JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Melanoma Cancer Detection using Artificial Intelligence.

Mandar Patil
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Pune,India
Email: mandar.patil20@pccoepune.org

Harshada Mhaske
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Pune,India
Email: harshada.mhaske@pccoepune.org

Ajaykumar Shendage
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Pune,India
Email: ajaykumar.shendage20@pccoepune.org

Rutuja Tallapalli
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Pune,India
Email: rutuja.tallapalli20@pccoepune.org

Jeevan Thote
Department of Computer Engineering
Pimpri Chinchwad College of Engineering
Pune,India
Email: jeevan.thote20@pccoepune.org

***Abstract* - The skin cancer melanoma is very dangerous. Because it is difficult for dermatologists to analyze each patient sample carefully, a decision support system is required to assess the risk involved with a particular sample. Skin cancer is one of the fatal diseases that people suffer from today. Skin cancer is primarily brought on by the abnormal proliferation of melanocytic cells. Due to genetic factors and UV exposure, melanoma appears on the skin as a brown or black color. With early detection, this skin cancer can be completely cured. The traditional, invasive, and painful method of finding skin cancer is biopsy. This type of laboratory testing takes longer. Computer-aided skin cancer diagnosis has been developed to address the problems. To find skin cancer, the proposed system employs four phases. Dermoscopy is first used to take a picture of the skin. The image must then be pre-processed. Following the pre-processing stage, the lesion is segmented, and then specific features from the segmented lesion are extracted. Finally, various classifiers was given these features to determine whether the given image was of normal skin or one with melanoma disease.**

***Index Terms*- Melanoma Detection, Image Segmentation, Melanoma Skin Cancer, Feature Extraction, CNN, SVM Classifier, Logistic Regression, Decision Trees, Gradient Boosting, Bagging Classifier, XGBoost Classifier.**

I. INTRODUCTION

According to the World Health Organization (WHO), cancer ranks among the top causes of global mortality, claiming approximately 10 million lives annually, with one in six deaths attributed to the disease. The WHO projects a 45% increase in global cancer deaths from 2008 to 2030. Skin cancer, the fifth most prevalent form of cancer, arises from the abnormal proliferation of skin cells, representing one of the most widespread cancer types worldwide. Skin cancer manifests in two primary tumor types: benign tumors, which grow without disrupting surrounding cells functions, and melanoma tumors, characterized by the excessive growth of abnormal cells that disrupt normal cellular operations.

In response to the increasing incidence of skin cancer, particularly melanoma, this project proposes the development of a computer-aided system utilizing transfer learning and data augmentation techniques for effective skin cancer detection. The approach involves a two-tier convolutional neural network (CNN) framework, with the baseline CNN identifying challenging samples using a cross variance score, followed by hair removal techniques and data augmentation for improved accuracy. The subsequent CNN extracts features from the new dataset, incorporating ABDC features based on melanoma rules. These features are then fused and utilized by various machine learning classifiers to predict the presence of melanoma in skin images. This integrated approach offers a cost-effective and accurate solution for early melanoma detection, especially beneficial in environments with limited access to specialized expertise and diagnostic tools, while minimizing exposure to radiation sources.

II. RELATED WORK

G.S.Jayalakshmi and V.Sathiesh Kumar [1], focus on the classification of dermoscopic images to classify the image of skin lesions into benign or malignant. CNN model is developed to classify the images for melanoma identification. For pre-processing of images, a filtering technique is used to avoid the pixel intensity distortion and images are resized. They have used batch normalization to improve the classification accuracy and reduce overfitting. There are 6 layers of convolutional blocks along with batch normalization, which is followed by a fully connected layer for the classification. Using Adam optimizer with a learning rate of 0.0001, the accuracy achieved is 89.30%.

The authors of the paper [2], Rika Rokhana et al. have presented a deep CNN for melanoma detection which is made up of several convolutional layers, max-pooling layer, dropout layer and a fully connected layer. They have used Adam optimizer and ReLU activation function. For the experimental purpose, they trained their CNN by using 10, 25, 50 and 100 epochs and obtained the highest accuracy of 84.76%, sensitivity of 91.97% and specificity of 78.71% on the model which was trained on 10 epochs.

R.S. Shiyam Sundar and M. Vadivel [3], proposes an automated diagnosis system that extracts various features from an image and classifies it into 5 types of melanoma. Multiclass Support Vector Machine, an extension of the Support Vector Machine (SVM) has been used in this system for classification with the help of Gray level co-occurrence matrix features like energy, entropy, contrast, homogeneity, and cross-correlation from an image for classifying the image.
.

The authors of the paper [4], D. Csabai et al. have presented a procedure for the detection of common skin lesions using features observed by the radiologist. They defined a total of 12 features which consist of 5 shaped based features and remaining texture-based features. They have used AdaBoost and Support Vector Machine (SVM) classifiers for classification purposes.

Majtner et al.[5], innovatively applied Convolutional Neural Network (CNN) techniques in their study, focusing on feature extraction and preprocessing for enhanced classification. They utilized the International Skin Imaging Collaboration (ISIC) dataset, consisting of 900 training samples and 379 test samples, categorized into malignant and benign classes. This approach deviates from traditional neural network methods, marking a significant advancement in dermatological image analysis. Preprocessing involved downsampling images and converting them to grayscale. AlexNet was utilized for feature extraction with bounding box and binary masking techniques. Feature dimensionality was reduced using LDA. KNN emerged as the most effective classifier, achieving 86% accuracy and 99.9% specificity, outperforming other classifiers assessed.

Vipin et al.[6] structured their system into segmentation and classification phases, aiming to automate melanoma diagnosis while reducing errors and time. They utilized a curated subset of 7,353 images from the ISIC dataset. Segmentation employed a symmetric U-Net design, comprising contracting/down sampling and expanding/up sampling paths, incorporating convolutional and pooling layers for feature extraction and abstraction. For the classification task, a deep residual network was employed, integrating recurrent and convolutional neural network (CNN) methodologies. Training entailed the utilization of a weighted binary cross-entropy loss function, culminating in a commendable accuracy of 88.7% and a recall rate of 91%.

Nasr-Esfahani et al.[7] developed a Convolutional Neural Network (CNN) framework encompassing preprocessing, feature extraction, and classification. Their study utilized a dataset comprising 170 non-dermoscopic images sourced from the dermatology department's digital collection at the University Medical Centre Groningen (UMCG). To augment training data, an additional 6,120 synthetic and genuine images were incorporated into the dataset. Pre-processing methodologies comprised mask construction, lighting adjustment, and Gaussian filter techniques. The CNN architecture consisted of an input layer, two stages alternating between max-pooling and convolution, culminating in a fully connected layer. Classification was performed to discern between nevus and melanoma classes. The resultant model demonstrated promising performance metrics, yielding an accuracy of 81% and a specificity of 80%, as recommended.

Arora et al.[8] developed a computer-aided detection and diagnostic method for classifying lesions into benign and malignant categories by employing precise feature extraction techniques. The study made use of one hundred images from the PH2 collection. The bag-of-features (BoF) method was utilised to extract features using the speeded-up robust features (SURF) approach as a local feature descriptor. Several models, including linear SVM, Fine tree, linear discriminant, and quadratic SVM, were utilised for classification. With an accuracy of 85.7%, sensitivity of 100%, and specificity of 60%, the recommended Quadratic SVM fared better than the other classifiers. Performance metrics showed a 3% improvement over the most sophisticated methods.

III. METHODOLOGY

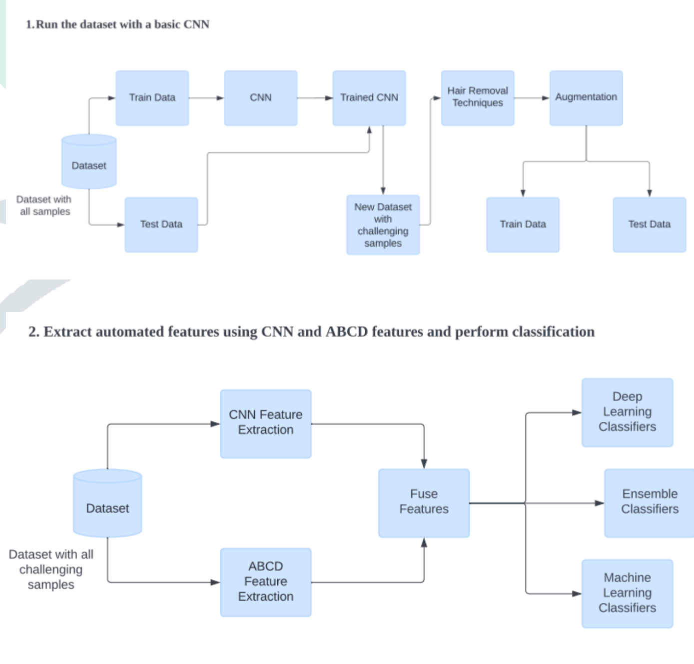


Fig. 1: System Architecture

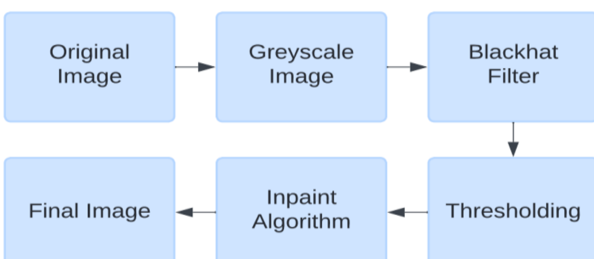
A. Dataset

The dataset used in this project is derived from the ISIC-2018 dataset, a benchmark dataset for melanoma detection. It consists of 1800 benign samples and 1500 malignant samples, totaling 3300 samples. The dataset is publicly available on Kaggle.This dataset is crucial for training and evaluating the proposed two-tier convolutional neural network (CNN) framework for malignant melanoma prediction. It serves as the foundation for developing and testing the computer-aided system aimed at detecting skin cancer, particularly melanoma, utilizing captured images. The dataset's diversity and size enable the implementation of transfer learning and data augmentation techniques, ensuring robustness and accuracy in skin cancer detection.

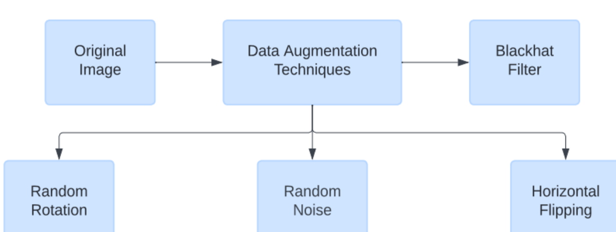
B. Pre-Processing

The preprocessing steps in this project play a crucial role in enhancing the quality of the images and preparing them for accurate analysis by the convolutional neural network (CNN). Below is an overview of the preprocessing steps involved:

1. Hair Removal: Hair on the skin lesions can interfere with the classification process. To address this, the images undergo hair removal techniques. This involves converting the images to grayscale and applying a Blackhat filter to enhance dark regions (hair) against a bright background (the lesion itself). Thresholding techniques are then used to isolate the hair, and an inpainting algorithm is applied to remove the masked hair regions from the original image.



2. Data Augmentation: Data augmentation techniques are applied to increase the diversity of the dataset and improve the generalization of the model. Techniques such as random rotation, adding noise, and horizontal flipping are commonly used. Random rotation helps the model become invariant to rotation, while adding noise introduces variations that mimic real-world conditions. Horizontal flipping increases the number of samples without changing the semantic content, thus improving the model's robustness.



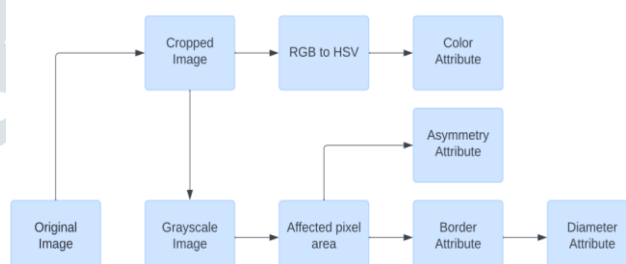
These preprocessing steps ensure that the images are appropriately cleaned and augmented, thereby enhancing the performance of the CNN in accurately detecting melanoma. By removing noise and irrelevant information and increasing the diversity of the dataset, the model becomes more effective in distinguishing between benign and malignant lesions, ultimately improving the accuracy of skin cancer detection.

C. Feature Extraction

Feature extraction is a critical step in the process of melanoma detection using convolutional neural networks (CNNs). In this project, feature extraction is performed at two levels: CNN feature extraction and ABCD feature extraction.

1. CNN Feature Extraction: After preprocessing the images, they are passed through a CNN to extract high-level features. These features capture the abstract representations of the images learned by the CNN during training. Typically, the CNN consists of several convolutional and pooling layers followed by fully connected layers. Features are extracted from the output of these layers, usually just before the final classification layer. In this project, a second-tier CNN is utilized to extract features specifically from challenging samples identified in the baseline CNN. The features extracted from the Dense layers of the CNN are stored separately for further analysis and classification.

2. ABCD Feature Extraction: Apart from CNN features, additional features are extracted based on the ABCD rule used in dermatology for melanoma diagnosis. The ABCD rule stands for Asymmetry, Border irregularity, Color variation, and Diameter. These features are derived from the characteristics of the skin lesions, such as the asymmetry of shape, irregularity of borders, variation in color, and diameter of the lesion. These features are computed from the preprocessed images and are fused with the CNN features for subsequent classification.



By combining CNN features, which capture deep representations of the images, with ABCD features, which encapsulate specific characteristics relevant to melanoma diagnosis, the model becomes more robust and capable of accurately distinguishing between benign and malignant lesions. This feature fusion approach enhances the overall performance of the melanoma detection system, making it more effective in real-world applications.

D. Classification

In the feature fusion and final classification phase of this project, features extracted from both the CNN and the ABCD parameters (asymmetry, border, color, and diameter) are amalgamated and stored for each sample in the new dataset. These combined features form a comprehensive representation of the skin lesions, capturing both visual characteristics and clinically relevant information associated with melanoma. Subsequently, this fused feature set is utilized as input for seven distinct machine learning classifiers, categorized into three classes: deep learning classifiers (Multi-layer perceptron), ensemble learning classifiers (Gradient boosting classifier, XG Boost classifier, and Bagging classifier), and traditional machine learning classifiers (decision trees, support vector machines, and logistic regression). By leveraging this diverse array of classifiers, the system ensures robustness and accuracy in melanoma prediction, facilitating early detection and intervention in skin cancer cases.

IV. RESULTS

In this study, various algorithms were evaluated for their effectiveness in predicting melanoma, a type of skin cancer, based on image features extracted from skin lesion images. The performance of algorithms including Convolutional Neural Networks (CNNs), Support Vector Machines (SVM), Decision Trees, Gradient Boosting, Bagging Classifier, XGBoost Classifier and Logistic Regression was compared using datasets such as the ISIC-2018 dataset.

Algorithm	Accuracy
SVM	90.48
Random Forest	91.01
Logistic Regression	89.95
Decision Trees	87.30
Bagging Classifier	92.59
XGBoost Classifier	91.53

V. CONCLUSION

In conclusion, this project presents a comprehensive approach to melanoma detection leveraging convolutional neural networks (CNNs) and machine learning classifiers. By integrating features extracted from CNNs and clinically relevant parameters based on the ABCD rule, the system achieves a holistic representation of skin lesions, enhancing accuracy in melanoma prediction. Through feature fusion and subsequent classification using a diverse set of classifiers including deep learning, ensemble learning, and traditional machine learning models, the system demonstrates robustness and reliability in identifying malignant melanoma. This integrated approach holds promise for early detection of skin cancer, offering a cost-effective and accurate solution, particularly in environments where specialized expertise and diagnostic tools may be limited. Ultimately, the system contributes to addressing the urgent need for effective melanoma detection methods, facilitating timely intervention and improved patient outcomes in the fight against skin cancer.

VI. REFERENCES

[1] G. S. Jayalakshmi and V. S. Kumar, "Performance analysis of Convolutional Neural Network (CNN) based Cancerous Skin Lesion Detection System," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), Chennai, India, 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862143.

[2] R. Rokhana, W. Herulambang and R. Indraswari, "Deep Convolutional Neural Network for Melanoma Image Classification," 2020 International Electronics Symposium (IES), Surabaya, Indonesia, 2020, pp. 481-486, doi: 10.1109/IES50839.2020.9231676.

[3] R. S. S. Sundar and M. Vadivel, "Performance analysis of melanoma early detection using skin lession classification system," 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), Nagercoil, India, 2016, pp. 1-5, doi: 10.1109/ICCPCT.2016.7530182.

[4] D. Csabai, K. Szalai and M. Gyöngy, "Automated classification of common skin lesions using bioinspired features," 2016 IEEE International Ultrasonics Symposium (IUS), Tours, France, 2016, pp. 1-4, doi: 10.1109/ULTSYM.2016.7728752.

[5] Majtner T, Yildirim-Yayilgan S, Hardeberg JY., "Optimised deep learning features for improved melanoma detection", May 2019 ,Multimedia Tools and Applications 78(9):11883-11903 DOI:10.1007/s11042-018-6734-6

[6] VipinV, NathMK, SreejithV, etal., "Detection of Melanoma using Deep Learning Techniques: A Review", June 2021,DOI:10.1109/ICCISc52257.2021.9484861,Conference : 2021 International Conference on Communication, Control and Information Sciences (ICCISc)

[7] E Nasr-Esfahani, S Samavi, N Karimi, S M R Soroushmehr, ., "Melanoma detection by analysis of clinical images using convolutional neural network", Annu Int Conf IEEE Eng Med Biol Soc2016 Aug:2016:1373-1376, doi: 10.1109/EMBC.2016.7590963.

[8] AroraG,DubeyAK,JafferyZA,etal., ""Melanoma detection by analysis of clinical images using convolutional neural network", June 2022 Neural Computing and Applications 34(2) DOI:10.1007/s00521-020-05212-y

[9] Scott E.Umbaugh, Randy H.Moss, and William V.Stoecker,"Applying Artificial Intelligence to the Identification of Variegated Coloring in Skin Tumors", IEEE Engineering in Medicine And Biology, Magazine, Vol.10, No.4, pp 57-62, 1991.

[10] O. Abuzaghleh, B. D. Barkana, and M. Faezipour, "SKINcure: A real time image analysis system to aid in the malignant melanoma prevention and early detection"in Proc. IEEE Southwest Symp. Image Anal. Interpretation (SSIAI),Apr. 2014, pp.85_88.

[11] Hadi, S., Tumbelaka, B. Y., Irawan, B., and Rosadi, R., "Implementing DEWA Framework for Early Diagnosis of Melanoma" International Conference on Computer Science and Computational Intelligence (ICCSCI 2015). Procedia Computer Science 59:410–418, 2015

[12] Mohd Afizi, Mohd Shukran, Nor Suraya Mariam Ahmad, Suzaimah Ramli, Farhana Rahmat, "Melanoma Cancer Diagnosis Device Using Image Processing Techniques", IJRTE, ISSN: 2277-3878, Volume-7, Issue-5S4, February 2019.

[13] Yang, J., Xie, F., Fan, H., Jiang, Z., & Liu, J. (2018). "Classification for dermoscopy images using convolutional neural networks based on region average pooling". IEEE Access, 6, 65130–65138.

[14] Vedanti Chintawar, Jignyasa Sanghavi, "A Review on Computer Aided Melanoma Skin Cancer Detection using Image Processing", EasyChair Print, No. 584, October 24,2018.

[15] Nidhal K. EL Abbadi, Zahraa Faisal, "Detection and Analysis of Skin Cancer from Skin Lesions", International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 19 (2017) pp. 9046-9052.