International Journal of Pure and Applied Mathematics
Volume 118 No. 20 2018, 623-633
ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

Automatic Identification and classification of Allergic skin disorders using DeepPattern Recognition neural Networks

Y.Vijaya
vijaya.vitw@gmail.com
Researchscholar, Department of CSE, Acharya Nagarjuna University, AP, India
Dr.E.Srinivas Reddy
esreddy67@gmail.com
Professor, Department of CSE, Acharya Nagarjuna University, AP, India

Abstract: Skin disease is mostly found in animals, humans and plants and it is a particular kind of illness caused by bacteria or an infection. These diseases mostly include ringworm, yeast infection, brown sports, allergies, etc. Early spotting of diseases will reduce the impact to a large extent. Our proposed system is usedto identify the skin disease where medical expertise is not available. This paper proposes automatic disease detection using convolution neural network. The traditional methods like GLCM,GTDM,Discrete haar wavelets, principal component analysis involves tedious calculations and their accuracy is limited. Performance of the of most classification systems depends on representation of the appropriate data and most of the attempts are devoted to feature engineering, Earlier expert domain systems used are most difficult and tedious processes to identify effective features whereas deep learning can able to identify and organize the discriminative information from the data,It does not require domain expert. This model uses AlexNet which is pretrained convolution neural network. In this paper, we have taken Dermnet skin image data base to test 100 categories of skin disease types. Our proposed model is providing accuracyupto 96% in identifying the skin disease.

Keywords: Image Classification, Deep Convolution Neural network, AlexNet, Dermnet

1 Introduction:

Skin is outer covering of our body which is largest organ of the body and protects us from bacteria, virusesand controls the body temperature. Different skin allergic symptoms such as itching, swelling, burning, irritation,rednesscan be temporary or permanent and they may be painful or pain less.These might be of genetic and may be caused by different situations. Skin conditions may be minor butcertain diseases may be life threatening. Immune system problems can cause dermatitis, hives, and other skin conditions. Many skin problems, such as acne, also affect skin appearance. Our skin can also develop several kinds of cancers.

Some of the common skin conditions are acne,moles, chickenpox, warts, eczema,psoriasis,impetigo,rosacea,eczema,hives, skin cancer, contact dermatitis and keratosis,pilarisare some of the temporary skin conditions.Some chronic skin conditions develop from the time of birth, some may appear later suddenly in life.The cause of disorders may not be known always. Examples of chronic skin conditions include rosacea, vitiligo and psoriasis. Early identification of skin disease is utmost important to avoid further damage. Our automatic skin disease identification is used to identify the skin disease in early state where medical expertise is not available.Deep learningtechnique is used in numerous applications. In the field of image classification deep learning technique is used extensively.

Convolution Neural Networks (CNNs) are used in the applications such as scene recognition and object classification because of its accuracy. The technique especially used in machine learning is the same feature space and distribution is used for training and test data. Most of the statistical methods are re build when the distribution is changed using newly collected data.In most of the real world applications, re-building the model using the newly collected data is very much difficult and expensive task. Recent developments in Deep Neural Networks have allowed researchers to drastically improve the accuracy of object detection and recognition systems.Since the success of AlexNetinLarge Scale Visual Recognition Challenged the deeper and deeper networks .These networks have proposed and accomplished best in classperformance on ImageNet and other benchmarks.

2 Related Work:

Several methods such as GLCM is used to extract the features are used in image recognition by several researchers before the deep learning. Different properties energy, correlation, contrast, Homogeneity and also other measures which include mean,median,mode,standard deviation,entropy as feature detectors to identify skin conditions.

Different features such as complexity ,strength, , busyness, coarseness and contrast are extracted by researchers using Neighbourhood Gray Tone Difference Matrix are used to identify skin condition. Some used Discrete Haar wavelet transform,Principal component analysis,Eigen values and Eigen vectors. Other Researchers used Scale Invariant Feature Transform, Histogram Orient Gradient as feature detectors to identify the skin conditions.These methods involves complex pre-processing ,feature extraction, and classification. These methods suffers with high computational cost and time consumption.Due to the recent advances in Machine Learning, the deep convolution networks are giving accurate results in solving real world problems. CNN is used in Natural Language processing,computer vision, Text classification, scene labelling, face recognition, image classification, human pose estimation,action recognition and document analysis.

3 Proposed System:

This System uses the method of unsupervised deep learning which uses human intelligence for skin disease detectionusing Pattern Recognition neural Networks.

Fig1. Shows the various steps in proposed system

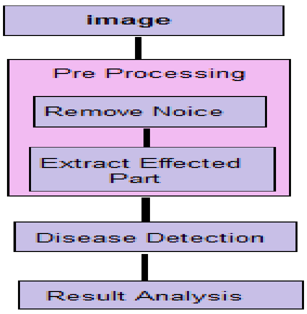


Fig1.various stages in proposed system

Pre-Processing: Pre-processing is used to perform operations onimages at lowest level ofabstraction, image preprocessing will improve the quality of feature extraction. Pre-processing aims at revamping of image data that enhance some features of image and suppress some unwanted distortions of image which are important for image processing .First operation is to remove the noise then extracting the infected portion of image.

Noise Removal: Image data may contain the noise. Noise removal is the primary thing to enhance and improve the quality of image is to remove the noise from image. To remove the noise from imageFirst thing is to load the pre trained denoising neural network

The Algorithm to remove noise is as follows

1.Convert the image into L*a*b* colour space. Dimension L* for lightness and a* and b* are the colour dimensions. This Lab colour space is 3-axis colour pace. All of colours in the spectrum are included while working with L*a*b* colour space, as well as colours outside of human perception.The L*a*b * colour space is the most exact means of representing colour and is device independent. And it is more accurate and portable

2. Noise is primarily in the luminance channel .Remove the noise from this channel only, Noise can be removed by the pretrained denoising neural network.

3. Convert the image back to the RGB colourspace.

Fig2:Original Image and Image after Noise removal

Extracting highlighted portion of the infected area:

First thing is to convert the colour image into gray scale image via principal component analysis,

Binary images should undergo different morphological operations such as opening, closing

The algorithm used to select the highlighted portion of the infected region is as follows

1. Convert the image from RGB to Gray using Principal component Analysis

2.Use morphological closing operations to enlarge the boundaries of foreground

3.Apply 2D wavelet decomposition using B-spline to find horizontal,vertical,diagonaldetails of the image

4.Calculate Otsuthreshold on each of the four outputs

5.Calculate new threshold from sum of the 4 Otsu thresholds and divide it by 2

6.Calculate the single level inverse discrete 2-D wavelet transform

7.Find Black and White segmentation

8.Use canny edge detector to detect the edges of the lesion

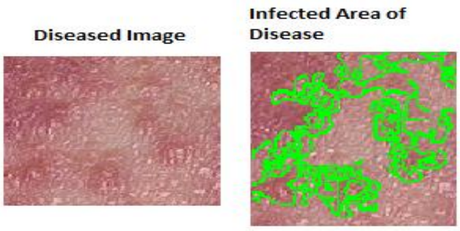


Fig3: Infected portion of Disease

ROI Segmentation: A Portion of image is filtered in order to perform some operations on it. This part of image is called Region of Interest(ROI) which is data set identified for a particular purpose which is a selected subset of samples within a dataset.

Binary mask is created by ROI .It has the size which is same as the input image size to be processed which is the binary image.Region based image segmentation method which is Region growing is used here.It uses initial seed points .To determines whether the pixel should be added to the region or not different neighbouring pixels of initial seed points are examined.

Seeded Region Growing (SRG) :The following stepsare illustrated

1. Neighbouring pixels are compared with the region which are not allocated and the region is grown accordingly

2.Similarity measure is computed by finding the difference between region's mean value and the pixel of interest intensity value. The pixels which have smallest difference measure are allocated to the particular region. When intensity difference between new pixel and the region mean exceeds the threshold value then the process stops

3.The element with radius of R = 3 which is morphological structuring element which has flat disk shape is used to erode and dilated the image

Pre-processed gray scale image is multiplied with resulting image to produce final image

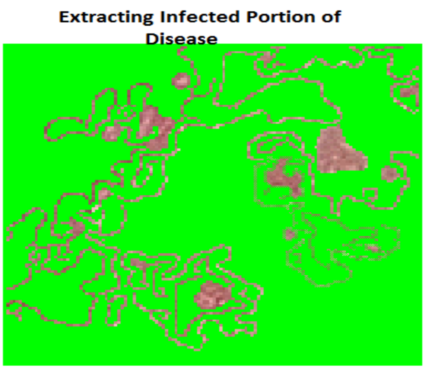


Fig4: Extracting Effected Portion of Disease

Detection Skin disease using deep pattern Recognition Neural Networks:

Traditional methods of pattern recognition techniques such as block-wise orientation histograms(SIFT or HOG) are used in standard datasets such as PASCAL VOC and they are unable to achieve high accuracy .These methods are not able to distinguish well among the different labels.Because they are able to extract very low level features from the objects .In the techniques of object detection by using deep learning concept, pattern recognition neural Networks have achieved success in recognizing patterns. The CNN's provide abstraction from lower to higher levels because of their hierarchical representation of construction. Deep learning can able to find these high level features more effectively in image classification problem which includes medical image analysis.CNN architecture consists of many trainable stages located one on the other.The different feature maps generated by each stage are classified by using classifier. The different inputs such as image,audio and video can be used

Possibility of learning features from input data directly can be explored by Deep learning. 2D array feature map is used in the input image to store the colour channels in colour images. The features are extracted from the associated input .Each feature map is set of arrays represents the image features

A deep net is trained to compute layer-by-layer by feeding it input .It generates final output .Correct answer is compared with output.Error is computed at the output.With backpropagation the error flows backward through the net. To process the data ,Model parameters are tuned at each step backward tries to reduce the error. Multiple passes of input are involved in training which iterative process until the model converges

There are three principle sorts of layers used to assemble CNN design: convolution layer, pooling layer, and fully connected layer. Usually a full CNN architecture is acquired by stacking these layers.

An example of typical CNN architecture is shown fig 5. Key computation In a CNN is convolution of input signal with feature detector. Convolution layer calculates the output of neurons associated to neighbourhood regions in the input. Neurons compute dot product with the weights to which they are connected.Neurons output is computed by the convolution layer .Filter or kernel that contains a set of weights convolved by using input Filter which is connected to the input volume in full depth .

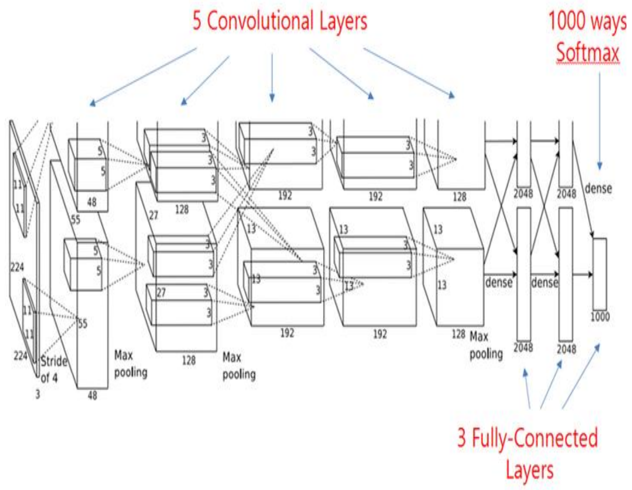


Fig5: An example of CNN Architecture

.

Typical filters are small areas (e. g., 3 × 3, 5 × 5, or 8 × 8) are used for inputs such as images and each neuron is connected to the previous layer of this area. Filters are applied to each neuron in network. The frequent patterns of the image are learned by these filters.**Stride** is distance between the applications of filters.

Different features which are simple primitives from pixels to horizontal and vertical lines, circles, and patches of colour, edges, lines, and corners etc are low-level features extracted by the first convolution layer.In all of the input channels CNN filters are computed In contrast to single-channel conventionalimage processing filters. Whenever a feature is detected high response is yielded by convolution filters Due to its translation-invariant property.Between two successive convolution layers poolingor sub sampling layer is inserted

The purpose of the pooling layer is to reduce the spatial size progressively. In order to control the over fitting problem in the network the pooling layer reduce the number of computation and parameters. Input volume is down sampled in each depth slice spatially by pooling layer. By discarding the activations the input size is resized along the width and height.In practice the window function is applied on the input patch which uses the max pooling function which computes the maximum value in the neighbourhood pixels. The better results are obtained by max function.

The other different functions performed by the pooling layer include average pooling or L2 –norm poolingin fully connected layer. The arrangement of the neurons is such that every neuron is connected to the previous layer neurons. In a regular neural network this type of layer is standard. Using a matrix multiplication followed by a bias offset the activations can be computed. The net output is generated by last fully connected layer such as probability distributions over classes

In our proposed system we have used Alex net which is pre trained Convolution neural network architecture which is used ImageNet data set that consists of more than 1.2M images in 1K categories.

In our experiments CNN architecture that provided the best results contains the following layers and parameters:

Input layer: This layer loads the input and produces the output used to feed convolution layers. The pre processed images of size 227X227X3 are fed to input layer

Some transformations such as mean-subtraction and feature-scaling can be applied. In our case, inputs are images and the parameters define the image dimension (32×32 or 64×64 pixels) and the number of channels (3 for RGB).

Convolution layers:In convolution layer set of learnable filters are convolved with the input image to produce output which feature map

Convolution Layer1:

Images: Input size of the image is 227x227x3, receptive field size (F) is: 11, stride(s) is 4, Convolution layer output: 55x55x96

55*55*96 = 290,400 neurons, each has 11*11*3 = 363 weights and 1 bias, 290400 * 364 = 105,705,600 are total number of parameters on the AlexNet architecture's first layer alone

Convolution Layer2:The output from the first convolution layer which is normalized and pooled

Will be fed as the input to the Convolution Layer2 .Its input size is [27x27x256],Number of filters it used is 256 of size: 5x5x48 with stride 1 pad 2

Convolution layer3: Input size [13x13x384] It has 384 kernels (size: 3x3) with stride 1 and pad 1

Convolution layer4: It has input size [13x13x384] It has 384 kernels (size: 3x3)

Convolution layer 5: It has input size [13x13x256] It has 256 Kernels (size: 3X3) with stride 1

Fully connected Layer: 'fc7' has 4096 neurons

Pooling layers or sub sampling layers:

This layer is used to subsamples features from the input this will reduce the resolution of the feature map and again it reduces the number of features. Effect of local distortions is also lowered by sub sampling. It is generally used as either mean or max pool and this layer is also referred as the pooling because of this the sub sampling. Sub sampling is applied after one or more convolution layers

. These layers are responsible for down sampling the spatial dimension of the input. There are 3 pooling Layer

MAX POOL1: Output volume size is [27x27x96] with 3x3 filters at stride 2

MAX POOL2: Output volume size is [13x13x256] with 3x3 filters at stride 2

MAX POOL3: Output volume size is [13x13x256] 3x3 filters at stride 2

All the max pool layers will use the filter size as 3x3 with stride 2.

All of them are set to use a 3×3 receptive field (spatial extent) with a stride of 2. The pooling layers use the most common max operation over the receptive field

• ReLU layers: After every Convolution operation ReLU operation is used .This is an additional operation. ReLU stands for Rectified Linear Unit .Non linear operation is performed by this layer

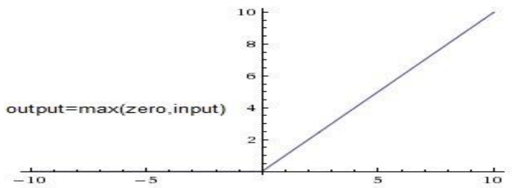


Fig6: Activation function of Rectified Linear Unit (ReLU)

The ReLU performs the function of replacing all the negative pixel values with zero in the given feature space. In our ConvNet the Non linearity is introduced by the ReLU. The data in real world applications the ConvNet learns is non linear. The activation function of ReLU is given by: f(x) = max (0, x) in other words the activation is simply threshold at zero.To add non linearity to the network instead of a linear activation function the rectifier activation function is used. f '(X) gives the derivative of the function which computes to 0 when x<0,1 when x>0 and N.Def when x=0

ReLU is used in convolution layers because of the faster convergence of the network due to non existence of vanishing gradient problem and to activate sparsity in features. The sparsity in the features is used to speed up the computation process by deleting the undesired features

Dropout: Counter measure for overfitting.

Overfitting occurs when fewer samples are used totrain the high number of weights on the training data set in the ConvNet then intrinsic noise of the training data is identified by the model. Dropping out the units in a neural network in both hidden and visible units are called “dropout” here units are not considered during a particular forward or backward pass.

Drop out prevents inter-dependencies from the nodes which will ignore some of the nodes randomly. By this nodes will not learn functions from dependent nodes for the input by this more robust relationship is learned by the network. In our model 50% drop out is used after relu6 and relu7.more useful features which are in conjunction with other subsets of neurons are learned by the dropout. Drop out converge the network by roughly doubling number of iterations. However each epoch uses less training time

Inner-product layers or fully connected layers:

The features which are extracted from the output of convolution and pooling layers are detailed and higher level features.

The fully connected layers classifying the input image into various classes by using these features in the training data set

This model consists of 3 fully connected layers:

‘fc6’ is a fully connected layer with 4096 neurons

‘fc7’ is a fully connected layer with 4096 neurons

‘fc10’ is the final classification layer with 1000 neurons

They treat the input as a simple vector and produce an output in the form of a single vector. There are two inner-product layers in this model. The last one is fully-connected output layer with softmax activation depends on the number of classes in the classification. ‘fc10’ layer takes the input from the output of ‘fc7’. The outputs of ‘fc7’ as its input to ‘fc10’, which generates the output through a fully connected architecture which is to the number of classes, a softmax function is performed on the output , the result of the softmax function is used as the negative log likelihood loss to generate the final result. Only one label is expected in each input image of the softmax loss layer.sigmoid function can be applied to outputs of ‘fc10’, then sigmoid cross-entropy loss layer replaces the soft max loss layer to produce predicted Probabilities,when sigmoid cross-entropy loss is used instead of softmax loss layer , then cross-entropy loss is computed. Here each input can use multiple label probabilities.

Choice of the Loss Function: In classification tasks using deep learning the most widely used loss functions are Softmax Loss and Sigmoid Cross-Entropy Loss .The Training of Alex net is effectively reduced the following cost or error function (negative log-likelihood): using soft max layer where N is number of training images,r isindexes across all traits ($r \in \{1, . . . , N\}$), With the deep net architecture the L2 regularization on weights W is LR = $\lambda||W||2$ where λ is a regularization parameter, by applying the softmax function in layer 'fc10' on M outputs the probability obtained is $\hat{p}_{r,yr,}$M is number of classes we wish to predict labels. $l_{r,m}$is m^{th}output for r^{th} image we have

If Sigmoid cross entropy loss is applied each image is annotated with label probabilities p_r with length M.To minimize the following loss the network is trained

sigmoid function is applied to obtain the probability vector $\hat{p}_r$ for the M outputs of layer 'fc10'.CNN can be trained forour prediction task on all our attributes sigmoid cross-entropy loss function is used because multiple traits are there in the image. Mutually exclusive attribute classes are there for each given image.We should train the net to find the predictions which provide high accuracy

The skin disease detection using deep learning include following implementation steps: input generation,Load deep Convolution Network, training the network, Extracting Features, Classify Features

Input Generation: Read the input image and generate pre processed Image.Here some images may be gray scale.The image is replicated for three times to create an RGB image. Resize the image as required for the CNN.Our image size is 227X227X3

Load Deep Convolution Network: Next step is to load the network. Here. We fed 100 categories of skin disease images and in each category we have used 50 images. Total 5000 images are fed to Alex Net Network.

Prepare the Training and Test set Images: Split the Entire image set into training and validation sets.Out of which Pick up the 30% of images as training data and pick up the remaining 70% of images as validation data. Avoid biasing of the results by randomizing the split. The CNN Architecture will process training and test sets.

Using CNN to Extract the Training Features: Response or activation to an input image is provided by each layer of the CNN. There are some layers in the CNN architecture that are suitable for feature extraction. The initial layers of the architecture extract low level features such as edges and blobs.Following figure shows first convolution layer network filter weights

Fig7: Weights of First Convolution Layer

Low level features like blob and edges are learned by features of first convolution layer, deeper network layers process these "primitive" features. Higher level image features are formed by combing these early features with primitive features.Better image recognition is done by higher level features. Richer image representation can be done combining all the primitive features

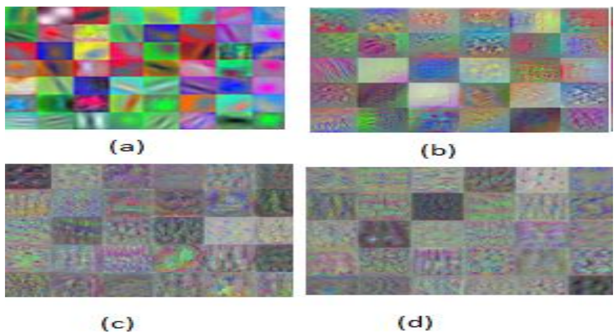


Fig8: Weights of Convolution Layers 2-5

Classify Features using Multi Class Support Vector Machine (MCSVM):

Once features are extracted to classify the test samples the classifier is trained to classify the data as member of one of the known classes. In this work Multi Class Support Vector Machine (MCSVM) are used.For binary classification problems Support Vector Machines (SVM) are well known methods in the machine learning community. By combining several binary SVMs the Multi-class SVMs (MCSVM) are implemented.

It needs to identify the core set of points to establish the boundary because most of the training data is redundant. These data points support the boundary so they are called support vectors

This classifier has been chosen due to its robustness, simplicity and does not tend to over fit training data.

Multi class SVM classifier includes hyper planes that maximizes the distance between classes of data

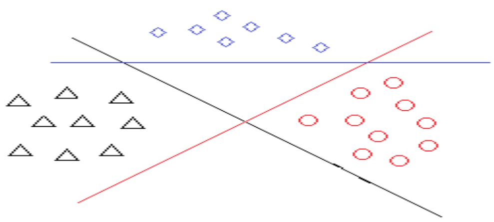


Fig 9: Multi Class SVM Classifier

4 RESULTS AND DISCUSSIONS:

There are many different publicly available skin disease databases. We have used Dermnet data base to test the methodology used in our work

Dermnet data base:Thomas Habif is founder of Dermnetwhich provides information on a wide variety of skin conditions through innovative media. It is photo dermatology source which is the largest independent database which provides articles, photos and video to online medical education.It contains more than 23,000 images. Largenumber of dermatology images is provided in online by Dermnet.com

Some of the Popular Image categories are:

Acne,Acnitic , Keratosis,Alopecia ,areata,Atopoc dermatitis,Basalcell,carcinoma,contact dermatitis,Cyst,Cystic, acne,Eczema,Erythema multiform,genitas,herpes,Genital warta,Garorrhrea,Hair loss,Herpersimplex,Herperzoster,Impetigo,Lice,Lichen, planus,Melanoma ,planus,Melanoma,Melasma,Mole,Mollusim,Nail disease,Nevi,Oraerpes,poison,ivi,Psoriasis,scabies,Kera tosis,skintag,spiderbite,aquamous cell,carcinoma,syphilis,tinea,urticaria,vitiligo.

In our work we have used 100 popular categories.

Some sample images are:

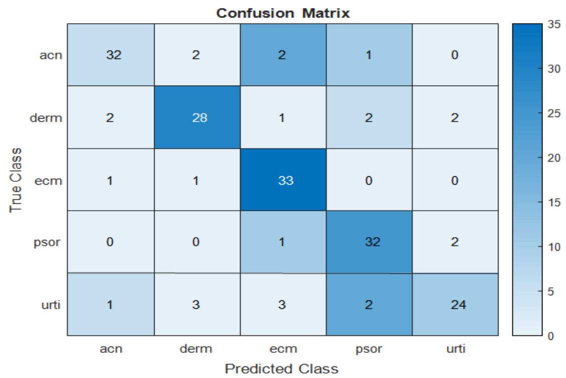
Fig10: Sample skin diseased images

Performance of the Deep Network: Different matrices are used for performance of the deep network are listed below.

Confusion Matrix:

Performance of a classification model is described by confusion matrix which represents instances in an actual class in row and instances in a predicted class in column in a set of test data.It provides the information which is commonly mislabelling one as another that the system is confusing between the two classes

Following figure shows confusion matrix



Performance Measures:Sensitivity,specificity, Accuracy and misclassification rate can be calculated from the confusion matrix.

Accuracy: Accuracy is the ratio of correct classifications to the total number of inputs.

Accuracy= (TP+TN)/ (TP+TN+FP+FN)

Sensitivity: Is the ability of the test to correctly identify patients with the disease

Sensitivity=TP/ (TP+FN)

Specificity: Is the ability of the test to correctly identify patients without the disease

Specificity=TN/ (TN+FP)

Misclassification Rate: Misclassification Rate is the ratio of wrong classifications to the total number of inputs

Misclassification Rate= (FP+FN)/(TP+TN+FP+FN)

Precision:It is the ability when the disease is predicted yes how often it is correct

Positive Predictive Value (PPV): It is the probability that positive test value is truly having the disease or not

PPV=TP/(TP+FP)

F1 score:This is a weighted average of true positive rate or recall and precision

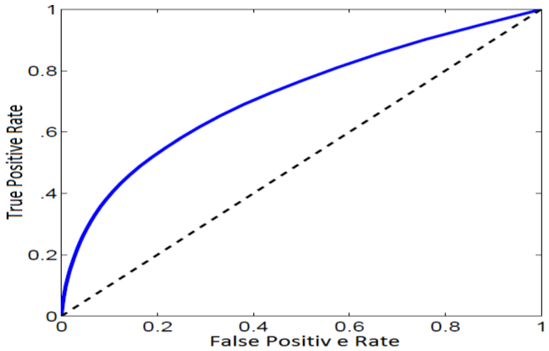
F1 score=2. (precision.recall)/(precision+recall)

Where TP is the True Positive, TN is the True Negative,FP is False Positive,and FN is False Negative

The table below shows the performance measures some sample diseases

	Accuracy	sensitivity	specificity	Miss classification rate	PPV	F1 score
acne	95.43%	76.92%	97.16%	0.07%	88.23%	82.17%
Derma	92.44%	80%	95.62%	0.06%	82.35%	81.15%
eczema	94.79%	94.28%	94.93%	0.05%	82.50%	87.99%
psoriasis	97.05%	91.43%	96.38%	0.05%	86.48%	88.87%
urticaria	92.57%	72.72%	97.18%	0.07%	85.71%	78.68%

ROC curve:In ROC curves the performance of a classifier is represented in the form of graph on all possible thresholds. False Positive Rate is plotted on the x-axis and True Positive Rate is plotted on the y-axis



Conclusion and future work:

In the proposed work, an unsupervised deep learning technique is used in skin disease detection. The images used in this work are obtained from publicly available database Dermnet. The images are first pre-processed to remove digitization noise. The proposed model has achieved an accuracy of up to 96.5% in classifying diseased skin images. Convolution Neural networks are used as feature extractors and feature detectors in effectively identify the skin diseases .We have used pre trained CNN architecturewhich is Alexnet. Experimental results indicate that CNN features easily outperform hand-craftedfeatures in terms of better sensitivity, specificity, and accuracy.

In future work we want to work with some more classification models. We want to work our self in some areas instead of using pre trained models such as edge detection, filter values and feature extraction .We want to use Increase the training data used for training the model, not only in term of quantity but also obtaining more data from different resources namely collecting data from hospitals and healthcare canters, to increase the learning model generalization. Apply better pre processing techniques to resolve the images distortions. Apply training data of more classes, that the model will be capable to recognize and diagnose more diseases. Develop a cross-platform application to work on different mobile platforms, which will increase the number of system users.

REFERENCES:

[1] I.Sturm, S. Lapuschkin, W.Samek, K.R. Müller Interpretable deep neural networks for single-trial EEG classification Methods 274 (Dec 2016) 141–145.

.[2] Y. Zhang, B. Zhang, F. Coenen, J. Xiau, and W. Lu, “One-class kernel subspace ensemble for medical image classification,” EURASIP Journal on Advances in Signal Processing, vol. 2014, no. 17, pp. 1–13, 2014

[3] Ranjanparekh ” Using Texture Analysis for Medical Diagnosis” Jadavpur University, India,1070-986x/12,published by IEEE computer society,vol. 177,May 2011

[4] T. Ravindra Babu, K. Mahalakshmi, and S.V. Subrah manya , “Optimal Skin Detection for Face Localization using Genetic Algorithms”, in Proc. of the IEEE, 2011.

[5] Cugu, I.Sener, E.Erciyes, C.Balci, B.Akin, E.Onal, I.Oguz-Akyuz, A.Treelogy: A Novel Tree Classifier Utilizing Deep and Hand-crafted Representations. arXiv 2017, arXiv: 1701.08291v1

[6]H. Liao, "A Deep Learning Approach to Universal Skin Disease Classification," University of Rochester, 2015

[7] S. S. K. a. V. S. Vinayshekhar Bannihatti Kumar, "Dermatological Disease Detection Using Image Processing And Machine Learning," IEEE, 2016.

[8] Antipov G, Berrani, Ruchaud N, Dugelay, 2015. Learned vs. hand-crafted features for pedestrian gender recognition. In: Proceedings of the 23rd ACM International Conference on Multimedia. ACM, pp. 1263–1266.

[9] Caicedo J, C Cruz, A Gonzalez, F.A., 2009. Histopathology image classification using bag of features and kernel functions. In: Proceedings of Conference on Artificial Intelligence in Medicine in Europe. Springer, pp. 126–135.

[10] Babenko A, Lempitsky V, 2015, “Aggregating local deep features for image retrieval” In: Proceedings

of the IEEE International Conference on Computer Vision, pp. 1269–1277.

[11] Leng C, Cheng J, Bai, Lu H, 2015. "Online sketching hashing" . In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2503–2511.

[12] Cai W, Liu S, Wen L, Eberl S, Fulham M.J, Feng D, 2010. 3rd neurological image retrieval with localized pathology-centric CMRGLC patterns. In: Proceedings of IEEE International Conference on Image Processing. IEEE, pp. 3201–3204.

[13] Tian G, Fu H, Feng, D, 2008. Automatic medical image categorization and annotation using LBP and MPEG-7 edge histograms. In: Proceedings of International Conference on Information Technology and Applications in Biomedicine. IEEE, pp. 51–53.

[14] Wang J, Xiao, Lin, W.Luo 2015. Discriminative and generative vocabulary tree: with application to vein image authentication and recognition. Image Vis Compute 34, 51–62.

[15] Wang, X., Yang, M., Cour, T., Zhu, S., Yu, K., Han, T.X., 2011b. Contextual weighting for vocabulary tree based image retrieval. In: Proceedings of International Conference on Computer Vision. IEEE, pp. 209–216.

[16] Jiménez-del Toro, O.A., Cirujeda, P., Cid, Y.D., Müller, H., 2015. Radlex terms and local texture features for multimodal medical case retrieval. In: Multimodal Retrieval in the Medical Domain. Springer, pp. 144–152

[17] Wang, J., Li, Y., Zhang, Y., Wang, C., Xie, H., Chen, G., Gao, X., 2011a. Bag-of-features based medical image retrieval via multiple assignment and visual words weighting. IEEE Trans. Med. Imaging

[18] Sze-To, A., Tizhoosh, H.R., Wong, A.K., 2016. Binary codes for tagging x-ray images via deep de-noising autoencoders. In: Proceedings of Neural Networks (IJCNN). IEEE, pp. 2864–2871

[19] Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D.D., Chen, M., 2014a. Medical image classification with convolutional neural network. In: Proceedings of the 13th International Conference on Control Automation Robotics & Vision (ICARCV), 2014. IEEE, pp. 844–848

[20] Shen, W., Zhou, M., Yang, F., Yang, C., Tian, J., 2015a. Multi-scale convolutional neural networks for lung nodule classification. In: Proceedings of International Conference on Information Processing in Medical Imaging. Springer, pp. 588–599

[21] Semedo, D., Magalhães, J., Novasearch at IMAGECLEF med 2016 subfigure classification task. In: Proceedings of Working Notes. CEUR Workshop CLEF 2016, Évora, Portugal, CEUR-WS.org (September 5–8).

[22] Salakhutdinov, R., 2015. Learning deep generative models. Ann. Rev. Stat. Appl. 2, 361–385. Salton, G., Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. Inf. Process Manag. 24 (5), 513–523.

[23]] ahbi, H., Audibert, J.-Y., Keriven, R., 2007. Graph-cut transducers for relevance feedback in content based image retrieval. In: Proceedings of IEEE 11th International Conference on Computer Vision. IEEE, pp. 1–8.

[24] Lim, J.-H., Chevallet, J.-P., 2005. Vismed: a visual vocabulary approach for medical image indexing and retrieval. In: Proceedings of Asia Information Retrieval Symposium. Springer, pp. 84–96.

[25] Liu, X., He, J., Lang, B., 2014c. Multiple feature kernel hashing for large-scale visual search. Pattern Recognition . 47 (2), 748–757.

[26] Liu, J., Zhang, S., Liu, W., Zhang, X., Metaxas, D.N., 2014. Scalable mammogram retrieval using anchor graph hashing. In: Proceedings of the IEEE 11th International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 898–901.

[27]Müller, H., Deserno, T.M., 2010. Content-based medical image retrieval. In: Biomedical Image Processing. Springer, pp. 471–494

[28] Müller, H., Menze, B., Langs, G., Montillo, A., Kelm, M., Zhang, S., Cai, W.T., Metaxas, D., 2014. Overview of the 2014 workshop on medical computer visionalgorithms for big data (mcv 2014). In: Proceedings of International MICCAI Workshop on Medical Computer Vision. Springer, pp. 3–10.

[29] Müller, H., Clough, P., Deselaers, T., Caputo, B., CLEF, I., 2010. Experimental evaluation in visual information retrieval. The Information Retrieval Series, 32, pp. 1–554.

[30] Markonis, D., Holzer, M., Baroz, F., De Castaneda, R.L.R., Boyer, C., Langs, G., Müller, H., 2015. User-oriented evaluation of a medical image

retrieval system for radiologists. Int. J. Med. Inform. 84 (10), 774–783

[31] Liu, J., Zhang, S., Liu, W., Zhang, X., Metaxas, D.N., 2014. Scalable mammogram retrieval using anchor graph hashing. In: Proceedings of the IEEE 11th International Symposium on Biomedical Imaging (ISBI). IEEE, pp. 898–901