International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I7P202

Volume 70 Issue 7, 10-23, July 2022

Original Article

Deep Scattering Convolutional Network for Cosmetic Skin Classification

Shwetambari Borade[1], Dhananjay Kalbande[2], Kristen Pereira[3], Rushil Patel[4], Sudhanshu Kulkarni[5]

[1]*Department of Information Technology, Thadomal Shahani Engineering College, Mumbai, Maharashtra, India*
[2]*Department of Computer Engineering, Sardar Patel Institute of Technology, Mumbai, Maharashtra, India*
[3,4]*Department of Information Technology, Sardar Patel Institute of Technology, Mumbai, Maharashtra, India*
[5]*Department of Electronics and Telecommunication, Sardar Patel Institute of Technology, Mumbai, Maharashtra, India.*

[1]shwetambari.borade@gmail.com

Received: 11 May 2022 Revised: 13 June 2022 Accepted: 20 June 2022 Published: 18 July 2022

Abstract *- The use of computer-based solutions or artificial intelligence in detecting face skin conditions has progressed considerably over the years. Combining cosmetic skin care concepts with technical advancements has shown noble outputs. In recent times, deep convolutional neural networks have been applied for image classification applications in various domains like healthcare, identification of objects in various aspects, control systems, machine visions, etc., and have shown remarkable results. The scattering network is commonly and consistently utilized on the initial layers of a supervised mixed convolutional model. This study demonstrates that layers in the first layer do not always have to be learned, with pre-defined models attaining the greatest performance to date when competing with Deep CNNs. ResNETs, in particular, have been adopted widely because they can solve the vanishing gradient problem that prevails in most deep networks. So, they have exploited the capacity of wavelet transforms, Scattering Networks, and the Residual Networks model for image classification to identify cosmetic skin problems like oily or dry skin. According to results obtained and statistical assessments, the ResNet model and Scattering Networks have the greatest validation accuracy of 98 percent for image classification of oily and dry skin pictures. They have also compared the model with others where scattering networks for SVM have an accuracy of 57 percent, whereas VGG16 gave an accuracy of 95.9 percent on validation sets. And DWT with ANN has an accuracy of 63.2 percent, and Gabor filters and support vector machine have the lowest accuracy of 48.8 percent.*

Keywords *- Deep Learning, Histogram Equalization, ResNet, Skin Classification, Wavelet Scattering.*

1. Introduction

The skin makes up about 15% of the total weight of the human body. A square inch of skin on the human body has approximately 300 sweat glands. Changes in the skin can indicate a medical condition or changes in body health. Skin type varies largely among individuals and can be broadly classified into two types dry skin and oily skin. Knowing what type of skin you can help narrow down the possible skin diseases you might be vulnerable to can also be instrumental in suggesting an effective remedy. The global skin care products market was worth $140.92 billion in 2020, and it's predicted to grow at a CAGR of 4.69% by 2026. This humongous figure is self-suggestive that people invest a lot in their skin. Normal skin, oily skin, dry skin, combination skin, and sensitive skin are different skin types. The skin of a normal person is neither too dry nor overly greasy. It has very few flaws and is not extremely sensitive. It has a beautiful complexion with barely discernible pores. It has a regular texture and is smooth to the touch. The skin tone is more or less even [1]. Excess oils and pores are related to oily skin produced by hyperactive sebaceous glands [2]. Oily skin is characterized by a glossy, thick, slippery, and heavy texture. It's more common in younger individuals, but as they become older, their skin becomes drier. Dry skin, commonly known as xerosis, is a common skin ailment that affects people [3]. It is usually not a significant condition, although it might sometimes be difficult to cure. It might be linked to other skin disorders or pathological situations, including diabetes, malnutrition, etc. Dry skin is a troublesome condition that causes itching, scaling, and cracking. It can occur during winter when moisturizing is neglected and might also be caused due to use of harsh skin products. With the increasing awareness about skincare and the increasing number of skin disorders, the need for prompt skin type classification [4] can help pharmacists and dermatologists narrow down to a few from the immense possible scope of potential skin ailments. But most people still don't use the right product according to their skin type as most people are simply unaware of their skin type [5]. So, knowing one's skin type could be very beneficial. This paper proposes a wavelet scattering-based neural network classifier which trains on a small dataset and gives promising results.

2. Related Works

In their paper [6], the authors discuss a deep learning-based procedure for classifying different skin types. Input training images are first pre-processed to lessen the computational power, which might have required much time to train. The process incorporated in this stage is reducing the image channels, i.e., converting an RGB image to grayscale, and after the conversion, the image is resized to a 32x32 pixel map. The image is then passed through convolution layers of filter size 3x3 for feature extraction. Rectifier Linear Unit (ReLu) is the activation function in the convolution layer. After the CNN layers, max-pooling is employed to reduce image size and finally passes through a fully connected layer after flattening the pixels. The model ran on a dataset of 20 images (12 training and 8 validation) for 150 epochs and reached an accuracy of 67\% and loss value of -9.2992.

Another technique for classification was developed based on a series of image preprocessing of the training data by authors in their paper titled [7]. In this study, the characteristics of the face are extracted using the local binary pattern (LBP) and the discrete wavelet transform (DWT). Small waves are called wavelets, and DWT aims to convert the image signal into a series of waves. An image is divided into four sub-images of size 1/4th of the original input image. The upper right, lower right, and lower left positioned sub-images roughly look the same as the original one. At the same time, the one on the upper left will be of finer version as it contains lower frequency components as opposed to the other three parts. LBP is used to identify textures in a face. LBP compares the binary values of pixels in the center of the image with 8-pixel values around it in which each pixel has a different value. Finally, a Support Vector Machine is then used to do the classification process of multiclasses [8]. The SVM classification runs as many times as the number of classes to classify the training data correctly. The median filtering method, which essentially replaces a given pixel with a median of all pixels in the window specified by size, was used as a pre-processing step to remove any noise, i.e., denoising the image. The suggested technique achieves an average classification accuracy of 91.66% while running in 31.571 seconds.

Researchers achieved an accuracy of about 95% in detecting facial skin type by performing DWT on the image and then training an ANN model on this data. Researchers showed in their paper [3] that after performing certain preprocessing steps like filtering and histogram equalization, the oily and dry skin images had different stage 4-connectivity and 8-connectivity region property values which could be used in further classification. However, their observation was based on a small dataset of 9 images per class.

Authors in their paper [9] made use of LBP, and the Scattering Transforms introduced in ScatNet to perform texture classification and achieved state-of-the-art results. This technique might prove beneficial in the skin type classification problem as the skin texture of oily and dry skin is fundamentally different. Further, the use of Wavelet Scattering Networks proves beneficial in data-constrain scenarios. Authors [10] have introduced a software package, kymatio, which is compatible with most contemporary deep learning frameworks and can be used as an initialization for convolutional networks.

The authors in the paper [11] provide a competitive strategy based on scattering networks for large-scale visual tasks. They developed a supervised Shared Local Encoder (SLE) that allows scattering networks to outperform existing local encoding approaches at the ILSVRC2012. This network of only three learned layers enables a thorough study of the processes. Their results also reveal that pre-defined qualities remain relevant and can provide important insights into deep learning algorithms while making them more interpretable. They offer stronger theoretical guarantees when combined with proper learning methods, which are required to construct superior deep models and durable representations.

Authors of the paper [12] state that they find scattering transform coefficients with adjusted parameters give improved feature representations for Speech Emotion Recognition. The enhanced wavelet averaged scale provides sufficient stability and invariance against unnecessary temporal changes to capture real-time emotional signals. The layer-wise analysis highlights the importance of time-domain averaging over standard scalogram/Mel-spectrogram coefficients. The findings obtained using second layer scattering coefficients show the significance of amplitude modulation of time series of various scalogram frequency bins. Their work is initial research into the Scattering transform for Speech Emotion Recognition.

In the paper [13], the authors claim that this study aims to overcome the drawbacks of traditional DSNs based on the FRWT, which is ideal for non-stationary texture analysis and may be conceived of as a library of linear translation-variant bandpass filters. The fractional wavelet scattering network was created after they started with fractional waveform scattering transforms connected to the FRWT (FrScatNet). They created a computational formula for the FrScatNet using the generated data. The fundamental elements of the FrScatNet were developed, and a successful deployment strategy was also given. The FrScatNet produced excellent statistical solutions to demonstrate the FrScatNet's use in image recognition.

According to the study's authors [14], scattering connectivity is a convolutional network composed of recurrent convolutions, which were before defined as wavelets accompanied by the modulus operator. The scattering network is a few mathematical techniques for defining convolutional neural networks (CNNs). On the other hand, the original scattering network lacks a pooling function, which is a critical component of CNNs. By modeling a continuous max-pooling network and applying it to the scatnet, authors have constructed a unique network termed the scattering-maxp network. Authors [15] show that the scattering-maxp network has various advantages over the scatnet, including translation invariance, and they have done many experiments to confirm their network's computational benefit.

In this work, the publication's authors [16] compared ScatNet features to conventional features based on its ability to distinguish between low- and high-quality FPs. The authors deliberately have trained each of the three complex classifiers of machine learning like SVM, PLS-LDA, and AdaBoost, using Scattering Networks and conventional characteristics, and demonstrated that the latter leads to improved performance with each classifier. They further observed that when Scattering Networks features are applied, PLS-LDA outperforms the other classifiers, but never when conventional features are used. Based on their present findings, they want to augment Scattering Networks features with traditional and several other stated characteristics to determine whether/how much additional performance improvement may be obtained.

In the article [17], the theoretical and experimental examination of beamforming based on scattering situations is indicated that by carefully regulating the scattering characteristics of intermediary devices, the incoming signal quality and strength may be enhanced. The study presented and tested three beamforming strategies for increasing signal strength. The approach based on Taguchi proposed here produces the optimal configuration for beamforming. This [18] was compared to two other approaches in the simulation: learning automata-based techniques and brute force. Each of these strategies had poor beam creation results. On the other hand, the brute-force method results in a rather speedy convergence on the bad choice. The primary benefits of using the Taguchi-based system over the LA-based scheme are lower power consumption and a shorter learning time.

In [19], the authors said that the work shows an improved version of Mallat's Scatter Network using dual-tree wavelets and parametric log nonlinearity. Regarding classification accuracy and computational efficiency, the DTCWT Scatter Network surpasses Mallat's Scatter Network on 2 different datasets. The network has been found to outperform unsupervised learning methods, showing evidence of the benefit of DTCWT Scatter Network over supervised methods (CNNs) approaches for applications with smaller training sets.

The paper's authors [20] have developed a scattering network design for epoch determination in a speech waveform based on the Gammatone wavelet. The major characteristic of the proposed scheme is the employment of a perceptually inspired Gammatone wavelet filter bank in a single layer. The 91-channel Gammatone filter bank was built using a discrete-time approximate of the wavelet-based transform. The local maxima of the obtained filter bank channels serve as a measure of the epoch positions. The proposed approach outperforms state-of-the-art algorithms for clean and telephone quality speech regarding false alarm rate and identification accuracy.

In [21], the authors have proposed using a translation-invariant scattering network for fingerprint identification in this research. Scattering characteristics are locally invariant and hold much high-frequency information that other descriptors, such as SIFT, lose. This is one of the earliest papers to investigate the use of deep architecture for fingerprint identification. Then, to decrease the dimensionality of the features, PCA is employed. Finally, multiclass SVM is employed to do template matching. It demonstrates the utility of scattering networks in biometric recognition systems.

In [22], offered the FrScatNet, which was incorporated into this study. To improve signal classification and segmentation, it extends Scattering Network to the comparatively small scattering area and provides a signal description in the time fractional-frequency plan. An automated approach for gland segmentation in histology pictures based on the FrScatNet was also demonstrated. The graph-cut technique was used to assess the average approximate error image and locate the nodes. The proposed method was contrasted with those reported in the 2015 Challenge of MICCAI Gland Segmentation. The experimental investigation yielded consistent and comparable results. When it came to dangerous goods, the results were considerably better. They might be improved further by using cutting-edge information.

The [23] authors created a great framework for automatically classifying HEp-2 cell stained patterns in this study. The exclusionary wavelet scattering transform generates rotation-invariant characteristics of single-cell pictures. A Random Forest classifier is then used to learn and assess the traits. The wavelet scatnet with RF classifier attained an overall accuracy of 90.38 % and 90.59%, respectively, on the ICPR 2012 and ICIP 2013 datasets. In a comparison study, their suggested strategy performs better than other state-of-the-art methods. The studies' findings show that the suggested framework effectively extracts texture information, particularly on the challenging image set of cells. According to the authors, the recent trend in designing a much more robust HEp-2 cell categorization system is to incorporate a range of features with a very well data augmentation approach. In the future, such techniques will improve the proposed framework further.

The authors of [24] have looked at the DQA of FP photos in this study and proposed utilizing MCS ScatNet features to enhance conventional features. Furthermore, they demonstrated that the proposed feature set outperforms competing feature sets in differentiating between low- and high-quality FPs that don't consider polar transform and MCS. The authors used PLS-LDA as a standard classifier and demonstrated the efficacy of MCS augmented features on 2 different datasets. Even though the project is to deliver a set that outperforms the present one, there is still room for improvement. Notably, their feature extraction approach is predictable. MCS-based learnable features and domain invariant deep features have recently been demonstrated to outperform on exceptionally large databases of FPs (training images over 10000).

This paper [25] [26] suggested a novel deep learning model for integrated age estimate and gesture recognition from facial images with changes in the expression. Their method fully utilizes deep learning, with several elevated fully convolutional networks combined with an age approximation that considers the ordering of labels. The model has shown good performance using cooperative classification and multi-level regression using deep learning methodologies. The use of cross coding for regression problems, which is similar to learning various classification algorithms, is largely responsible for the enhanced performance. Their studies proved that their deep model outperformed state-of-the-art techniques on two facial expression datasets. Overall, their research highlights the possibilities of a very well deep learning model, and they trust that the outcomes will spur extra study.

This paper [27] introduces a unique MSCCN for improving the classification accuracy of complicated images. MSCCN successfully expresses multiscale directional features by integrating curvelet transformation and Scattering techniques. The curvelet-scattering module can improve the learning process of networks. The results of the experiments and evaluations demonstrate that MSCCN has a lot of beneficial qualities. On the other hand, the scattering and curvelet processes have predetermined parameters rather than being data adaptive. As a result, a dynamic and adaptable multiscale network is expected to be established in the future. Furthermore, dynamic multiscale representation has yet to be mathematically proven.

The study [28] is concerned with the difficulties in identifying finger selfies taken with a smartphone camera. A Deep Scattering Network (DSN)-based finger-selfie representation is matched using an RDF classifier. A finger-selfie categorization and improvement strategy is also offered to aid the recognition process. Each of the four groupings in the database has 19,456 images from 304 distinct categories. The experimental findings show that the proposed method performs better than the state-of-the-art algorithms for finger picture to finger selfie matching and finger selfie to live fingerprint recognition. The researchers want to investigate partial finger-selfie matching for unrestricted applications, akin to subconscious fingerprint matching.

A novel affine invariant finger photo representation based on ScatNet is suggested in [29]. Matching is carried out using an RDF classification-based methodology compared to minutiae-based and CompCode-based approaches. A finger photo segmentation and enhancement technique facilitates the matching procedure. A publicly accessible IIITD smartphone database is being built to examine and address various difficulties. The database is made freely available for research purposes and consists of three sets with 128 classes and over 5100 photos. The results reveal a significant performance gain over the existing methods in several experimental conditions.

This research [30][31] offers a palmprint identification technique based on a deep convolutional scattering network. These scattering characteristics are spatially invariant and convey much high-frequency information that other descriptors, such as SIFT, ignore. The features are then subjected to PCA to reduce their dimensionality. Finally, multiclass SVM and closest neighbor classifier are utilized to conduct recognition. Authors believe that the scattering network's multilayer representation has high discriminating power, which can help with biometric recognition tasks. In the future, the authors will research the scattering network's application to various biometrics and medical picture categorization challenges.

This research [32] proposes a deep convolutional neural network for face attractiveness prediction termed region-aware scattering convolution networks (RegionScatNet). The RegionScatNet is combined with the scattering transform and region-aware face attribute decomposition. It generates a face representation that is discriminative for attractiveness assessment and invariant to transformational variations such as rotation. The most current SCUT-FBP benchmark database, particularly built for face beauty analysis, yielded cutting-edge facial beauty evaluation findings.

This paper [33] offers a competitive method for large-scale visual recognition based on scattering networks, specifically for ILSVRC2012. The authors have got state-of-the-art results compared to unsupervised modeling on CIFAR-10 or limited data environments on STL-10 and

CIFAR-10. On ILSVRC2012, they have built a supervised Shared Local Encoder that allowed scattering networks to outperform previous local encoding approaches. This network of only three learned layers allows for investigation of the operation done. Their findings also show that pre-defined characteristics are still important since they may shed light on deep learning approaches and make them more interpretable. Combined with proper learning methods, they may provide additional theoretical guarantees required to construct stronger deep models and stable representations.

In this article [34], the authors discuss self-supervised training for limited dataset problems. They especially investigate if the sophisticated encoding system used for contrastive training can be partially substituted by a simpler custom network while retaining geometrical invariance. They demonstrated that the proposed approach, based on a mathematically invariant Scattering Network and with less learnable parameters, can achieve state-of-the-art performance on the CIFAR100-20 and STL-10 datasets. They demonstrate that including pretext task normalization based on augmentation transform estimation improves the performance of the planned ScatSimCLR systems and SimCLR with ResNet.

As part of [35] an SCHMT framework, a hidden Markov tree model and a scattering transform have been provided. The scattering transformation, combined with the data's invariants, turns the information into actual details with higher dimensions but reduced volume. Then, a generative adversarial model — a Markov tree — was used to fit a predictive model to the dataset representation's distribution. As a result, the proposed model employs how the scattering transformation introduces mappings into the representation and how Markov models preserve parameter dependencies. Trials have found that the simulated distribution can perform good classification tasks even with small training volumes.

The proposed challenge provides a one-of-a-kind opportunity to evaluate algorithms related to bioacoustic categorization [36]. This task involves all the significant difficulties and issues that arise when working with a sizable dataset in the real world. Although the scattering transform provides a brilliant new data format, the authors have found that it is not yet apparent how to extract information from its coefficients. They certainly did not test every feasible method, but the concept of thresholding/combination in the 2 dimensions appears to be intriguing. The ability to record a feature using a time-invariant method in terms of its location in a signal is [37] significant. However, the authors must maintain a "time" element to study frequency evolution within the features.

The authors of [38] describe a skeleton graph scatter network (SGSN), which uses graph scatter to retrieve motion information from graph available spectrum to predict 3D human movement using skeletons. At the heart of their SGSN is the adaptive graph scattering block (AGSB), which includes I graph scattering degradation, which breaks down information into different graph available spectrum and upgrades the learnable features, and ii) graph frequency band interest, which incorporates those characteristics using learnable attention weights. Extended experiments indicate that their SGSN outperforms others on the 3DPW, CMU Mocap, and Human3.6M datasets for long-term and short-term movement prediction.

In [39], the authors stated a 2-stage method for identifying anomalies in ECG wearable technology. By merging the spectral residual and wavelet scattering networks, an SR-ScatNet-based Fast Fourier Transform (FFT) approach requires much less processing. It can be executed on a smart cloth ECG sensor platform. Because the training model is trained on-the-fly at the device containing the data, such on-device applications can ease network constraints, reduce power consumption, boost privacy, and enhance data protection. The detection may be done on the device directly by removing the requirement to transfer all personal ECG information to the server.

[40] this paper introduces a novel unconstrained ear identification approach that uses Scattering Network to collect translation and deformation consistent local ear biometric information. The recovered features are resilient; they outperform other state-of-the-art local features extracted in R1 and EER. EER has a larger value than most other approaches, although R1 has a value that is comparable to other local features. The scale factor is crucial for ear recognition, and super-resolution and image refocusing technologies can be incorporated in the image preparation phases. The Scattering network can be concatenated with CNN networks to improve recognition accuracy.

In [41], the authors develop a new network model termed weight changeable scattered convolution systems based on wavelet scattered convolution systems. Two electromagnetic signal categorization tests are designed to evaluate the network's performance, and actual ACARS and ADS-B signals transmitted and over the air are used in the research. The results of the experiments indicate that a significant approach based on wavelet scattered convolution systems and SVM classifier is especially sensitive to feature dimension selection and Classifier kernel function. The choice of digital filter type affects classifier performance as well. These difficulties must be extensively investigated in real-world applications, and parameter selection is tough.

3. Dataset Used

The dataset used to train the models mentioned consists of 200 dry and oily skin images, each making up 400 images. The images were specially and carefully obtained by consulting local dermatologists where the patient's privacy was maintained, and none of the personal details were revealed about the patient. These images were manually verified and segregated as per the two conditions by the dermatologists before using them as input for their machine learning algorithms.

4. Proposed Methodology

In the research method, they have used Scattering Networks over Residual Networks (ResNets), which are an advancement over the traditional Convolutional Neural Networks (CNN) for image classification. Image pre-processing is an antecedent step to model building and evaluation. It helps in data cleaning and removes any inconsistencies that could adversely affect the model's performance.

4.1. Image Pre-Processing

The image resolution in image recognition tasks plays an important role because the right image quality can boost the classification results. Before training the model, images are pre-processed [35]. The major goal is to remove extraneous information from the images, which increases the detectability of important and pertinent details in the image, and greatly simplifies the data to improve the model's feature extraction capability and recognition efficiency. First, in image pre-processing, they have applied median filtering followed by histogram Equalization. Then various image augmentation [36] methods are applied, and the final output is fed to the training model.

4.1.1. Cleaning of Data

The purpose of cleaning the data is primarily to ensure the standardization of data characteristics in infected skin images. The most often utilized data cleaning strategy in dermatological image identification is to eliminate noise to lessen the effect of minute particles in the frame and shadow on the classification. The type of skin, the location, the equipment, and the lighting conditions all have an impact on the image's quality. A poor-quality image can reduce the accuracy leading to poor diagnosis. Denoising algorithms often include transforming domain filtering, partial differential equation, and spatial domain filtering. Many previous works done by many researchers have carried out the removal of hair from facial images, if there are any [37]. Filters were utilized to lessen the impact of noise by various other researchers [38]. Various image augmentation methods like zooming, rotating, shearing, flipping, etc., are employed to increment the number of data samples.

4.1.2. Median Filtering

Noise is a random variation in image density, brightness, or texture that adds inconsistency to images. Median filtering is a widely used image processing technique to remove noise from the image. It helps in preserving the image edges while removing random noises. In this method, a window of a fixed size slides over the image, and the pixel values are changed to the median of all pixels in the particular window arranged in ascending order. Like a low pass filter, the median processing smoothens the image yet preserves the boundary information. Different median filtering techniques remove noisy pixels in an image [39]. Like Adaptive Median Filtering technique which uses impulse detection technique [40]. In this process, impulse noise is removed, and the remaining pixels are smoothened. To reduce corruption of nearby pixels, thinning and thickening of object boundaries are also performed. Other median filtering techniques include Tristate-Median Filter (TMF), Progressive Switching Median Filter (PSMF), Two-Stage Iterative Median Filter, and Two-Phase Median Filter. They have incorporated the use of Simple median filtering on their dataset. It can be seen in Fig. 1 given below.

Fig. 1 Median filtering done on a training image. (Left is the original image. Right is filtered image)

4.1.3. Histogram Equalization

A histogram gives a graphical representation of the distribution of the image intensity. It shows the number of pixels in an image for each intensity value. The X-axis represents the intensity value from black on the left to white on the right. The number of pixels at each intensity is shown on the Y-axis. Histogram equalization is a technique used in image processing to boost visual contrast. It accomplishes this by extending an image's intensity value range. This approach aids in enhancing the contrast of low-contrast sections of an image. Each channel can be isolated, and equalization performed independently for a three-channel picture (such as RGB).

The emergence of pixels with a particular greyscale value is expressed as the probability of a total number of pixels [41]. The corresponding cumulative probability function is calculated using these probabilities. Now a transformation is applied that corresponds to each value in the original image. The pixels of the grayscale image is mapped to a scale from 0 to 1. These mapped pixels transform to retrieve the original domain of the image. For simplicity of the coding, they have used various functionality provided in the NumPy module of python. Histogram equalization can be seen in Fig. 2 below.

Fig. 2 Histogram equalization performed on a training image. (Left one is the original image. Right is the histogram image)

Table 1. Pseudo Code for Image Pre-processing

Input: RGB image of any dimension (m,n,3) such that m,n >=128
Output: Grayscale image with dimension 128 x 128 x 1
Begin:
Image = Histogram_Equalization(Image) Image = Median_filter(Image) Image = convert_to_graysacale(Image) Resized_Image = Image_resize(Image ,(128,128,1)) Return Resized_Image
End

Table 2. Pseudo Code for Image Augmentation

Input: Preprocessed Image
Output: Augmented Image
Begin:
Initialize Rotation Range(R), Initialize Mean(M), Standard deviation(SD)
Set Horizontal Flip to True Set Vertical Flip to True Image = Normalize(Image,M,SD) Augmented_Image = ImageDataGenerator(R) Return Augmented_Image
End

4.2. Prediction Model (Scattering Networks + ResNet)

They present the scattering networks, which use the idea of scattering transformations and their usage as a generalized input for supervised learning problems. A scattering network enabled deep neural network is a CNN in which the filters are specified as wavelets. This network's structure has a solid mathematical foundation, so it is easily comprehended and requires few variables. Geometric changes also don’t affect it much. In general, the variables of this representation do not need many adjustments. Hence the outcome has an acceptable and general representation. The usage of ResNet constructed on top of the scattering network is proposed. It can be seen in Fig. 3 That the proposed model is given below.

4.2.1. Wavelet Scattering Networks

A deep network usually involves Filtering the input data, applying nonlinearities, and pooling or averaging the output over a set of epochs. It can be seen below in Fig. 4. This drawback requires large data. Secondly, Tuning the hyper-parameter is a cumbersome task. The filters in fully trained networks are analogous to wavelets; these are trained in traditional CNNs or RESNETs but are fixed in wavelet scattering networks. In the first layer, also called layer 0 Input signal (in their case, input image) is averaged using a wavelet filter (from a particular wave family) which acts as a low pass filter. In subsequent layers, a wavelet transform with the preceding outputs is performed. Finally, the modulus operator is applied, which adds nonlinearity, after which it is again filtered with a wavelet filter (a low pass filter). The same process is repeated, and they get coefficients for different layers. Wavelet Scattering Network [25] resembles the deep convolutional neural network in its multilayer structure (CNN). Every layer performs both linear and non-linear computations.

To summarize, the convolution operation of the already defined complex filter with the input image is performed first on each layer. Then the modulus operation, which introduces nonlinearity, is applied to the result. Finally, a lowpass filter is used to calculate the local average.

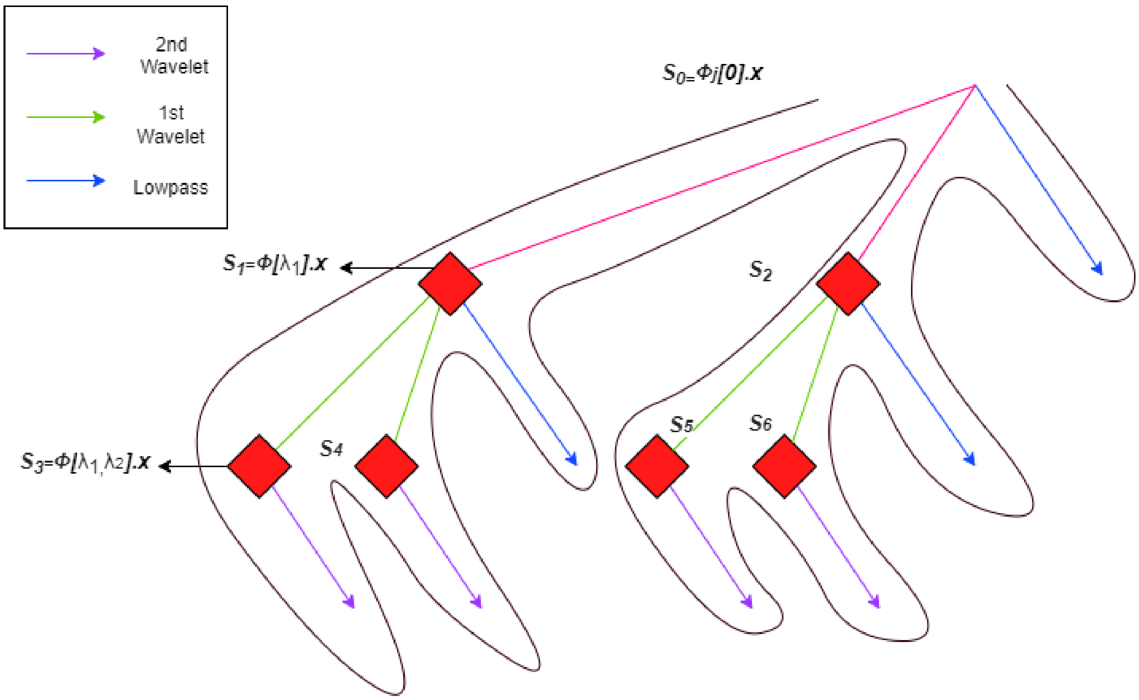

Fig. 4 Wavelet Scattering Network

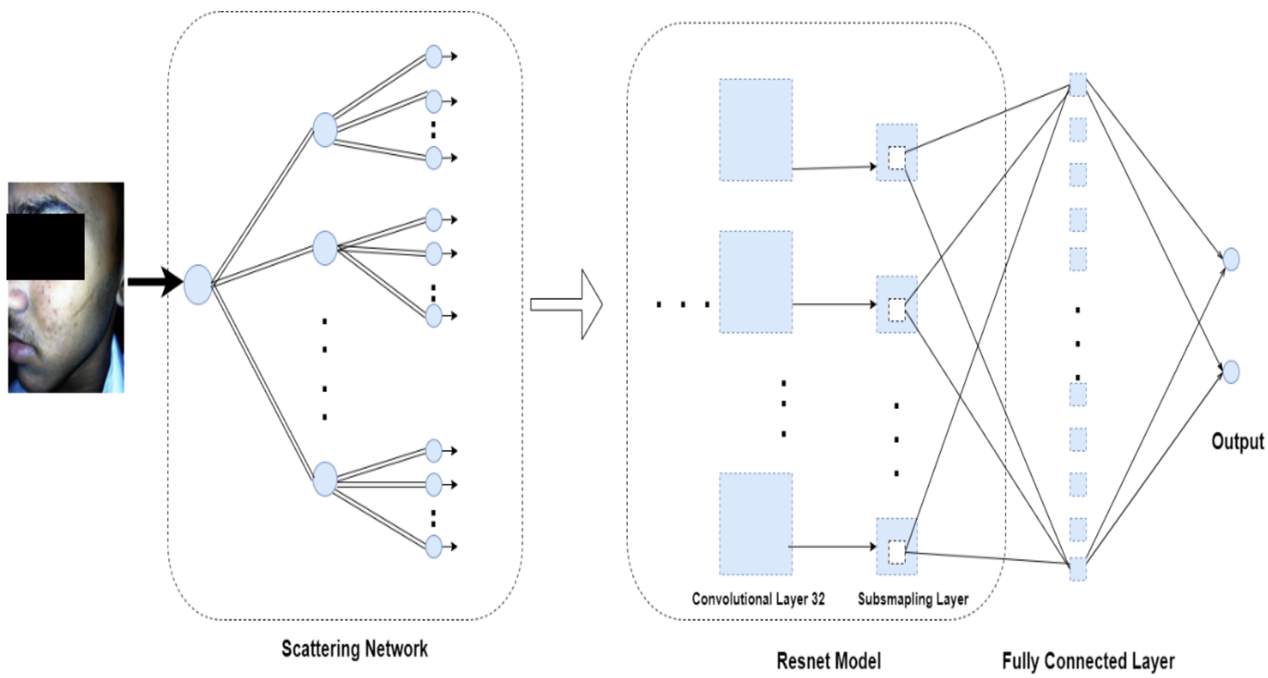

Fig. 3 Overview of the proposed model.

4.2.2. ResNet and its Architecture

Residual Network, also known as (ResNet) is the most well deep neural learning model introduced by authors in their paper [42]. The ResNet model is among the most widely used and effective deep learning models. Residual Blocks: This model's basic building blocks, which solve the problem of training deep networks, can be seen in Fig. 5 below.

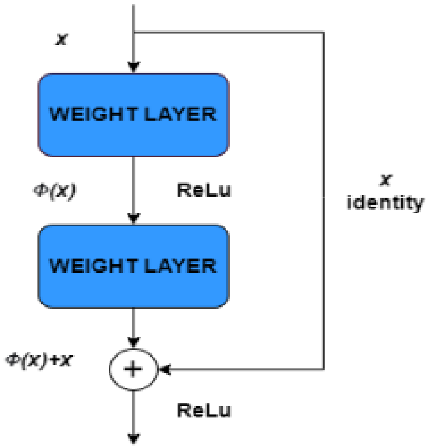

Fig. 5 Residual Blocks

The residual block consists of what is called a "Skip Connection." It indicates that there is indeed a direct connection that bypasses the model's layers, a characteristic non-existent in traditional convolution neural networks. Due to these connections, the output is different. Earlier the input 'X' would be multiplied by the layer weights, and then the bias term would have been added. To which they would have applied the activation function. It would look like this to mathematically get the output H(x).

$$(x) = f(wx + b) \text{ or } H(x) = f(x)$$

Now, due to the new skip connections, the output is H(x) is changed to

$$H(x) = f(x) + x$$

However, it may happen so that the dimension of the input and output may vary due to the convolution and pooling operation. To solve this, they can either zero pad the skip connections to increase the dimension, which can compensate for the reduction in a dimension that occurs during the operations of a convolution layer of size 1x1 is added to match the input size. In such a case, the output will become

$$H(x) = f(x) + w1.x$$

W1 is a distinct parameter that was created to accommodate the 1x1 convolution.

Introducing these skip connections solves the issue of vanishing gradients in deep Convolution Neural Networks (CNN). In the Vanishing gradient, due to a large number of layers in CNN, the gradient, which is calculated as a product of many gradients (of different layers), tends to decrease with each multiplication to the point where it becomes so small that it vanishes due to which they are unable to train. The skip connection provides an alternate or short path for the gradient to flow due to a smaller number of computations and fewer chances of vanishing. If any layer isn't favourable for the model's performance, it is skipped through regularization. In the design, VGG-19 influences a 34-layer plain network with a shortcut or skip connection. These residual blocks or skip connections are then employed to transform the concept further into a residual network.

Table 3. Pseudo Code for Model Training

Input: Input Images with labels INPUT_DATA
Begin:
Initialize Model ScatteringResnet with random weights
Train_images, Train_Labels, Test_images, Test_Labels = split(INPUT_DATA, train_split = 66%) For epoch = 1 to epoch = 60: Train_output = ScatteringResnet.predict(Train_images) Training_Loss = Calculate_Loss(Train_output,Train_Labels) Training_Accuracy = Calculate_Accuracy(Train_output,Train_Labels) Compute Gradients and Update Model Parameters Test_output = ScatteringResnet.predict(Test_images) Validation_Loss = Calculate_loss(Test_output,Test_Labels) Validation_Accuracy = Calculate_Accuracy(Test_output,Test_Labels) Display the calculated loss and accuracy for the current epoch
End

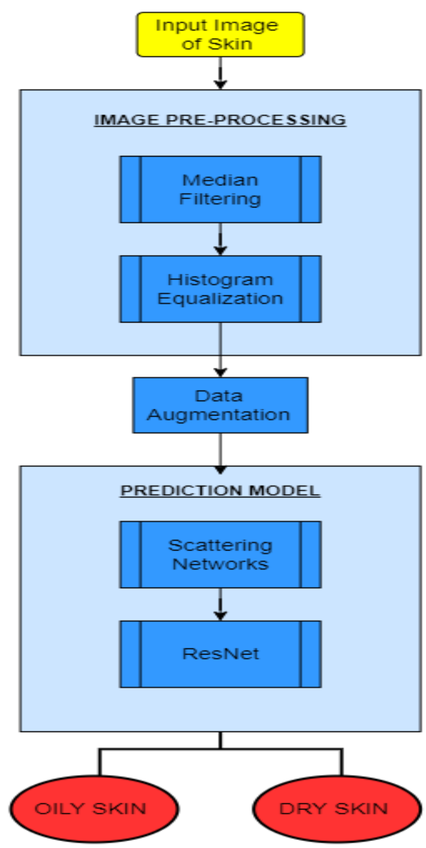

Fig. 6 Architecture of Proposed ScatNet Model

5. Experimental Results

To illustrate the effectiveness of the suggested algorithm, they used an image dataset collected from dermatologists and divided the data in 66-33 ways, wherein they used 66% of the images for training the model, and the remaining 33% were used to test the model. When the same data were compared with other models, it was observed that Gabor Features with Support Vector Machine (SVM) was classified poorly, giving an accuracy of 48.8%. VGG-16 has given the accuracy of 95.9%, followed by Scattering along with Support Vector Machine (SVM) & Discrete Wavelet Transform combined with Artificial Neural Network (ANN) have given results less than 65%. Scattering ResNet has outperformed the existing models and given an accuracy of 98%. Below is the Accuracy metrics graph, which gives the details in Fig. 7.

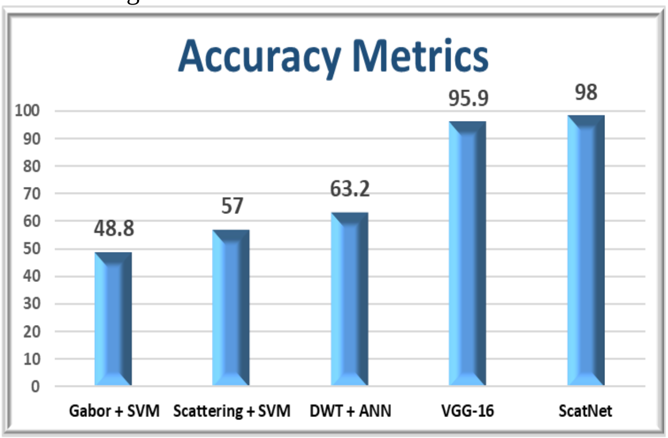

Fig. 7 Accuracy Metrics for the different models trained

From the table above, they can conclude that scattering networks performed better than other tested models. It achieved 98% accuracy on the validation set. Support vector machines do not perform well on image classification problems as they have only one layer called a kernel. It becomes a tricky task to design a kernel in SVM so that it can linearly separate the images. Hence the best approach for image classification is using a deep neural network. It is because, in a deep neural network (consisting of a large number of neurons), only a few neurons (generally the last layer) will do the task of classification that the SVMs do, the rest of them act as kernel and work towards feature extraction, and the main advantage is that these neurons learn the weights themselves with every passing epoch. Hence neural network-based architecture performs better. The high performance of Scattering ResNets as compared to other models can be attributed to the fact that these networks use filters that have already trained weights as compared to traditional RESNETs or VGG-16 [43]. The Scattering Network has a scattering layer before the RESNET. In supervised learning, an early layer of CNNs is not always required to be trained and can be substituted with scattering networks. As a result, when they use a hybrid structure like this, they tend to achieve the best outcomes.

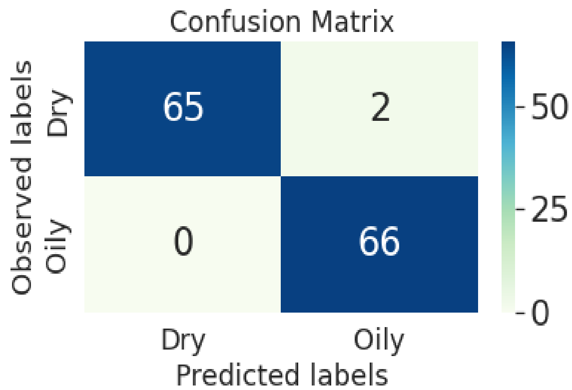

Fig. 8 Confusion Matrix for the trained model

The confusion matrix in Fig. 8 above for the said dataset shows that their model has achieved an accuracy of 98%.

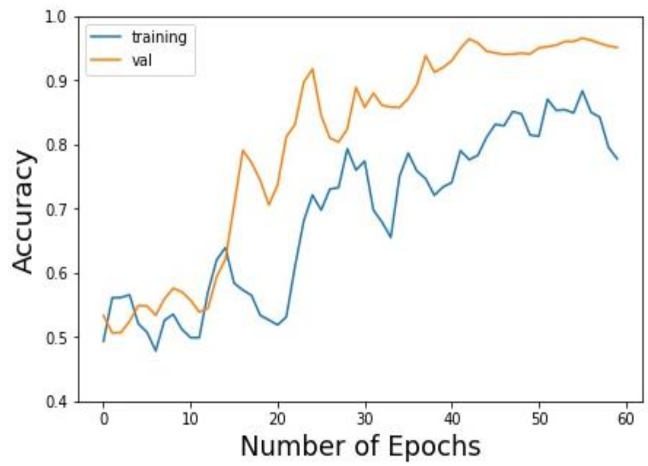

Fig. 9 Accuracy of ScatNet model for training & validation for each epoch

One parameter for assessing classification models is accuracy. Unofficially, accuracy is the percentage of correct predictions made by their model.

$$Accuracy = \frac{Total\ number\ of\ correct\ estimates}{Total\ number\ of\ estimates}$$

Accuracy in binary classification may also be measured, as shown below:

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP}$$

Where,

FP = False Positive: A test result that incorrectly implies the presence of a specific condition or feature,

FN = False Negative: A test result that incorrectly suggests the lack of a specific condition or feature,

TN = True Negative: A test result that accurately identifies the lack of a condition or feature,

TP = True Positive: A test result that accurately identifies the presence of a condition or feature.

The above graph shows the training system's accuracy over 60 epochs, where you can see around the 40th epoch that accuracy was achieved over 80%.

Loss of the trained system over epochs is shown in the graph below.

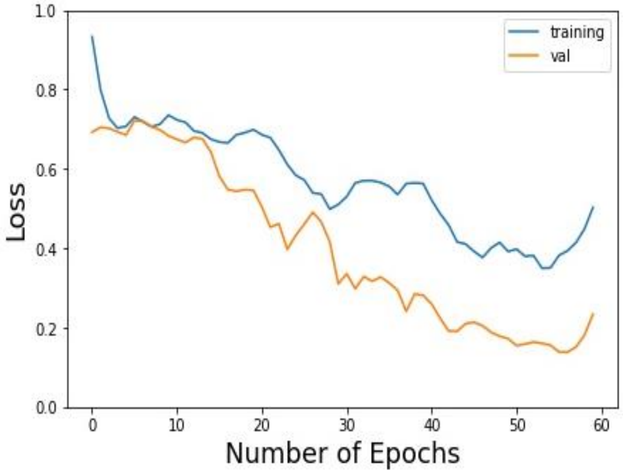

Fig. 10 Loss for ScatNet model for training & validation for each epoch

Similarly, Precision and recall are two numbers used to assess the effectiveness of categorization or information retrieval systems. Precision is used to gauge the number of correct classifications out of all the positive predictions. Recall, also known as sensitivity, is the proportion of recovered occurrences among all relevant instances. Precision and recall are both equal to one for an ideal classifier.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

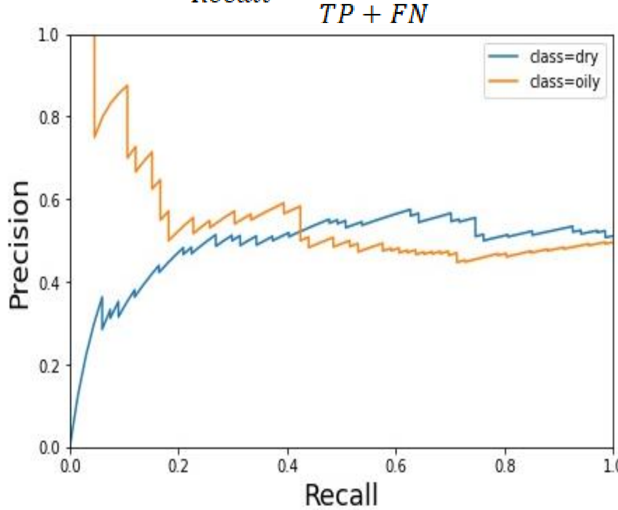

Fig. 11 Loss for ScatNet model for training & validation for each epoch

It is frequently feasible to calibrate a model's number of outcomes and enhance Precision at the recall price, or vice versa. Precision and recall should be stated jointly at all times. If a single numerical measurement of a system's performance is necessary, Precision and recall are frequently combined into the F-score. The graph depicts the systems, Precision, and Recall for both the oily and dry classes.

The following graph shows a comparison of the different systems that show accuracy. It can be seen clearly that Scattering ResNet has outperformed VGG & DWT + ANN.

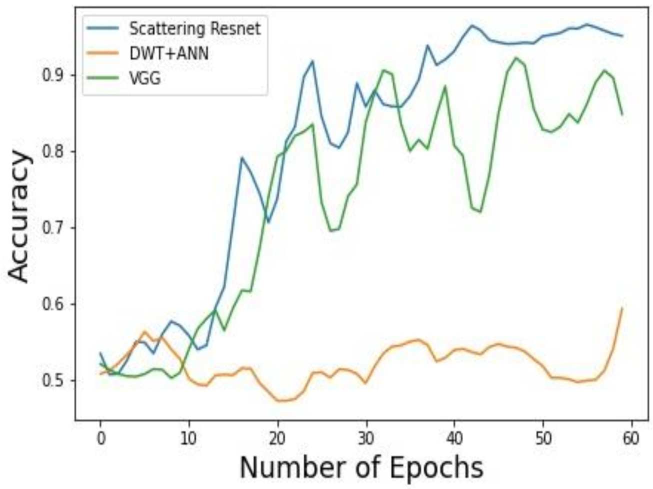

Fig. 12 Comparative Accuracy of the models trained

Similarly, the comparative loss of all the systems can be seen in the graph below.

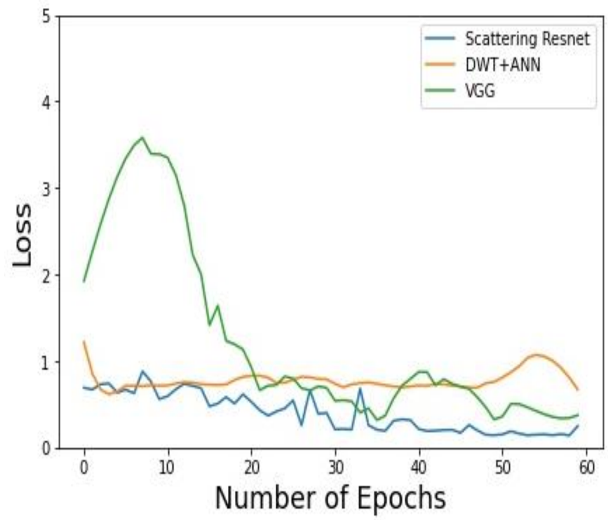

Fig. 13 Comparative Loss of the models trained

It can be seen from the output given by the trained system that on trained epoch 48, test accuracy achieved is 98.50% & loss is 0.076.

```
Train Epoch: 48 [0/268 (0%)]    Loss: 0.427322
Train Epoch: 48 [200/268 (75%)] Loss: 0.387069

Test set: Average loss: 0.0762, Accuracy: 131/133 (98.50%)

Train Epoch: 49 [0/268 (0%)]    Loss: 0.279051
Train Epoch: 49 [200/268 (75%)] Loss: 0.390186

Test set: Average loss: 0.0965, Accuracy: 130/133 (97.74%)

Train Epoch: 50 [0/268 (0%)]    Loss: 0.020944
Train Epoch: 50 [200/268 (75%)] Loss: 0.136528

Test set: Average loss: 0.1014, Accuracy: 127/133 (95.49%)
```

Fig. 14 Output of the system for ScatNet model trained

6. Conclusion

Skin is the largest organ of one's body, so the chances of ailments and other problems affecting it becomes high. There is a plethora of treatments, both palliative and curative. Hence it becomes very important to understand an individual's skin type before one can proceed with the treatment. The study aims to solve this challenge by which they intend to classifying a person's skin type as either oily or dry, which are the two most distinctive features of human skin. After research on the existing methodologies and recent development in bio-medicinal deep learning, they came up with a deep learning architecture that performs well and addresses this classification problem. As stated earlier, the model was trained and tested on medical images from dermatologists. Their current model using scattering network & ResNet works well compared to the other existing deep learning architecture, which was also tested to solve the problem. The accuracy it achieved is 98% which is more than other models tested. The major boost in the performance can be attributed to the series of image processing techniques they employed which are centric on skin and identification of its texture. Once the textures are detected, working on the classification becomes convenient and can be achieved by a deep network. They intend to enhance the performance of this model in the future by designing a skin detection system that will first segment the human skin from an image and then go for the classification of skin. This way, they aspire to reduce any error generated due to bad camera quality or environment.

Acknowledgments

The author would like to acknowledge Skinzy Software Solutions Pvt. Ltd, Mumbai, India, for helping in data processing, cleaning, and validating results in consultation with Dermatologists. The authors wish to thank Dr. Uday Khopkar, Dermatologist, for his valuable guidance and support for the study. This image data set are proprietary and copyrighted by Skinzy Software Solutions Pvt. Ltd, Mumbai, India.

References

[1] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, and J. Jatakia, “Human Skin Detection Using RGB, HSV and YCbCr Color Models,” vol. 137, pp. 324–332, 2017, doi: 10.2991/iccasp-16.2017.51.

[2] D. Kalbande, Nikhil Barhate, Parth R. Dedhia, Arnab Ghorai, Darshan Savalia, Dr.Uday Khopkar, “AI Enabled Method for predicting Acne-Non Acne.” National Conference on Artificial Intelligence in Health Informatics & Virtual Reality, 2019.

[3] P. M. Arabi, G. Joshi, R. N. Reddy, A. S. R, and A. P. S, “Categorising Normal Skin, Oily Skin and Dry Skin using 4-Connectivity and 8-Connectivity Region Properties,” Int. J. Adv. Netw. Appl., vol. 4, no. 4, pp. 2016–2018, 2017, [Online]. Available: http://www.ijana.in/Special Issue/A4.pdf.

[4] L. Baumann, “Understanding and Treating Various Skin Types: The Baumann Skin Type Indicator,” Dermatol. Clin., vol. 26, no. 3, pp. 359–373, 2008, doi: 10.1016/j.det.2008.03.007.

[5] M. I. Razzak, S. Naz, and A. Zaib, “Deep learning for medical image processing: Overview, challenges and the future,” Lect. Notes Comput. Vis. Biomech., vol. 26, pp. 323–350, 2018, doi: 10.1007/978-3-319-65981-7_12.

[6] A. Anton, N. F. Nissa, A. Janiati, N. Cahya, and P. Astuti, “Application of Deep Learning Using Convolutional Neural Network (CNN) Method For Women’s Skin Classification,” Sci. J. Informatics, vol. 8, no. 1, 2021, doi: 10.15294/sji.v8i1.26888.

[7] Indriyani and I. Made Sudarma, “Classification of facial skin type using discrete wavelet transform, contrast, local binary pattern and support vector machine,” J. Theor. Appl. Inf. Technol., vol. 98, no. 5, 2020.

[8] T. J. Jebaseeli, D. J. David, and R. Venkatesan, “Prediction of COVID’19 through multiple organ analysis using iot devices and machine learning techniques,” Int. J. Eng. Trends Technol., vol. 69, no. 8, 2021, doi: 10.14445/22315381/IJETT-V69I8P213.

[9] R. D. Amelia, I. I. Tritoasmoro, and N. Ibrahim, “Klasifikasi Jenis Kulit Wajah Menggunakan Metode Discrete Wavelet Transform dan Backpropagation,” e-Proceeding Eng., vol. 6, no. 2, 2019.

[10] M. Thanjaivadivel and R. Suguna, “Leaf disease prediction using fast enhanced learning method,” Int. J. Eng. Trends Technol., vol. 69, no. 9, 2021, doi: 10.14445/22315381/IJETT-V69I9P205.

[11] E. Oyallon et al., “Scattering Networks for Hybrid Representation Learning,” IEEE Trans. Pattern Anal. Mach. Intell., 2018, doi: 10.1109/TPAMI.2018.2855738.

[12] P. Singh, G. Saha, and M. Sahidullah, “Deep scattering network for speech emotion recognition,” Eur. Signal Process. Conf., vol. 2021-August, pp. 131–135, 2021, doi: 10.23919/EUSIPCO54536.2021.9615958.

[13] J. Shi, Y. Zhao, W. Xiang, V. Monga, X. Liu, and R. Tao, "Deep Scattering Network with Fractional Wavelet Transform," IEEE Trans. Signal Process., 2021, doi: 10.1109/TSP.2021.3098936.

[14] T. Ki and Y. Hur, "Deep Scattering Network with Max-Pooling," in Data Compression Conference Proceedings, 2021, vol. 2021-March, doi: 10.1109/DCC50243.2021.00052.

[15] T. Karim et al., "Cnn applied in public transport for the protection against the covid-19 spread," Int. J. Eng. Trends Technol., vol. 69, no. 10, 2021, doi: 10.14445/22315381/IJETT-V69I10P205.

[16] C. Dev et al., "Diagnostic Quality Assessment of Ocular Fundus Photographs: Efficacy of Structure-Preserving ScatNet Features," in Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2019, doi: 10.1109/EMBC.2019.8857046.

[17] V. R. Surendra and M. Zawodniok, "Distributed beamforming using a scattering network," in CNSR 2010 - Proceedings of the 8th Annual Conference on Communication Networks and Services Research, 2010, doi: 10.1109/CNSR.2010.61.

[18] S. Srivastava, A. Nigam, and M. Muthaiyan, "Low-cost Driver Assistance System for drivers suffering from Dyslexia or Color-blindness using Machine Learning," Int. J. Eng. Trends Technol., vol. 69, no. 12, 2021, doi: 10.14445/22315381/IJETT-V69I12P207.

[19] A. Singh and N. Kingsbury, "Dual-Tree wavelet scattering network with parametric log transformation for object classification," in ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2017, doi: 10.1109/ICASSP.2017.7952631.

[20] P. Kulkarni, J. Sadasivan, A. Adiga, and C. S. Seelamantula, "Epoch Estimation from a Speech Signal Using Gammatone Wavelets in a Scattering Network," in ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2020, vol. 2020-May, doi: 10.1109/ICASSP40776.2020.9052986.

[21] S. Minaee and Y. Wang, "Fingerprint recognition using translation invariant scattering network," in 2015 IEEE Signal Processing in Medicine and Biology Symposium - Proceedings, 2016, doi: 10.1109/SPMB.2015.7405471.

[22] L. Liu, J. Wu, D. Li, L. Senhadji, and H. Shu, "Fractional wavelet scattering network and applications," IEEE Trans. Biomed. Eng., vol. 66, no. 2, 2019, doi: 10.1109/TBME.2018.2850356.

[23] B. H. Li, J. Zhang, and W. S. Zheng, "HEp-2 cells staining patterns classification via wavelet scattering network and random forest," in Proceedings - 3rd IAPR Asian Conference on Pattern Recognition, ACPR 2015, 2016, doi: 10.1109/ACPR.2015.7486535.

[24] S. R. Manne, S. Bin Bashar, K. K. Vupparaboina, J. Chhablani, and S. Jana, "Improved Fundus Image Quality Assessment: Augmenting Traditional Features with Structure Preserving ScatNet Features in Multicolor Space," in Proceedings - 2020 IEEE EMBS Conference on Biomedical Engineering and Sciences, IECBES 2020, 2021, doi: 10.1109/IECBES48179.2021.9398757.

[25] H. F. Yang, B. Y. Lin, K. Y. Chang, and C. S. Chen, "Joint estimation of age and expression by combining scattering and convolutional networks," ACM Trans. Multimed. Comput. Commun. Appl., vol. 14, no. 1, 2018, doi: 10.1145/3152118.

[26] E. Ortega-Espinoza, M. Flores-Cruz, D. C. Loayza, J. Velarde-Cuadros, A. Delgado, and E. L. Huamaní, "Use of Machine Learning to Predict the Occurrence of Deaths in the Departments Most Affected by Covid-19 in Peru," Int. J. Eng. Trends Technol., vol. 70, no. 3, 2022, doi: 10.14445/22315381/IJETT-V70I3P206.

[27] J. Gao, L. Jiao, F. Liu, S. Yang, B. Hou, and X. Liu, "Multiscale Curvelet Scattering Network," IEEE Trans. Neural Networks Learn. Syst., 2021, doi: 10.1109/TNNLS.2021.3118221.

[28] A. Malhotra, A. Sankaran, M. Vatsa, and R. Singh, "On Matching Finger-Selfies Using Deep Scattering Networks," IEEE Trans. Biometrics, Behav. Identity Sci., vol. 2, no. 4, 2020, doi: 10.1109/TBIOM.2020.2999850.

[29] A. Sankaran, A. Malhotra, A. Mittal, M. Vatsa, and R. Singh, "On smartphone camera-based finger photo authentication," in 2015 IEEE 7th International Conference on Biometrics Theory, Applications, and Systems, BTAS 2015, 2015, doi: 10.1109/BTAS.2015.7358782.

[30] S. Minaee and Y. Wang, "Palmprint recognition using deep scattering network," in Proceedings - IEEE International Symposium on Circuits and Systems, 2017, doi: 10.1109/ISCAS.2017.8050421.

[31] S. K. Bandyopadhyay, P. Bose, A. Bhaumik, and S. Poddar, "Machine Learning and Deep Learning Integration for Skin Diseases Prediction," Int. J. Eng. Trends Technol., vol. 70, no. 2, 2022, doi: 10.14445/22315381/IJETT-V70I2P202.

[32] L. Liang, D. Xie, J. Xu, M. Li, L. Lin, and L. Jin, "Region-aware scattering convolution networks for facial beauty prediction," in Proceedings - International Conference on Image Processing, ICIP, 2018, vol. 2017-September, doi: 10.1109/ICIP.2017.8296805.

[33] E. Oyallon, E. Belilovsky, and S. Zagoruyko, "Scaling the Scattering Transform: Deep Hybrid Networks," in Proceedings of the IEEE International Conference on Computer Vision, 2017, vol. 2017-October, doi: 10.1109/ICCV.2017.599.

[34] V. Kinakh, O. Taran, and S. Voloshynovskiy, "ScatSimCLR: Self-supervised contrastive learning with pretext task regularization for small-scale datasets," in Proceedings of the IEEE International Conference on Computer Vision, 2021, vol. 2021-October, doi: 10.1109/ICCVW54120.2021.00129.

[35] J. B. Regli and J. D. B. Nelson, "Scattering convolutional hidden Markov trees," in Proceedings - International Conference on Image Processing, ICIP, 2016, vol. 2016-August, doi: 10.1109/ICIP.2016.7532685.

[36] R. Balestriero and H. Glotin, “Scattering Decomposition for Massive Signal Classification: From Theory to Fast Algorithm and Implementation with Validation on International Bioacoustic Benchmark,” in Proceedings - 15th IEEE International Conference on Data Mining Workshop, ICDMW 2015, 2016, doi: 10.1109/ICDMW.2015.127.

[37] Yulianto, Nurhasanah, R. Yulistiani, and G. P. Kusuma, “Face Image Super-Resolution Using Combination of Max-Feature-Map and CMU-Net to Enhance Low-Resolution Face Recognition,” Int. J. Eng. Trends Technol., vol. 70, no. 3, 2022, doi: 10.14445/22315381/IJETT-V70I3P201.

[38] M. Li et al., “Skeleton Graph Scattering Networks for 3D Skeleton-based Human Motion Prediction,” in Proceedings of the IEEE International Conference on Computer Vision, 2021, vol. 2021-October, doi: 10.1109/ICCVW54120.2021.00101.

[39] H. Y. Feng, P. Y. Chen, and J. Hou, “SR-ScatNet Algorithm for On-device ECG Time Series Anomaly Detection,” in Conference Proceedings - IEEE SOUTHEASTCON, 2021, vol. 2021-March, doi: 10.1109/SoutheastCon45413.2021.9401872.

[40] P. Birajadar, M. Haria, S. G. Sangodkar, and V. Gadre, “Unconstrained Ear Recognition Using Deep Scattering Wavelet Network,” in 2019 IEEE Bombay Section Signature Conference, IBSSC 2019, 2019, vol. 2019January, doi: 10.1109/IBSSC47189.2019.8973055.

[41] H. Zhou et al., “Weight-Variable Scattering Convolution Networks and Its Application in Electromagnetic Signal Classification,” IEEE Access, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2957519.

[42] M. Vashisht and B. Kumar, “A Survey Paper on Object Detection Methods in Image Processing,” 2020 Int. Conf. Comput. Sci. Eng. Appl. ICCSEA 2020, pp. 5–8, 2020, doi: 10.1109/ICCSEA49143.2020.9132871.

[43] X. Shen, J. Zhang, C. Yan, and H. Zhou, “An Automatic Diagnosis Method of Facial Acne Vulgaris Based on Convolutional Neural Network,” Sci. Rep., vol. 8, no. 1, 2018, doi: 10.1038/s41598-018-24204-6.

[44] N. Hameed, A. M. Shabut, M. K. Ghosh, and M. A. Hossain, “Multiclass multi-level classification algorithm for skin lesions classification using machine learning techniques,” Expert Syst. Appl., vol. 141, 2020, doi: 10.1016/j.eswa.2019.112961.

[45] A. Singhal, R. Shukla, P. K. Kankar, S. Dubey, S. Singh, and R. B. Pachori, “Comparing the capabilities of transfer learning models to detect skin lesion in humans,” Proc. Inst. Mech. Eng. Part H J. Eng. Med., vol. 234, no. 10, 2020, doi: 10.1177/0954411920939829.

[46] G. George, R. M. Oommen, S. Shelly, S. S. Philipose, and A. M. Varghese, “A Survey on Various Median Filtering Techniques For Removal of Impulse Noise From Digital Image,” in Proc. IEEE Conference on Emerging Devices and Smart Systems, ICEDSS 2018, 2018, doi: 10.1109/ICEDSS.2018.8544273.

[47] T. Chen and H. R. Wu, “Adaptive impulse detection using center-weighted median filters,” IEEE Signal Process. Lett., vol. 8, no. 1, 2001, doi: 10.1109/97.889633.

[48] S. Patel, K. P. Bharath, S. Balaji, and R. K. Muthu, “Comparative Study on Histogram Equalization Techniques for Medical Image Enhancement,” in Advances in Intelligent Systems and Computing, 2020, vol. 1048, doi: 10.1007/978-981-15-0035-0_54.

[49] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, vol. 2016-December, doi: 10.1109/CVPR.2016.90.

[50] S. Borade and D. Kalbande, “Survey paper based critical reviews for Cosmetic Skin Diseases,” in Proceedings - International Conference on Artificial Intelligence and Smart Systems, ICAIS 2021, 2021, doi: 10.1109/ICAIS50930.2021.9395803.