

# Automated Diagnosis of Melanoma from Dermoscopic Images using Convolution Neural Network

R. Ragumadhavan<sup>1</sup>, P. Karthikeyan<sup>2</sup>, M. Arivalagan<sup>3</sup>, J. Joseph Durai Selvam<sup>4</sup>, K. R. Aravind Britto<sup>5</sup>

<sup>1</sup>Assistant Professor, Department of Electronics and Communication Engineering, PSNA College of Engineering and Technology, Dindigul, India.

<sup>2</sup>Associate Professor, Department of Electronics and Communication Engineering, P.S.N.A College of Engineering and Technology, Dindigul, India.

<sup>3</sup>Assistant Professor, Department of Electronics and Communication Engineering, Nadar Saraswathi College of Engineering and Technology, Theni, India.

<sup>4</sup>Associate Professor, Christ Deemed University, Bengaluru, India.

<sup>5</sup>Professor, Department of Electronics and Communication Engineering, PSNA College of Engineering and Technology, Dindigul, India.

## ABSTRACT

Artificial intelligence and machine learning lead to a frenetic pace of change in healthcare and medical imaging. In recent years deep learning algorithms play an important role in designing clinical decision support systems using images. Melanoma is the fifth most widespread cancer among men and the sixth most common cancer among women. Melanoma is 20 times more common in white people than in black people. In the present work, the novel attempt has been made to classify melanoma images from normal images using a convolution neural network. Images are collected from various online data repositories, namely Human against Machine (HAM), Ph2 dataset, mednet, dermis and dermnet dataset. We explored various pre-trained models, namely alexnet, Visual Geometry Group VGG16, and VGG19, using the concept of transfer learning. Pre-trained models used for binary classification melanoma images. Further deep convolution neural network architecture is built exclusively for the binary classification of input images. The hyper parameters like learning rate, activation function, and batch normalization of the network are adjusted to improve the performance. The performance is analyzed in terms of sensitivity, specificity, accuracy, precision, recall, and F-score. Further, it is observed that the built architecture outperforms the pre-trained model.

**Key words :** Dermoscopy Images, Deep Learning, Convolution Neural Network, Transfer Learning, Hyper Parameters.

## 1. INTRODUCTION

Melanoma is a type of skin cancer that develops when melanocytes in the skin start to grow out of control. A mole is a benign skin tumor, and it is not detrimental. Melanoma stages are designated based on the size and thickness of the

tumor. The melanoma staging system is known as TNM staging where T= tumor thickness, N= node, and M= Metastasis. Stages of melanoma include from zero to four. Phase zero indicates that cancer cells are confined to the epidermis. Stage one indicates cancer cells have grown deeper but not spread to the lymph nodes. Stage 2 depicts still further deep growth than 1. Stage 3 means cancer cells spread to other lymph nodes. In stage four, cancer cells grow beyond the skin and regional lymph node.

## 2. RELATED WORKS

Image preprocessing is an important step to exercise decision-based on images. The artifacts present in the dermoscopy image include skin hair, reflection, and ruler marking. Color normalization, Color correction step, and filtering are used in the preprocessing stage. Thresholding is a very common method of image segmentation. Color optimization and clustering-based histogram thresholding is used to perform lesion segmentation [1]. Clustering, fuzzy learning, supervised learning, and graph theory-based methods are explored and used for lesion segmentation. Threshold fusion methods such as Huang and Wang's fuzzy similarity method (Huang and Wang, 1995), Kapur et al.'s maximum entropy method [2], Kittler and Illingworth's minimum error thresholding method (Kittler and Illingworth, 1986), and the Otsu's method (Nobuyuki, 1979) is used for lesion segmentation. Features extracted after segmentation can be grouped under three heading, namely shape, texture, and color. Famous ABCD (A-Asymmetry, B-Border, C-Color, D-Diameter) provides a scoring scheme, which is used to classify the input image. Total Dermoscopic Score (TDS) value is calculated by using equation(1). TDS value of less than 4.75 is benign lesion, value between 4.75 and 5.45 is suspicion and above 5.45 indicates malignant lesion.

$$TDS = [(A \times 1.3) + (B \times 0.1) + (C \times 0.5) + (D \times 0.5)] \quad (1)$$

Menzies scoring scheme [3] uses both positive and negative features. Negative features contain two items, namely symmetry, and color. Positive traits include blue-white veil, multiple brown dots, pseudopods, radial streaming, scar-like

depigmentation, peripheral black dots/globules, multiple colors (five or six), multiple blue/grey dots, and broadened network.

In the 7-point agenda technique [4], three significant rules (atypical color organize, blue-white cover, atypical vascular example), each with a score of 2 are utilized. Four minor standards (sporadic streaks, unpredictable pigmentation, sporadic specks/globules, relapse structures), each with a score of 1, are recognized. A total score of 3 ascertains the lesion to be melanoma. CASH algorithm (C- color, A-Architecture, S-Symmetry and H- Homogeneity) is a widely used method for distinguishing normal image from melanoma [5]. Solar scan developed by polartechnics extracted 103 features based on color, pattern and geometry. The developed system was tested with a total of 2430 images including 382 melanoma images. Segmentation of the skin lesion can be grouped under four heading namely (1) histogram thresholding [6] (2) color clustering method [7] (3) edge based[8] and region based methods. Active-contour methods such as snakes which use curve evolution techniques to determine the contours of the shape , as discussed in[9, 10]. Median cut algorithm[11] used for color quantization , which divides the color space based on the color distribution of colors in the image.

Using shape and color features, with a help of k- nearest neighbor classifier, a robust system is developed by Ganster et al [12]. The developed system classifies the input image into three classes: benign, dysplastic, and malignant. In feature extraction step, while studying about asymmetry several methods were used in literature to determine symmetry axis namely major axis of the best fit ellipse[13], Fourier transform[14] and longest or shortest diameter. Then the difference between areas on both the sides is calculated [15]. Shape asymmetry, color and pattern asymmetry is used as metric for evaluating symmetry [16].

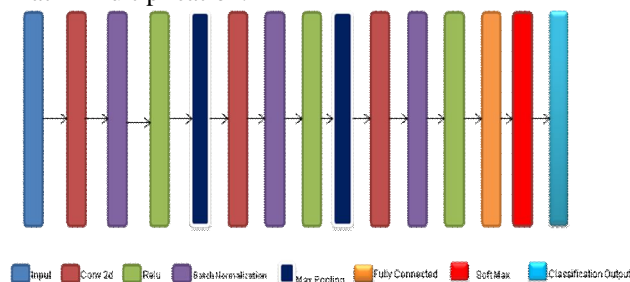
### 3. MATERIAL AND METHODS

**Material:** Images are collected from various online resources. Dermquest dataset has two types of images, namely nevus and melanoma. Med-node dataset has 100 nevi and 70 melanoma images collections. .From the International Skin Imaging Collaboration(ISIC) dataset images under benign(19373 images) and malignant images (2285 images) were taken for the present work. DermIs dataset is also used. Ph2 dataset has 160 nevi and 40 melanoma images. A collection of 13000 images (10,000 benign and 3000 malignant) were used for the present work.

#### 3.1 Methodology

The CNN is designed with deep 15 layers of architecture with learning parameters to perform binary classification. The layer by layer DCNN design is shown in figure 1. The whole input images are resized (28×28) with eight-bit depth RGB channels. The convolution layer extracts the features from the input layers. The convolved weights with input are known as the kernel. The amount of filter shifting is called stride, usually lower than the kernel. The size of the kernel is 3×3,

and the stride value is 2. The weights are passed by using the activation layer Relu. The pooling layer performs the dimensionality reduction using downsample the features with stride 1. The fully connected layer has a complete connection with neurons. The activation is computed with bias offset and matrix multiplication.



**Figure 1:** Constructed Architecture

The learning parameters are implemented to develop the DCNN model, based on the optimizer stochastic gradient descent (SGD). SGD does frequent updates of weights, hence converge faster for the larger dataset. For the training process, the learning rate is chosen as 0.001, which gives better results. The maximum epochs terminated with 300 and further no improvement in accuracy after 300 epochs. The validation frequency is selected as 10. DCNN is the feed-forward neural network includes a series of convolution operation and subsampling layers. The function of layers depends on three parameters, namely, height, width, and the number of channels.

Relu is the activation function utilized between the pooling layer and the convolution layer. The activation function produces an activation map (M):

$$M = F(\text{Input image} * \text{width} + \text{Bias})$$

The dimension of the activation map is given as,

$$1 + \frac{l-f}{s} \times 1 + \frac{l-f}{s} \times 1$$

By applying different kernels (K) in the activation map, the dimension can be changed as

$$1 + \frac{l-f}{s} \times 1 + \frac{l-f}{s} \times 1 \times K$$

From the above equations, the spatial dimensions can be reduced based on the factors input image dimensions (i), filter size (f), and stride (Gallego-Posada, Montoya-Zapata, & Quintero-Montoya, n.d.). The padding is used to extract the low-level features by adding pixels outside the image.

The convolution layer (l) is the input layer comprises feature maps  $x_1$  of size  $x_2 \times x_3$ . The convolution layer receives the raw input images and sends feature maps to the output layer.

The output feature maps  $y_1$  of size  $y_2 \times y_3$ . The  $l^{th}$  feature map in the convolution layer is denoted by,

$$y_j^l = B_i^l + \sum_i k_{ij}^l \times y_j^{l-1}$$

Where  $B_i^l$  is the bias parameter for training data? The kernel size is denoted by  $k_{ij}^l$ .

The max-pooling function computes maxima in the neighborhood given as

$$y_j = \max_{M \times M} v_i^{M \times M} w(x, y) \tag{1}$$

Where  $w(x, y)$  is the window function and  $v_i^{M \times M}$  maxima variable calculated from the image patch?

The softmax activation layer noticed as a normalized exponential function and the equation gives out

$$p(y) = \frac{e^{x \cdot w_j}}{\sum_{q=1}^Q e^{x \cdot w_q}} \tag{2}$$

Where  $p(y)$  is the probability vector. The input and weight are denoted by  $x$  and  $w$ .

The convolution layer width and height can be determined by using the following equation (3) and (4)

$$C_w = \frac{i_w - f_w + (2 \times \text{Zero padding})}{s_w} + 1 \tag{3}$$

$$C_h = \frac{i_h - f_h + (2 \times \text{Zero padding})}{s_h} + 1 \tag{4}$$

Where  $C_w$  and  $C_h$  represents convolutional layer width,  $i_w$  and  $i_h$  denoted as the input image width and height,  $f_w$  and  $f_h$  refers to the convolution filter width and height. Similarly  $s_w, s_h$  and represents the stride width and height.

The mathematical representation of DCNN architecture with the input size  $28 \times 28$  is illustrated by following calculations.

- In the first convolution layer ( $C_1$ ) we used the filter size of  $3 \times 3$ , stride 1 and kernel 8. The feature map utilized 6292 ( $28 \times 28 \times 8$ ) neurons in  $C_1$

$$C_1 = \frac{28 - 3 + (2 \times 1)}{1} + 1 = 28$$

- The first max pooling layer ( $MP_1$ ) used 1568 neurons for feature mapping

$$MP_1 = \frac{28}{2} = 14$$

- The second convolution layer ( $C_2$ ) employs the filter size  $(3 \times 3)$ , stride 1 and kernel 16.  $C_2$  exploits 3136 ( $14 \times 14 \times 16$ ) neurons for feature mapping

$$C_2 = \frac{14 - 3 + (2 \times 1)}{1} + 1 = 14$$

- The second max pooling ( $MP_2$ ) implements 784 ( $7 \times 7 \times 16$ ) neurons for feature mapping

$$MP_2 = \frac{14}{2} = 7$$

- The final convolution layer ( $C_3$ ) utilized the filter size  $(3 \times 3)$ , stride 1 and kernel 32.  $C_3$  used 1568 ( $7 \times 7 \times 32$ ) neurons for feature mapping.

$$C_3 = \frac{7 - 3 + (2 \times 1)}{1} + 1 = 7$$

### 3.2 Alex-net

Transfer learning is the process of fine-tuned parameters utilized to perform a different task from the network trained with one task. AlexNet is the deep neural network architecture with 4096 fully connected layers. By using the pretrained network, we analyze the strength of the network when transferring network to the medical image domain. Training of AlexNet had taken more than 4 hours, depending on the specification of the system and size of the datasets. In this experiment, the dataset is divided into two sections, one for training (80%) and another 20% is utilized for validation. The termination of the epoch is based on the observation of the highest accuracy.

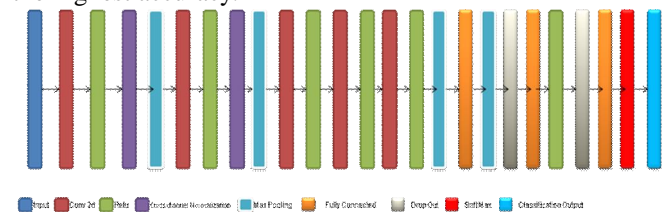


Figure 2: The arrangement of layers in Alexnet

In our experiment, we change the input image size  $227 \times 227$ . Alexnet architecture is shown in figure 2. The fully connected layer is changed with dimensionality reduction. Finally, for binary class, the fully connected layer is 2. The training is limited up to 3 epochs, and the learning rate is  $10^{-5}$ .

### 3.3 VGG-16

The VGG-16 is a multi-layered deep convolution neural network. The layers used in VGG-16 depicted in figure 3. The fully connected layer comprised of 4096 neurons. For feature extraction and dimensionality reduction, convolution, and max-pooling layer are used. The entire dataset is divided into binary and multiclass for training and validation. The 80% of data is utilized for training, and remaining is used for validation. The validation frequency is 20 with three epochs, and the learning rate is  $10^{-5}$ . The training time depends on the size of the dataset and the specification of the system. The VGG-16 fine-tuned pre-trained model for the histology image classification is shown in figure 4.

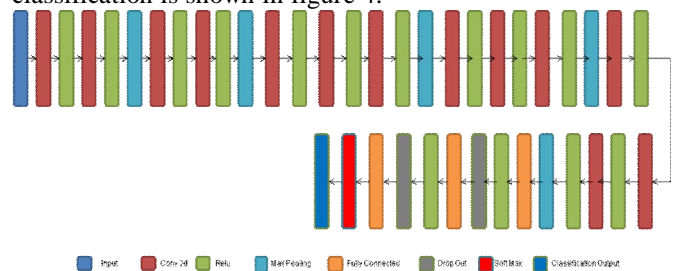


Figure 3: VGGnet-16 Architecture

#### 4. EXPERIMENTAL RESULTS AND DISCUSSION

Using Alex with the following hyper parameter specification, better results were obtained. Epoch – 7, Iterations – 1001, Iterations per Epoch – 143, Validating at every 50 Iterations, 90% Data - Training and 10% Data- Testing, Learning Rate - 0.00001 and Batch processing - 50.

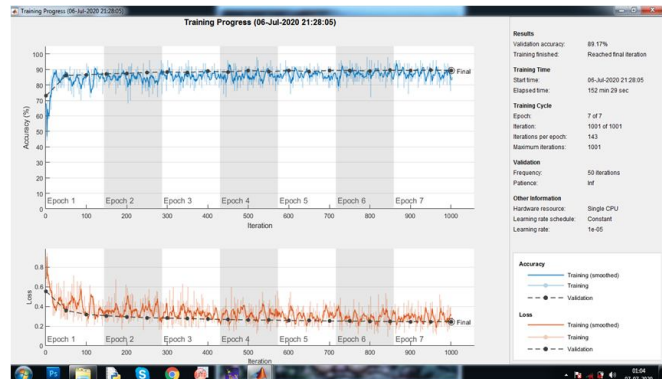


Figure 4: Alexnet Performance Profile

The performance of the pretrained model and developed deep convolution neural network is assessed by various metrics name true positive, true negative false positive, false negative. Using the above said four parameter, further sensitivity, specificity, accuracy, precision, recall and F score is evaluated. Further from receiver operator characteristic graph, area under the curve also evaluated. Table 1 indicate the obtained accuracy of classification in case benign and malignant model for the pre trained model and constructed architecture. The developed DCNN architecture performs well than pre trained model.

Table1: Results of obtained Accuracy of different models

Model	Class	Accuracy
Constructed DCNN model	Benign	98%
Constructed DCNN model	Malignant	96%
Alexnet	Benign	89%
Alexnet	Malignant	86%
VGG16	Benign	90%
VGG16	Malignant	91%
VGG19	Benign	94%
VGG19	Malignant	95%

#### 5. CONCLUSION

Early detection of melanoma reduces the spreading of cancer cells to other lymph nodes and almost eliminates mortality. In the present work, three pretrained models have explored namely alexnet, VGG16, and VGG19, for the binary classification using the idea of transfer learning. Further, the DCNN architecture is constructed with 15 layers, with

optimized hyper parameters gives better results than pretrained models. The obtained accuracy with the constructed architecture is 98 and 96%, respectively, for benign and malignant classes. The achieved accuracy outperforms the existing method in the literature. Further the computation time for the constructed architecture is less.

#### REFERENCES

- Garnavi, R., Aldeen, M., Celebi, M.E., Varigos, G., Finch, S., “Border detection in dermoscopy images using hybrid thresholding on optimized color channels”. *Comput. Med. Imaging Graph*, 2011, 35 (2), 105–115.
- Kapur, J.N., Sahoo, P.K., Wong, A.K.C., “A new method for gray-level picture thresholding using the entropy of the histogram”. *Comput. Vis. Graph. Image Process*, 1985, 29 (1), 273–285.
- S. W. Menzies, C. Ingvar, and W. H. McCarthy, “ A sensitivity and specificity analysis of the surface microscopy features of invasive melanoma”. *Melanoma Research*, 6(1):55–62, February 1996.
- G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino. Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: Comparison of the ABCD rule of dermatoscopy and a new 7-point checklist based on pattern analysis. *Archives of Dermatology*, 134:1563–1570, 1998.
- J. Henning, S. Dusza, S.Wang, A. Marghoob, H. Rabinovitz, D. Polsky, and A. Kopf. The CASH (colour, architecture, symmetry, and homogeneity) algorithm for dermoscopy. *Journal of the American Academy of Dermatology*, 56(1):45–52, 2007.
- M. Hintz-Madsen, L. K. Hansen, J. Larsen, and K. Drzewiecki. A probabilistic neural network framework for detection of malignant melanoma. In *Artificial Neural Networks in Cancer Diagnosis, Prognosis and Patient Management*, pages 141–183, CRC Press, 2001.
- G. A. Hance, S. E. Umbaugh, R. H. Moss, and W. V. Stoecker. Unsupervised colour image segmentation: With application to skin tumor borders. *IEEE Engineering in Medicine and Biology Magazine*, 15:104–111, 1996.
- W. E. Denton, A. W. G Duller, and P. J. Fish. Boundary detection for skin lesions: an edge focusing algorithm. In *5th IEEE International Conference on Image Processing and its Applications*, pages 399–403, 1995.
- B. Erkol, R. Moss, R. Stanley, W. Stoecker, and E. Hvatum. Automatic lesion boundary detection in dermoscopy images using gradient vector flow snakes. *Skin Research and Technology*, 11:17–26, 2005.
- T. Mendonca, A. Marcal, A. Vieira, J. C. Nascimento, M. Silveira, J. S. Marques, and J. Rozeira. Comparison of segmentation methods for automatic diagnosis of dermoscopy images. In *29th IEEE EMBS Annual Int. Conference*, volume 1, pp. 6572–6575, 2007.
- T. Donadey, C. Serruys, and A. Giron. Boundary detection of black skin tumors using an adaptive

- radial-based approach. In SPIE Medical Imaging, volume 3379, 2000.
12. H. Ganster, A. Pinz, R. Rohrer, E. Wildling, M. Binder, and H. Kittler. Automated melanoma recognition. IEEE Transactions on Medical Imaging, 20:233–239, March 2001.
  13. Y. Chang, R. Joe Stanley, R. H. Moss, and W. V. Stoecker. A systematic heuristic approach for feature selection for melanoma discrimination using clinical images. Skin Research and Technology, 11:165–178, 2005.
  14. K. M. Clawson, P. J. Morrow, B. W. Scotney, D. J. McKenna, and O. M. Dolan. Determination of optimal axes for skin lesion asymmetry quantification. In IEEE International Conference on Image Processing (ICIP 2007), volume 2, pages 453-456, 2007.
  15. Sasi Bhanu J, Baswaraj D, Sunitha Devi Bigul, JKR Sastry, “Generating Test cases for Testing Embedded Systems using Combinatorial Techniques and Neural Networks based Learning Model”, International Journal of Emerging Trends in Engineering Research, Volume 7, No. 11, November 2019, pp.417-429.
  16. Anilkumar B, P.Rajesh Kumar, “Tumor Classification using Block wise fine tuning and Transfer learning of Deep Neural Network and KNN classifier on MR Brain Images”, International Journal of Emerging Trends in Engineering Research, Volume 8, No. 2, February 2020, pp.574-583.