SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I11P125

Volume 11 Issue 11, 264-274, November 2024

Original Article

A Fine Tune CNN Model for Human Skin Type Classification

Ruchika Chouhan[*1], Sachin Patel[2]

[1,2]Sage University, Indore, India.

**Corresponding Author : katariya.15@gmail.com*

Received: 14 September 2024 Revised: 15 October 2024 Accepted: 13 November 2024 Published: 30 November 2024

Abstract - *Cosmetic consumers must know their skin type when choosing particular products. The contemporary lifestyle can be quite hectic, often leaving minimal time for self-care. Nevertheless, prioritizing self-care, particularly skincare, remains essential despite busy schedules. Simply depending on popular products or in-store suggestions may not effectively determine whether a skincare item suits an individual's specific skin condition. Determining these types can be challenging, especially when different skin areas present various conditions, such as oily or dry. Skin specialists can provide more accurate assessments. Recently, Artificial Intelligence (AI) and Machine Learning (ML) have been utilized in various fields, including healthcare, to assist in identifying and predicting different conditions. This study aimed to develop a skin type classification model using Convolutional Neural Networks (CNN), a deep learning approach. The dataset consisted of 3,152 images representing normal, oily, and dry skin, with 1,274 images for normal skin, 1,120 for oily skin, and 758 for dry skin. Several CNN architectures were optimised and evaluated, including AlexNet, VGG16, and ResNet50. However, these models did not meet the expected performance levels. The results revealed that the proposed fine-tuned CNN architecture achieved the best performance, with a validation accuracy of 99.62% and an average loss of 22.74%. Following hyperparameter tuning, the accuracy increased to 94.57%, with a validation loss of 0.0113. This indicates significant improvements in the model's ability to classify skin types accurately.*

Keywords - *Cosmetic product, Makeup, Recommendation, Skin type, Deep Learning techniques.*

1. Introduction

Skin care plays a vital role in maintaining a balanced living. Over time, the overall age of consumers using cosmetic products dropped, with increasing interest among young boys and girls and professionals. These products help enhance youthful appearances, thereby boosting self-esteem. Numerous brands have introduced various skin care items in response to this growing demand. However, it is crucial to assess individual skin types before purchasing. Experts emphasize the importance of understanding one's skin type for effective product selection. While some skin types can be easily identified, others may require careful consideration, as some individuals may experience a combination of oily and dry areas. The skin care market has expanded significantly over the last twenty years, with the combination of Artificial Intelligence across several fields, including healthcare and cosmetics. Deep Learning techniques enable computers to grasp examples and are increasingly applied to tackle real-world challenges like skin cancer classification.

Today, cosmetics play a significant role in enhancing people's appearances. Using e-commerce platforms, customers can access no different beauty products online [27]. However, selecting the appropriate product for a particular skin type can be challenging. Predictions help consumers make better choices, allowing them to select products well-suited for their skin type.

This process benefits greatly from deep learning technology, which helps simplify complex decision-making in the cosmetics and beauty industry. As the beauty sector evolves and expands, with more products and a growing customer base, finding the right cosmetics has become increasingly user need and requirement. Cosmetics significantly impact everyone's appearance, and individuals must choose products based on personal features, such as skin type. Since everyone has unique skin conditions like oily, dry, or normal, selecting the right product can be complicated. Additionally, some users may have particularly difficult skin conditions, making the selection even more challenging. AI algorithms can handle these complexities by analyzing vast amounts of unstructured data to provide valuable insights and recommendations [28].

In deep learning, Convolutional Neural Networks (CNN) act as the core neural part of a system, where data is processed using a pre-trained model to produce output based on the given input. The process is integrated into the system so that when input is provided, the application automatically utilizes the trained datasets to deliver precise results. The primary advantage of using a CNN is its ability to achieve high accuracy in the proposed method. It operates on datasets to extract meaningful information from them [29]. However, it is limited to processing inputs matching the data the neural network has been trained on. This task is tackled using an AI system that efficiently manages large volumes of unorganized information and yields encouraging outcomes. CNNs, or Convolutional Neural Networks architecture, can easily do this. The system analyses human facial images and identifies the skin type (such as dry, oily, combination, or normal). Based on that, the system will suggest a product suitable for that specific type [31]. Initially, image recognition is performed, and the recognized image is then fed into a model for further processing (image classification), resulting in the final output. Essentially, this involves "analyzing and manipulating images."

The process consists of three primary steps:

- First, import all images from various devices, including cameras, sensors, scanners, or even system-generated photos.
- Then, evaluate, manipulate, and enhance the images while providing a descriptive data summary.
- Finally, the image processing results could involve minor image modifications, report generation, image prediction, and more.

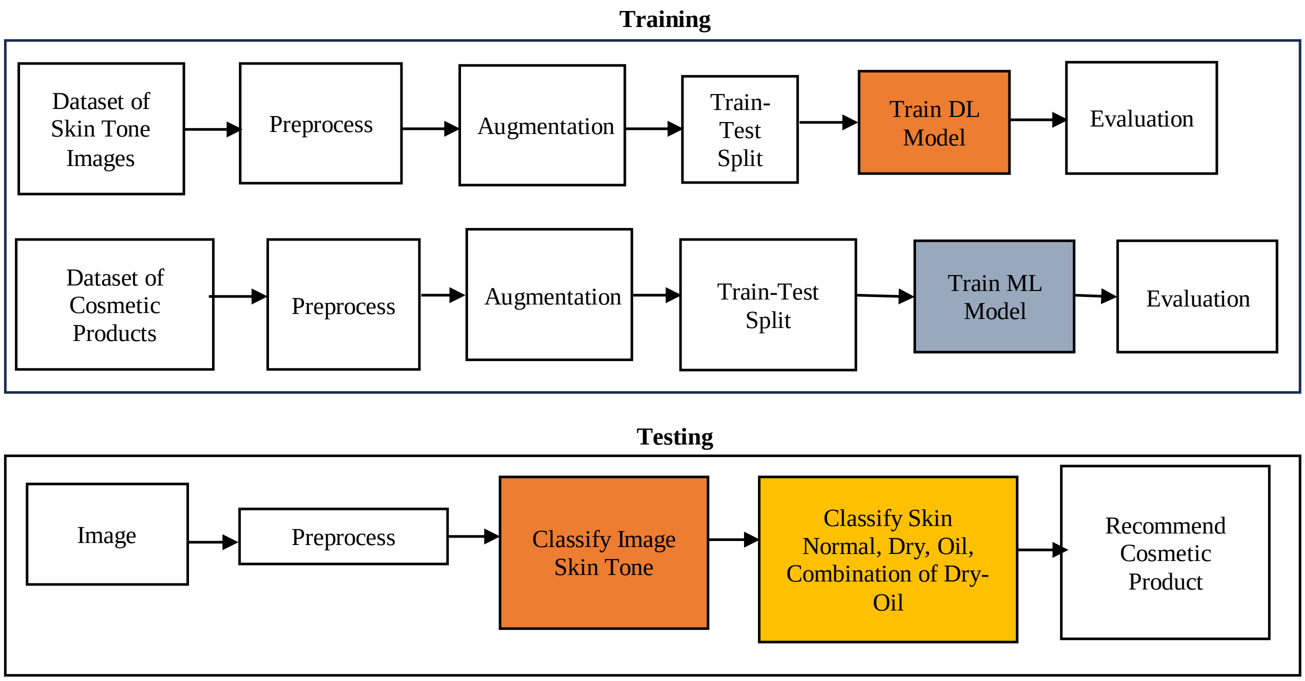


Fig. 1 Proposed overall system

The main concept involves gathering human facial images, preprocessing them, and training a model to classify different skin types. Once the model accurately classifies the skin types, it can recommend the most suitable cosmetic products for each skin type. This approach utilises advanced image processing and machine learning techniques to analyze facial images and extract relevant features that determine skin type. The process starts with collecting and preprocessing facial images to ensure the model receives high-quality input. After training, the model can accurately identify the skin type, crucial for recommending appropriate skincare products. The accompanying figure illustrates this process, primarily focusing on classifying skin types to enhance personalized cosmetic recommendations.

Choosing the best skin product is difficult because it depends on skin type and user preferences. Users need a simple and effective system that can classify skin type and tone and, based on that, provide suggestions for cosmetic products. That is all a user wants when purchasing online cosmetic products. The proposed model can solve this problem of end users.

2. Related Work

This section discusses different research papers based on cosmetic product suggestions, skin type classification, skin tone, and face shape identification using a deep neural network model, specifically CNN.

S. Umer (2020) [10] uses 40 different cosmetic products and clicks the product's image using a mobile phone. The proposed algorithm takes input images and recognizes the product's brand, product and availability. They also perform some analytical tasks for brand and retail recognition.

In another research, the author uses a deep neural network model [16]. That system suggests the makeup style on the user's facial image and shows how that particular makeup looks on their face. Unlike other research, they used 961 females in two image databases, one image before makeup and one after makeup. This extensive database makes this system more efficient.

R. Iwabuchi (2017) [18] On the number of websites consumers review. For example, when a person tries to search for skin products on the @cosme website, users read reviews from users with the same characteristics (such as age, skin type, etc.) to find products that align with the skincare routine she follows. For example, a user with dry skin may seek out a lotion known for its hydrating properties. In contrast, someone interested in skin brightening will look for products with positive reviews regarding their whitening effects. Users often assume their compatibility with a skincare product depends on its ingredients.

H. Gunaasighe (2016) [9] proposed a technique based on machine learning to identify face shapes for beauty-related activities like hairstyle, makeup, and eyeglasses. It is required to know about face shape. The author proposed a neural network that identifies a person's facial shape by taking a personal image. The author considers some types of face shapes, i.e. diamond, triangle, inverted triangle, and oval, as well as oblong, square, and round face shapes.

In C. L. Chin (2018) [14], the author suggested a deep learning approach using Convolution Neural Networks (CNNs) for the classification of facial skin images. It builds a system for skin image classification that identifies some skin problems, such as whether skin quality is good or not, bad facial skin quality, and different facial makeup for the face. It does this by using a smartphone to detect facial skin type. As per the results, a model with three convolution layers, three pooling layers, and four fully connected layers gets the maximum recognition rate. When the system is finished, anyone can use it to improve facial skin problems.

C.J. Holder (2019) [13], in this paper, proposed a Siamese convolution neural network. The system aims to propose a vision-based system that can recommend cosmetic products. In the era of big data and deep learning, a recommender system like this is important to predict a person's preferences. Here, the author creates data from 91 persons, taking images of them and cropping the eye area because they mainly focus on mascara products. The system takes the eye area as input and, as a result, produces an output of a product preferred by the persons whose eyes are visually similar.

According to C. H. Hsia (2020), appearance matters to everyone. The primary researcher of the investigation recommended a system that evaluates the degree of acne and the skin on the face to provide product recommendations to customers [24]. In this instance, a camera takes a picture labeled and shows it in a window so that the customer can confirm that it shows the image in the right location. This paper's suggested method accurately determines skin type and detects acne on both cheeks. In addition, the system counts the quantity of acne scars after determining if the skin is normal, oily, or dry. In the future, we may include a chatbot to interact with clients. This aids customers in selecting the ideal cosmetic item.

In modern times, cosmetics have a significant impact on the personal appearance of individuals. However, selecting the best skin care product is becoming challenging. Therefore, there is a need for a predictive approach that helps to understand which product is the best for which skin type. To solve this problem, an AI algorithm will be used because it performs well under high volumes of unstructured data and generalizes with excellent results [27, 30, 31].

3. Dataset

3.1. Human Facial Image Dataset

The dataset consists of 3,152 frontal facial images of humans aged between 20 and 60. It is broken into three sections: training dataset, testing dataset, and validation purposes. For the training dataset, 1,000 images of oily skin, 1,104 images of normal skin, and 652 images of dry skin were used.

The testing set includes 40 images of oily skin, 59 images of normal skin, and 35 images of dry skin. For validation, 84 images of oily skin, 111 images of normal skin, and 71 images of dry skin were utilized. The dataset comprises 1,274 for normal skin types images, 1,120 for oily skin types images, and 758 for dry skin types images. This dataset was applied to transfer learning models such as AlexNet, VGG16, and ResNet50, which were then used to train our proposed fine-tuned model for skin type classification. By leveraging transfer learning, the models could quickly adapt and accurately classify skin types, making them a robust choice for this task.

This structured approach allows for comprehensive training, testing, and validation, ensuring that the models are well-equipped to handle different skin types' diverse characteristics and provide a strong foundation for classification.Data augmentation strategies like rotation, scaling, flipping and color adjustment enrich the dataset. An augmented dataset consisting of 3408 images of human faces.

Oily Skin **Normal Skin** **Dry Skin**

Fig. 2 Different skin types

4. Transfer Learning Model

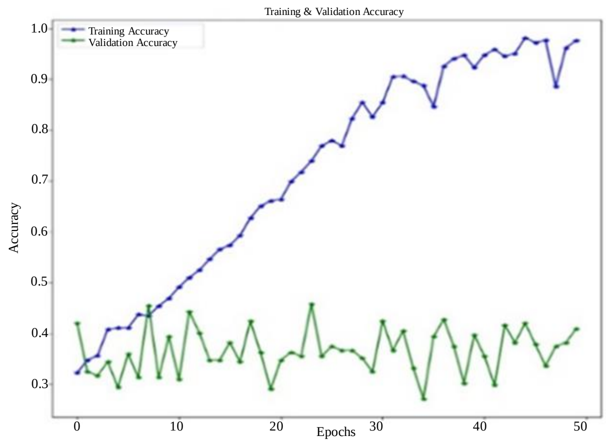
Deep learning is a significant aspect of Artificial Intelligence (AI), enhancing computers' ability to classify using data like images and voice. By using machine learning, complex calculations can now be executed more efficiently. Between many machine learning algorithms, Convolutional Neural Networks (CNNs) stand out for their effectiveness. These networks can take raw images as input, using them as foundational elements for the learning process. AlexNet, VGGNet 16, and ResNeSt 50 are landmark Convolutional Neural Network (CNN) architectures that notably influence deep learning. AlexNet, introduced in 2012, marked a breakthrough with its deep structure and use of ReLU activation functions, winning the ImageNet competition and demonstrating the power of deep learning. VGGNet 16, developed by the Visual Geometry Group in 2014, further pushed the boundaries by emphasizing depth with 16 weight layers, consistently using 3x3 convolutional filters to enhance feature extraction. This model showcased the importance of network depth in achieving higher accuracy. ResNeSt 50, a more recent innovation, reduces the fading gradient problem. This model achieves state-of-the-art performance by splitting convolutional feature maps into cardinal groups and applying channel-wise attention. Despite their differences, these models share common principles: they all employ multiple layers to process images, utilize pooling layers to reduce dimensionality, and benefit from advancements in GPU technology for training. Their success in the ImageNet challenge and their foundational influence on subsequent models underscore their critical role in advancing computer vision and deep learning. AlexNet, VGGNet 16, and ResNeSt 50 are transfer learning models that work better on image classification. Alexnet is delicate, VGG 16 is simple and extensive, and ResNet 50 reduces the fading gradient problem. That is why AlexNet, VGGNet 16, and ResNeSt 50 were selected over other potential models.

4.1. AlexNet

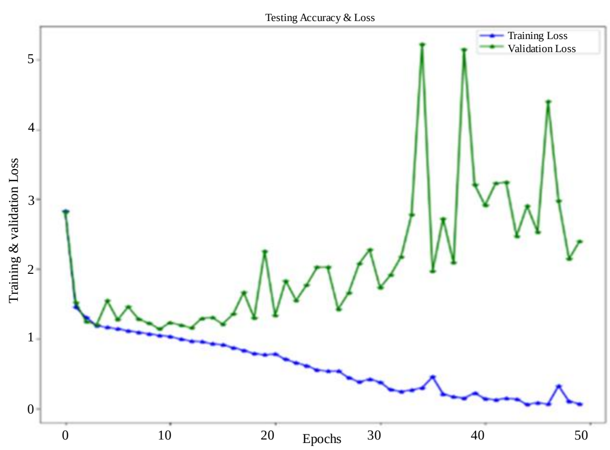
The AlexNet study demonstrated that a large perceptron Recurrent Neural Network (RNN) can perform strongly on a challenging dataset using purely supervised learning methods. AlexNet enriches the dataset using data augmentation strategies like rotation, scaling, flipping, and color adjustment. The architecture also leveraged two GPUs for parallel processing, a groundbreaking approach at the time, to handle the significant computational load required for training deep networks. AlexNet comprises eight layers: five are convolutional, and 3 are fully linked. It uses Rectified Linear Unit (ReLU) activations, accelerating the training task to resemble conventional activation functions.

- Input Layer: Input to AlexNet is a 224x224x3 image.
- 1st Convolutional Layer: This layer uses filters around 96 having 11x11 size and stride of 4, resulting in a 54x54x96 output volume.
- Max-Pooling Layer 1: A pooling layer of size 3x3 with 2 stride 2 decreases the output volume to 26x26x96.
- 2nd Convolutional Layer: In these filters, 256 of size 5x5, generating a 26x26x256 output volume.
- Max-Pooling Layer 2: Another max-pooling 3x3 layer with a stride of 2 decreases the output to 12x12x256.
- 3rd Convolutional Layer: This layer uses 384 filters of size 3x3, maintaining the output size 12x12x384.
- 4th Convolutional Layer: This layer applies 384 filters of size 3x3, resulting in a 12x12x384 output volume.
- 5th Convolutional Layer: In this final layer, filters 256 have a 3x3 size, resulting in a 12x12x256 output.
- Max-Pooling Layer 3: A 3x3 pooling layer with a stride of 2 reduces the output to 5x5x256.
- 1st Fully Connected Layer: In this layer, a total of 4096 neurons are used.
- 2nd Fully Connected Layer: Another fully Linked layer with 4096 neurons.
- 3rd Fully Connected Layer: This Last fully Linked layer has 3 neurons for classification, equal to the 3 skin type categories in the dataset.
- Final Output Layer: The softmax function produces the probability distribution over the classes.

Despite its groundbreaking success, AlexNet has several drawbacks. Its architecture, while innovative, is relatively large and computationally intensive, requiring significant hardware resources, particularly GPUs, for efficient training. The model also tends to overfit due to its high number of parameters despite the use of dropout. Additionally, the large memory footprint can be problematic for deployment on devices with limited resources. Finally, AlexNet's structure, with its fully connected layers, contributes to a lack of flexibility and adaptability compared to more modern, streamlined architectures that followed, such as those utilizing residual connections or more sophisticated layer designs.



(a)



(b)

Fig. 3 AlexNet performance graph

We implemented the AlexNet architecture on our skin type dataset in our research. However, the model's performance did not meet our expectations. Despite its proven effectiveness and pioneering design, the results revealed that AlexNet faced challenges in achieving adequate performance metrics for our specific dataset.

	Precision	Recall	F1-Score	Support
Normal	0.36	0.23	0.28	35
Dry	0.40	0.53	0.46	59
Oily	0.26	0.23	0.24	40
Accuracy			0.36	134
Macro Avg	0.34	0.33	0.33	134
Wighted Avg	0.35	0.36	0.35	134

Fig. 4 Classification report using AlexNet

This finding implies that while AlexNet has shown strong capabilities in different applications, it may not be the best fit for the characteristics of our data and the classification tasks at hand. Therefore, exploring other architectures or optimizing the model further may be essential to improve performance.

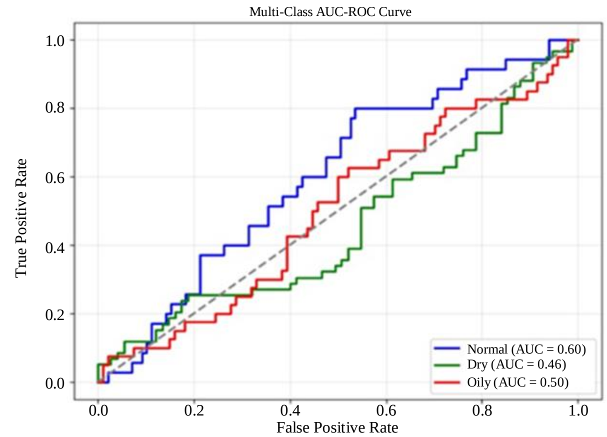


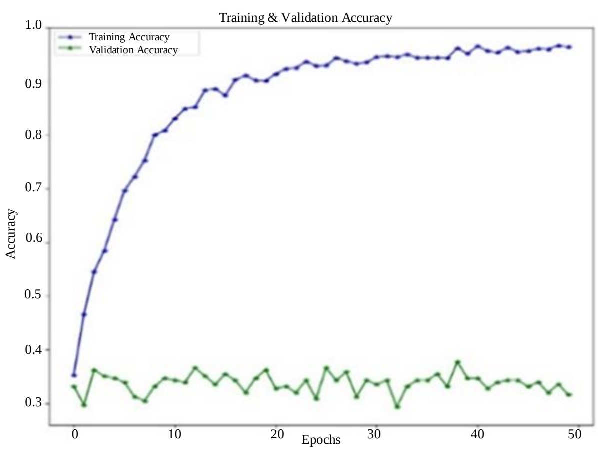
Fig. 5 ROC-AUC curve AlexNet

4.2. VGG16

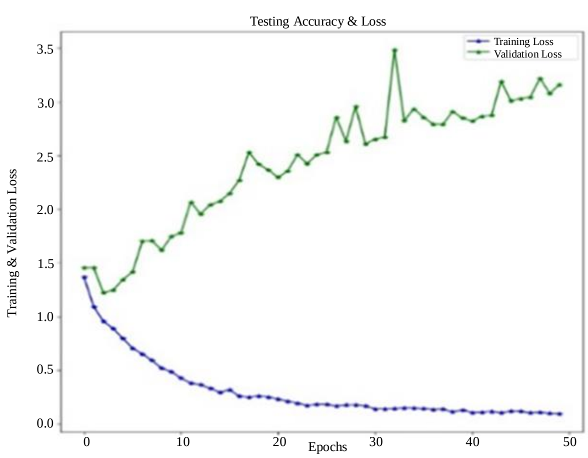
The model has sixteen weight layers, comprising three fully connected layers and thirteen convolutional layers. Its purity and uniform architecture characterize VGG16. It uses small 3x3 convolutional filters filled on each other, allowing the network to learn more complex features while maintaining a manageable number of parameters. VGG16 architecture is designed to progressively reduce the dimensions of the input data with max-pooling layers, which are applied after particular convolutional layers. This pooling operation helps retain important features while reducing computation. VGG16 employs five max-pooling layers, which downsample the feature maps, ultimately leading to a final output that can be classified into 1000 categories. VGG16 can transfer learned features to different tasks, making it a popular choice for transfer learning. Its architecture is a foundation for many subsequent models, influencing the design of more profound and complex networks in deep learning.

- Input Layer: VGG16 accepts input data images that are 224x224 pixels in size and include 3 color channels (RGB).
- Convolutional Layers: VGG16 uses a series of convolutional layers with small 3x3 filters, allowing for deeper networks while maintaining computational efficiency. The architecture is structured as follows:
- Conv Layer 1: 64 filters, followed by another Conv Layer with 64 filters.
- Max-Pooling Layer 1: Reduces the feature map dimensions.
- Conv Layer 2: 128 filters, followed by another Conv Layer with 128 filters.
- Max-Pooling Layer 2: Further reduces the dimensions.

- Conv Layer 3: 256 filters, followed by 2 more Conv Layers with 256 filters each.
- Max-Pooling Layer 3: Continues dimensionality reduction.
- Conv Layer 4: 512 filters, followed by 2 more Conv Layers with 512 filters each.
- Max-Pooling Layer 4: Further pooling.
- Conv Layer 5: 512 filters, followed by 2 more Conv Layers with 512 filters each.
- Max-Pooling Layer 5: Final pooling operation.
- Fully Connected Layers: Three fully linked layers make up the architecture after the convolutional and pooling layers
  1st Fully Connected Layer: 4096 neurons.
  2nd Fully Connected Layer: 4096 neurons.
  3rd Fully Connected Layer: 1000 neurons, corresponding to the number of classes in the ImageNet dataset.
- Output Layer: In the final layer, the softmax activation function creates probability distribution across the classes.



(a)



(b)

Fig. 6 VGG16 performance graph

In our research, we applied the VGG16 architecture to our sample dataset. However, the model did not achieve the expected performance. Despite its proven effectiveness in various applications, the results indicated that VGG16 struggled to deliver satisfactory accuracy on our specific dataset.

	Precision	Recall	F1-Score	Support
Normal	0.35	0.51	0.42	35
Dry	0.43	0.22	0.29	59
Oily	0.34	0.45	0.39	40
Accuracy			0.37	134
Macro Avg	0.38	0.39	0.37	134
Wighted Avg	0.38	0.37	0.35	134

Fig. 7 Classification report using VGG16

This suggests that, while VGG16 is a robust model, it may not be the most appropriate choice for the characteristics of our data and the classification tasks involved. Further exploration of alternative architectures or optimization techniques may be necessary to enhance performance.

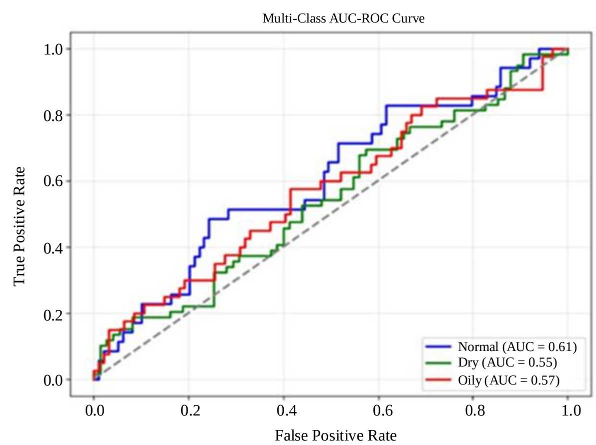
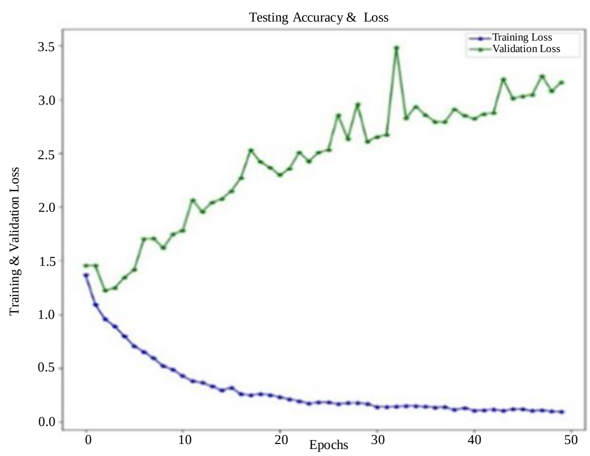


Fig. 8 ROC-AUC curve VGG16

4.3. ResNet50

It is a groundbreaking deep learning architecture introduced by Kaiming He et al. in 2015, designed to address the challenges of training very deep neural networks. The key innovation of ResNet50 is to use residual connections, allowing the input of a layer to bypass one or more subsequent layers and be added to their output. This structure reduces the fading gradient problem, facilitating the training of networks with hundreds or even thousands of layers without degradation in performance. ResNet has several variants, with ResNet50 being one of the most popular, featuring 50 layers. It employs 3x3 filters convolutional and max-pooling layers to extract and downsample features effectively. It primarily focuses on residual learning through the use of skip connections.

(a)



(b)

Fig. 9 ResNet50 performance graph

- Input Layer: Accepts data in image format, 224x224 pixel size with three color channels (RGB).
- Convolutional Layer 1: This layer uses 7x7 convolutional filters with 64 filters and 2 strides, producing feature maps of size 112x112.
- Max-Pooling Layer: A 3x3 max-pooling layer using 2 stride reduces the size to 56x56.
- Residual Blocks: ResNet50 contains 16 residual blocks, each consisting of:
- Convolutional Layer: 3x3 convolutional filters (64, 128, 256, or 512 filters depending on the block).
- Batch Normalization: Applied after each convolution to stabilize training.
- ReLU Activation: The ReLU activation function is used for non-linearity in image classification.
- Skip Connection: Inserts the input of the block to the result, allowing gradients to flow more easily during backpropagation.
- Fully Connected Layer: The final residual block result is compressed and inserted into a fully connected layer containing 1000 neurons, each representing one of the 1000 classes in the ImageNet dataset.
- Output Layer: A softmax function is applied to produce class probabilities.

In our study, we implemented the ResNet50 architecture on our sample dataset and got the result shown in Figure 9. However, the model did not achieve the expected performance outcomes. Despite its effectiveness in various applications, the results indicated that ResNet50 struggled to deliver satisfactory accuracy on our specific dataset.

	Precision	Recall	F1-Score	Support
Normal	0.38	0.34	0.36	35
Dry	0.50	0.42	0.46	59
Oily	0.37	0.47	0.41	40
Accuracy			0.42	134
Macro Avg	0.41	0.41	0.41	134
Wighted Avg	0.43	0.42	0.42	134

Fig. 10 Classification report using ResNet50

This suggests that, while ResNet50 is a robust architecture, it may not be ideally suited for the characteristics of our data and the classification tasks at hand. Further exploration of alternative models or optimization strategies may be necessary to improve performance.

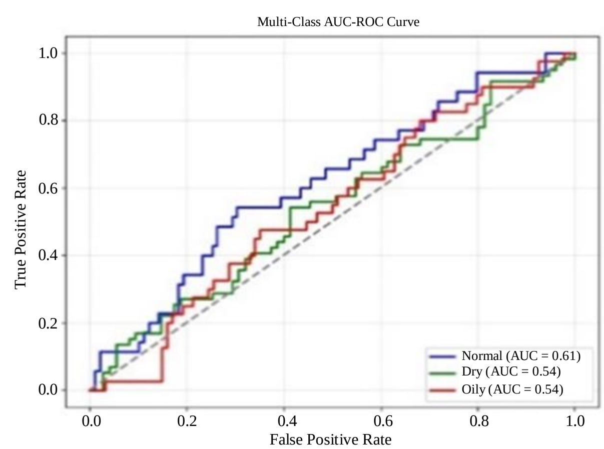


Fig. 11 ROC-AUC curve ResNet50

5. Proposed Fine Tune Model

The suggested architecture uses 3-channel RGB images of size 224x224 pixels with a Convolutional Neural Network (CNN) intended for image classification for various skin types (oily, dry, and normal). The final output of this architecture is three classes.

Let us break down each layer and its role within the network:

1. Input Layer
   Shape: 224x224x3
   This is the input layer for images of 224x224 pixel size with 3 color RGB channels.
   First Convolutional Block

2. Conv2D:
   Output Shape: 112x112x128
   Filter Size / Stride: 7x7 filters, stride 2
   This layer applies 128 convolutional filters of size 7x7, effectively reducing the spatial dimensions of the input by using a stride of 2.
   LeakyReLU:
   In order to prevent dead neurons, the activation function permits a slight, non-zero gradient when the unit is not in use.
   MaxPooling2D:
   Output Shape: 56x56x128
   This layer performs max pooling, reducing the spatial dimensions further by taking the maximum value over a 2x2 window.

3. Second Convolutional Block
   Conv2D:
   Output Shape: 56x56x256
   Filter Size / Stride: 5x5 filters, stride 2
   Applies 256 filters of size 5x5.
   LeakyReLU:
   Applies non-linearity to introduce complexities.
   MaxPooling2D:
   Output Shape: 28x28x256
   Reduces dimensions by taking maximum values over a window, effectively performing downsampling.

4. Third Convolutional Block
   Conv2D (1x1):
   Output Shape: 28x28x256
   Filter Size / Stride: 1x1 filters, stride 2
   This layer uses 1x1 convolutions to control the depth of the network, enhancing learning of cross-channel interactions.
   Conv2D (3x3):
   Output Shape: 28x28x256
   Filter Size / Stride: 3x3 filters, stride 2
   Applies more complex feature extraction.
   LeakyReLU:
   Further non-linear transformation.
   MaxPooling2D:
   Output Shape: 14x14x256
   Downsamples to capture essential features.

5. Fourth Convolutional Block
   Conv2D (1x1):
   Output Shape: 14x14x512
   Filter Size / Stride: 1x1 filters, stride 2
   Again, control depth and focus on channel-wise interactions.
   Conv2D (3x3):
   Output Shape: 14x14x512
   Filter Size / Stride: 3x3 filters, stride 2
   Complex pattern extraction.
   Conv2D (1x1):
   Output Shape: 14x14x512
   Filter Size / Stride: 1x1 filters, stride 2
   Deepening channel interactions.
   Conv2D (3x3):
   Output Shape: 14x14x512
   Filter Size / Stride: 3x3 filters, stride 2
   Final convolution to further refine learned features.
   MaxPooling2D:
   Output Shape: 7x7x512
   to bring features into a manageable size for classification.

6. Flatten Layer
   Output Shape: 25088
   3D feature maps are converted into vectors of 1D features for making fully connected layers

7. Fully Connected Layers
   Dense Layer:
   Output Shape: 4096
   Large fully connected layer allowing for high-capacity learning.

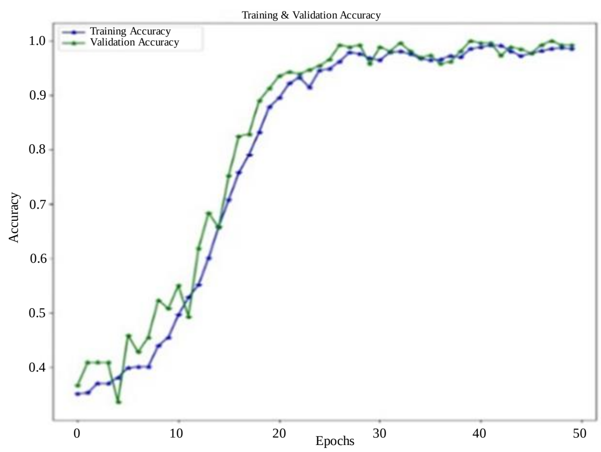
   - Dropout:
     Output Shape: 4096
     Dense Layer:
     Output Shape: 512
     Further fully connected layer to distill features.
     Dense Layer:
     Output Shape: 3
     The final layer with 3 neurons for classification into one of the three target classes.

5.1. Key Characteristics

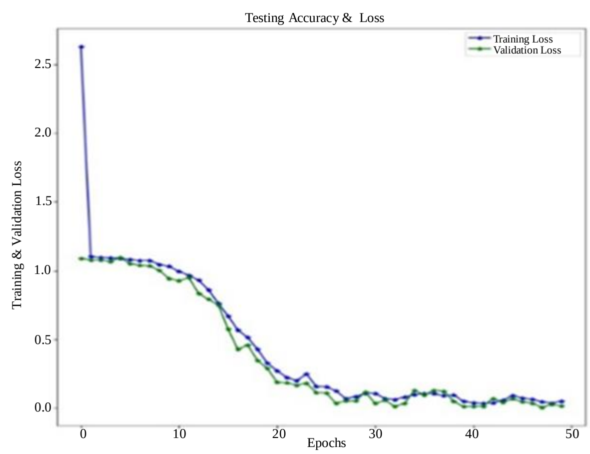
- Activation Functions: Architecture having the LeakyReLU activation function throughout helps maintain a non-zero gradient for inactive neurons, reducing the risk of dead neurons and gradient vanishing issues.
- Pooling Layers: In the pooling layer, max-pooling decreases the dimensionality of feature maps, thereby controlling overfitting and computational costs by condensing spatial information.
- Convolutional Layers: The network uses convolutional layers with different filter sizes (1x1, 3x3, 5x5, 7x7) to capture a wide range of features and spatial hierarchies within the images.
- Fully Connected Layers: Architecture includes a large, fully connected layer for high-capacity learning, followed

by a smaller one, and finally, a classification layer that outputs predictions for three classes.

- Hyper Tune Parameter: Architecture having learning rate 0.0001, 50 epoch, filter 7x7, four convolution layers, Optimizer is Adam, beta_1=0.9, beta_2=0.999, epsilon=None, decay=0.00



(a)



(b)

Fig. 12 Proposed fine tune model performance graph

Based on the results, our proposed fine-tuned model performs well compared to existing transfer learning models. The figure above demonstrates that our model achieves high accuracy while maintaining loss is significantly less. This ensures that the model effectively learns the data and makes accurate predictions.

The fine-tuning process has enhanced the model's performance, surpassing the benchmarks set by traditional transfer learning approaches. Consequently, the model demonstrates strong generalization capabilities and improved precision, making it a more robust solution for the task at hand. Overall, the model shows promising results in accuracy and loss metrics.

	Precision	Recall	F1-Score	Support
Normal	1.00	1.00	1.00	35
Dry	1.00	1.00	1.00	59
Oily	1.00	1.00	1.00	40
Accuracy			1.00	134
Macro Avg	1.00	1.00	1.00	134
Wighted Avg	1.00	1.00	1.00	134

Fig. 13 Classification report using the proposed model

The parameters in the classification report clearly illustrate the models' overall performance. We can assess the effectiveness of the models by analysing the metrics provided.

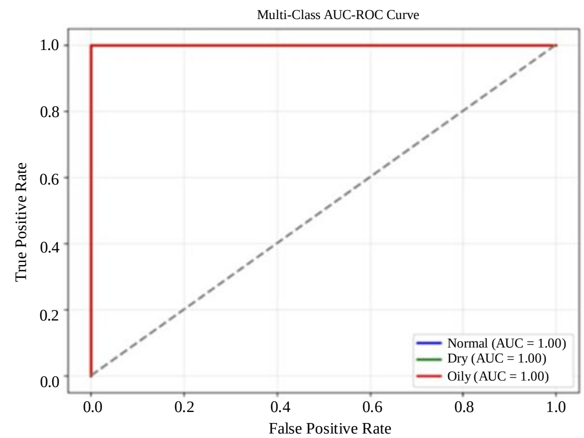


Fig. 14 ROC-AUC curve proposed model

6. Result Analysis

Table 1. Results comparision

Method_Name	Accuracy	Skin Type	Precision	Recall	F1_Score
AlexNet	0.36	Normal	0.36	0.23	0.28
		Dry	0.40	0.53	0.46
		oily	0.26	0.23	0.24
VGG16	0.37	Normal	0.35	0.51	0.42
		Dry	0.43	0.22	0.29
		oily	0.34	0.45	0.39
ResNet50	0.42	Normal	0.38	0.34	0.36
		Dry	0.50	0.42	0.46
		oily	0.37	0.47	0.41
Fine_Tune_Model	1.00	Normal	1	1	1
		Dry	1	1	1
		oily	1	1	1

The table above compares the results obtained from our proposed fine-tuned model against existing models.

7. User Experience

In today's world, everyone goes online to make purchases using various e-commerce sites. This contains lots of information, and users purchase the items using that information. They found it difficult to choose the best skin product because it totally depends on skin type and user personal preferences. The end user can easily interact with the proposed system, and the user inserts a face image based upon that proposed model to identify skin type as normal, dry and oily. After determining the skin type, the user can select the cosmetic product according to skin type and tone. Unlike other models, the proposed model works well for female and male users. Also, AI-powered skin care solutions often require sensitive personal data such as user skin images, health history, lifestyle information, etc.

8. Conclusion

Identifying and predicting skin type is crucial for recommending the best skincare products. This research explores various transfer learning models for skin type classification. The proposed fine-tuned model demonstrated significantly improved accuracy and reliability compared to existing models, showcasing potential in effectively classifying human skin types (normal, dry, oily). Using the proposed model, which offers more personalized skincare recommendations, serves individual needs and preferences. This research highlights the importance of model refinement in achieving superior results. In future work, additional algorithms will be integrated into proposed fine-tuned models to further improve the precision of cosmetic product recommendations. This will enable a more comprehensive understanding of individual skincare needs and provide users with optimized product suggestions. Ultimately, the research advances personalized skincare solutions and enhances the overall user experience in cosmetic product selection.

8.1. Future Research Direction

In future models, more added features like skin texture and hydration level can improve the model's accuracy. Also, one can include different demographical factors. Advanced models like vision transformers or researchers can implement hybrid models to improve the overall model's efficiency and performance.

References

[1] Zhengya Xu, and Hong Ren Wu, "Shape Features Based Extraction for Face Recognition," *2009 4th IEEE Conference on Industrial Electronics and Applications*, Xi'an, China, pp. 3043-3048, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[2] Kristina Scherbaum et al., "Computer-Suggested Facial Makeup," *Computer Graphics Forum*, vol. 30, no. 2, pp. 485-492, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[3] Antitza Dantcheva, Cunjian Chen, and Arun Ross, "Can Facial Cosmetics Affect the Matching Accuracy of Face Recognition System?," *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Arlington, VA, USA, pp. 391-398, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[4] Miroslaw Loboda, and Aleksandra Lopaciuk, "Global Beauty Industry Trends in the 21st Century," *Management, Knowledge and Learning International Conference*, pp. 19-21, 2013. [Google Scholar] [Publisher Link]

[5] Marie-Lena Eckert, Neslihan Kose, and Jean-Luc Dugelay, "Facial Cosmetics Database and Impact Analysis on Automatic Face Recognition," *2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP)*, Pula, Italy, pp. 434-439, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[6] Cunjian Chen, Antitza Dantcheva, and Arun Ross, "Automatic Facial Makeup Detection with Application in Face Recognition," *2013 International Conference on Biometrics (ICB)*, Madrid, Spain, pp. 1-8, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[7] H.N. Gunasinghe, T.G.L. Fernando, "A Computer Aided System to Identify the Face Shape of a Person Using Machine Learning and Linage Processing Techniques," *Proceedings of 16th Conference of the Science Council of Asia*, 2016. [Google Scholar] [Publisher Link]

[8] Aparna Bharati et al., "Detecting Facial Retouching Using Supervised Deep Learning," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 1903-1913, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[9] Gozde Yolcu Oztel, and Serap Kazan, "Virtual Makeup Application Using Image Processing Methods," vol. 1, no, 5, 401-404, 2015. [Google Scholar] [Publisher Link]

[10] Saiyed Umer et al., "Machine Learning Method for Cosmetic Product Recognition: A Visual Searching Approach," *Multimedia Tools and Applications*, vol. 80, pp. 34997-35023, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Swati Solanki, and Gayatri Jain, "Deep Learning Methods for Selecting Appropriate Cosmetic Products for Various Skin Types: A Survey," *International Research Journal of Engineering and Technology*, vol. 6, no. 12, pp. 72-75, 2019. [Google Scholar] [Publisher Link]

[12] Yuki Matsunami, Mayumi Ueda, and Shinsuke Nakajima, "Tag Recommendation Method for Cosmetics Review Recommender System," *Proceedings of the 19th International Conference on Information Integration and Web-Based Applications & Services*, pp. 166-170, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[13] Christopher J. Holder, Boguslaw Obara, and Stephen Ricketts, "Visual Siamese Clustering for Cosmetic Product Recommendation," *Computer Vision-ACCV 2018 Workshops*, pp. 510-522, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] Chiun-Li Chin et al., "Facial Skin Image Classification System Using Convolutional Neural Networks Deep Learning Algorithm," *2018 9th International Conference on Awareness Science and Technology (iCAST)*, Fukuoka, Japan, pp. 51-55, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[15] Victor Martin et al., "Understanding the Visual Effects of Cosmetic Products on Beauty via Deep Learning," *IFSCC Congress*, 2018. [Google Scholar] [Publisher Link]

[16] Taleb Alashkar et al., "Examples-Rules Guided Deep Neural Network for Makeup Recommendation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, pp. 941-947, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[17] Sayaka Yabe, Mayumi Ueda, and Shinsuke Nakajima, "A Comparative Method Based on the Visualization of Cosmetic Items Using their Various Aspects," *2021 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, pp. 1-6, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Rio Iwabuchi et al., "Proposal of Recommender System Based on User Evaluation and Cosmetic Ingredients," *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*, Denpasar, Indonesia, pp. 1-6, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[19] Bendjillali Ridha Ilyas et al., "Enhanced Face Recognition System Based on Deep CNN," *2019 6th International Conference on Image and Signal Processing and their Applications (ISPA)*, Mostaganem, Algeria, pp. 1-6, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[20] Syed Afaq Ali Shah, Mohammed Bennamoun, and Michael K. Molton, "Machine Learning Approaches for Prediction of Facial Rejuvenation Using Real and Synthetic Data," *IEEE Access*, vol. 7, pp. 23779-23787, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] Se-Won Kim, and Sang-Woong Lee, "Deep Learning Approach for Cosmetic Product Detection and Classification," *Computers, Materials & Continua*, vol. 69, no. 1, pp. 713-725, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yuri Takashima, and Masaki Aono, "Predicting the Usefulness of Cosmetic Reviews," *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*, Denpasar, Indonesia, pp. 1-6, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[23] Hsiao-Hui Li et al., "Based on Machine Learning for Personalized Skin Care Products Recommendation Engine," *2020 International Symposium on Computer, Consumer and Control (IS3C)*, Taichung City, Taiwan, pp. 460-462, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Chih-Hsien Hsia et al., "System for Recommending Facial Skincare Products," *Sensors and Materials*, vol. 32, no. 10(1), pp. 3235-3242, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[25] Toan H. Vu, An Dang, and Jia-Ching Wang, "Learning to Remember Beauty Products," *Proceedings of the 28th ACM International Conference on Multimedia*, pp. 4728-4732, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[26] Gyeongeun Lee, Xunfei Jiang, and Natalie Parde, "A Content-Based Skincare Product Recommendation System," *2023 International Conference on Machine Learning and Applications (ICMLA)*, Jacksonville, FL, USA, pp. 2039-2043, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[27] Kavyashree N. et al., "Artificial Intelligence Based Smart Cosmetics Suggestion System Based on Skin Condition," *2022 International Conference on Automation, Computing and Renewable Systems (ICACRS)*, Pudukkottai, India, pp. 797-801, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] R. Abishek et al., "Advanced Skin Category Prediction System for Cosmetic Suggestion Using Deep Convolution Neural Network," *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, pp. 627-631, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[29] A. Kavitha et al., "Cosmetic Suggestion Based on Skin Condition Using Artificial Intelligence," *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*, Tuticorin, India, pp. 1026-1031, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[30] Samrat Ray et al., "Cosmetics Suggestion System Using Deep Learning," *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, Tashkent, Uzbekistan, pp. 680-684, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] S. Bhuvana et al., "Cosmetic Suggestion System Using Convolution Neural Network," *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, pp. 1084-1089, 2022. [CrossRef] [Google Scholar] [Publisher Link]