eISSN No. : 2976-2278

UTM Computing Proceedings
Innovations in Computing Technology and Applications
Volume: 6 | Year: 2024 | ISBN No. : 2976-2278

CLASSIFICATION OF ACNE VULGARIS AND ROSACEA USING CONVOLUTIONAL NEURAL NETWORK

Arina Zafira Amran, Noor Hidayah Zakaria
Faculty of Computing,
Universiti Teknologi Malaysia,
Johor, Malaysia
arinazafira@graduate.utm.my, noorhidayah.z@utm.my

***Abstract*— This study aimed to develop a classification model using convolutional neural networks (CNN) with VGG and ResNet architectures to differentiate between acne vulgaris and rosacea, two common skin conditions. The researchers utilized a dataset of photographs showing various stages and symptoms of the two conditions. The dataset underwent preprocessing to enhance image quality and normalize color fluctuations. Training, validation, and testing sets were created for model evaluation. The VGG and ResNet architectures, known for their strong performance in computer vision tasks, were used to extract meaningful features from the images. The pre-trained models were fine-tuned on the training set to capture distinctive characteristics of acne vulgaris and rosacea. Hyperparameters were optimized to prevent overfitting, and model performance was evaluated using metrics like accuracy, precision, recall, and F1 score. The experimental results revealed that both the default CNN and VGG models achieved high accuracy rates of 83% and 88% respectively. The ResNet model, after hyperparameter tuning, achieved moderate results with an accuracy of 74%.**

Keywords- Acne Vulgaris, Rosacea, CNN, VGG, ResNet, Classification, Deep Learning

I. INTRODUCTION

Deep learning has made important contributions across many sectors over the past year and has kept making progress. Innovating new methods, improving current models, and obtaining amazing outcomes have been the goals of researchers and practitioners.

Consistent convolutional neural network (CNN) success in computer vision tasks is one noteworthy development. In image classification, object detection, and segmentation, CNNs have successfully demonstrated their efficacy. These jobs are now much more accurate and productive thanks to

CNNs' capacity to automatically learn and extract pertinent information from photos.

Some of the factors that affect acne include increased sebum production, keratinocyte hyperproliferation, inflammation, and bacterial colonization with Propionibacterium acnes. Hormones, the environment, inflammatory and neurologic mediators, and lipid metabolism all have an impact on sebum production. Acne is brought on by P. acnes' interactions with keratinocytes, sebocytes, and innate cutaneous immunity. A number of factors, including androgen hormones, growth factors, the insulin/insulin growth factor-1 signaling pathway, and diet, affect sebaceous gland activity and the development of acne. By making dietary changes, such as eating low-glycemic-load meals, acne lesions that are both inflammatory and non-inflammatory may be decreased. Rosacea is a chronic inflammatory condition that mostly impacts the face. Because some of its clinical signs are similar to those of acne and other skin disorders, diagnosis can be challenging. An all-encompassing strategy is required to correctly diagnose rosacea.

II. EASE OF USE

According to a detailed study from (2) on a computer-aided diagnosis (CAD) system for classifying multiple skin diseases using machine learning. The proposed approach involves pre-processing, segmentation, feature extraction, and classification phases.

The dataset used in the research work consists of 1800 images of various skin diseases. The images were collected from different sources, including the International Skin Imaging Collaboration (ISIC) and the Dermatology Image.

Database (DID). The proposed system achieved an impressive 83% accuracy for six-class classification. The accuracy of the proposed CAD system is a significant outcome of this study. The system achieved an accuracy of 83% for six-class classification, which is a promising result considering the diversity of the images used in the dataset. The authors attribute the high accuracy to the use of color and texture features and the selection of SVM with a quadratic kernel for classification. The results of the experiments are compared to existing literature on skin lesion classification. The paper discusses the significance of CAD systems in dermatology and future directions for research. The authors suggest that future studies are required regarding the segmentation of inflammatory skin lesions. They also propose that the skin lesion classification results can be improved by using deep neural networks, which is the work under development.

Five convolutional layers and three fully connected layers of a large, deep convolutional neural network are used in this study (4) to investigate ImageNet categorization. With top-1 and top-5 test set error rates of 37.5% and 17.0%, respectively, the network outperforms previous state-of-the-art methods and achieves greater performance. The strategy using six sparse-coding models on average that were trained on various features resulted in the best performance during the ILSVRC-2010 competition at 47.1% and 28.2%. Deep convolutional neural networks may be used for tasks other than picture classification, according to the study, which also claims that larger networks and more labeled data can produce even better results.

Gray Level Co-Occurance Matrix (GLCM) filters with Naive Bayes and Convolutional Neural Network (CNN) were used by the authors of research (1) to classify Rosacea and Acne Vulgaris, two types of face skin illnesses. The dataset was compiled via dermnet and kaggle.com sources, and classification was carried out in consultation with a dermatologist. While the CNN method's accuracy was 74.24%, the GLCM method's accuracy was just 45%. With an accuracy of 96.67% versus 93.33%, the CNN model surpassed the Nave Bayes model. The study showed how image processing and teledermatology techniques might be used to accurately and effectively classify skin disorders. Future study should aim to increase training data in terms of quantity and better qualities, according to the scientists.

The study (3) describes an automated convolutional neural network (CNN) technique for diagnosing face acne vulgaris. The technique employs binary-classifiers to distinguish between skin and non-skin and seven-classifiers to distinguish between acne vulgaris and healthy skin. It extracts features from photos. The VGG16 neural network that has already been trained extracts features for subsequent classifiers. Due to the limited information in the 50x50 box, the approach performs well in detecting healthy skin, blackheads, whiteheads, and nodules but performs poorly in detecting papules, cysts, and pustules.

Acne and rosacea are two frequent skin conditions that are experienced by a large number of people globally, according to this study's authors (5). Similar symptoms between these two skin disorders can cause a case to be misdiagnosed. People who suffer from these two skin rashes typically do not go to dermatologists for a medical diagnosis, instead turning to over-the-counter drugs and cosmetics for self-care. Although both rosacea and acne are typically thought of as being non-dangerous, treating rosacea with acne medication (and vice versa) might aggravate symptoms. In this study, we offer a deep learning model that, when trained on photos of sick skin, can automatically identify between Rosacea and acne instances. We used picture augmentation to increase the data set because there were not many images available. The results of the experiments demonstrate that our model performs well, with an average testing accuracy of 87.1% (almost 10-folds) and 91.2% on the validation set. The model's capacity to classify novel, unheard-of instances is demonstrated by its strong prediction performance. We think that a model like this can be a useful starting point for developing software that automatically distinguishes between acne and rosacea.

III. METHODOLOGY

CNN's performance in classifying acne vulgaris and rosacea involves three phases. The goal is to ensure that the research is carried out correctly. The initial study and research planning phases are followed by the data preparation, design and development of an algorithm phase, then the testing and assessment phases. The flow of this project's research process is represented in Figure 1.

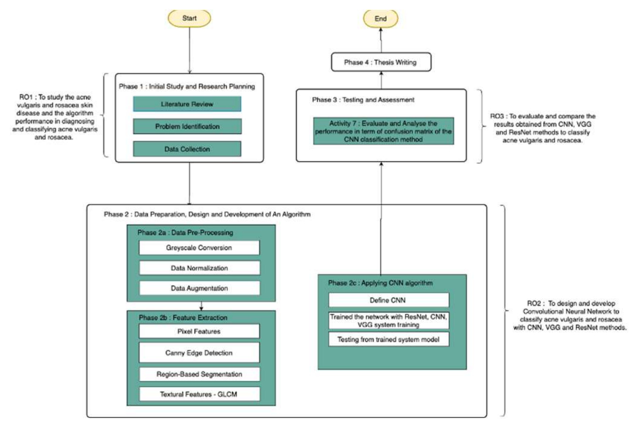

Figure 20. Flowchart of the research methodology

A. *Maintaining the Integrity of the Specifications*

This study uses a Kaggle dataset of 19,500 images from 23 skin illnesses, including acne vulgaris and rosacea, from the dermatology source Dermnet. The dataset includes 312 photos of acne rosacea in the test set and 840 photos in the training set. The resolutions vary between images and categories.

TABLE I. TOTAL DATASETS

Set	Images
Train Set	840 Images
Test Set	312 Images

B. *Data Pre-processing*

Data pre-processing is a process that transforms raw data into a clean and tidy dataset before statistical analysis. It assesses and enhances data quality for confidence in statistical analysis. Pre-processing involves cleansing, integration, transformation, and reduction, and may be repetitive. It is crucial to avoid bias by modifying the dataset to ensure accurate statistical results (6).

eISSN No. : 2976-2278

Data pre-processing, including grayscale conversion, data standardization, and data augmentation, is crucial for computer vision and machine learning applications. Grayscale conversion reduces complexity and resource usage, while data normalization ensures uniform pixel values and prevents feature dominance. Data augmentation increases model diversity, robustness, and generalization capacity, reducing overfitting and improving model generalization.

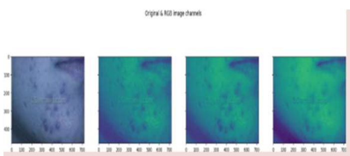

Figure 21. Original and RGB Image

Figure 22. Grayscale Image

C. Feature Extraction

Feature extraction in computer vision tasks involves pixel characteristics, edge detection, region-based segmentation, and textural aspects. Edge detection locates boundaries, while pixel features gather low-level data. Region-based segmentation separates images based on color similarity, texture, or intensity uniformity. GLCM extracts textural information by quantifying spatial interactions between pixel pairs, benefiting tasks like texture classification, material detection, and medical picture analysis.

The following stage is to extract Canny edge detection to obtain a character's characteristic value after the color feature extraction is complete. Figure 6 showing the result of canny edge detection on the random image from the dataset that highlights its features.

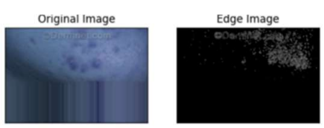

Figure 23. Canny Edge Detection

TABLE II. FEATURES OBTAINED FROM THE GRAY LEVEL CO-OCCURRENCE MATRIX (GLCM)

Feature	Value
Contrast	[[75.74973343 110.61960041 66.5026125 85.89230868]]
Dissimilarity	[[5.69453523 6.72538988 4.33898457 6.31313208]]
Homogeneity	[[0.22303546 0.20270577 0.51203332 0.20035982]]
Energy	[[0.02856315 0.0274911 0.05960075 0.02727717]]
Correlation	[[0.95412602 0.93299851 0.95968835 0.94798322]]
ASM	[[0.00081585 0.00075576 0.00355225 0.00074404]]

IV. RESULT AND DISCUSSION

It is simple to separate the dataset into training and testing data because it has already been sorted into these categories with default ratios of 70% and 30%, respectively. In a simple two-part data split, the training data set is used to build models and train them. The estimation of various parameters or the comparison of the performances of several models both typically employ training sets. The testing data set is utilized after the training is finished.

TABLE III. RATIO OF DATASET SPLITTING

Training (%)	Testing (%)
70	30

In order to discover the best configuration for better performance for the ResNet model, hyperparameter optimization tuning algorithm entails modifying the hyperparameters of a machine learning model.To achieve reliable classification results, parameter selection for a ResNet model is crucial. Grid search strategies can be applied to identify the optimal hyperparameters for the

TABLE IV. PARAMETER OF PARAMETERGRID

Parameter	Params	Value
Epochs	num_epochs	[5, 10, 15]
Batch Size	batch_size	[32, 64, 128]
Learning Rate	lr	[0.001, 0.01, 0.1]

The performance metrics of various models (CNN, VGG, ResNet, and ResNet with ParameterGrid) in terms of precision, specificity, and sensitivity (which can be referred to as recall) are shown in Table 5.4. The accuracy of the CNN model is 83%, meaning that 83% of the cases were classified correctly. Additionally, the specificity and sensitivity (recall) are correspondingly 83% and 80%. These numbers show that the model does a fair job of identifying both positive and negative events, with a balanced performance.

TABLE V. PARAMETER OF PARAMETERGRID

Method	Accuracy	Specificity	Sensitivity/ Recall
CNN	0.83	0.83	0.80
VGG	0.94	0.98	0.92
ResNet	0.60	0.69	0.67
ResNet with ParameterGrid	0.74	0.76	0.73

The findings show that the VGG model performs the best overall, having the highest accuracy, specificity, and sensitivity/recall scores. Although it performs better than the standard ResNet model, the ResNet model with the ParameterGrid implementation still falls short of the VGG model in terms of performance. The CNN model performs almost as well as the other models under consideration, whereas the simple ResNet model performs the least well.

V. CONCLUSION

At the end of this research, all three objectives to classify the acne vulgaris and rosacea stated had been achieved. First, acne vulgaris and rosacea skin disease and the algorithm performance in diagnosing and classifying acne vulgaris and rosacea had been studied. Second, a Convolutional Neural Network to classify acne vulgaris and rosacea with CNN, VGG and ResNet methods had been developed. Third, the results obtained from CNN, VGG and ResNet methods to classify acne vulgaris and rosacea were compared. As a result, Visual Geometry Group (VGG) chosen as the most suitable Convolutional Neural Network (CNN) architecture in classifying the acne vulgaris and rosacea with the accuracy 94% and ResNet achieve better performance with hyperparameter tuning with accuracy of 74%.

ACKNOWLEDGMENT

The Universiti Teknologi Malaysia (UTM) also supported this work through UTM Fundamental Grant (PY/2023/01586) with Cost Centre No. Q.J130000.3828.23H44.

REFERENCES

[1] Cahya Rahmad, Rosa Andrie Asmara, Alvionitha Sari Agstriningtyas. Acne Vulgaris and Rosacea Skin Diseases Image Classification using Gray Level Co-Occurance Matrix and Convolutional Neural Network. 2021 International Conference on Electrical and Information Technology (IEIT). 2021 Sep 14;4.

[2] Hameed N, Shabut A, Hossain MA. A Computer-aided diagnosis system for classifying prominent skin lesions using machine learning [Internet]. IEEE Xplore. 2018 [cited 2023 Jun 13]. p. 186–91. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8 6741832.

[3] 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).

[4] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM. 2012 May 24;60(6):84–90.3.

[5] Gerges F, Shih FY, Azar D. Automated Diagnosis of Acne and Rosacea using Convolution Neural Networks. 2021 Sep 24;

[6] .Malley B, Ramazzotti D, Wu JT. Data Pre-processing. Secondary Analysis of Electronic Health Records. 2016;115–41.