

Detection of Skin Cancer using Convolutional Neural Network

Vishwajeet Jadhav¹, Shivani Mane², Pranay Allepally³, Neha Sonawane⁴, Prof. Santosh Kale⁵

Students, Department of Computer Engineering^{1,2,3,4}

Guide, Department of Computer Engineering⁵

N.B.N Sinhgad School of Engineering, Pune, Maharashtra, India

Abstract: *Humanity faces a serious threat from skin cancer. Because of Melanoma skin cancer's quick development rate, high treatment costs, and high mortality rate, the importance of early skin cancer diagnosis has increased. Cancer cells are carefully found, and in most cases, treatment takes time. Using image processing and machine learning, this paper suggested an artificial skin cancer diagnosis system. After segmenting the dermoscopic pictures using the feature extraction approach, the characteristics of the damaged skin cells are retrieved. The retrieved features are stratified using a convolutional neural network classifier based on deep learning. After using the publicly accessible data set, an accuracy of 89.5 percent and a training accuracy of 93.7 percent were reached.*

Keywords: Melanoma, Feature Extraction, Machine Learning, Convolution Neural Network, Information Search and Retrieval

I. INTRODUCTION

One of the most common malignancies in the world is skin cancer. Skin cancer comes in a variety of forms. The most common types of skin cancer are basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) [1]. BCC and SCC begin in the epidermis, or outer layer of the skin, and because they are often brought on by sun exposure, they commonly appear on the face, ears, neck, head, arms, and hands. SCC can spread to neighbouring organs or lymph nodes, although BCC seldom spreads to other parts of the body. BCC and SCC are commonly referred to as keratinocyte cancers since they begin in keratinocytes (the most prevalent type of skin cell). AKIEC stands for Actinic Keratoses (Solar Keratoses) and Intraepithelial Carcinoma (Bowen's disease), which are prevalent non-invasive lesions that are precursors to SCC.

If left untreated, they can progress to invasive SCC [2]. Melanoma is a type of skin cancer that starts in the melanocytes [1]. Melanocytes are epidermal skin cells that create melanin, the dark pigment that gives skin its colour. Melanin also serves as a natural sunscreen, protecting the skin's deeper layers from UV damage. Melanoma accounts for 5% of skin malignancies in the United States, but it is responsible for 75% of skin cancer mortality [3]. Because of melanoma's high fatality rate, skin cancer is commonly divided into two types: melanoma and non-melanoma. A dermatologist's ability to diagnose skin cancer from a dermoscopy image of a skin lesion is limited [3]. A biopsy and pathology investigation may be required to diagnose cancer in some circumstances. Previous research has established computer-based systems for detecting skin cancer from photographs of skin lesions. These models [4 - 6] were previously based on traditional machine learning techniques that required segmentation of the lesion from the surrounding skin in the image, followed by extraction of valuable information from the lesion area. The shape, texture, and colour of the lesion are examples of these characteristics. Finally, to detect cancer, the features are given into a classifier. This strategy, on the other hand, is inconvenient because defining and extracting traits that would be effective in identifying cancer is challenging.

Deep learning has emerged as a potent technique for feature learning, because to recent breakthroughs in software and hardware technology. Feature engineering, or the process of a human expert defining and extracting features, is a time-consuming and tedious task. Deep learning eliminates the need for feature engineering by learning and extracting meaningful features from raw data automatically. Many fields, particularly computer vision, have been transformed by deep learning. Deep learning has recently shown considerable promise in biomedical engineering [7 - 9]. Deep learning has been used to identify skin cancer in multiple research since 2016. A deep learning convolutional neural network (CNN) model built with 125000 clinical pictures of skin lesions was used in a study [3] done by Stanford University researchers in 2017. This algorithm was then tested on new skin lesion photos and found to be as good as trained dermatologists at

detecting malignancy. This study shows that deep learning is effective at detecting skin cancer. However, the study's database is not open to the public, preventing other researchers from developing models for future improvements.

Tschandl et al. published HAM10000 [2] in 2018, which is a public dataset of 10,000 dermoscopy images collected from Austrian and Australian patients. The photos in this dataset contain around 8000 benign lesions and the rest are malignant lesions. Pathology, expert consensus, and confocal microscopy were used to corroborate the ground truth for this dataset. A deep learning algorithm is suggested in this article to determine the malignancy of skin lesion photos from the HAM10000 dataset.

II. LITERATURE SURVEY

Melanoma is a type of skin cancer that causes a malignant tumour to form on the skin. Dermatological pictures are used to detect skin cancer. Skin cancer was detected using machine learning based on a high-performance image, and the detection rate was high (Srividhya, Sujatha, Ponnagall, Durgadevi, Madheshwaran, et al., 2020). However, by extracting more characteristics, the model's accuracy can be improved, and the sensitivity can be raised. The author presented a system that uses image processing processes to improve skin cancer diagnosis accuracy (Hoshyar, Al-Jumaily, & Hoshyar, 2014). They were unable to describe a precise model that can effectively identify cancer.

In another work, the author proposed an architecture-driven model for skin cancer diagnosis that used a DL algorithm. Because DL based on model-driven architecture can be developed so quickly, the model can anticipate the outcome almost instantly. It had a greater detection rate for skin cancer (Kadampur & Al Riyae, 2020). However, in order to improve the medical profession, the methodology requires real-time interface with medical images. The author proposed CNN-based skin cancer diagnosis in (Hasan, Barman, Islam, & Reza, 2019), where the feature is retrieved from dermoscopic pictures utilising feature extracting algorithms. During the testing phase, they achieved an accuracy of detection of 89.5 percent.

However, the detection accuracy was insufficient and needs to be improved. Furthermore, there was overfitting between the testing and training phases, which was a flaw in that study. The author suggested a lesion indexing network (LIN) based on DL to identify and classify skin cancer in (Li & Shen, 2018). By extracting more features, they were able to achieve good results using DL-based LIN. However, in order to improve the results even further, segmentation performance needs to improve.

III. CHALLENGES OF SKIN CANCER DETECTION

There are several challenges in detecting skin cancer due to differences in image kinds and sources. The variability in human skin tone makes skin cancer detection more difficult and complex. These difficulties are depicted in Fig. 1, and the most visible features of skin lesions photographs are detailed below:

1. The main challenges in skin cancer are the various sizes and shapes of the pictures, which make reliable identification impossible. Pre-processing is necessary for accurate analysis in this case.
2. A few unused signals that were not initially part of an image but can interfere with a good result will be compromised. All of these noise and artefacts should be removed during the pre-processing phases.
3. In some cases, low contrast from nearby tissues adds to the difficulty of accurately analysing skin cancer.
4. Color illumination creates challenges due to aspects such as colour texture, light beams, and reflections.
5. Some moles on the human body may never form cancer cells, but they make it difficult to effectively diagnose skin cancer from malignant photos.
6. Another problem in skin cancer diagnosis is the present bias, which alters the performance of the models to reach a better outcome.

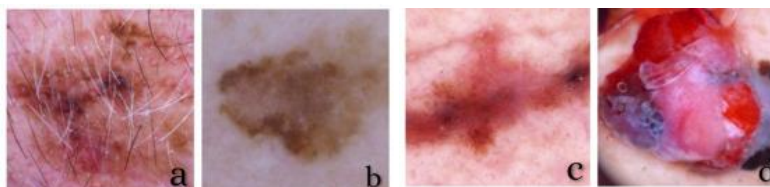


Figure 1: System Architecture of Skin Cancer

Challenges of skin lesions detection

1. Hair artifacts
2. Low contrast
3. Irregular boundaries
4. Colour illumination

IV. METHODOLOGY

4.1 Dataset

Our dataset consists of a collection of various skin cancer photos. To use deep learning techniques, you'll need a lot of data to get a solid result. The collecting of skin cancer photographs, on the other hand, is quite important. Furthermore, the lack of training data is one of the major difficulties when using DL algorithms.

To address these issues, we used the HAM10000 dataset, which contains 10015 dermoscopy images collected from Australian and Austrian patients. In this collection, there are 6705 benign photos, 1113 malignant images, and 2197 images with undetermined lesions. Pathology, master agreement, or confocal microscopy were used to confirm the ground truth for this dataset. The HAM10000 data used in our work are meticulously crafted from biopsy-proven melanocytic tumours that are classified as malignant or benign.

4.2 Proposed Approach

This method employed the tagged images "benign" and "malignant." Because the photos in the "other and unknown" groups could not be diagnosed, they were not used. Images were included to the dataset based on their analysis mark, which was derived from the photographs' information.

The dataset has been divided into two categories: one for harmful dermoscopic images and another for positive dermoscopic images. For the experimental portion, photos from the ISIC dermoscopic archive were picked at random. There are three layers in our proposed system. The input layer is the first layer, which is where the data sets are trained. The input layer collects data and adds some weight to it before sending it to hidden levels.

To detect a pattern, the neurons in the hidden layer separate the characteristics from the data. After that, the pattern is utilised to create output layers that select relevant classes. Finally, binary classification is utilised to identify appropriate classes 1 and 0. In our situation, class 0 denotes the absence of hazardous cells, whereas class 1 denotes the presence of malignant cancerous cells. The implementation of our system utilising convolutional neural networks is shown.

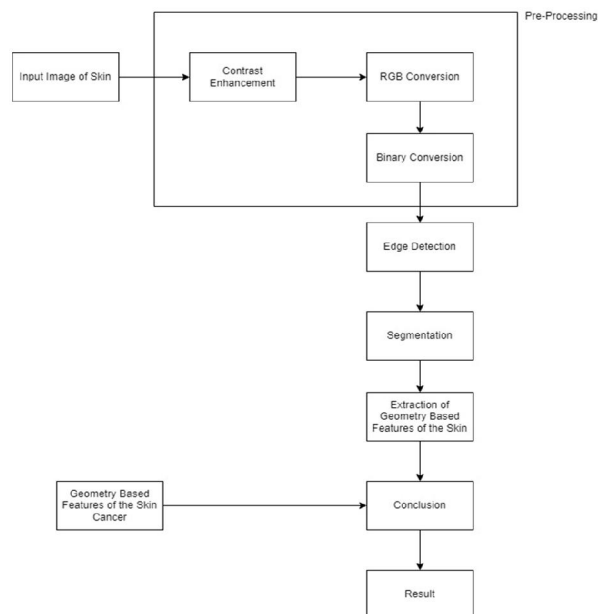


Figure 4.2: System Architecture

V. THE SYSTEM STEPS

The steps below are used to determine whether a dermoscopic image has cancer or not:

- Step-1: Initialize all of the photos and settings required by the system.
- Step-2: The system receives a training image as input and saves the images.
- Step-3: The system determines the prediction using a convolutional neural network.
- Step-4: Train using the convolutional neural network created in step-3.
- Step-5: Save the model into the system for test data prediction.
- Step 6: Assess the outcome using common assessment criteria like as precision, recall, and f1 score.

Step-1: Preprocessing data

One of the most significant challenges in computational vision is the enormous size of the images. The data intake can be quite large. When the inputted photos are 70703, the input feature dimension can be 14700. If the image size is 102410243, the feature size for calculation to transmit it to a deep neural network, particularly a convolutional neural network, will be enormous (depending on the number of hidden units). There are three image channels.

RGB has three channels (Red, Green, Blue). We must attempt to define a single channel when reading the picture due to a lack of processing resources. Another issue is the image's length. The data set contains photos with unusually large widths and heights.

The picture's width is 1022 pixels and its height is 767 pixels, making it extraordinarily large to analyse and necessitating significantly greater computer capability to register many images, which takes a long time and wastes memory. In this vein, we must resize the information pictures so that our machine can process them with less memory and computing power. To address these two issues during image reading, it will be defined in such a way that only one colour channel remains. Gray scale images are generated from original photographs in our situation since they are easier for the CPU to process.

Step-2: Save the preprocessed file

Each of the preprocessed photos, together with their classifications, is preserved in the record. For further processing, benign and malignant photos are selected from the dataset. The photos that do not have a class label must be discarded. Finally, the recorded images are fed into a convolutional Neural Network for processing.

Step-3: Feeding the preprocessed data to convolutional neural network (CNN)

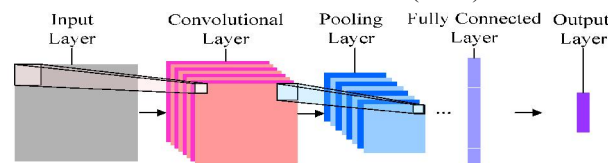


Figure 5: Convolutional neural network with its multiple layers

Three types of layers are present in a convolutional Neural Network.

That are given in following part-

- Convolution layer
- Pooling layer
- Fully connected layer

Step-4: Train

Our model will need to be trained 200 times. Every time, the system's loss drops to a set amount. We didn't find any substantial differences in loss while training epochs were around 180. As a result, we must stop iterating at 200.

Step-5: Saving the model

The model is kept for further testing. The model is then used to forecast which photos are potentially cancerous or benign.

Step-6: Prediction

Using the final output layer, we must anticipate the images. We evaluate our system using the accuracy, precision, recall, and f1 score measures after we forecast the testing images.

5.1 Convolution Layer

By using an example, our system are described here. Suppose we have a 6×6 gray-scale image (i.e. only one channel) as figure 3. Again, We have 3×3 filter. Firstly, 3×3 matrix were taken from the 6×6 image and accumulate the filter with it. As a result, the sum of the element-wise product of these values equals to the first element of 4×4 output, for examples $5 \times 1 + 0 \times 2 + -1 \times 3 + 1 \times 5 + 0 \times 8 + -1 \times 2 + 1 \times 5 + 0 \times 6 + -1 \times 1 = -6$. The second element of 4×4 output were calculated again by the sum of the element-wise product via shifting the filter one unit at the right. Similarly, the entire image were convoluted to produce a 4×4 output as figure 5.5.

In general, it can be stated as convolving an input of $x \times x$ with a $y \times y$ filter will results in $(x - y + 1) \times (x - y + 1)$:

5	3	2	1	7	4
3	5	8	9	1	3
2	5	6	0	1	4
1	6	7	1	0	2
6	2	4	0	8	2
2	5	4	2	3	9

1	0	-1
1	0	-1
1	0	-1

Fig 5.1.1 : 6×6 image with 3×3 filter.

-6	3	7	-1
-15	6	19	1
-8	12	8	-7
-6	10	4	-10

Fig 5.1.2: 4×4 image after applying 3×3 filter to 6×6 image.

- Input: $x \times x$
- Filter size: $y \times y$
- Output: $(x - y + 1) \times (x - y + 1)$

One major disadvantage of the convolution operation is the shrinkage of the size of the image. Compare to the pixel at the center of an image, the pixels at the corner are utilized only a few number of times to overcome the information loss. It has been done by padding the image by adding an extra border (i.e. adding one pixel all around the edges) which makes the input of size an 8×8 matrix (instead of a 6×6 matrix). Now, convolution of 8×8 input with a filter of size 3×3 matrix will result the original image of a size of 6×6 matrix which can be generalized as:

- Input: $x \times x$
- Padding: p
- Filter size: $y \times y$
- Output: $(x + 2p - y + 1) \times (x + 2p - y + 1)$

CNN has a tool that allows you to minimise the image size significantly. Convoluting the image with a stride of 2 will, for example, capture both vertical and horizontal directions individually.

The dimensions for stride s can be stated as:

- Input: $x \times x$
- Padding: p
- Stride: z
- Filter size: $y \times y$
- Output: $[(x + 2p - y)/z + 1] \times [(x + 2p - y)/z + 1]$

As a result, after adding the bias, the equation will be 1. The rectified linear unit activation function 2 is then used. Bi is the biased phrase here. The input image is xi, and the filter is wi. $z_i = b_i + x_i \times w_i$
 $Relu(z_i) = \max(0, z_i)$.

5.2 Pooling Layers

To reduce the image size and increase the computation speed, pooling layers are typically used. Consider a 4×4 matrix as shown below:

-6	3	7	-1
-15	6	19	1
-8	12	8	-7
-6	10	4	-10

Fig 5.2.1: Images for pooling layer

-6	3	7	-1
-15	6	19	1
-8	12	8	-7
-6	10	4	-10



6	19
12	8

Fig 5.2.2: Result after applying max pooling.

For every consecutive 2×2 block, the maximum number were taken and 2 unit size of both filter and stride were applied. If the input of the pooling layer is $x_h \times x_w \times x_c$, the output will be $\lfloor \frac{x_h - y}{z} + 1 \rfloor \times \lfloor \frac{x_w - y}{z} + 1 \rfloor \times x_c$. Then, We again apply convolutions and pulling for extract more complex features. The features are flattened to a single layer so that we can feed the model to a fully connected neural network. Then after applying the softmax as shown in equation 3, the desired result that is benign or malignant is found.

$$\text{Output} = Z_i \ln_{i=1} (Z_{i,k})$$

VI. EXPERIMENTAL SETUP

6.1 Data Set:

The ISIC Archive contains around 23907 photos [8]. Cancer is predicted using these photos.

6.2 Metrics:

- True Positives (TP): An incident in which the expected yield matched the actual yield.
- True Negatives (TN): When we predicted a false result and the actual result was also false
- False Positives (FP): This is when we expect something to be true but it turns out to be false.
- False Negatives (FN): This is when we expect something to be false but it turns out to be true.

To evaluate the model, accuracy, recall, precision, specificity, and the f1 score are used to quantify its performance. Recall is the number of threatening cases that can be distinguished from a set of all dangerous cases. $\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$. Positive Specificity is what number of benign cases can recognize out of complete given favorable cases.

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}$$

The number of potentially dangerous situations that the model could effectively predict out of the total number of cases it predicted as harmful is known as precision.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}$$

F1-score is a consolidation of precision and recall to admit the fundamental concept on how this system works.

$$\text{FMeasures} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

VII. RESULT AND DISCUSSION

The primary goal of our proposed model is to classify benign and malignant skin lesions from a DCNN-extracted dataset. The results come from the total number of photos acquired following the data reduction stage (<http://www.isic-archive.com>). Figure 6 depicts some benign and cancerous appearances.

In this work, we tested our suggested DCNN model in two ways: one with 70% of training photos, and the other with 80% of training images, which showed the best accuracy. We also tested current DNN models such as AlexNet, ResNet, VGG-16, DenseNet, and MobileNet on the same dataset, but our suggested DCNN model yielded the greatest classification rate. Figures 7 and 8 show the accuracy and loss acquired from our proposed DCNN model during the training and testing processes, respectively.

VIII. CONCLUSION

A Convolutional Neural Networks-based technique for melanoma classification is proposed in this research. A technique is being created to assist patients and doctors in detecting and identifying skin cancer classifications, whether benign or malignant. The model can be considered a benchmark for skin cancer identification by supporting healthcare practitioners, according to the experimental and assessment part. Any doctor may identify accurate findings by obtaining few random photos, but the usual approach takes far too long to recognise instances correctly.

REFERENCES

- [1]. Geoffrey E. Hinton Alex Krizhevsky, Ilya Sutskever. 2012. ImageNet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems (2012).
- [2]. Spencer Shawna Bram Hannah J, Frauendorfer Megan and Hartos Jessica L. 2017. Does the Prevalence of Skin Cancer Differ by Metropolitan Status for Males and Females in the United States? Journal of Preventive Medicine 3, 3:9 (2017), 1–6. <https://doi.org/10.21767/2572-5483.100019>
- [3]. Koby Crammer and Yoram Singer. 2005. Online ranking by projecting. Neural Computation 17, 1 (2005), 145–175.
- [4]. Swati Srivastava Deepti Sharma. 2016. Automatically Detection of Skin Cancer by Classification of Neural Network. International Journal of Engineering and Technical Research 4, 1 (2016), 15–18.
- [5]. A. Goshtasbya D. Rosemanb S. Binesb C. Yuc A. Dhawand A. Huntleye L. Xua, M. Jackowskia. 1999.
- [6]. Shweta V. Jain Nilkamal S. Ramteke1. 2013. ABCD rule based automatic computeraided skin cancer detection using MATLAB. International Journal of Computer Technology and Applications 4, 4 (2013), 691–697.
- [7]. World Health Organization. 2019. Skin Cancer. Retrieved March 16, 2019 from <http://www.who.int/en/>
- [8]. ISIC project. 2018. ISIC Archive. Retrieved March 16, 2019 from <https://www.isic-archive.com>
- [9]. Sibi Salim RB Aswin, J Abdul Jaleel. 2013. Implementation of ANN Classifier using MATLAB for Skin Cancer Detection. International Journal of Computer Science and Mobile Computing (2013), 87–94.
- [10]. Alexander Wong David A. Clausi Robert Amelard, Jeffrey Glaister. 2014. Melanoma Decision Support Using Lighting-Corrected Intuitive Feature Models. Computer Vision Techniques for the Diagnosis of Skin Cancer, Series in Bio Engineering (2014), 193–219. https://doi.org/10.1007/978-3-642-39608-3_7
- [11]. Wild CP Stewart BW. 2014. World Cancer Report. Retrieved March 16, 2019 from <http://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports/World-Cancer-Report-2014>
- [12]. Cancer Research UK. 2012. Cancer World Wide - the global picture. Retrieved March 16, 2019 from <http://www.cancerresearchuk.org/cancer-info/cancerstats/world/the-global-picture/>
- [13]. Xin Yao. 1999. Evolving artificial neural networks. Proc. IEEE 87, 9 (1999), 1423 – 1447. <https://doi.org/10.1109/5.784219>
- [14]. Mi Zhang, Jie Tang, Xuchen Zhang, and Xiangyang Xue. 2014. Addressing cold start in recommender systems: A semi-supervised co-training algorithm. In Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval. ACM, 73–82.