Skin Lesion Classification using Deep Learning and Image Processing

Atharva Jibhakate
Department of Computer Engineering and Technology
Dr. Vishwanath Karad MIT World Peace University
Pune, India
atharvajibhakate@gmail.com

Pranav Parnerkar
Department of Computer Engineering and Technology
Dr. Vishwanath Karad MIT World Peace University
Pune, India
pranav13300@gmail.com

Sahil Mondal
Department of Computer Engineering and Technology
Dr. Vishwanath Karad MIT World Peace University
Pune, India
sahilmondal@gmail.com

Vastav Bharambe
Department of Mechatronics Engineering
Symbiosis Skills and Professional University
Pune, India
vastavbharambe@gmail.com

Shamla Mantri
Department of Computer Engineering and Technology
Dr. Vishwanath Karad MIT World Peace University
Pune, India
shamla.mantri@mitwpu.edu.in

Abstract— Skin Cancer is the most common (accounting for 40% of cancer cases globally) and potentially life-threatening type of cancers. It was diagnosed in about 5.6 million individuals last year. Automated classification of skin lesions through images has been a challenge throughout the years because of fine variability in their appearance. Deep Learning techniques exhibit potential in tackling fine-margined image-based analysis and manage to provide accurate results. The three modelling stages include data collection and augmentation, model architecture and finally prediction into 7 different types of skin cancer namely actinic keratoses, basal cell carcinoma, benign keratosis-like lesions, dermatofibroma, melanoma, melanocytic nevi and vascular lesions. A Convolutional Neural Network was fabricated (using TensorFlow) obtaining an accuracy of 81.24%. Further Transfer learning Approach was implemented in PyTorch, which yielded accuracies of 96.40%, 98.20%, 98.70% and 99.04% respectively for Wide Resnet101, Resnet50, Densenet121 and VGG19 with batch normalization, which are all trained end-to-end from images directly, to proliferate the scalability of these models and curtail initial diagnostic costs. The aim of this research paper is to render non-invasive skin cancer screening a common norm, making it simpler.

Keywords— Deep Learning, CNN, VGG19, Batch Normalization, Densenet121, Wide Resnet101, Resnet50, Transfer Learning

I. INTRODUCTION

Human skin is immensely complex in nature. It is divided into two major membranes, the Epidermis and the Dermis. Epidermis is the outermost layer which serves as a defensive shield, retaining fluids in the body and blocking bacteria. Dermis consists of connective tissues, providing the skin with tensile strength and elasticity. Skin is subjected to dust, pollution, microbes and UV radiations [10]. Exposure to these tends to be the primary reason why people end up contracting skin diseases like Skin Cancer. Skin Cancer is the abnormal growth of cells in the dermis or epidermis that has the ability to spread to certain parts of the human body [12]. Heavy exposure to ultraviolet radiation, weak immune system in certain individuals and light skin colour generally contribute to the rise in cases of skin cancer. It affects 2.5 million people annually in countries such as Australia, New Zealand, South Africa and the US [15].

Skin lesions, as shown below in Figure 1 [2], are broadly classified into Malignant (cancerous type) and Benign (non-cancerous type). Though benign skin lesions are generally harmless, some of its forms are pre-cancerous conditions and need treatment. Benign Skin lesions are mainly classified into actinic keratoses, benign keratosis-like lesions, dermatofibroma, melanocytic nevi and vascular lesions [1]. Malignant type skin lesions need intensive care and are potentially lethal. Malignant skin lesions are further classified into Non-melanoma Skin Cancer (NMSC) and Melanoma Skin

Cancer (MSC) [9]. NMSC is classified into Basal Cell Carcinoma and Squamous Cell Carcinoma [8]. Any type of cancer targets the normal functioning of the immune system and alters it. Hence, it becomes paramount to diagnose cancers as soon as possible and get proper treatment.

Statistical data from all over the world shows that skin cancer, if detected early, is curable, but often is diagnosed very late. At this point of time, traditional treatments are ineffective and the cancer cells mutate and spread to the other internal parts of the body. Primary diagnosis is done visually, followed by clinical screening, dermoscopic analysis, biopsy and histopathological examination. Despite monumental advancements in skin cancer treatment, early and accurate diagnosis rate stands at an average of 65% (subject to experience of the specialist) [5]. That is where advanced and intelligent deep learning techniques come to help. Deep Learning as a field looks promising when it comes to image-based detection and analysis [3]. In modern times, deep learning is assisting medical practitioners and researchers to discover hidden data opportunities thereby providing doctors with objective evaluation of every condition and help them treat it better, contributing to life-saving medical decisions.
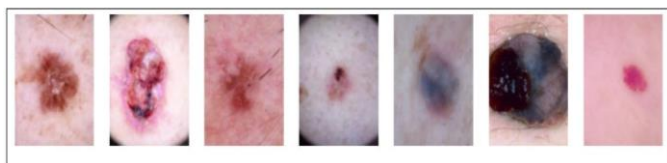

Figure 1: Types of Skin Lesions: Actinic Keratoses, Basal Cell Carcinoma, Benign Keratosis-like lesions, Dermatofibroma, Melanocytic Nevi, Melanoma and Vascular Lesions going from left to right.

II.   RELATED WORKS

Considerable amount of reputable research has been done in image processing and it is being used in medical sciences day in and day out using different machine learning and deep learning techniques. Clinicians and doctors are understanding the positives of using cutting edge technology in every part of their treatment. Following research works have primarily inspired the approach presented in this paper.

Andre Esteva et al. [3] classified skin lesion images into 23 distinct classes using GoogleNetInception v3 network. The accuracy achieved with this approach was 72.1%. 129,450 clinical images were used which consisted of 2032 diseases. t-SNE algorithm was applied for dimensionality reduction.

Haseeb Younis et al. [1] employed MobileNet and CNN consisting of 93 layers out of which 5 layers were dropped and the remaining 88 were considered to develop a skin lesion classification system. The weights of all layers except the last 25 were frozen and were used for training. Using the HAM10000 Dataset an accuracy of 97.1 % was achieved with a 70-30 train test split in the dataset.

V. Pomponiu et al.[4] used the ISBI 2016 Challenge dataset for Skin Lesion Analysis 399 [19] to classify 3 9 RGB

images augmented to 10000 images into 2 classes namely, benign nevi and MM. A pretrained CNN was implemented with the aid of Caffe deep learning library. The features extracted from the last 3 layers of the CNN were fed to a custom classifier built using k nearest neighbor(kNN with 10 fold cross validation to estimate generalization error and k=2). Maximum accuracy of 93.64±1.9 was achieved . Adria Romero Lopez et al. [5] also used the same dataset, and employed transfer learning with the pretrained VGG-16 Network. During fine tuning of the ConvNet network, only the higher level portion of the convolutional layers were trained, keeping the lower level layers frozen. The RMSProp Optimizer function was used for binary classification of images into malignant or benign. The system achieved an accuracy of 81.33%, sensitivity of 78.66% and precision of 79.74%. Despite attaining reputable results, formerly published works were not incorporated in real-world systems because of lower accuracies as compared to manual histopathological examinations.

III.   PROPOSED WORK

The novel system proposed has primarily 5 modules which are pre-processing, modelling, Deep Neural Network, training the model and classification unit as shown in Figure 2 for outputting the requirement. The solution is presented after rigorous testing with five models, four of them being modified transfer learning models (framework used: PyTorch) and one being a CNN model trained from the ground up.
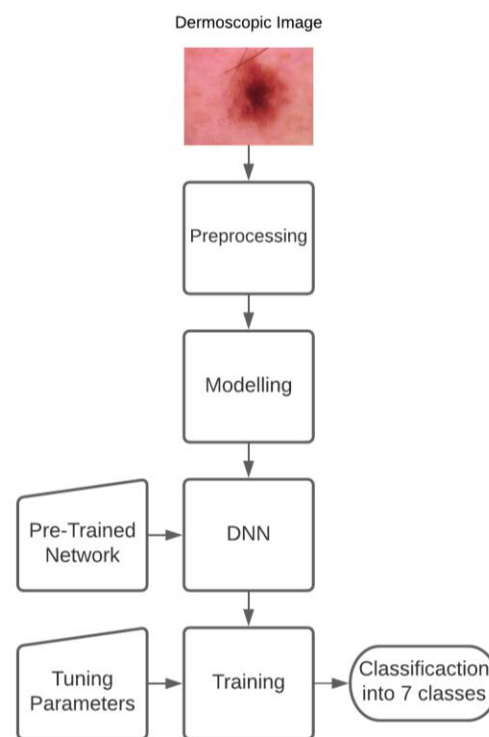

Figure 2: Flowchart of Modelling Process

A. Dataset

Dataset employed for training and testing the models is "Skin Cancer MNIST: HAM10000 (Human Against Machine with 10000 training images)" by "Harvard Dataverse " [2]. It is a large collection of multi-sourced dermatoscopic images of pigmented lesions. In total, the dataset has 10,015 dermatoscopic images, which are sub grouped into seven classes namely, Actinic keratoses and intraepithelial carcinoma / Bowen's disease (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratosis, bkl), dermatofibroma (df), melanoma (mel), melanocytic nevi (nv) and vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and haemorrhage, vasc) [2]. More than 50% of lesions are confirmed through histopathology (histo), the ground truth for the rest of the cases is either follow-up examination (follow_up), expert consensus (consensus), or confirmation by in-vivo confocal microscopy (confocal).

B. Pre-processing

The implemented approach keeps the number of pre-processing steps minimum to maximize the adaptability of the system which means to improve the generalisation ability of the models when feeding-in dermatoscopic images from other skin lesion image data. Thus, standard pre-processing steps required for transfer learning are applied to the dataset. The first step is to eliminate duplicates (images having the same lesion id) from the dataset. The second step is splitting the dataset into training and validation sets. A factor of 0.1 was taken for all the transfer learning models and a factor of 0.2 was taken for the CNN model as the validation set, after considerable permutations of various train-test splits. The third step is normalization of images. This is done by calculating the mean and standard deviation of RGB values followed by application of these values on the dataset to scale all images to the same range. This is done to decrease biasing which in turn yields better results.

$$mean = \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right) \qquad (1)$$

$$std = \sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}(x_i - mean)^2\right)} \qquad (2)$$

Here, n represents the total no. of pixels in the image, and x_i represents a particular pixel value. Lastly, noise in the images has been handled by the transforms.compose() function from the torchvision library wherein the images in the training set are augmented by rotating and flipping them horizontally and vertically as shown in Figure 3 [2]. They are also transformed by modifying their brightness, contrast and hue. The sole reason for augmentation is to make all the networks aware, which implies, training it in a way so as to yield accurate results in cases of incompetent or unreliable imaging.
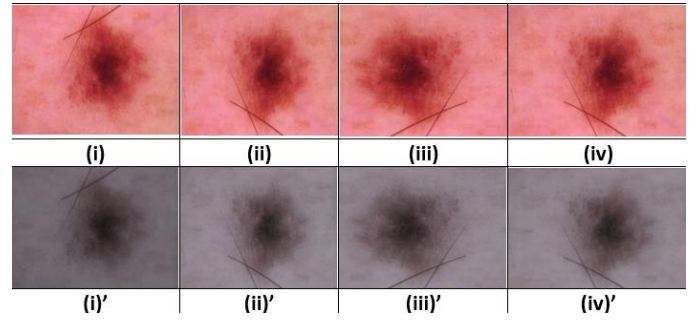

Figure 3: Preprocessed Images, after rotating, flipping and varying contrast, hue and brightness were fed-in to curtail biasing.

C. Modelling

This section presents the implemented Deep Learning Models for skin lesion classification. The architecture of the networks have also been discussed.

1. Convolutional Neural Network

A Convolutional Neural Network (CNN) [20] is a deep learning algorithm that involves a combination of serial convolutional and pooling layers followed by fully connected layers along with a softmax layer at the end, rendering it into a multilayer neural network. CNN is a class of algorithms which is motivated to take advantage of any 2d structure in data. The number of features utilised are 22,500 for each of the resized images in the dataset, with a resolution of (75x100) ppi.

The network presented comprises 6 convolutional layers given by :

$$S[m,n] = (a * b)[m,n] = \sum_j \sum_k b[j,k]\, a[m-j, n-k] \qquad (3)$$

Here, S[m,n] is the feature map, where m,n are the indexes of rows and columns of the feature map. 'a' represents the input image, 'b' represents the kernel and 'j', 'k' are summation variables.

The first two with 128 and 64 kernels respectively of size 3x3. The output of the second convolutional layer serves as an input to the third convolutional layer and filters it with 256 kernels of size 3x3. The fourth, fifth and sixth layers have 128, 256 and 128 kernels of size 3x3 respectively. A pooling layer of 2x2 with equal padding is applied after every even convolutional layer. The process of batch normalisation was carried out after each convolutional layer. Activation function used for the convolutional layer is the Rectified Linear Unit (ReLU) function. ReLU [21] is the standard activation function used when developing multilayer Perceptron and convolutional neural networks. It curbs the vanishing gradient problem thereby aiding the model to learn faster and perform better. After the sixth convolutional layer, the layers are flattened and

attached to a fully connected layer of size 512 which is further connected to a softmax layer having 7 output classes. After toggling with the learning rate, steady and expected results were obtained when it was set to 0.06737947. Adam optimizer [22], an extension to stochastic gradient descent, is used in the system to calculate the gradients effectively to help perform backpropagation. It is computationally efficient and is invariant to diagonal rescale of the gradients. Hyperparameters in it have an intuitive interpretation, setting them is easy and typically needs little tuning.

2. Transfer Learning

Transfer learning is the technique to start with and train a pre-trained model for a new related problem domain [23]. However, there are many options that can be used, including feature transfer and fine-tuning (depending on the similarity of the problems at hand), as well as freezing certain layers of the network and retraining others. If there is insufficient training data, an existing model (from a related problem domain) can be used with additional training to support the new problem domain. The number of features utilized (in all the transfer learning networks) are 150,528 for each of the resized images in the dataset, with a resolution of (224x224) ppi. The networks used in this approach are mentioned below.

2.1 Densenet-121

A characteristic trait of this network is that each layer acquires additional inputs from all preceding layers and tags on its own feature-maps to all subsequent layers. Employed pre-trained densenet network consists of 4 dense blocks which in turn comprise of 6, 12, 24 and 16 dense layers respectively. Every adjacent dense block is amalgamated with a transition block. So, in total, three transition blocks were used. Lastly, the dense neural network with 1024 input features and 512 output features in the first layer, was integrated with the pretrained densenet network. Subsequently, two additional linear layers with a drop out factor of 0.3 were also added. Ultimately the modified network was followed by the output layer (softmax) with seven (output) classes. An accuracy of 98.72% was observed with a learning rate of 0.0497870684, adam optimizer and cross entropy loss function with the modified Densenet-121 network. The ReLu activation function is used for activating the neurons, given by:

$$R(z) = max(0,z) \qquad (4)$$

Here, z can be any real number. Any z<0, returns R(z) = 0 and for any z >= 0, R(z) = z.

2.2 Resnet-50 and Wide Resnet-101

Resnet and Wide Resnet networks consist of BottleNeck blocks. These are similar to BasicBlocks. A BottleNeck block uses a 1x1 convolution to minimize input channels before performing an expensive 3x3 convolution. It then uses another 1x1 convolution to project it back into the original shape. Resnet has four sequential blocks. Each sequential block has a certain number of bottleneck layers. The reason for using bottleneck blocks is to optimally utilize the GPU RAM and not squander it with those expensive 3x3 convolutions. The number of bottleneck blocks in the four sequential blocks are 3, 4, 6 and 3 respectively. Wide Resnet's architecture is heavily inspired by Resnet's architecture. It differs in the number of bottleneck blocks in the third sequential block, which are 23 in its case. The number of filters in the convolutional layer have also increased when compared to resnet. Lastly, the novel dense neural network with 2048 input features and 1024 output features in the first layer, was integrated with both of the pretrained Resnet and Wide Resnet networks. Subsequently, two additional linear layers with a drop out factor of 0.3 were also added. Ultimately the modified networks were followed by the output layer (softmax) with seven (output) classes. Again, the Relu activation function was utilised for activating the neurons.

Accuracies of 98.2% and 96.4% were observed with Resnet and Wide Resnet networks respectively. Adam optimizer and cross entropy loss function were utilised as well with the two modified networks. The weight updation of the adam optimizer is given by:

$$w_t = w_{t-1} - \eta \frac{\widehat{m_t}}{\sqrt{\hat{v}}_t + \epsilon} \qquad (5)$$

Here w_t refers to the weight of the t_{th} training example, $\widehat{m_t}$ refers to the expected value of the moving average of gradient, $\hat{v_t}$ refers to the expected value of the moving average of squared gradient and ϵ refers to the bias.
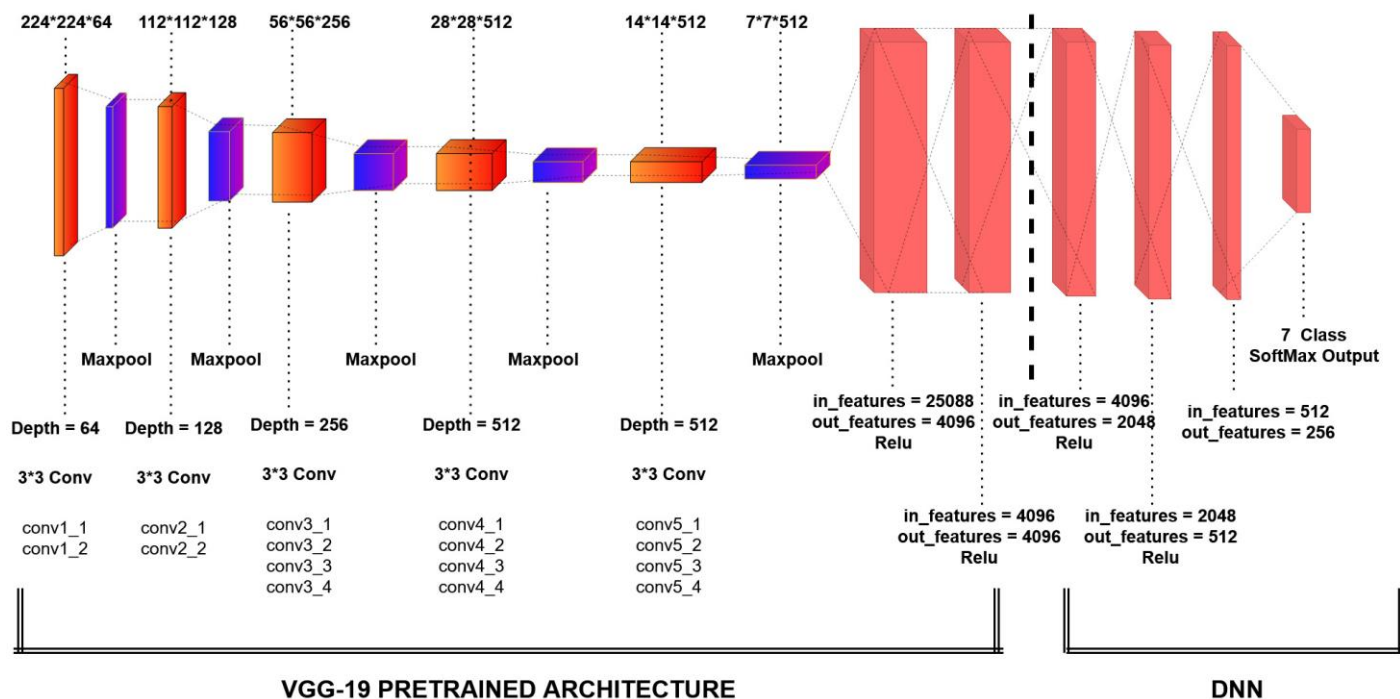
2.3 *VGG-19*

Figure 4: Modified VGG-19 Architecture

VGG-19 DCNN was constructed and the model parameters were fine-tuned, as per the model training and test results.VGG network is structured primarily with multiple connected convolutional layers and fully connected layers. The number of convolutional layers and fully connected layers are sixteen and three respectively for the VGG-19 network. The size of the convolutional kernel is 3x3 and the input size is 224x224x3. Sixty-four kernels from the first convolutional layer, were utilized for feature extraction from the input images. Since it is built with an alternating structure of multiple convolutional and non-linear activation layers, it thumps over a single convolution. It can better extract image features, utilise max-pooling for down sampling and modify the linear unit (ReLU) as the activation function. This implies that it has the ability to select the largest value in the image area as the pooled value of that area. The main purpose of implementing the down sampling layer is to better the anti-distortion ability of the network while retaining the main features of the image sample and to reduce the number of parameters. All the layers in the network, except the last five, are freezed during training. After which the updated weights were fed into the new deep neural network which comprises a sequential block having four linear layers. The modified network was activated using the ReLU function and connected to the softmax output layer with seven (output) classes.

An accuracy of 99.04% was recorded with the modified VGG-19 network as shown in Figure 4, which is the best among all the models used for classification.

The classification report table for all the seven types of skin lesions classified by the network are provided in Table 1.

TABLE I. Classification Report Table for Modified VGG-19

Types of Skin Lesions	Precision	Recall	F1-Score
Melanocytic Nevi	1	1	1
Melanoma	0.99	1	0.99
Benign Keratosis	0.98	0.98	0.98
Basal Cell Carcinoma	1	1	1
Actinic Keratosis	0.98	0.95	0.97
Vascular Lesions	1	1	1
Dermatofibroma	0.98	1	0.99

IV. RESULT ANALYSIS

This section contains trends and key findings that were identified after analyzing the implemented models and extracting valid information from their classification reports.

1. *Age vs Lesion Plot*

From the age vs. lesion plot shown in Figure 5, it can be inferred that a significant spike in Melanoma cases (most lethal) is noticed among individuals in their late 50s, vascular lesions have the highest occurrence rates among individuals of all ages and occurrence of Basal Cell Carcinoma and Actinic Keratosis (pre-cancer) is observed among individuals over 40 years of age.
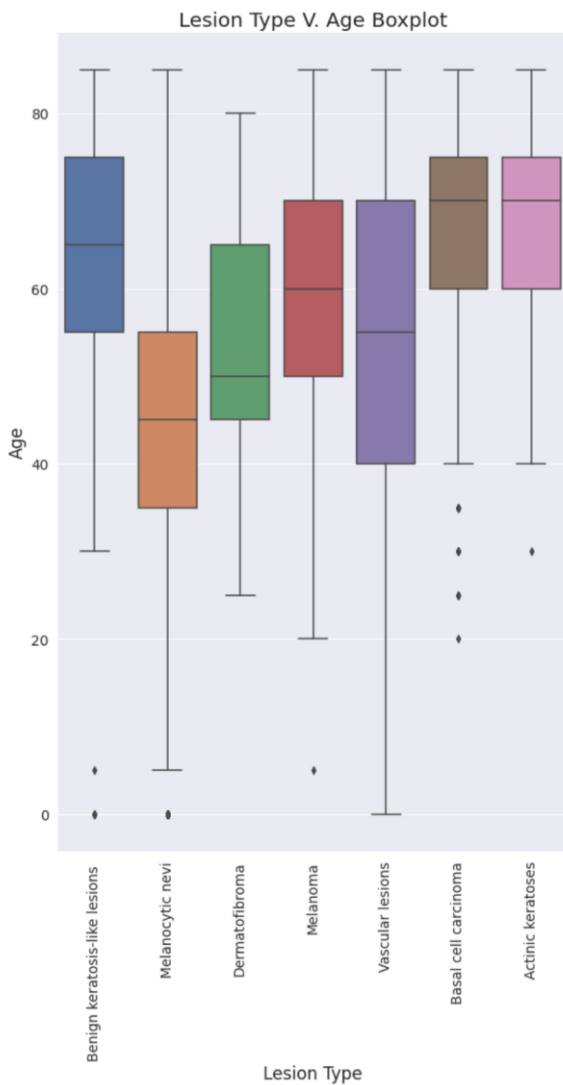

Figure 5: Lesion Type Vs. Age Boxplot

2. *Performance Analysis*

The reason behind the increased performance from a deeper network, is that a more complex and nonlinear function can be learned. If sufficient training data is provided, then it enables the networks to easily differentiate between classes.The motive behind extending the pretrained networks into a much deeper neural network is to maximize the feature extraction potential of the system.

By analyzing the barplot given in Figure 6, it can be concluded that any pretrained network is improved by adding an external DNN. This increases the overall accuracy and precision of the model. The VGG-19 Network, when combined with a DNN, performed exceptionally well, achieving an accuracy of 99.04%.
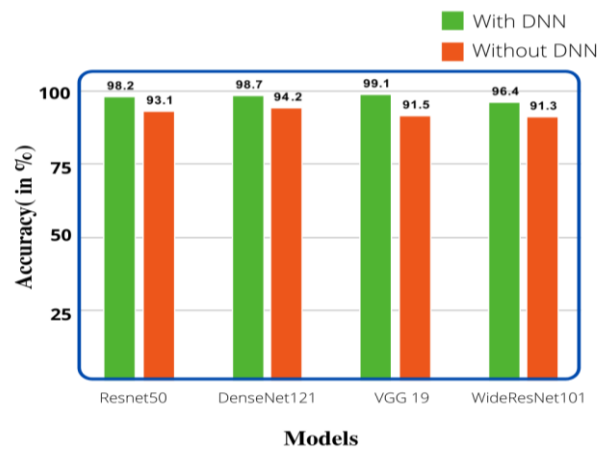

Figure 6: Accuracy comparison with and without DNN
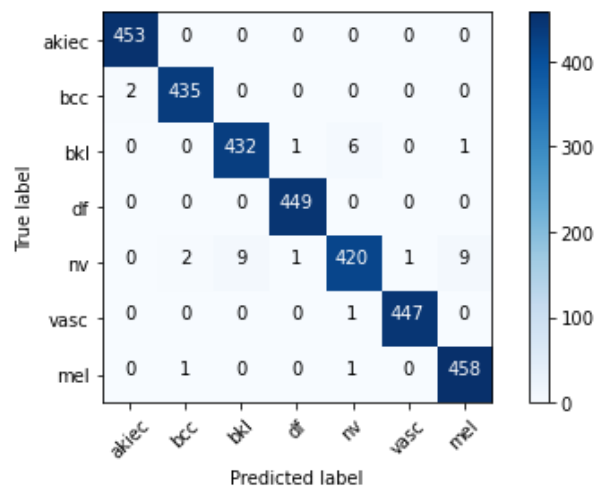
3. *Confusion Matrix*

Figure 7: Confusion Matrix for VGG-19 Network

Confusion Matrix is a table layout for visualising performance. It is also called an error matrix. The diagonal

(highlighted in dark blue) represents the number of true positives. Confusion matrix lends us various quantitative figures like precision, recall and accuracy. For example, in Figure 7, for the 'bkl' class, precision computed from the matrix is 0.98 as shown in Table 1. In this case, the number of true positives and false positives are 432 and 8 respectively, thus the precision is (432/[432+8]) which corresponds to 0.98. The confusion matrix presented in Figure 7 is the outcome of execution of the VGG19 network.

Trends extrapolated from the confusion matrix are as follows: classes of Actinic Keratosis and Dermatofibroma were classified with 100% precision and the class of Melanocytic nevi has the highest amount of inaccuracy when compared to other classes.

V. CONCLUSION

Remarkable research carried out on skin lesion classification paved the way to move forward with this comparative study. The novelty of these networks when compared to previously done research is that they have surpassed the current benchmarks and their performance was additionally boosted as a result of adding the extended dense neural network.

After analysing all the 5 models, the most promising result was obtained using VGG-19 Network which delivered an accuracy of 99.04%. Accuracies of CNN, Wide Resnet101, Resnet50 and Densenet121 are 81.24%, 96.40%, 98.20% and 98.70% respectively. The results show that using these modified networks, better visual diagnostic precision is obtained as compared to a human pathology expert. Additionally, these models take care of histological sampling bias. These algorithms run without any physical hindrance and transform histopathology from a qualitative art to a quantitative science. They can evaluate thousands of features and compare them efficiently to a library of millions of known samples which has been a monumental challenge for experts up until now. The ease of implementation clearly exhibits the potential of these models being used in dermatoscopic systems and modern smartphones in the near future. The ultimate research objective is to build a state-of-the-art application for non-invasive skin cancer detection using computer vision and image processing, which would stand to be a global standard and be inexpensive as compared to other systems which are already in place for the same.

REFERENCES

[1] H. Younis, M. H. Bhatti and M. Azeem, "Classification of Skin Cancer Dermoscopy Images using Transfer Learning," 2019 15th International Conference on Emerging Technologies (ICET), Peshawar, Pakistan, 2019, pp. 1-4, doi: 10.1109/ICET48972.2019.8994508.

[2] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Sci. Data, vol. 5, no. 1, p. 180161, Dec. 2018

[3] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, Feb. 2017.

[4] V. Pomponiu, H. Nejati and N. -. Cheung, "Deepmole: Deep neural networks for skin mole lesion classification," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016, pp. 2623-2627, doi: 10.1109/ICIP.2016.7532834.

[5] Romero-Lopez A, Giro-i-Nieto X, Burdick J, Marques O. Skin lesion classification from dermoscopic images using deep learning techniques. Calgary: ACTA Press. 2017. https://www.actapress.com/PaperInfo.aspx?paperId=456417. Accessed 18 Dec 2019.

[6] Han SS, Kim MS, Lim W, Park GH, Park I, Chang SE. Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm. J Invest Dermatol. 2018;138(7):1529-1538. doi:10.1016/j.jid.2018.01.028

[7] Nahata H., Singh S.P. (2020) Deep Learning Solutions for Skin Cancer Detection and Diagnosis. In: Jain V., Chatterjee J. (eds) Machine Learning with Health Care Perspective. Learning and Analytics in Intelligent Systems, vol 13. Springer, Cham. https://doi.org/10.1007/978-3-030-40850-3_8

[8] P. Dubal, S. Bhatt, C. Joglekar and S. Patil, "Skin cancer detection and classification," 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI), Langkawi, 2017, pp. 1-6, doi: 10.1109/ICEEI.2017.8312419.

[9] A. Demir, F. Yilmaz and O. Kose, "Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3," 2019 Medical Technologies Congress (TIPTEKNO), Izmir, Turkey, 2019, pp. 1-4, doi: 10.1109/TIPTEKNO47231.2019.8972045.

[10] M, Vijayalakshmi. (2019). Melanoma Skin Cancer Detection using Image Processing and Machine Learning. International Journal of Trend in Scientific Research and Development. Volume-3. 780-784. 10.31142/ijtsrd23936.

[11] Lei, B, Jinman, K, Euijoon A, and Dagan F.: ISIC 2017- Automatic Skin Lesion Analysis using Large-scale Dermoscopy Images and Deep Residual Networks. https://arxiv.org/ftp/arxiv/papers/1703/1703.04197.pdf. Accessed 27 June 2018

[12] S. N. Hasan, M. Gezer, R. A. Azeez and S. Gülseçen, "Skin Lesion Segmentation by using Deep Learning Techniques," 2019 Medical Technologies Congress (TIPTEKNO), Izmir, Turkey, 2019, pp. 1-4, doi: 10.1109/TIPTEKNO.2019.8895078.

[13] A. Namozov and Y. I. Cho, "Convolutional Neural Network Algorithm with Parameterized Activation Function for Melanoma Classification," 2018 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, 2018, pp. 417-419, doi: 10.1109/ICTC.2018.8539451.

[14] Hao Chang. Skin cancer reorganization and classification with deep neural network, arxiv.org/ftp/arxiv/papers/1703/1703.00534.pdf

[15] Wikipedia contributors. "Skin cancer." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 19 Jul. 2020. Web. 8 Sep. 2020.

[16] Fujisawa Yasuhiro, Inoue Sae, Nakamura Yoshiyuki, The Possibility of Deep Learning-Based, Computer-Aided Skin Tumor Classifiers, Frontiers in Medicine,Volume 6, 2019, https://www.frontiersin.org/article/10.3389/fmed.2019.00191, DOI=10.3389/fmed.2019.00191, ISSN=2296-858X

[17] Porcia T. Bradford, Dermatol Nurs. Author manuscript; available in PMC 2010 Jan 1. Published in final edited form as: Dermatol Nurs. 2009 Jul–Aug; 21(4): 170–178. PMCID: PMC2757062

[18] Barata, Catarina & Marques, Jorge. (2019). Deep Learning For Skin Cancer Diagnosis With Hierarchical Architectures. 841-845. 10.1109/ISBI.2019.8759561.

[19] Gutman, David & Codella, Noel & Celebi, M. Emre & Helba, Brian & Marchetti, Michael & Mishra, Nabin Kumar & Halpern, Allan. (2016). Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC). arXiv:1605.01397 [cs.CV].

[20] Lecun, Yann & Haffner, Patrick & Bengio, Y.. (2000). Object Recognition with Gradient-Based Learning.

[21] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," Haifa, 2010, pp. 807–814. [Online]. Available: https://dl.acm.org/citation.cfm

[22] Kingma, D.P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. CoRR, abs/1412.6980.

[23] Tan C., Sun F., Kong T., Zhang W., Yang C., Liu C. (2018) A Survey on Deep Transfer Learning. In: Kůrková V., Manolopoulos Y., Hammer B., Iliadis L., Maglogiannis I. (eds) Artificial Neural Networks and Machine Learning – ICANN 2018. ICANN 2018. Lecture Notes in Computer Science, vol 11141. Springer, Cham. https://doi.org/10.1007/978-3-030-01424-7_27