IJNRD.ORG ISSN : 2456-4184

INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG

An International Open Access, Peer-reviewed, Refereed Journal

SKIN CANCER DETECTION USING CNN

Yasaswi[1], K.Pavani[2], B.Tanuja Sri[3],P.Kushal Sai[4], B.Ajit[5], L.V.Sri Vandana[6]

1Assistant Professor, Department of Computer Science and Engineering, Anil Neerukonda Institute of Technology and Sciences, Visakhapatnam, India

2,3,4,5Department of Computer Science and Engineering, Anil Neerukonda Institute of Technology and Sciences, Visakhapatnam, India.

ABSTRACT

In comparison to the identification of skin lesions, skin cancer detection is delicate due to remains, minimal disparity, and comparable visuals as intelligencers, scars, etc. Skin cancer is frequently detected in its early stages because it spreads slowly to other body parts and is therefore easier to cure. There are increasingly more instances of terrible carcinoma, the skin cancer that is most fatal. Skin cancer may be challenging to distinguish from skin lesions because of leftovers, little disparity, and similar visualization to an operation, scar, etc. Hence Skin lesions are automatically detected using methods for lesion detection that take into account performance, efficacy, and delicacy requirements. The proposed approach uses point birth using the ABCD principle, GLCM, and overeater point birth as the goal in the early diagnosis of skin lesions. By removing residues, skin colour, hair, and other impurities, pre-processing is used in the proposed study to enhance the skin lesion's appearance and clarity. Segmentation was done using Geodesic Active Contour (GAC), a tool that splits the lesion apart into sections and is also efficient for point birth. The harmony, border, colour, and perimeter features were rated using the ABCD scale. The textural elements were rooted by using Overeater and GLCM. In identifying 7 different types of skin cancer, classifiers use a variety of machine learning methods, including the classifiers SVM, CNN, KNN, and Naive Bayes. For this design, a total of 10015 pictures of malignant skin lesions, benign skin lesions and other types are downloaded from the HAM10000 dataset. Effective and precise bracketing is achieved. They include ABCD, overeater, GLCM, SVM, CNN, KNN, and naive Bayes.

1. INTRODUCTION

The largest organ in a mortal is the skin, which serves as the body's outermost layer. Touch, cold, and heat are all sensations we may perceive thanks to our skin. It also regulates our body temperature and keeps bacteria and other contaminants out of our bodies. Skin lesions are aberrant skin sections that differ from other skin pathways. Infections that happen in or generally initially lead to skin lesions. Two categories of skin lesions can be distinguished generally: Those that are primary and secondary (those that are inherited or change with time) (which are brought on by improper handling of the primary skin lesion). Both of these categories can contribute to the more than three million skin cancer diagnoses that occur annually in the US. India has around 5000 new cases of skin cancer annually, and over 4000 individuals pass away as a result of the illness. Skin excrescences fall into three main categories: rigid cell melanoma

(BCC), scaled cell melanoma (SCC), and melanoma. Excrescences are considered cancer if they transform into lethal skin cancer that quickly spreads to further skin-contact sites. Due to benign excrescence's ability to expand without spiking, it is not a particularly harmful variety. Hence, self-diagnosis of skin cancer is unsuitable since it is impossible to accurately distinguish features when the skin lesion is seen with unaided eyes, which can lead to ineffective treatment and ultimately death. Increased survival rates for skin cancer can result from early, accurate identification. With preprocessing techniques, all of this can be minimized. The segmented skin lesion image from the pre-processing stage shows the precise location of the skin lesion. One of the segmentation methods with similarities is the Sea algorithm. Examples include simple global thresholding, region-based segmentation, the snake system, the Otsu system, active silhouettes, and geodesic active silhouettes. We make use of the geodesic active figure for segmentation. The segmentation skin lesion image is currently utilized for point birth, and there are numerous styles for rooted features, including the CASH rule, ABCD rule, ABCDE rule, GLCM, overeater, etc. Asymmetry, color, borders, and peripheral attributes are not considered when applying the ABCD rule, which establishes how points are granted.

2. LITERATURE SURVEY:

To segment and analyse skin lesions automatically, this five research employ a variety of methods and algorithms.

According to the first paper by Jaisakthi et al., "Automated skin lesion segmentation of dermoscopic photos using GrabCut and k-means algorithms," there is a method for automating the segmentation of skin lesions using the GrabCut and k-means algorithms. With the use of a collection of dermoscopic images, the authors test their approaches, and the results they presentare positive.

The authors of "Segmenting skin lesions with partial-differential-equations-based algorithms," Chung and Sapiro, offer a method for segmenting skin lesions in their second paper. The authors evaluate their technique and demonstrate its accuracy using a small sample of skin lesion photos.

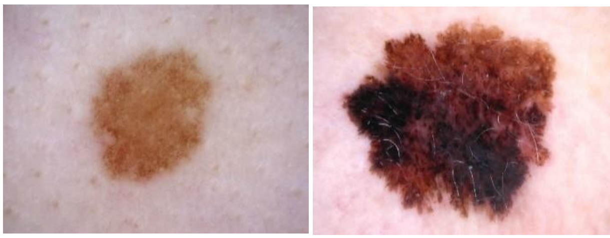
The third work by Hemalatha et al., titled "Active Contour-Based Segmentation Methods for Medical Image Analysis," examines active contour-based segmentation approaches for this field of study. Segmenting skin lesions will help with medical image analysis. The authors give a summary of various active contour-based approaches, including their benefits and drawbacks.

For their fourth publication, Salih and Al-Raheym compared several segmentation algorithms using a skin lesion image. It investigates various skin lesion segmentation algorithms and assesses each one's performance based on a variety of criteria. The GrabCut approach allegedly outperforms earlier algorithms in terms of accuracy and processing speed.

Li and Shen present a deep learning-based approach to skin lesion analysis for melanoma identification in their fifth article, "Skin Lesion Analysis Toward Melanoma Diagnosis Using Deep Learning Network." The authors evaluate their method and claim successful results usinga substantial collection of photos showing skin lesions.

3. DATASET USED

Dermatoscopic images of skin lesions can be seen by anyone by visiting the collection HAM10000. It features 10015 high-quality images of pigmented skin lesions, such as benign and malignant melanocytic lesions, seborrheic keratoses, dermatofibromas, and vascular lesions. Using a high-resolution scanner, the images were digitally transformed after being gathered from several clinical facilities. The patient's age, gender, kind of skin cancer, and the lesion's unique identification number are all included in the metadata that is attached to each image.



(a) (b)

Fig 1 An example of Skin lesion images (a) Benign image

(b) Melanoma image

4. PROPOSED MODEL

The method that is being proposed for identifying skin lesions entails multiple processes, including data collection, pre-processing, segmentation, feature extraction, and classification. The process of gathering skin lesion photos for the purpose of testing and refining the suggested algorithm is known as data acquisition. Pre-processing improves the clarity and quality of images of skin lesions by removing artefacts and alterations to features like skin tone and hair colour. an important step in the detection of skin lesions is segmentation, which separates the lesion's territory from the skin around it. The lesion portion is successfully divided and helped with feature extraction using the Geodesic Active Contour (GAC) segmentation technique in the proposed methodology. The procedure's crucial phase is feature extraction, in which the properties are used to locate skin lesions. Verify the skin lesions to check which type of cancer it is. The symmetry, border, colour, and diameter attributes are extracted using the ABCD scoring system. The Gray-Level Founder Matrix (GLCM) and HOG methods are also a method for obtaining textural information. The acquired characteristics are then used to train machine learning classifiers such as the Naive Bayes classifier, Convolutional Neural Network(CNN), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) to categorize the skin lesions.

4.1 CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN stands for convolutional Neural Network. It is a type of deep learning algorithm used primarily for image and video recognition and processing. CNNs are inspired by the structure and function of the human visual cortex. They are designed to automatically and adaptively learnspatial hierarchies of features from raw data.

The fundamental building blocks of a CNN, convolutional layers, are applied as filters to the input image to extract characteristics like edges, textures, and forms. The most crucial information is then preserved and the spatial dimensions are decreased by down-sampling the feature maps using pooling layers. These layers then communicate the results of the convolutional layers.

One or more fully connected layers are then used to predict or categorize the image after relaying the pooling layers' output via them. The output layer uses the softmax activation function to construct the probability distribution over the seven potential classes. The model is trained with the Adam optimizer and sparse categorical cross-entropy loss function, using accuracy as the evaluation metric. Early pause, which prevents overfitting, preserves the best model.

CNNs are powerful machine learning algorithms that can learn and extract complex features from input data. They are trained using supervised learning and can be used for a variety of computer vision tasks, such as image classification, object detection, and segmentation techniques.

CNNs have achieved state-of-the-art results in a variety of computer vision tasks, such as object detection, image segmentation, and facial recognition. They have also been applied to other domains, such as natural language processing and speech recognition.

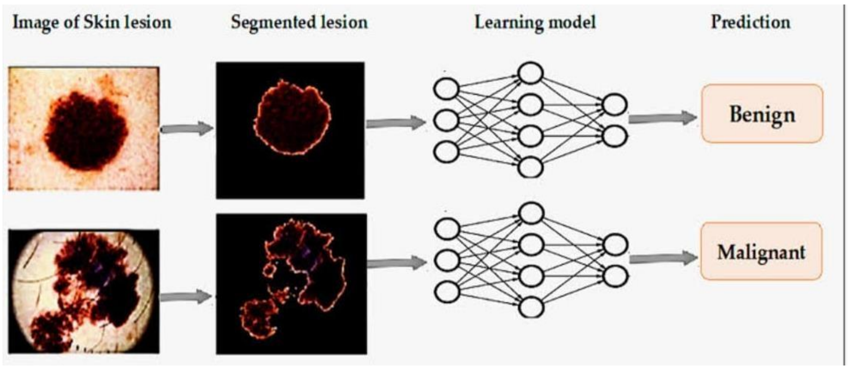


Fig 2 image showing the process of prediction of cancer using CNN

4.2 SYSTEM ARCHITECTURE

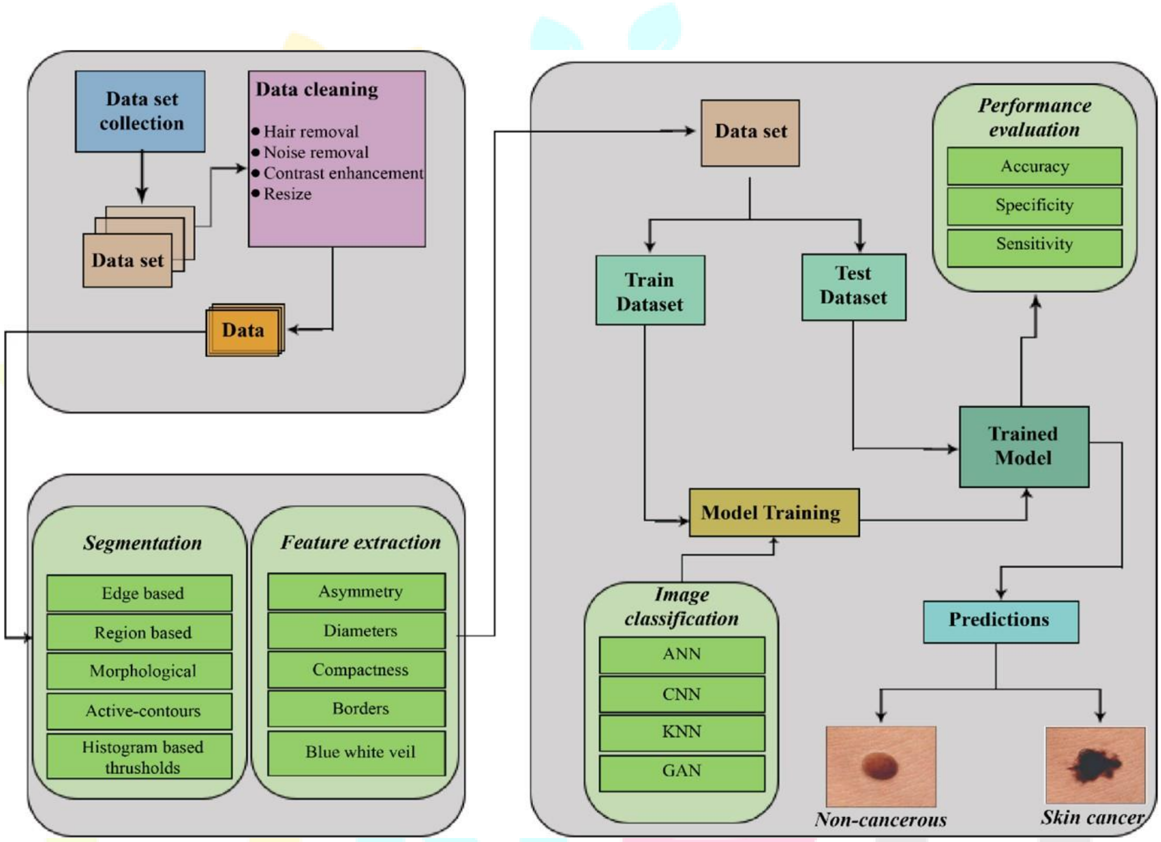


Fig 3 System architecture of skin cancer detection

4.2.1 DATA PREPROCESSING

To be effective, modelling and analysis require that data be kept clean and accurate. The two sets must be joined in order to organize the data effectively, and a special column must be added to indicate the image's source. The dataset will become more manageable and simpler if extraneous columns are removed.

The original dimensions and data for each image must be kept in order to preserve the integrity of each one and ensure that the analysis is founded on reliable data. The inclusion of DICOM files in the collection is a helpful first step towards attaining this goal because they are a typical medical imaging format for collecting images and metadata.

In order to safeguard the validity and integrity of the data, the cleaning and organization of medical picture collections calls for meticulous attention to detail and a methodical approach. The dataset collected can be utilized to create efficient models that can help in the diagnosis and treatment of diseases by using the right formats, like DICOM files, and by following best practices.

4.2.2 SEGMENTATION

The last step is to segment the previously edited images. The skin lesions' exact location is determined by using the segmentation approach. The data in this investigation was segmented using Geodesic Active Contours (GAC). The highest variations in the entire skin lesion are frequently detected by GAC close to the skin lesions edges. After pre-processing the skin image with Otsu thresholding, GAC is applied to the binarized image. This example displays the segmented final image. The segmentation procedure in the suggested technique was assessed using the Jaccard Index (JA) and Dice Index (DI), as well as the JA framework that was reached,and the outcomes were compared to prior work.

4.2.3 FEATURE EXTRACTION

The fourth phase's beginning will be the segmented skin lesion. In order to extract detailed information about the skin lesion, such as its border, colour, width, symmetry, or textural composition, this returned data was employed in the feature extraction process. Skin cancer is easy to spot. We employed the feature extraction techniques ABCD, GLCM, and HOG.

ABCD (ASYMMETRY, BORDER, COLOR, DIAMETER)

Skin lesions are classified according to the ABCD rule based on their diameter, color, border, and symmetry.

GLCM (GRAY LEVEL CO-OCCURRENCE MATRIX)

In order to collect an item's dispersed intensity, textural analysis is done using GLCM. A neighbour pixel and a reference pixel are the two pixels that GLCM takes into account.

HOG (HISTOGRAM OF ORIENTED GRADIENTS)

The HOG method is used to extract form and edge data. To determine the edge intensity, a lesion's orientation histogram is generated. This calls for the use of the unit cell and block. The gradient of the cells is depicted below.

4.2.4 CLASSIFICATION

It is possible to distinguish between cancerous and non-cancerous skin lesions using a variety of models. The majority of the time, machine learning methods are used to categorize various lesion types. Naive Bayes, KNN, SVM, CNN and neural networks are some of the most well- liked classifiers. CNN, the classifiers employed in this study, automatically and adaptively learn spatial hierarchies of features from raw input data such as images. CNNs are powerful machine learning algorithms that can learn and extract complex features from input data. The training process of a CNN involves optimizing the parameters of the network to minimize a loss function that measures the difference between the predicted outputs and the actual labels of the training data. Once the CNN is trained on skin cancer images like benign and malignant, it can be used to make predictions on new, unseen data.

4.3 MODULE DIVISION

This module handles data loading, pre-processing, and segmentation. Data loading from a CSV file and photo resizing, scaling, and grayscale conversion are all included in the preparation. After that, Otsu and GAC algorithms are used for segmentation.

This module, which makes use of the Keras API, describes the CNN model's architecture.

The role of the module for model training is to validate and train models. It is in charge of both training and validating the defined model using the training set. Overfitting and early termination callbacks are prevented, and the model's optimal weights are restored.

The Keras Image Data Generator class is used by this module to create batches of enhanced images for training and validation.

This module connects your trained model to the web application created in Flask. It manages user image uploads, prepares the image for prediction, sends it to the trained model for prediction, and then returns the anticipated diagnosis along with a bar chart showing the predicted likelihood for each type of skin cancer detected in the images and datasets.

4.4 INPUT

The input to our model is a skin lesion image of size 224x224x3 pixels in JPEG format.

4.5 EXPECTED OUTPUT

The goal of our model is to produce a probability distribution over the seven possible types of skin cancer that the input image might correspond to.

5. CONCLUSION

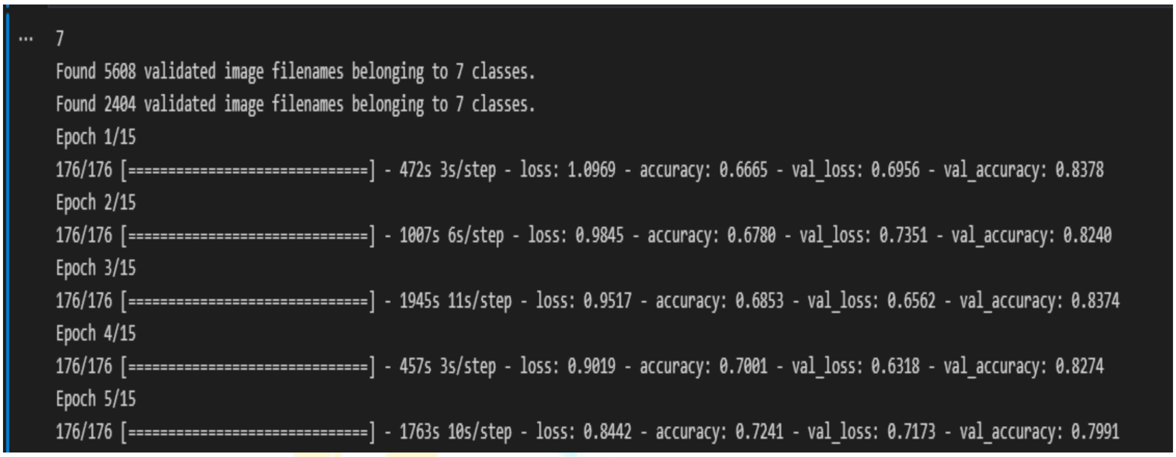
Our proposed CNN architecture can accurately classify photos of skin cancer thanks to its testset's 65% overall accuracy rate. The findings imply that dermatologists may find our method helpful in identifying the existence of skin cancer.

6. FUTURE SCOPE

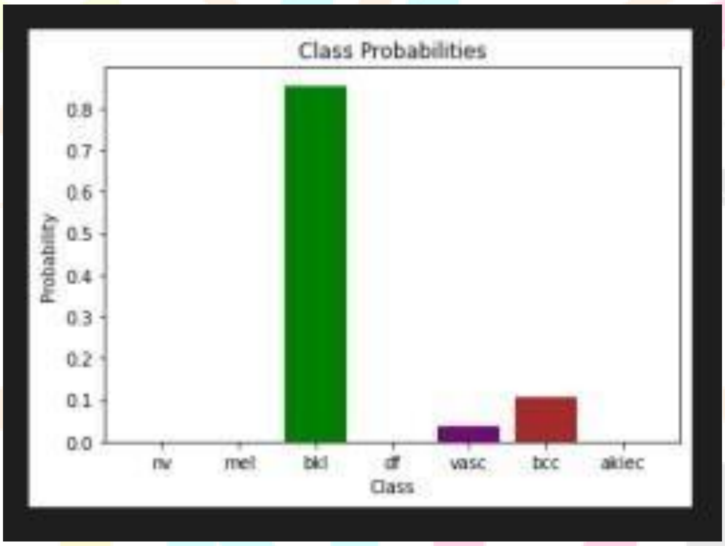
We intend to research the use of transfer learning techniques in the future in order to incorporate more datasets into our model and improve its accuracy and generalizability. Also, we plan to develop a smartphone application that would make it simpler for people to access our tool for categorizing skin cancer.

7. RESULT

7.1 TRAINING


```
... 7
Found 5608 validated image filenames belonging to 7 classes.
Found 2404 validated image filenames belonging to 7 classes.
Epoch 1/15
176/176 [==============================] - 472s 3s/step - loss: 1.0969 - accuracy: 0.6665 - val_loss: 0.6956 - val_accuracy: 0.8378
Epoch 2/15
176/176 [==============================] - 1007s 6s/step - loss: 0.9845 - accuracy: 0.6780 - val_loss: 0.7351 - val_accuracy: 0.8240
Epoch 3/15
176/176 [==============================] - 1945s 11s/step - loss: 0.9517 - accuracy: 0.6853 - val_loss: 0.6562 - val_accuracy: 0.8374
Epoch 4/15
176/176 [==============================] - 457s 3s/step - loss: 0.9019 - accuracy: 0.7001 - val_loss: 0.6318 - val_accuracy: 0.8274
Epoch 5/15
176/176 [==============================] - 1763s 10s/step - loss: 0.8442 - accuracy: 0.7241 - val_loss: 0.7173 - val_accuracy: 0.7991
```

7.2 TESTING



8. REFERENCES

1. Chandrabose Aravindan, Seetharani Murugaiyan, Jaisakthi, and Palaniappan Mirunalini. Dermoscopic picture segmentation of skin lesions automatically utilising the GrabCut and k-means algorithms. 1088–1095 in IET Computer Vision 12, no. 8 (2018).

2. Chung, Do Hyun, and Guillermo Sapiro are number two. "Using partial differential equations-based image processing techniques to segment skin lesions". IEEE Transactions on Medical Imaging, Volume 19, No 7, 2000, Pages 763–767.

3. Chandrasekaran R, Hema Latha RJ, Thamizhvanan TR, Dhivya AJ, Joseph JE, Babu

A. Active contour-based segmentation methods for analysing medical images.2018 July 4:17. Medical and biological image analysis.

4. SALIH, SHAKAR H., and S. H E R E N AL-RAHEYM. Skin lesion image comparison using various

segmentation algorithms. 96, no. 18 of Journal of Theoretical and Applied Information Technology (2018).
5. "Skin cancer lesion analysis towards melanoma detection using a deep learning network," by Yuexiang Li and Linlin Shen. (2018) Sensors 18, no. 2: 556.
6. In 2016, Kasmi, R., and Mokrani, K. Use of the automatic ABCD rule for the classification of benign and malignant skin lesions. Pages. 448–455 in IET Image Processing, 10(6).
7. "Skin cancer detection using glcm matrix analysis and back propagation neural network classifier," by Rashi Goel and Saranjeet Singh. 112, No. 9 of the International Journal of Computer Applications (2015).