RESEARCH ARTICLE OPEN ACCESS

Deploying Artificial Intelligence techniques for detection of skin cancer in earlier stage through Mobile App

Madhavi R K

Senior Scale Lecturer, Dept. of CSE, Government Polytechnic, Chintamani

ABSTRACT

Cancer poses a significant global health challenge, contributing to a rising number of patients and fatalities. Notably, skin cancer has emerged as one of the most prevalent forms of cancer. Enhancing the survival rates of individuals grappling with skin cancer and curbing treatment expenses can be achieved through timely diagnosis. However, current diagnostic methods within healthcare systems have limitations, including their dependence on extensive human resources, prolonged turnaround times, and limited accessibility for all individuals. Hence, there is a pressing need for systems that are user-friendly, yield precise outcomes based on scientific methodologies, and are accessible to a broad demographic for early detection of skin cancer. Leveraging artificial intelligence techniques can facilitate the early identification of skin cancer. This research aims to categorize images of benign and malignant skin conditions and communicate the findings to users via a mobile application. To achieve this, various algorithms, including CNN, KNN, and Decision Tree, were employed for image classification. The study applied augmentation techniques to the "Benign" dataset within the ISIC: Skin Cancer: Malignant vs. Benign experiments. The Transfer Learning algorithm emerged as the most successful, achieving a 94.89% accuracy rate. The proposed system affirms the potential of artificial intelligence in detecting skin cancer at an early stage, with the added benefit of notifying users through a mobile application. The researchers anticipate that these findings will inspire further investigations into innovative approaches for early skin cancer detection.

Keywords: - CNN, KNN, Skin cancer, Data augmentation, Dermoscopy

I. INTRODUCTION

Cancer, a disease responsible for the highest number of fatalities, is becoming increasingly prevalent globally. Ranking second only to heart disease in terms of causes of death, cancer represents a universal health challenge that requires comprehensive attention [1]. Described as a non-lethal "genetic disorder," cancer is characterized by the uncontrolled proliferation of these disorders throughout the body. While there are various types of cancer, the common initiation of this ailment involves the unregulated expansion of cells. If not addressed, cancer has the potential to cause significant degeneration and irreversible consequences, ultimately leading to death within the human body [2].

As per the World Health Organization [3], the incidence of skin cancer has risen in recent years within the spectrum of cancer types. Additionally, the escalation of environmental issues, such as the depletion of the ozone layer, has contributed to the heightened prevalence of skin cancer [4]. The onset of skin cancer is associated with the deterioration and unregulated proliferation of the human DNA structure, a consequence of exposure to ultraviolet rays (UV light) in the epidermal layer of the skin [5]

There exist three types of skin cancer, with Basal cell carcinoma being the most prevalent, constituting 80% of cases worldwide. Basal cell carcinoma typically carries a low mortality rate, leading to fatal outcomes only in the absence of early diagnosis. Squamous cell carcinoma represents 16% of skin cancer cases and is generally treatable, yet it can result in death if not detected early due to the nature of the disease. While melanoma is the least common type of skin cancer, it is the most lethal. Stemming from the deterioration of melanocyte cells in the skin, melanoma can become life-threatening if not identified promptly, potentially spreading throughout the body and complicating the treatment process [6], [7].

Implementing early detection and diagnostic methods plays a crucial role in the recovery or prolonged survival of individuals grappling with skin cancer. Examining the rates of skin cancer and its incidence, it becomes evident that detecting skin cancer at an advanced stage imposes a substantial economic burden on patients. Although obtaining data on global cancer-related expenditures is challenging, European Union Countries spent €57.3 million in 2017 to combat this disease. Early detection has the potential to reduce one-third of the expenses associated with cancer treatment [8].

Dermoscopy is one of the contemporary techniques employed for skin cancer detection. In this method, the tumor area undergoes analysis through imaging approaches. Various non-surgical imaging procedures, including photography, dermoscopy, ultrasound, fluorescence spectroscopy, optical tomography, and multi-photon scanning, are utilized. Dermoscopy, involving microscopic examination of skin lesions, magnifies moles on the skin 10-20 times, allowing for the differentiation between benign and malignant conditions. However, this process necessitates access to specialized devices and institutions, particularly those with field expertise [9].

Dermoscopy entails a comprehensive and long-term imaging and analysis process. Consequently, delivering prompt services to all at-risk patients and conducting ongoing investigations on vulnerable age or patient groups using traditional imaging techniques pose challenges to health institutions [11]. Additionally, considering factors such as the universal COVID-19 pandemic, it has become imperative to safeguard the health workers, optimize the capacity of health institutions, and ensure patient health. Despite the increased patient frequency in health institutions during the pandemic, the battle against skin cancer remains crucial for public health [11].

The challenges associated with prolonged imaging procedures in the early detection of skin cancer have spurred the exploration of innovative methodologies. Additionally, the health sector's digital transformation has facilitated the swift integration of artificial intelligence techniques. Consequently, numerous services in the health sector have experienced enhanced efficiency in terms of transaction cost and speed through the application of artificial intelligence techniques [12].

II. LITERATURE REVIEW

In [14], the authors explored the application of image processing and artificial intelligence techniques for early detection of skin cancer. The model underwent training on a research dataset using Support Decision Machine, K-NN, and Ensemble Learning algorithms, with MATLAB employed as the programming language. The Support Decision Machine algorithm demonstrated the highest performance, achieving an accuracy rate of 83.90%.

Reference [15] employed image processing and machine learning approaches for early skin cancer detection. The study utilized the ISIC-ISBI 2016 dataset, incorporating OTSU image thresholding and Gray Co-occurrence Matrix techniques in data pre-processing. Support Vector Machine, Discriminant Analysis, and Random Forest algorithms were applied, with the Random Forest algorithm attaining the highest performance at an accuracy rate of 88.17%.

In another investigation [16], AI techniques were utilized for the early diagnosis of melanoma cancer, employing 23,000 images from the International Skin Imaging Collaboration (ISIC) dataset. The Support Vector Machine algorithm yielded an accuracy rate of 83%.

A study by Ahsraf [17] utilized actual patient data from Pakistan DHQ hospital for early skin cancer detection. The Convolutional Neural Networks (CNN) algorithm, a deep learning model, was employed, achieving an accuracy rate of 93.29%.

Arasi's research [18] focused on early detection of melanoma cancer using data from artificial intelligence techniques (DermIS) and DermQuest databases. Naive Bayes and Decision Tree algorithms were implemented, resulting in a Decision Tree algorithm accuracy rate of 92.86% and a Naive Bayes algorithm accuracy rate of 98.8%.

The analysis of existing literature reveals a growing number of studies focused on the early detection of skin cancer through skin image analysis. However, an examination of the study outcomes highlights variations in the utilized datasets, data pre-processing methods, algorithms for feature extraction and classification, as well as the applicability and accuracy performance across different studies. The reviewed studies predominantly assess algorithm performance and exhibit limitations in terms of practical applicability.

It becomes evident from the surveyed research that the emphasis is generally placed on evaluating algorithmic performance, with a relatively limited focus on real-world applicability. To enhance the development of high-performing algorithms and provide valuable support to health institutions, there is a recognized need for the creation of applications featuring a visual interface. These applications can serve to foster the advancement of systems recommended for the early diagnosis of skin cancer.

III. METHODOLOGY

This section provides comprehensive information about the dataset employed in the proposed system for early skin cancer detection through a mobile application utilizing artificial intelligence techniques. It outlines the dataset features, operations conducted on the dataset, the programming language could be utilized, and the libraries employed. Additionally, Figure 1 offers a concise overview of the working method of the system

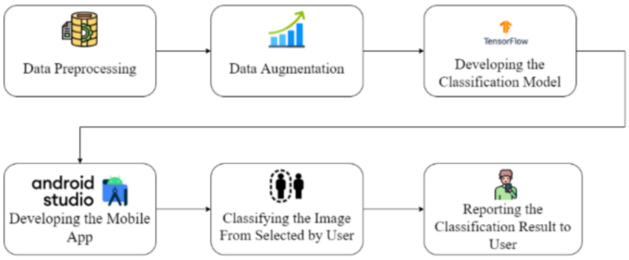

Fig. 1: System Architecture

The proposed application seeks to enhance awareness regarding skin cancer, simplify the diagnosis of various skin diseases, including skin cancer in humans, and directly guide patients to seek medical attention.

A. **Information of the Dataset**

This research utilized the "ISIC: Skin Cancer: Malignant vs Benign" dataset (ISIC, 2019). The selection of this dataset was guided by the objective of leveraging algorithms with proven high performance, as demonstrated in available academic publications.

The ISIC (International Collaboration on Skin Imaging) database represents a collaborative effort between academia and industry. Its overarching goals include reducing mortality rates associated with late detection of malignant skin cancer, minimizing the need for prolonged procedures like dermoscopy and biopsy in the skin cancer diagnostic process, and promoting the adoption of faster techniques. Since 2016, ISIC has consistently organized competitions focused on the classification of malignant skin cancer. The dataset in the archive is freely accessible, and it has undergone review by dermatologists, incorporating valuable clinical metadata. Consequently, the dataset in the archive is diverse, encompassing features of various disease types and images.

B. **Features of Dataset**

The dataset employed for model training comprises 3,297 color images of skin cancer tissue, each having dimensions of 224x224 pixels. The dataset is split into 80% for training and 20% for testing.

Within the training phase dataset, there are 1,440 image data instances for benign skin cancer type and 1,197 image data instances for malignant skin cancer type. In the test dataset, there are 360 image data instances for benign skin cancer types and 300 image data instances for malignant skin cancer types.

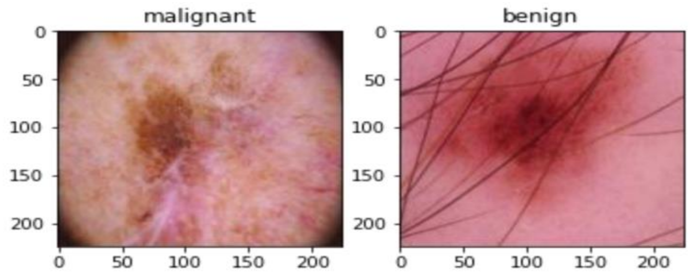

Fig.2: Original pictures in the dataset

C. **Pre-processing of data**

Initially, all chosen images undergo standardization to a consistent size of 224x224 pixels for model training. Subsequently, following the resizing of the images, data augmentation is implemented to enhance the accuracy performance of the model.

D. **Augmentation of Data**

The efficacy of data mining techniques in health studies is closely tied to the quality of both images and content utilized. To optimize the performance of the trained model, we augmented the dataset, ensuring that the model learned the intended features comprehensively during training. The augmented images were incorporated in the training phase to enhance the model's overall performance.

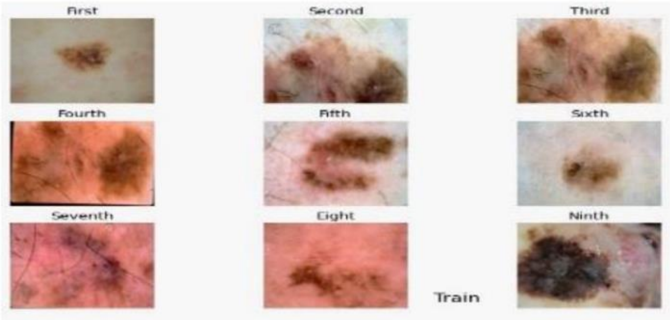

Fig.3: Pictures used in training phase

Figure 3 displays the images from the dataset utilized during the training phase, encompassing both benign and malignant skin images.

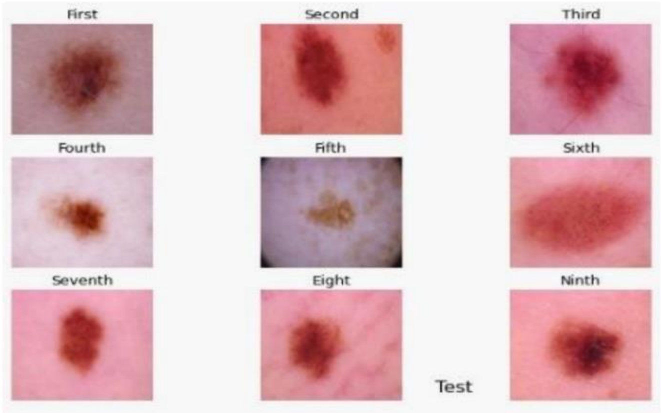

Fig.4: Pictures used in training phase

Figure 4 illustrates the dataset images employed during the classification phase, encompassing both benign and malignant skin images.

E. **Classification of the disease**

The classification phase involves both feature extraction and learning from these features. In this research, skin images were classified using the K-NN (K-Nearest Neighborhood or K-Nearest Neighbor), Decision Trees, and Transfer Learning algorithms.

F. **CNN Architecture**

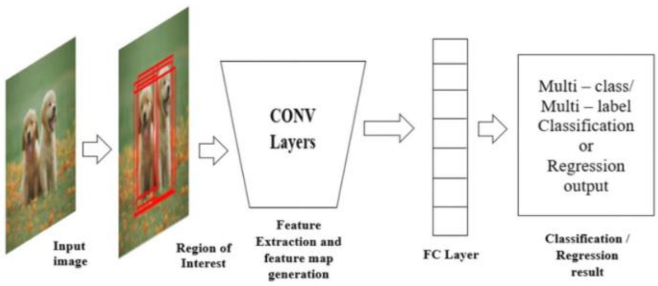

Fig.5: Architecture of CNN

CNN (Convolutional Neural Network) is a type of neural network commonly used for the image recognition and the classification tasks, although it is also applicable to the other tasks like natural language processing and signals processing. CNN reduces the human effort by automatically detecting the features. CNNs are the classes of Neural Networks that capture images and classify the features, this features are widely used for analyzing visual images. Their applications range from image and video recognition, image classification, medical image analysis, computer vision and natural language processing. CNN has high accuracy.CNN has a image recognition so this technique is used in various industries such as medical image analysis, phone, security, recommendation system, etc

G. **KNN Algorithm**

KNN is one of the most basic essential classification algorithms in the machine learning. It belongs to supervised learning domain and it finds intense application in the pattern recognition, data mining, and intrusion detection.KNN is a versatile algorithm with its own strengths and the weaknesses, and its effectiveness depends on specific characteristics of the dataset and the problem at hand. KNN calculates the distance between the input data point and the other data points in training set.

1) Step1: Selecting the optimal value of K
2) Step2: Calculating the distance
3) Step3: Finding the Nearest Neighbors

Step4: taking the average

IV. CONCLUSION

The performance of the classification achieved in the research aligns with the outcomes reported in the literature, affirming the consistency of the proposed system. While existing literature primarily focuses on classification performances and comparisons, our proposed system, equipped with a mobile application designed for Android operating systems, demonstrates enhanced applicability and holds the potential for approval by health institutions. This innovative approach, distinct from other studies, suggests that the proposed system could streamline the efforts of healthcare professionals by minimizing procedures like biopsy and dermoscopy, thereby aiding in the early detection of skin cancer.

REFERENCES

[1] A. Ilgaz and S. Gözüm, "Kanser Taramalarinda Öncelikli Hedef Birinci Derece Akrabalar," 2014. [Online]. Available: http://www.deuhyoedergi.org

[2] C. Wood and J. Ward, "A general overview of the cancer education needs of non-specialist staff," European Journal of Cancer Care, vol. 9, no. 4, pp. 191–196, Dec. 2000, doi: 10.1046/j.1365- 2354.2000.00216.x.

[3] WHO, "Radiation: Ultraviolet (UV) radiation and skin cancer," 2017.

[4] M. B. Siegel et al., "Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer," Journal of

Clinical Investigation, vol. 128, no. 4, pp. 1371–1383, Apr. 2018, doi: 10.1172/JCI96153.

[5] L. Rey-Barroso, S. Peña-Gutiérrez, C. Yáñez, F. J. Burgos-Fernández, M. Vilaseca, and S. Royo, “Optical Technologies for the Improvement of Skin Cancer Diagnosis: A Review,” Sensors, vol. 21, no. 1, p. 252, Jan. 2021, doi: 10.3390/s21010252.

[6] E. de Vries, F. I. Bray, J. W. W. Coebergh, and D. M. Parkin, “Changing epidemiology of malignant cutaneous melanoma in Europe 1953-1997: Rising trends in incidence and mortality but recent stabilizations in Western Europe and decreases in Scandinavia,” International Journal of Cancer, vol. 107, no. 1, pp. 119–126, Oct. 2003, doi: 10.1002/ijc.11360.

[7] C. Garbe et al., “Diagnosis and treatment of melanoma: European consensus-based interdisciplinary guideline,” European Journal of Cancer, vol. 46, no. 2, pp. 270–283, Jan. 2010, doi: 10.1016/j.ejca.2009.10.032.

[8] Cancer Atlas, “The Economic Burden of Cancer,” 2017.

[9] V. Narayanamurthy et al., “Skin cancer detection using non-invasive techniques,” RSC Advances, vol. 8, no. 49, pp. 28095–28130, 2018, doi: 10.1039/C8RA04164D.

[10] L. Krueger, A. Saizan, J. A. Stein, and N. Elbuluk, “Dermoscopy of acquired pigmentary disorders: a comprehensive review,” International Journal of Dermatology, vol. 61, no. 1, pp. 7–19, Jan. 2022, doi: 10.1111/ijd.15741.

[11] M. A. Pizzichetta et al., “Dermoscopic Features of Difficult Melanoma,” Dermatologic Surgery, vol. 33, no. 1, pp. 91–99, Jan. 2007, doi: 10.1111/j.1524- 4725.2007.33015.x.

[12] M. H. van Velthoven, C. Cordon, and G. Challagalla, “Digitization of healthcare organizations: The digital health landscape and information theory,” International Journal of Medical Informatics, vol. 124, pp. 49–57, Apr. 2019, doi: 10.1016/j.ijmedinf.2019.01.007.

[13] H. Beenish and M. Fahad, “Skin Cancer Prediction using Data Mining and its Techniques – A Review,” in 2020 International Conference on Computing and Information Technology (ICCIT-1441), Sep. 2020, pp. 1–4. doi: 10.1109/ICCIT144147971.2020.9213800.

[14] M. Babar, R. T. Butt, H. Batool, M. A. Asghar, A. R. Majeed, and M. J. Khan, “A Refined Approach for Classification and Detection of Melanoma Skin Cancer using Deep Neural Network,” in 2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2), May 2021, pp. 1–6. doi: 10.1109/ICoDT252288.2021.9441520.

[15] A. Javaid, M. Sadiq, and F. Akram, “Skin Cancer Classification Using Image Processing and Machine Learning,” in 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Jan. 2021, pp. 439–444. doi: 10.1109/IBCAST51254.2021.9393198.

[16] MA. A. Thaajwer and UA. P. Ishanka, “Melanoma Skin Cancer Detection Using Image Processing and Machine Learning Techniques,” in 2020 2nd International Conference on Advancements in Computing (ICAC), Dec. 2020, pp. 363–368. doi: 10.1109/ICAC51239.2020.9357309.

[17] R. Ashraf, I. Kiran, T. Mahmood, A. Ur Rehman Butt, N. Razzaq, and Z. Farooq, “An efficient technique for skin cancer classification using deep learning,” in 2020 IEEE 23rd International Multitopic Conference (INMIC), Nov. 2020, pp. 1–5. doi: 10.1109/INMIC50486.2020.9318164.

[18] M. A. Arasi, E.-S. M. El-Horbaty, and E.-S. A. E.-D. El-Dahshan, “Classification of Dermoscopy Images Using Naïve Bayesian and Decision Tree Techniques,” in 2018 1st Annual International Conference on Information and Sciences (AiCIS), Nov. 2018, pp. 7–12. doi: 10.1109/AiCIS.2018.00015.

Author Profile

Madhavi R K - the Senior Lecturer, she has awarded BE degree in Computer Science & Engieering from UBDT college of Engineering, Davangere Karnataka India in 2004 and M.Tech degree in Computer Science & Engineering from SJC Institute of Technology affiliated to Vishweshwaraiah Technological University in the year 2010. She worked as an Asst. Professor, department of Computer Science & Engineering in SJCIT Chikkaballapur India from 2004 to 2006 and in TJohn Institute of Technology Bangalore India from 2008 to 2010. She is working as a Senior Lecturer, department of Computer Science & Engineering in Government Polytechnic, Chintamani, Karnataka, India from the year 2010. Her Area of interests on latest Technologies such as cloud computing, Internet Of Things, Artificial Itellegence, Machine Learning etc.