

Skin Cancer Detection Using Machine Learning Algorithm

Sanjana M¹, Dr. V. Hanuman Kumar²

Department of Computer Science^{1,2}, New Horizon College of Engineering^{1,2}

Bangalore-560103, Karnataka, India

Email: sonusanjana2016@gmail.com¹, vhanumankumar@gmail.com²

Abstract—The medical industry is advancing with the innovation of new technologies; newer healthcare technology and treatment procedures are being developed. Biotechnology is the base for all these advancement in technology. With the advent of several pollutants, cosmetics and chemicals into our day to day lives, the health of individuals has been deteriorating every day. These not only effects physical or mental health, but also change our lifestyles. This project work concentrates on identification of skin cancer, caused by one of the above-mentioned conditions. The images are processed using combinations of machine learning and image processing to detect the stage of cancer. Images of the affected area are captured with the help of derma scope. Several algorithms have been proposed to detect skin cancer but most of the inputs are fed manually. The main objective of this project is to develop a machine learning algorithm which requires minimal intervention of human.

Index Terms— Machine learning, Skin Cancer, Image processing, derma scope

1. INTRODUCTION

The revolution of computer in the past decades, have unfolded the use of computers, be it any field. The Era of maintaining huge records for data storage has evolved. Computers have become most used tool in almost each field. It has made the life of humans easier. Also, Computers have made research very easier.

Biomedical has offered modern medical devices for diagnostic and preventive purposes, which include diagnostic test kits, vaccines, antibodies and radio-labeled biological therapeutics used for imaging and investigation purpose.

It has played a prominent role in improving the challenges regarding to human health as it has flexibility to reduce global health differences by the provision of promising technologies. Many types of cancers are being detected and has become very common disease with the evolution of the technology. Skin cancer is the most explored disease. With increased use of cosmetics, pollution & radiations, cancer is becoming a common disease in the modern era. The images of the affected area termed as “lesion” are captured with the help of derma scope and are fed as input to the algorithm. Several algorithms have been proposed which requires input to be fed manually. The aim of this work is to propose an algorithm which require minimal intervention of doctors. In the recent years, due to the increased use of

cosmetics, and pollution & radiations, cancer is becoming a common disease in the modern era.

1.1. Stages of skin cancer

As soon as the disease is discovered, next task would be determining in which stage the cancer is. The stage in which the cancer is can be determined by various factors such as thickness, the depth of penetration, and the extent to which the melanoma has spread. Based on the stage determined, the patients are treated.

The first stage of the skin cancer, that is the early stage of melanoma (Stage 0 and Stage 1) are insular. Stage 0 tumors are in situ, which means they are noninvasive and have not entered beneath the external layer of the skin (the epidermis). Stage I tumors have attacked beneath the epidermis into the skin's next layer (the dermis) yet are little and have no different characteristics, for example, ulceration that put them at high danger of spreading (metastasizing) to close-by lymph hubs or beyond.

Stage II tumors, however limited, are bigger (for the most part 1 mm. thick or bigger) and additionally may have different characteristics, for example, ulceration that put them at high danger of spreading to the close-by lymph hubs or beyond. They are viewed as transitional or "high-chance" melanomas. Further developed melanomas (Stages III and IV) have metastasized to different parts of the body. There are additionally subdivisions inside stages.

1.2. Machine learning

Machine learning is not bond to one field it consists of many things, and it is growing rapidly. The whole objective of machine learning is to program, so as to make the computer perform tasks accordingly without explicitly programming each time. A computer learns from the past inputs given by the user, records the patterns and improves with the experience. When a computer is provided with data and then recognize those data with the help of an algorithm, it is called machine learning, the process by which the machine learns is called as training and output processed is the model. The objective of the learner is that to generalize from its experience. Generalization means the ability of the machine to perform accurately on unseen or new data.

In the past decade machine learning has amazed us with the invention of automated cars, practical speech recognition, effective web search, and a vast understanding of human genome. Machine learning today, is widely spread that you probably use it dozens of times a day without knowing it.

Machine learning in health care industry are taking the industry by storm. Machine learning in medicine has recently made buzz. With regards to adequacy of machine adapting, more information quite often yields better outcomes—and the human services area is perched on an information goldmine. Calculations can furnish quick advantage to disciplines with procedures that are reproducible.

Likewise, those with expansive picture datasets, for example, radiology, cardiology, and pathology, are solid applicants. Machine learning can be prepared to take an insight at pictures, distinguish anomalies, and point to regions that need consideration, consequently enhancing the exactness of every one of these procedures.

Long haul, machine learning will profit the family expert or internist at the bedside. Machine learning can offer a target assessment to enhance productivity, dependability, and exactness.

1.3. Regression

Regression examination is a type of predictive displaying method which researches the connection between a reliant (target) and autonomous variable (s) (indicator). This system is utilized for estimating, time arrangement displaying and finding the causal impact connection between the factors. For instance, connection between rash driving and number of street mischances by a driver is best concentrated through relapse. Regression investigation is a critical

instrument for displaying and breaking down information. Here, we fit a curve/line to the information focuses, in such a way, to the point that the contrasts between the separations of information focuses from the curve or line is limited.

Regression investigation gauges the connection between at least two factors. We should comprehend this with a simple precedent:

Suppose, you need to gauge development in offers of an organization dependent on current monetary conditions. You have the ongoing organization information which demonstrates that the development in deals is around more than two times the development in the economy. Utilizing this knowledge, we can foresee future offers of the organization dependent on current and past data.

There are various advantages of utilizing relapse investigation. They are as per the following:

- It shows the noteworthy connections between ward variable and free factor.
- It shows the quality of effect of numerous autonomous factors on a reliant variable.

Regression examination likewise enables us to look at the impacts of factors estimated on various scales, for example, the impact of value changes and the quantity of limited time exercises. These advantages help economic specialists/information experts/information researchers to wipe out and assess the best arrangement of factors to be utilized for building prescient models.

Types of Regression Technique used:

- (1) Linear Regression
- (2) Logistic Regression
- (3) Polynomial Regression
- (4) Stepwise Regression
- (5) Ridge Regression
- (6) Lasso Regression
- (7) Elastic Net Regression

2. LITERATURE SURVEY

Image handling systems give a productive instrument to characterize the malignancy from the images. Different creators have utilized diverse approaches to consolidate these innovations to accomplish better end. Different works have been done in the location of

Skin malignancy utilizing picture handling and mix of neural system.

Paper [1] proposes image segmentation scheme based on an algorithm, Support vector Machine (SVM) and snake active counter. To help finding the appropriate parameters for snake algorithm, SVM is used. The initial curve and the snake parameters must be chosen appropriately to perform the snake algorithm efficiently. Hence, to choose the appropriate initial curve and parameters we apply Support vector Machine (SVM). The initial curve is predicted to be of the following shapes: curve, eclipse and rectangle. To reduce the complexity in the implementation of SVM model, and to keep up the SVM implementation without any degradation, these shapes are chosen [1].

Based on the results of the SVM, the attributes can fit in the image. The images used as testing set were chosen as the use for template circle and to determine the accuracy of the edges. The experimental results show that the Snake algorithm is the basis to determine the edge, in comparison to the edge found by the experts.

The segmentation technique applied by the algorithm found the edges, which are closer to the results determine by the experts. But the undesired areas are also segmented. To obtain accurate results of segmentation, better classification or the refinements of the algorithm is required [1].

Paper [2] proposes an automated method for detecting skin cancer from the photographs which are plain, captured from the affected area. Melanoma, which is the last stage of the skin cancer, should be diagnosed as early as possible for the patients to survive. Since, the skin cancer occurs on the body externally, Computer aided diagnoses using digital image processing can help the skin doctors in assisting to identify the skin cancer.

The method used to detect the melanoma is as follows, using the Grab Cut algorithm first segment an input image into lesions of interest appeared to be melanoma, next features such as the shape, color, and geometry by using image processing techniques are extracted [2].

The extracted features are they classified as cancerous and non-cancerous. Cancerous named as "malignant" or non-cancerous mole named as "benign" by applying support vector machine with Gaussian radial basis kernel (SVM-RBF). The authors conducted experiments with nearly 200 images (100 of melanoma and 100 of benign) and found from the

experimental results that only six features are enough to detect melanoma [2].

The disadvantage of this method is that the database doesn't consists classified images for dark skinned images to experiment with. They also are planning to apply ensemble learning method in the future [2].

Paper [3] proposes uses image segmentation using an automatic method for lesion detection using algorithm based on deep learning for pixel-wise labelling. The experiment has been conducted on two network architectures by testing on public data and use the ISIC database for training the network and the PH2 database to show the used method is not data set related. The experimental results show that the proposed approach is very accurate and performs segmentation even in presence of hair and air/oil bubbles. Data annotation that generates many test images, and the implementation of a GUI which is semi-automatic is the additional contribution of this paper.

In Paper [4] prevention of melanoma and early detection of skin cancer are proposed. An analysis system to determine lesion and to prevent the melanoma are proposed. To prevent the people from skin burns a real-time system has been developed which alerts the users. An alert reaches the users, when the users is exposed to the sun for a longer time and when the rays are harmful. An equation was introduced for the alert to be caused.

In this work, the components of a portable real-time noninvasive skin lesion analysis system to assist in the melanoma prevention and early detection are proposed. The first component is a real-time alert to help users to prevent skin burn caused by sunlight; a novel equation to compute the time for skin to burn is thereby introduced. The second part is a robotized picture examination including picture procurement, hair recognition and prohibition, injury division, highlight extraction, and characterization. The structure has been produced in a PDA application. The test results demonstrate that the proposed framework is effective, accomplishing high grouping exactness's [4].

Paper [5] propose a system that recognizes and distinguishes skin sores as kindhearted or threatening dependent on pictures taken from general cameras. The pictures are portioned, highlights separated by applying the ABCD rule and a Neural Network is prepared to order the injuries to a high level of

precision. The prepared Neural Network accomplished a general arrangement precision of 76.9% on a dataset of 463 pictures, partitioned into six classes.

The general exactness rate and execution of the framework can be enhanced via preparing the neural system on a much bigger and differing dataset with high intra-class changeability. This would diminish the misclassification and emphatically affect the precision rate. An option in contrast to broadening the dataset is expanding the quantity of highlights removed from the pictures [5].

In this paper [6], a consequently skin tumor characterization framework is produced and the relationship of skin growth picture over the neural system are examined with various sort of pre-preprocessing. The gathered picture is feed into the framework and picture pre- preprocessing is utilized for clamor evacuation. Pictures are section utilizing thresholding. There is sure component special in skin growth area these elements are separate utilizing highlight extraction system. Multilevel 2-D wavelet disintegration is utilized for highlight extraction system. These highlights are given to the info hubs of neural system. Back spread neural system and outspread essential neural system are utilized for characterization reason, which classifications the given pictures into destructive or non-carcinogenic.

Smartphones are assuming real job in e-wellbeing so that m-wellbeing is assuming a critical job in human services industry. Picture handling systems are instrumental in medicinal services industry to distinguish irregularities in human body. Skin growth (Melanoma) is a standout amongst the most dangerous diseases, yet when analyzed early, it very well may be restored. Reports tell that in excess of millions of people expire because of Skin tumor itself. This paper [7] talks about how skin disease can be distinguished in beginning times utilizing cell phone application by dissecting properties of the disease, Asymmetry, Border, Color variety, Diameter and Expansion (ABCDE). These properties are dissected utilizing distinctive picture preparing methods like Gray scale change, Segmentation, shape following and histogram examination.

In this paper [8], distinctive advanced pictures have been breaking down considering unsupervised division methods. Highlight extraction strategies are then connected on these portioned pictures. After this, an extensive exchange has been investigated dependent on the acquired outcomes. In automated diagnosis of skin lesions, feature extraction is based

on the so-called ABCD-rule of derma scope. ABCD represent the asymmetry, border structure, color variation, and dermascope structure so called diameter of the lesion and define the basis for a diagnosis by a dermatologist.

At the point when a skin sore is suspected as melanoma, it must go through every one of the four examinations. On the off chance that the presumed skin sore go through just the three of these, it may demonstrate wrong results about its being melanoma or not. Therefore, all the four measures must be considered to choose whether a skin sore is melanoma or not. All things considered, the most ideal approach to bring down the danger of melanoma is to confine the presentation to solid daylight and other wellspring of Bright light. Deal with all the important estimates, for example, ensuring skin with dress, wearing cap, utilizing sunscreen, remaining in the shade (and so forth). In addition, dependably remain caution about skin and do month to month skin-self exams to diminish the shot of getting any skin tumor which a hazard to human life is [8].

3. CONCLUSION

This paper focuses on determining the stage of the skin cancer, based on various feature such as the area of the spread, diameter, color of the lesion, etc. The analysis can be made with the help of machine learning algorithm, in with we train the system based on the history of the images stored in the database, and test the current image to determine, whether the test image comes in the category of the melanoma or not, if it does, then to determine its stage. A comparison can be made with the existing systems, machine learning reduces the computational time. Hence, the treatment can begin faster.

REFERENCES

- [1] Detection skin cancer using SVM and snake model. Bumrungkun, P., Chamnongthai, K., & Patchoo, W. 2018 International Workshop on Advanced Image Technology
- [2] A SVM based diagnosis of melanoma using only useful image features, Suleiman Mustafa; Akio Kimura , International Workshop on Advanced Image Technology (IWAIT) Year-2018
- [3] Deep Convolution Pixel-wise Labeling for skin lesion Image Segmentation ,Ali Youssef ;Domenico D. Bloisi ;Mario Muscio ; Andrea Pennisi ; Daniele Nardi ; Antonio Facchiano IEEE International Symposium on Medical Measurements and Applications (MeMeA) Year: 2018

- [4] Early Detection of Melanoma Skin Cancer Using Classifiers, VS. Sabeera, P. Vamsi Krishna PG Scholar, Dept. of ECE, R. K. College of Engineering, Vijayawada, Andhra Pradesh Assistant Professor, Dept. of ECE, R.K. College of Engineering, Vijayawada, Andhra Pradesh, Year-2016
- [5] Skin Cancer Detection and classification Pratik Dubal, Sankirtan Bhatt, Chaitanya Joglekar, Dr. Sonali Patil, Department of Information Technology K. J. Somaiya College of Engineering Vidyavihar, Year-2017
- [6] Skin Cancer Detection using Artificial Neural Network, Ekta Singhal M.Tech II Year, Dept of Computer Science Engineering, MUST – FET, Lakshmangarh, India Shamik Tiwari, Assistant Professor, Dept of Computer Science Engineering, MUST – FET, Lakshmangarh, India, Year-2015
- [7] Detection of Skin Cancer Using Image Processing Techniques, Chandrasasa M, Varun Vadigeri and Dixit Salecha, Computer Science and Engineering, The National Institute of Engineering, Year-2016
- [8] Image Processing for Skin Cancer Features Extraction Md.Amran Hossen Bhuiyan, Ibrahim Azad, Md.Kamal Uddin, Department of Computer Science & Telecommunication Engineering, Noakhali Science & Technology University, Bangladesh, International Journal of Scientific & Engineering Research Volume 4, Issue 2, February-2013