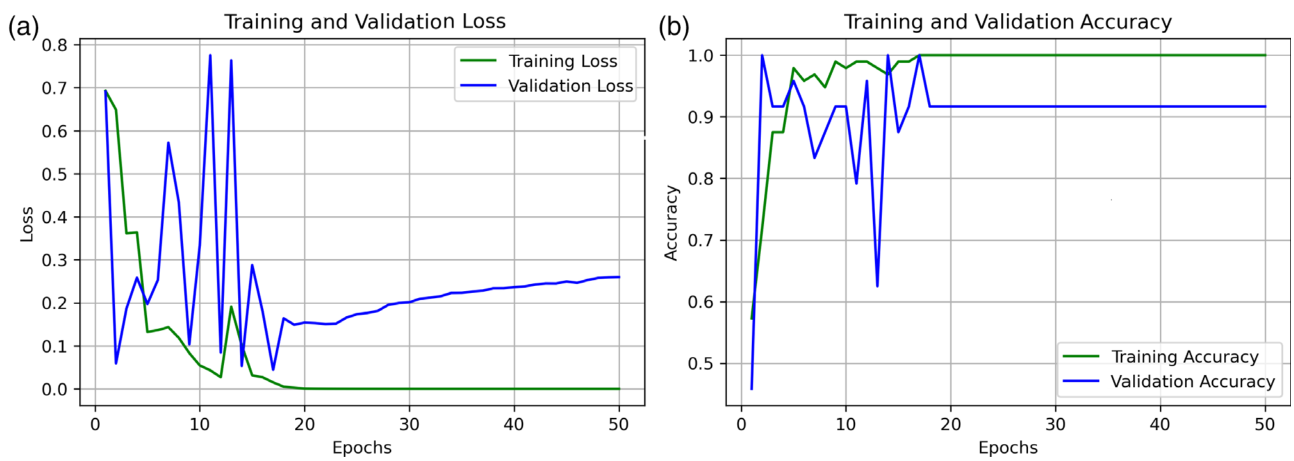

FIGURE 15 (a) Training and validation loss (LeakyReLU) (b) training and validation accuracy (LeakyReLU)

TABLE 3 Result composite report table

Methods	SVM	CNN	CNN
Activation Function	N/A	ReLU	LeakyReLU
Learning Rate	N/A	0.001	0.001
Batch Size	N/A	1	1
Epoch	N/A	50	50
Loss	N/A	10.08%	6.36%
Accuracy	88%	95.58%	97.54%
Prediction (0 = Acne, 1 = Normal)	True: 0,1,0,0,1 Predict: 1,1,0,0,1	True: 0,1,0,0,1 Predict: 0,1,0,0,1	True: 0,1,0,0,1 Predict: 0,1,0,0,1

TABLE 4 Comparison with other methods

Methods	Training images used	Accuracy
Alarifi's GoogLeNet (NAG) Model	164	89.90%
Jung's CNN Model	15,594	96.26%
Hameed's NBC Method	40	93.42%
Lobo's SSD MobileNet V2	3024	91.00%
Proposed CNN Model	120	97.54%

6 | CONCLUSION AND FUTURE SCOPE

In this research, the goal was to propose a CNN model using very less amount of images or datasets, which can detect and classify facial acne images very accurately. In this paper, a CNN model with LeakyReLU has been proposed. For image preprocessing or image segmentation, three methods, K-Means clustering, Texture Analysis, and HSV Model-based method, have been applied. While the Texture Analysis approach applied in the analysis is not adequate, but HSV and K-Means techniques had satisfactory performance in many cases. Using the output images of HSV and K-Means algorithms, SVM with Scikitlearn, CNN with ReLU, and CNN with LeakyReLU models have been trained. Out of these three models, the proposed CNN model with LeakyReLU is able to reach an accuracy of 97.54%. The accuracy of the proposed model is also better than many previous methods proposed in the past by various researchers. This proposed model is also better than the various previous models because this model gives very high accuracy even-though the number of images used to train the model is very less. In the future, we wish to extend our work to improve and optimize our proposed model using parameter fine-tuning.

CONFLICT OF INTERESTS

On behalf of all authors, the corresponding author states that there is no conflict of interest.

ACKNOWLEDGMENTS

Dr. Babita Pandey is thankful to UGC -BSR start-up grant no F.30-469/2019 (BSR), India under which the current research work is carried out. Dr. Dang N.H. Thanh also thankful to University of Economics Ho Chi Minh City, under which research work was funded and carried out.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Aditya Khamparia https://orcid.org/0000-0001-9019-8230

REFERENCES

Aishwaria S. (2020). Demystifying the Mathematics Behind Convolutional Neural Networks (CNNs). https://www.analyticsvidhya.com/blog/2020/02/mathematics-behind-convolutional-neural-network/.

Alamdari, N., Tavakolian, K., Alhashim, M., & Fazel-Rezai, R. (2016). Detection and classification of acne lesions in acne patients: A mobile application. In *2016 IEEE international conference on electro information technology (EIT)* (pp. 739–743). IEEE.

Alarifi, J. S., Goyal, M., Davison, A. K., Dancey, D., Khan, R., & Yap, M. H. (2017). Facial skin classification using convolutional neural networks. In *International conference image analysis and recognition* (pp. 479–485). Springer.

Banerjee, P., & Bandyopadhyay, S. K. (2018). *Modern trends on facial skin disease detection system*. Springer.

Batool, N., & Chellappa, R. (2014). Detection and inpainting of facial wrinkles using texture orientation fields and Markov random field modeling. *IEEE Transactions on Image Processing*, *23*(9), 3773–3788.

Bharati, S., Podder, P., & Mondal, M. R. H. (2020). Hybrid deep learning for detecting lung diseases from X-ray images. *Informatics in Medicine Unlocked*, *20*, 100391.

Chen, B. H., Cheng, H. Y., Tseng, Y. S., & Yin, J. L. (2021). Two-pass bilateral smooth filtering for remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters*, *32*, 26–32.

Chen, B. H., Tseng, Y. S., & Yin, J. L. (2020). Gaussian-adaptive bilateral filter. *IEEE Signal Processing Letters*, *27*, 1670–1674.

DermNet NZ. (2007). Facial acne images. https://dermnetnz.org/topics/acne-face-images/.

Doukas, C., Stagkopoulos, P., Kiranoudis, C. T., & Maglogiannis, I. (2012). Automated skin lesion assessment using mobile technologies and cloud platforms. In *2012 annual international conference of the IEEE engineering in medicine and biology society* (pp. 2444–2447). IEEE.

Erika, D. (2019). Accuracy, recall & precision. https://medium.com/@erika.dauria/accuracy-recall-precision-80a5b6cbd28d

Hameed, A.Z., Awad, W.K., Irsan, N.A. and Abdulbaqi, A.S. (2019). Hybrid technique for skin pimples image detection and classification. 20-32.

Jung, C., Yeo, I., & Jung, H. (2019). Classification model of facial acne using deep learning. *Journal of the Korea Institute of Information and Communication Engineering*, *23*(4), 381–387.

Kwon, Y. H., & da Vitoria Lobo, N. (1999). Age classification from facial images. *Computer Vision and Image Understanding*, *74*(1), 1–21.

Lobo, H., Raut, S., Augustin, R. and Padate, R. (2020). *A Mobile application for Detection & Classification of facial acne*. Elsevier.

Maroni, G., Ermidoro, M., Previdi, F., & Bigini, G. (2017). Automated detection, extraction and counting of acne lesions for automatic evaluation and tracking of acne severity. In *2017 IEEE symposium series on computational intelligence (SSCI)* (pp. 1–6). IEEE.

Mayo Clinic. (2020). Acne, diseases & conditions. https://www.mayoclinic.org/diseases-conditions/acne/symptoms-causes/syc-20368047.

Narkhede, U. P., & Adhiya, K. P. (2014). Evaluation of modified K-means clustering algorithm in crop prediction. *International Journal of Advanced Computer Research*, *4*(3), 799.

Nitin. (2014). What does the numbers in the classification report of sklearn mean?. https://stats.stackexchange.com/questions/117654/what-does-the-numbersin-the-classification-report-of-sklearn-mean.

Padilla-Medina, J. A., León-Ordoñez, F., Prado-Olivarez, J., Vela-Aguirre, N., Ramírez-Agundis, A., & Diaz-Carmona, J. (2014). Assessment technique for acne treatments based on statistical parameters of skin thermal images. *Journal of Biomedical Optics*, *19*(4), 046019.

Pixels, S. P. O. R. (2017). Application of color based image segmentation paradigm on rgb color pixels using fuzzy c-means and k means algorithms.

Shen, X., Zhang, J., Yan, C., & Zhou, H. (2018). An automatic diagnosis method of facial acne vulgaris based on convolutional neural network. *Scientific Reports*, *8*(1), 1–10.

Zheng, X., Lei, Q., Yao, R., Gong, Y., & Yin, Q. (2018). Image segmentation based on adaptive K-means algorithm. *EURASIP Journal on Image and Video Processing*, *2018*(1), 68.

AUTHOR BIOGRAPHIES

Neha Yadav, School of Computer Science and Engineering, Lovely Professional University, Jalandhar, India.

Sk Md Alfayeed, School of Computer Science and Engineering, Lovely Professional University, Jalandhar, India.

Aditya Khamparia, Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Satellite Centre, Tikarmafi, Amethi, India.

Babita Pandey, Department of Computer Science, Babasaheb Bhimrao Ambedkar University, Lucknow, India.

Dang N.H. Thanh, Department of IT, University of Economics Ho Chi Minh City, Vietnam.

Sagar Pande, School of Computer Science and Engineering, Lovely Professional University, Jalandhar, India

How to cite this article: Yadav, N., Alfayeed, S. M., Khamparia, A., Pandey, B., Thanh, D. N. H., & Pande, S. (2022). HSV model-based segmentation driven facial acne detection using deep learning. *Expert Systems*, *39*(3), e12760. https://doi.org/10.1111/exsy.12760