IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CLASSIFICATION OF SKIN-LESION BASED ON DEEP LEARNING TECHNIQUES AND IMAGE PROCESSING

[1]P. Pujitha, [2] M. S. N. L. Narasimha Rao, [3] S. Pallavi, [4] K. Sai Venkat, [5] T. Harish, [6] R. Gopi

1Student, Department of Electrical and Electronics Engineering, Sir C.R Reddy College of Engineering

2Assistant Professor, Department of Electrical and Electronics Engineering Sir C.R Reddy College of Engineering

3Student, Department of Electrical and Electronics Engineering, Sir C.R Reddy College of Engineering

4Student, Department of Electrical and Electronics Engineering, Sir C.R Reddy College of Engineering

5Student, Department of Electrical and Electronics Engineering, Sir C.R Reddy College of Engineering

6Student, Department of Electrical and Electronics Engineering, Sir C.R Reddy College of Engineering

ABSTRACT: Skin diseases are mostly caused by fungal infection, bacteria, allergy, viruses etc. The lasers advancement and photonics based medical technology is used in diagnosis of the skin disease quickly and accurately. The medical equipment for such diagnosis is limited and most expensive. Skin cancer's early detection has very much importance to save the life of a victim but it is a challenging task for dermatologists. Deep learning techniques helps in detection of skin disease at an initial stage. The feature extraction plays a key role in classification of skin lesion. The usage of Convolution Neural Networks has reduced the need for human labour, such as manual feature extraction and data reconstruction for classification purpose. In this project we are going to use transfer learning with mobile net (A convolution neural network architecture) architecture for better speed and accuracy.

I. INTRODUCTION:

The skin of the human body acts as the organ and also acts as the cover to the human body. In the integumentary system of the human, the skin is found as the biggest organ. The skin is treated as one of the humans five sense organs which covers and protects the internal systems of a human body. It has three layers: Epidermis, Dermis and the hypo dermis. Epidermis of the skin is the top layer which protects as waterproof shield and plays a major role in defining the colour of the skin. The middle layer of the skin is known as dermis which contains hair follicles, tough connective tissues and sweat glands. The last layer of the skin is hypo dermis, this is the inner layer of fat and other connective tissues which is also known as Subcutaneous tissue. The skin functions in metabolic functions, sensation, protection and thermoregulation. Skin always acts as a protection shield against the infections. It protects the body from pathogens and from more loss of water that causes dehydration by playing a immunity part of the system. It helps to the protection of vitamin B float and synthesis the vitamin D. A variety of factors from sun hydration to repetitive face movement, can cause skin damage. From sun burn to skin cancer, the skin can be damaged to a large extent. Keeping our skin moist helps to strengthen the barrier. The injured skin will self-cure by creating tissue which is not pigmented beside discoloured.

II. STATEMENT OF THE PROBLEM:

Accurate classification of skin lesions is critical for early detection of skin cancer, yet it remains a challenging task due to high visual similarity among lesion types and dependence on expert diagnosis. Conventional diagnostic methods are time-consuming, costly, and not easily accessible in all regions. Existing machine learning approaches often require large datasets, involve high computational complexity, and may not achieve sufficient accuracy. Therefore, there is a need for an efficient and automated system that can reliably classify skin lesions with high accuracy and reduced computational cost using deep learning techniques.

III. PROPOSED METHODOLOGY:

The proposed system aims to automatically classify skin lesions using deep learning techniques. The overall workflow includes dataset preparation, preprocessing, feature extraction, model training, and classification.

The HAM10000 dataset is used in this study, which consists of 10,015 dermoscopic images categorized into seven classes of skin lesions. The dataset provides sufficient diversity for training and evaluating the model.

The input images are preprocessed to ensure consistency and improve model performance. All images are resized to 224×224 pixels to match the input requirements of the model. Pixel values are normalized to enhance training efficiency. Data augmentation techniques such as rotation, flipping, and scaling are applied to increase dataset diversity and reduce overfitting.

A Convolutional Neural Network based on the MobileNet architecture is used for feature extraction and classification. Transfer learning is applied by initializing the model with pre-trained weights and fine-tuning it for the skin lesion classification task. Additional layers such as dense and dropout layers are added to improve classification accuracy and prevent overfitting.

The model is trained using categorical cross-entropy as the loss function and evaluated using accuracy metrics. The dataset is divided into training and validation sets to assess model performance. The trained model is then used to classify input skin lesion images into their respective categories.

IV. IMPLEMENTATION:

The proposed skin lesion classification system is implemented using Python with deep learning libraries such as TensorFlow and Keras. The system follows a structured pipeline including data collection, data preprocessing, model development, training, evaluation, and deployment.

A. DATA COLLECTION:

Data collection is a technique which we get all the data that is required for project. The collected data is stored in a data base which is known as data set. Data set that is stored is used for training and testing the artificial neural networks and algorithms. Data is collected from different locations and sources contains raw and high noise content. So, the data needs to be pre-processed. The data we have collected id from HAMS10000 dataset from Harvad dataverse. HAMS10000 dataset consists of 10015 dermatoscopic images which can serve as a training set for academic machine learning purposes.

Cases include a representative collection of all important diagnostic categories in the realm of pigmented lesions:

1. Actinic keratoses and intraepithelial carcinoma (akiec)
2. Basal cell carcinoma (bcc)
3. Benign keratosis-like lesions melanoma (bkl)
4. Melanocytic nevi (nv)
5. Vascular lesions (vasc)
6. Dermatofibroma(df)
7. Melanoma(mel)

B. Data Preprocessing:

The dermoscopic images are obtained from the HAM10000 dataset. The dataset is preprocessed to ensure consistency and improve model performance. All images are resized to 224×224 pixels to match the input requirement and the other preprocessing techniques which are present in the most required and important library keras.applications. mobilenet.preprocess_input, this library edits the image according to the requirement of the mobilenet architecture, and some other techniques are rotating an image flipping an image etc. And a popular technique known as image data argumentation which is a technique that can be used to artificially expand the size of a training dataset by creating modified versions of the images which are present in the dataset . Image data augmentation is used to expand the training dataset in order to improve the performance and ability of the model.

C. MODEL DEVELOPMENT:

A pre-trained Convolutional Neural Network based on the MobileNet architecture is utilized as the base model for feature extraction. The use of transfer learning allows the model to leverage pre-learned features from large-scale datasets, significantly reducing training time and improving performance.

The top layers of the pre-trained network are removed and replaced with custom layers tailored to the classification task. These include fully connected dense layers followed by dropout layers to prevent overfitting. The final output layer uses a softmax activation function to classify the input images into one of the predefined skin lesion categories.

This architecture effectively balances computational efficiency and classification accuracy, making it suitable for real-time applications.

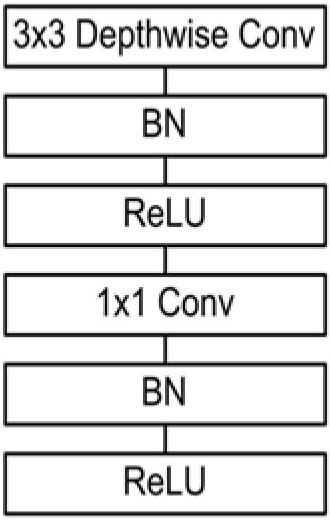


Fig 1: View of Mobile Net

D. Model Training and Evaluation:

The model is trained using categorical cross-entropy as the loss function, which is suitable for multi-class classification problems. An appropriate optimization algorithm is used to update model weights during training. The training process is conducted over multiple epochs to ensure convergence.

Model performance is evaluated using matrix such as accuracy and validation loss. During the training, matrix like accuracy and categorical accuracy are used to monitor the model performance.

THE TRAINING GRAPHS FOR TOP_K_CATEGORICAL ACCURACY METRICS:

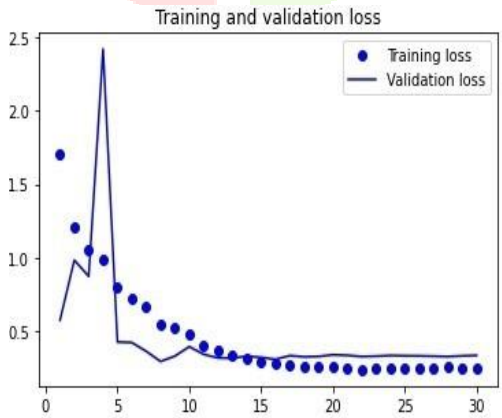


Fig 2: Training and validation loss

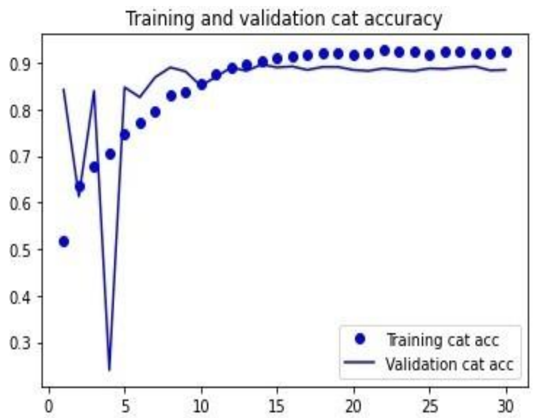


Fig 3: Training and validation cat accuracy

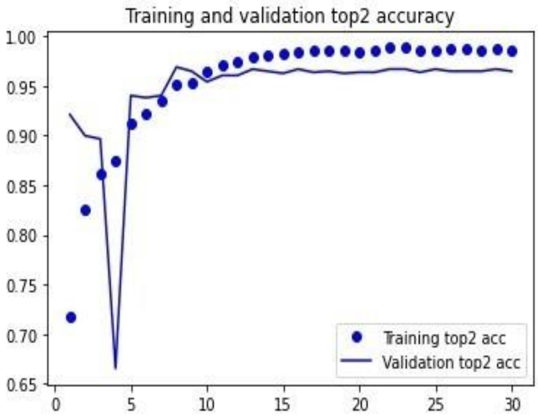


Fig 4: Training and validation Top2 Accuracy

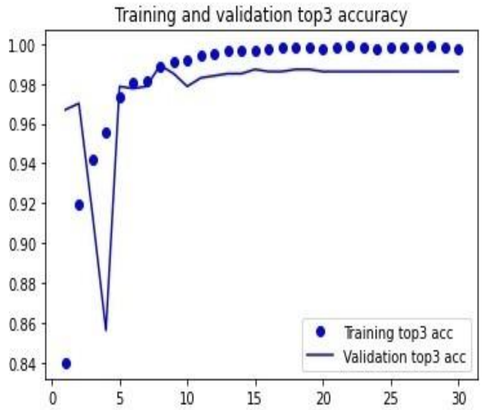


Fig 5: Training and validation Top3 Accuracy

THE TRAINING GRAPHS FOR ACCURACY MATRIX:

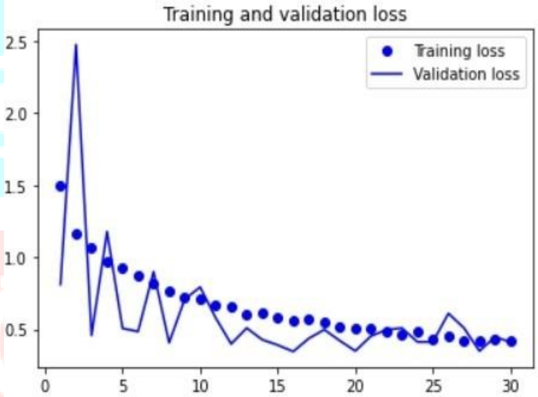


Fig 6: Training and validation loss

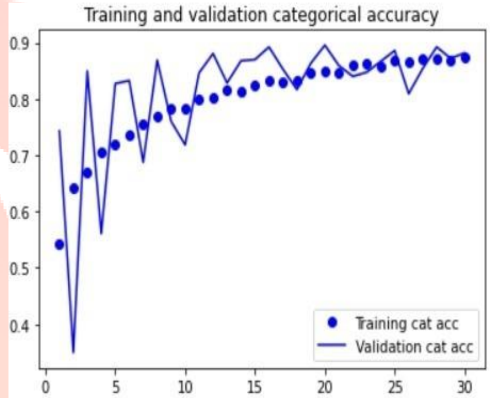


Fig 7: Training and validation categorical accuracy

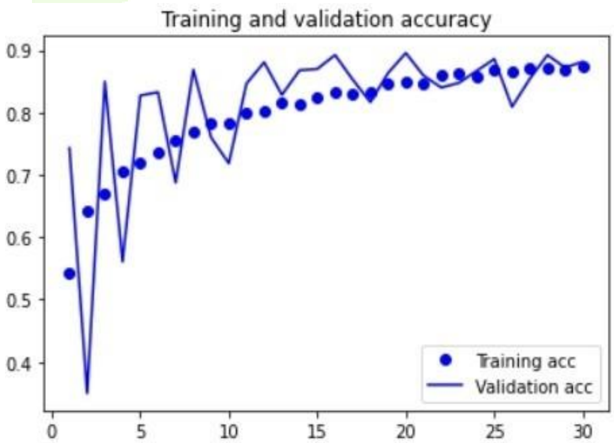


Fig 8: Training and validation accuracy

E. DEPLOYMENT:

To facilitate practical usage, the trained model is deployed using a web-based application developed with the Flask framework. Flask is a web framework that provides some tools, libraries, and technologies that allow us to build a web application. Flask provides us a templates folder, a static folder where we can include our HTML file and CSS file needed by the web application. By using these folders, tools, libraries provided by the Flask framework, we can create a web API. We are using werkzeug.utils.secure_filename module in back-end which helps in returning a secure version of a file when a particular filename is passed to this module. This secure version of the file can then safely be stored on a regular file system and passed to the path using os.path.join(). For achieving maximum portability, this filename will be returned only in an ASCII only string. The module also ensures that the file is not named after one of the particular device files on windows systems.

V. RESULTS:

The performance of the proposed skin lesion classification model is evaluated using the validation dataset derived from the HAM10000 dataset. The model achieves an overall classification accuracy of approximately 85%, demonstrating its effectiveness in classifying different types of skin lesions.

The classification performance is analyzed using a confusion matrix.

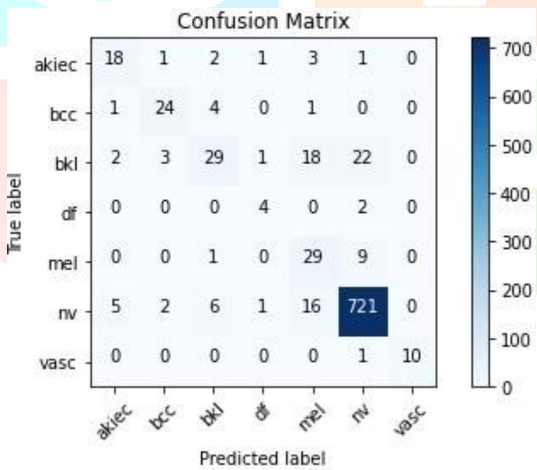


Fig 9: confusion matrix for accuracy matrix

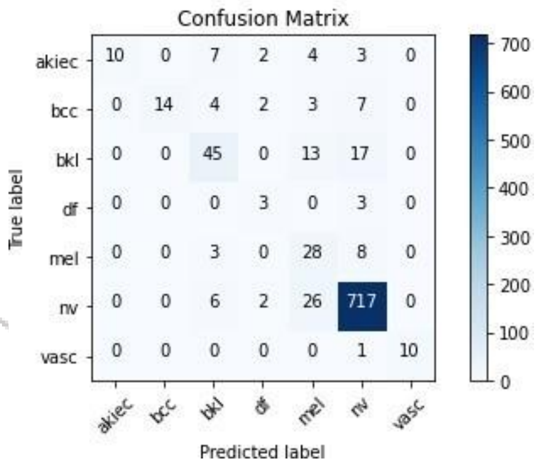
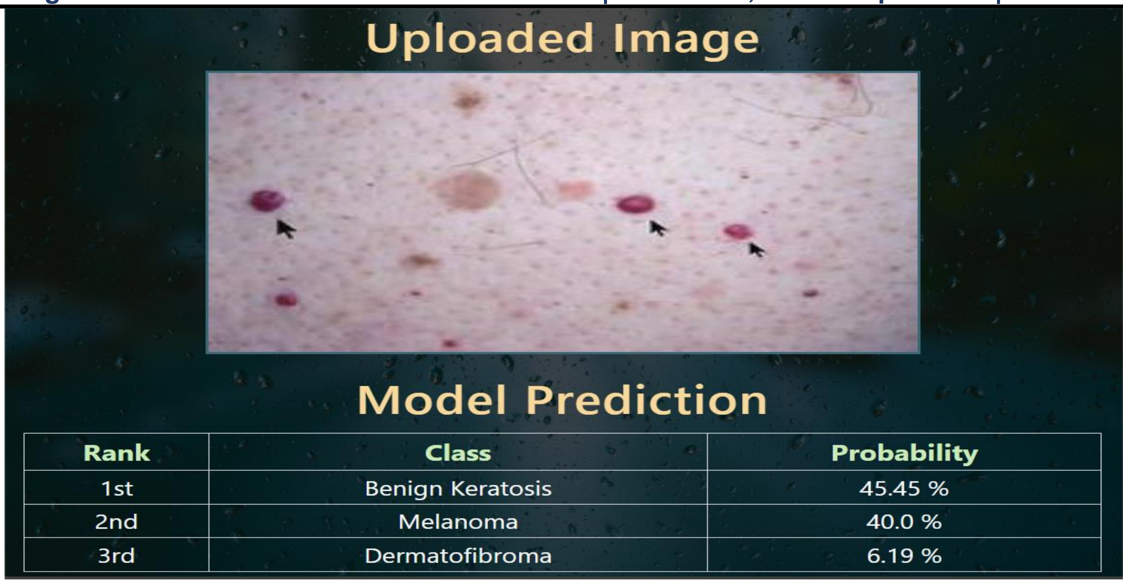


Fig 10: Confusion matrix for top k categorical accuracy

From the confusion matrix, it is observed that the majority of the samples are correctly classified across all seven classes. High prediction accuracy is achieved for well-represented classes, while minor misclassifications occur between visually similar lesion types. These misclassifications are primarily due to the similarity in texture and color features among certain categories.

The model maintains consistent performance across training and validation datasets, indicating good generalization capability. The achieved accuracy of approximately 85% confirms that the proposed approach is effective for automated skin lesion classification.

Overall, the results demonstrate that the use of deep learning with transfer learning provides reliable and efficient classification performance, making the system suitable for practical applications in medical image analysis.



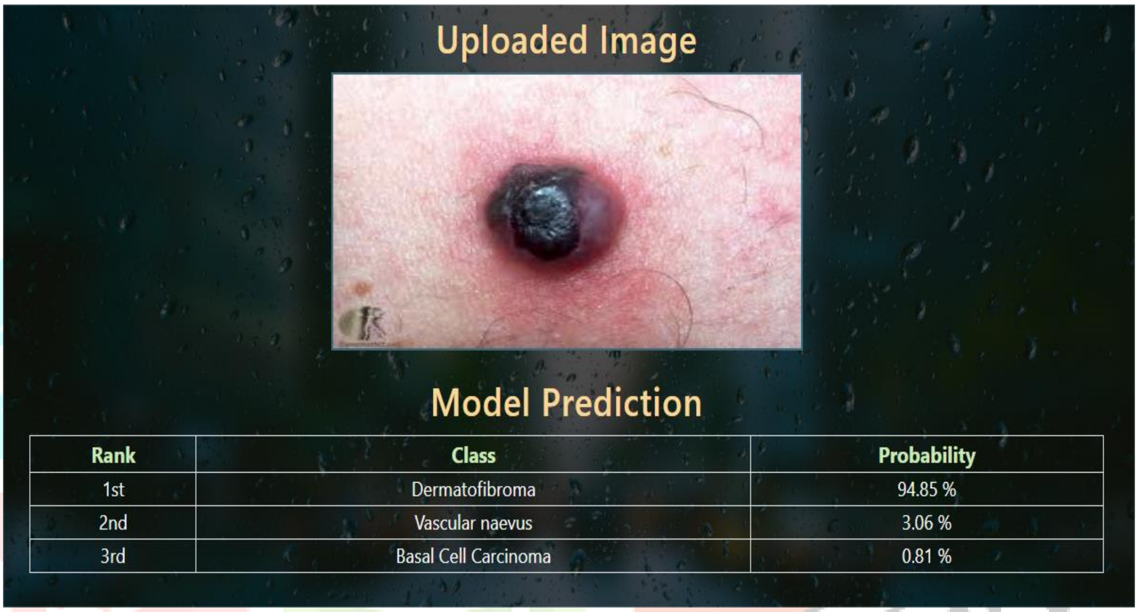


Fig 11: Final outputs after classification

VI. FUTURE SCOPE:

- The proposed skin lesion classification system can be further enhanced in several ways to improve its performance and real-world applicability. Future work may focus on utilizing advanced deep learning architectures such as EfficientNet or ResNet to achieve higher classification accuracy. Increasing the size and diversity of the dataset can also improve model generalization and reduce bias.
- In addition, the system can be extended to support real-time analysis through mobile or web-based applications, enabling easier access for users and healthcare professionals. Integration with clinical data and patient history may further enhance diagnostic reliability.
- Further improvements can include incorporating explainable artificial intelligence techniques to provide interpretability of model predictions, which is crucial for medical applications. The model can also be optimized for faster inference and deployment on low-resource devices, making it suitable for use in remote or resource-limited settings.

VII. CONCLUSION:

Skin lesion (skin cancer) has become a serious problem and there are many methods and techniques that helps in predicting and classifying the skin lesion. The overall scope of project is to develop a method to predict and classify the skin cancer with more accuracy and less complexity when compared to existing techniques and that is achieved**.** The project is also developed in a effective way that ensure better understanding of doctors and patients. In order to achieve this, we used transfer learning with mobilenet model, which is a pre-trained Convolution Neural Networks model, and various python packages. At present, the model is still in development stage and we got 85% accuracy and there is a way and path to be continue to do research in this field of Neural Networks. To conclude this paper we got results in best possible way to predict the skin lesions.

VIII. REFERENCES:

1. Soniya Mane, Dr. Swati Shinde, "A Method for Melanoma Skin Cancer Detection Using Dermoscopy Images",2018.
2. Wilson F. Cueva, F. Muñoz, G. Vásquez., G. Delgado, "Detection of skin cancer Melanoma through Computer Vision",2017.
3. Shalu, Aman Kamboj, "A Colour-Based Approach for Melanoma Skin Cancer Detection",2018.
4. Sertan Kaymak, Parvaneh Esmaili, and Ali Serener, "Deep Learning for Two-Step Classification of Malignant Pigmented Skin Lesions,2018.
5. Kyamelia Roy, Sheli Sinha Chaudhuri, Sanjana Ghosh, Swarna Kamal Dutta "Skin Disease detection based on different Segmentation Techniques,2019.
6. Omar Abuzaghleh, Buket D. Barkana and Miad Faezipour, "Automated Skin Lesion Analysis Based on Colour and Shape Geometry Feature Set for Melanoma Early Detection and Prevention.2014.
7. Reinaldi Nur Pranata, Andrew Brian Osmond, Casi Setianingsih, "Potential Level Detection of Skin Cancer with Expert System Using Forward Chaining and Certainty Factor Method,2018.
8. Giulia Mansutti, Ahmed Toaha Mobashsher, and Amin M. Abbosh, "Millimeter-wave Substrate Integrated Waveguide Probe for Near-Field Skin Cancer Detection,2018.
9. He Huang, Pegah Kharazmi, David I. McLean, Harvey Lui, "Automatic detection translucency using a deep learning method from patches of clinical basal cell carcinoma images,2018.
10. A.Prabhu Chakkaravarthy, Dr. A.Chandrasekar, "An Automatic Segmentation of Skin Lesion from Dermoscopy Images using Watershed Segmentation,2018.
11. Alper ARIK, Mesut GÖLCÜK, "Derin Ögrenme Yöntemleri ile Cilt Kanseri Te¸shis Learning Based Skin Cancer Diagnosis,2017.
12. Ms. Amulya P M, Asst Prof Mr. Jayakumar T V, "A Study on Melanoma Skin Cancer Detection Techniques",2017.

13. Jemal A, Siegel R, Ward E, Murray T, Xu J, Smigal C, et al. Cancer statistics, 2006. CA Cancer J Clin. 2006:56: 106–30. doi: 10.3322/ canjclin.56.2.106.
14. Renuka Devi M N, Ambehwar Kumar , Swetga G, U Shambhavi Chavanand Vaishnavi M Davasam "Cancer Detection Using Image Processing and Machine Learning",2022
15. Celebi ME, Kingravi HA, Iyatomi H, Aslandogan YA, Stoecker WV, Moss RH, et al. Border detection in dermoscopy images using statistical region merging. Skin Res Technol. 2008;14:347-53. doi.org/10.1111/j.1600-0846.2008.00301.x. PubMed PMID: 19159382. PubMed PMCID: 3160669