Detection of Melanoma Skin Cancer using Deep Learning

Nanthini N

Department of ECE
Sri Krishna College of Engineering and Technology
Coimbatore, India
nandhini36268@yahoo.co.i n

Aishwarya D

Department of ECE
Sri Krishna College of Engineering and Technology
Coimbatore, India
20euec010@skcet.ac.in

Angelin Simon

Department of ECE
Sri Krishna College of Engineering and Technology
Coimbatore, India
20euec013@skcet.ac.in

Baby Vishnupriya N

Department of ECE
Sri Krishna College of Engineering and Technology
Coimbatore, India
20euec025@skcet.ac.in

***Abstract*— The skin cancer diagnosis system is a foretelling model that programmatically denotes the presence of melanoma cancer in the skin by assessing the dermoscopic images with the assistance of deep learning. It is caused due to unrepaired de-oxy ribonucleic acid[DNA] that triggers a mutation. These mutations cause the skin cells to proliferate rapidly and thereby forming malignant tumours. The primary purpose behind this analysis is to spot the presence of melanoma skin cancer at its preliminary stages by acquiring quick results with high accuracy. Our proposed model put forward a remedy for the trouble of exorbitant expenses behind the analysis is the least accuracy rate in identification and manual detection system which faces portability problem. In our model, dermatoscopic images are categorized to prognosticate skin cancer by making use of the multi-layered CNN approach. As an outcome, our model has produced an accuracy of 95.24% which is greater than the existing model by 1.66 per cent.**

Keywords—melanoma skin Cancer, Convolution Neural Network, SVM, Deep Learning, Google Colab.

I. INTRODUCTION

In present circumstances, computer-aided diagnosis using machine learning and deep learning has become an essential tool for the detection and diagnosis of fatal diseases at their early stages[1]. The present approach is colour normalization, adaptive curvature and feature extraction using the ABCD rule[2]. The deep learning approach is a subset of the machine learning approach which uses the multilayers concept to continuously take out higher-level characteristics of the raw input data[3]. Considering a case in point, in the processing of the image, underneath layers may detect the edges, whereas the above layers may detect the approaches related to the human being such as letters or digits or faces[4]. Currently, models of deep learning are related to the artificial neural network, especially convolutional neural network, even though they can also be incorporated into a thesis or concealed variables are is organized in a layer-wise manner in generative models of deep learning like the nodes in deep belief networks[5]. In a deep learning concept, each and every level studies to metamorphose its data as input into a kind of intellectual and synthesized representation[6]. Usually, for recognition of image application, the raw data i.e the input, maybe pixels arranged in matrix format[7].

II. NEED FOR THE MODEL

The main objective of our system is to diagnose, categorize and segment the melanoma skin cancer images using deep learning with the support of algorithms like Convolutional Neural Network[5][8]. The dataset which consists of nearly 300 images is used. With the help of deep learning code, the images which are found to be affected by the skin cancer case will be detected more effectively and accurately than the existing model and results will be obtained with less complexity than any other conventional technique[9].

III. RELATED WORKS

The existing melanoma skin cancer analysis system is a prognostic system which interactively foresees melanoma skin cancer by estimating dermatoscopic images with the support of deep learning approaches[6][10]. The foundational goal on the side of this analysis is to estimate skin cancer at its preliminary stages by attaining instantaneous results with better accuracy[11][12]. The point on the side of the goal indicates the case of the increasing rate of melanoma skin cancer patients

worldwide, high medical expenses and exponentially increasing death rate cases due to melanoma skin cancer because of lack of awareness in its early stages[14][15]. In the existing model, Convolutional Neural Network is used as an algorithm in order to achieve an accuracy of 93.58 percentage that is 1.66 per cent less than that of our proposed model[5].

Accuracy and the computational speed of the output are one of the limitations of the existing model. To overcome this limitation of the existing model we came up with an idea of a model in which the accuracy of the system will be high and the computational speed of the code will be low with less complexity of code.

IV. PROPOSED WORK

A. Software Description:

1)Algorithm - Convolutional Neural Network [CNN]

The convolutional neural network (CNN) is a particularized kind of artificial neural network that utilizes the mathematical operation known as a convolution under matrix in general multiplication in at least the minimum of their one layers. They are particularly planned for the operation of processing data of pixels and are utilized in the image recognition and examination process. In the deep learning model, each and every level get the hang to change its input data into a sort of intellectual and synthesized presentation.
In the deep learning concept, the approach called convolutional neural network (ConvNet) is the subset of an artificial neural network, which is most commonly utilized for the analysis of visual imagery. Contradictorily, many of the convolutional neural network approaches are equally variant, as opposite to invariant\. They have four applications in image and video recognition, image classification, recommender systems, image segmentation, medical image analysis, brain-computer interfaces, financial time series and natural language processing.

Need for Convolutional Neural Network:

We use Convolutional Neural Network because it has dimensionality reduction and feature parameter sharing which makes that system independent of human supervision.

2) Steps involved

a) Preparation of Data.
b) Selection of a suitable Algorithm.
c) Fit a Model.
d) Choose a Validation Method.
e) Examine Fit and Update Until Satisfied.
f) Use a Fitted Model for Predictions.

3) Software used

a) Google collab
b) Google Drive

4) Programming language

The programming language used in our project is python, which is an object-oriented programming language. Python is a general-purpose programming language which is also referred to an as high level and is most widely used in producing deep learning algorithms and data science. Python is a programming language that aids in the creation of a massive range of applications. Developers consider it a great choice for Artificial Intelligence (AI), Machine Learning(ML) and Deep Learning(DL) models. Python and its libraries like **Numpy, Pandas, Matplotlib, TensorFlow, Keras etc** are used**.**

5)Libraries used

The Python library can be found **in the site-packages folder within the Python directory** which is installed in our system.

a)NumPy

Numerical Python also referred as NumPy is an open-source library in the python programming language that makes it easy to solve the complex numerical operations. All the applications of machine learning and deep learning models involve complex numerical operations with large sets of data. Numerical python is a very popular library of the python programming language for matrix processing and large multi-dimensional array.

b) Pandas

Pandas is an open-source package, mostly used for data analysis and machine learning functions it is a two-dimensional data structure with rows and columns, it provides a flexible way to merge and reshape the data. It is built on top of another package named Numpy, which provides help for multi-dimensional arrays and easily handles the missing data. Pandas have series and data frames in which the series is for shallow data structures and the data frame is for complex data structures. The statement ‘import pandas’ of the code demands python to bring the data analysis library of pandas into the current environment. The portion ‘as pd’ of the code then demands python to provide pandas with the alias of pd. This allows us to use pandas library functions by just typing pd.function_name rather than using pandas.function_name.

c) Matplotlib

Matplotlib is one of the most popular and ancient plotting libraries in Python that is used in Deep Learning. In the Deep learning model, this is helpful to understand the various visualizations through the huge amount of data. ‘matplotlib.pyplot’ is a plotting function of the library that is used for two-dimensional graphics in python programming. ‘matplotlib’ is not a part of the Standard Libraries in python which is installed by default. There are several tool kits which are readily available, that extend python matplotlib functionalities. Some of them are downloaded separately while others can be shipped

along with the matplotlib's source code but has external dependencies.

d) TensorFlow

TensorFlow is a data frame of python library which is used in numerical computation, created by Google. TensorFlow is a foundational library which can be utilized for creating Deep Learning models by using wrapper libraries which can simplify the process built on top of the TensorFlow data frame or directly. TensorFlow is an open sourced library function that can be used for swift numerical computing. Unlike other numerical libraries in python intended for use in Deep Learning such as Theano, TensorFlow was created for use both in development, research and production systems. TensorFlow also produces a collection of workflows to train and develop systems using Python language or JavaScript; and to smoothly deploy in the cloud, in the browser, or on-device, no matter what language we use. It can also run on single CPU systems, mobile devices and as well as GPUs and large-scale distributed systems of several machines.

e) Keras

Keras is one of the easiest and most powerful to use python libraries, that is built on top of popular libraries in deep learning like Theano, TensorFlow, etc., for designing deep learning models. Keras is built on a minimal structure which provides an easy and clean way to manufacture deep learning models based on Theano or TensorFlow. Keras is created to quickly detail deep learning models. Thus Keras is a classic option for deep learning applications. Keras is a highly dynamic and powerful framework and provides us with the following advantages.

- Large community support.
- Easy to experiment.
- Keras neural networks are written using Python which makes everything simpler.
- Keras supports both recurrent and convolution networks.
- Deep learning models are distinct components, so, we can combine them in several ways.

f) CV2

OpenCV is a useful tool for processing images and performing tasks of computer vision. This is an open-source library function which can be used for performing tasks like objection tracking, face detection, landmark detection and many more. This supports multiple languages inclusive of python. This also allows us to get our models trained using customized deep learning libraries or tools and then reliably use them directly in our OpenCV scripts.

B. Dataset

Data will be stored by using google drive, in which two sets of data are stored in two different folders that are used as the sample data for the model. One set of data contains the images with melanoma skin cancer whereas other sets of data containing the symptoms which are not actual skin cancer but some other allergic reactions will be stored in the train folder and another set of the same data will be stored in the validation part.

1)Flow Diagram

Fig 1 describes the flow of the model of detection of Melanoma Skin Cancer using Deep Learning.

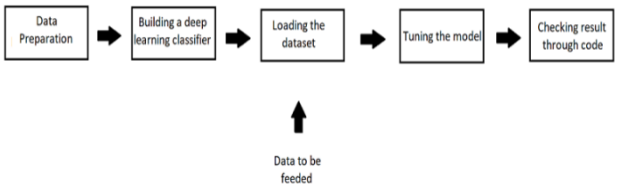

Fig. 1 Flow Diagram

a) Data preparation

Two sets of data will be prepared for the case of skin cancer and not skin cancer and they will be stored for future use.

b) Building a Deep Learning classifier

A code will be created such that it is free from all types and errors and used readily.

c) Loading the datasets

We use the drive to save the data set. Drive will be mounted to the code so that, it can access all the datasets.

d) Tuning the model

Dataset will be fed to the code from training and the validating module and code will be trained again and again for better accuracy.

e) Checking the result

Pictures to be checked will be used on the test folder so that, they can be compared with both the data and results will be provided accordingly.

2) Working

The process of the system starts with the preparation of the data set from available sources like Kaggle. Kaggle is the website where we can find data sets for the process which involves Artificial Intelligence applications like Deep Learning and Machine learning. Two sets of data will be created from the available sources. One of the data sets which is created involves the pattern in which the actual problem correlates with the given input, which means the previous collection of data for the same

problem that occurred in different forms. Another set of data involves the opposite of the first set of data which means, the case given as input is appeared to be the same as the actual problem but it is not actually the same one which is indeed difficult to figure out. Here, we will be considering the data with the symptoms of melanoma skin cancer which is the first set of data and skin problems other than cancer like allergy will be treated as the second set of data. We will create a folder in **google drive** which consists of three sub-folders namely train, validation and test in which, the sample pictures that are collected will be placed in the train and validation folder. The data sets which are stored in the train and validation folder will be used in the code for the process of comparison. Code for the working of a deep learning system is created in such a way that it is free from all kinds of error and one which can be readily used. Here we will be using the python programming language which is easy to write, interpret, understand and debug the error. The algorithm which is selected to execute our system is Convolutional Neural Network, which extracts images as input, processes them and produces the output. This neural network attempts to stimulate the behaviour of the human brain, though far from matching, its ability is to learn from a large amount of data. Sample pictures which are placed in train and validation folders will be fed to our system which is referred to as the process of tuning the model. In the tuning process, the images will be sampled for better accuracy. As the amount of sample images increases the accuracy of the system of finding melanoma cancer will also increase. Libraries like **Matplotlib, Numpy, Pandas and TensorFlow, Keras** will be imported into the code. We will mount our drive in the google colab so that it will be easy for us to access the data set. The location of required folders will be placed at the correct position in the code. The input images which should be checked for the case of melanoma skin cancer will be placed in the test folder. The images in the test folder will be compared with the sample data which was stored in the validation folder. The given input image will be compared with each and every data that is available in the given source in order to give the most accurate output. In our system, images will be sampled 30 times thereby increasing the accuracy of the output by 1.66% from the existing model. This is an added advantage of the proposed model.

3)Overall accuracy

we have proposed this system of skin melanoma cancer detection as it has higher accuracy of about **95.24%** in performance than most other conventional detection systems.

4) Advantages

a) In our proposed model, we came up with an idea to provide a solution with less code complexity which is referred to as a lightweight system.
b) Our proposed solution uses simple coding techniques that result in high accuracy.
c) Since we use a deep learning technique, the results of our model will be more accurate than any other machine learning model.

5) Disadvantages

This model requires a large amount of dataset for better accuracy which is not possible in all cases

V. RESULT

A. Graphs

Fig 2 and Fig 3 represent the relationship between loss of the model corresponding to epoch and the accuracy with respect to epoch and it represents the parameter that takes place in our system like loss, accuracy and validation accuracy.

```
plt.plot(model_fit.history['accuracy'])
plt.plot(model_fit.history['val_accuracy'])
plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train', 'val'], loc='upper right')
plt.show()
```

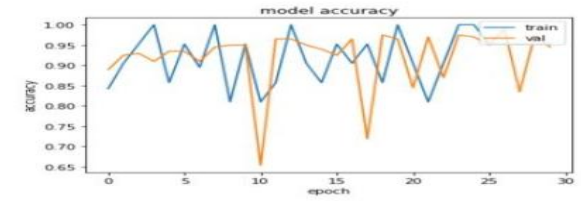

Fig. 2 Accuracy of model

```
plt.plot(model_fit.history['loss'])
plt.plot(model_fit.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'val'], loc='upper right')
plt.show()
```

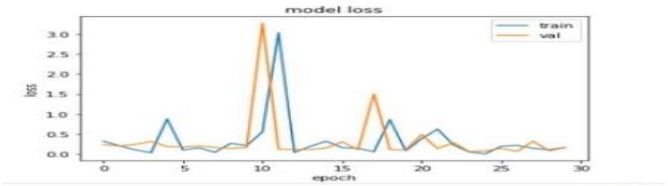

Fig.3 Loss in model

TABLE I Comparison table

S.NO	PARAMETER	PROPOSED MODEL	EXISTING MODEL
1	ACCURACY	95.24%	93.58%
2	COMPUTATIONAL SPEED	15 sec	21 sec
3	LOSS	16.33%	18.73%
4	VALIDATION ACCURACY	94.47%	92.62%

B. Output

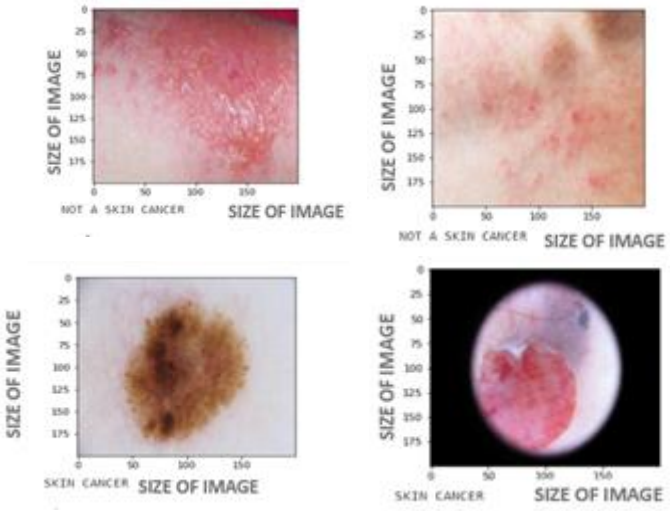

Fig. 4 Output of the model

The above figure describes the output of our system which predict whether the given input pictures as the symptoms of cancer or not. The input picture will be compared with the data set available in our source. If the input picture has the symptoms of cancer means, our system will produce the output as **"SKIN CANCER"**. If the input picture does not contain any symptoms of cancer means, our system will produce the output as **"NOT A SKIN CANCER".**

VI. CONCLUSION

A simple, painless, economical, configurable, adaptable software guidance system is proposed as a support for detecting skin cancer cases and to impart constructive assistance. Our system has been designed, implemented, tested, and verified. This system has high accuracy in performance and is also cost-effective which can fill up the need for affording rapid tests thereby helping people in all walks to detect skin cancer at the initial stages itself. We have proposed this system of skin melanoma cancer detection as it has higher accuracy of about 95.24% in performance than most other conventional detection systems. Moreover, the effects and changes in the accuracy of using multi-layered CNN in skin melanoma cancer detection systems have been observed in this research implementation. Additionally, it also reduces the time in finding the result, as our system can provide the output almost instantly. The computational speed of the code is observed to be 15 Sec which is less than that of the existing models.

The future aim of our project is to detect the input image and produce the output that describes the type of the skin cancer and the level of the skin cancer using deep learning model with improvised accuracy and with the decremented value of loss in the model .

References

[1]Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challengesT Saba - Journal of Infection and Public Health, 2020 – Elsevier

[2]Melanoma skin cancer detection method based on adaptive principal curvature, colournormalisation and feature extraction with the ABCD ruleDNH Thanh, VB Prasath, LM Hieu, NN Hien - Journal of Digital Imaging, 2020 – Springer

[3]A novel end-to-end deep learning approach for cancer detection based on microscopic medical images M Hammad, M Bakrey, A Bakhiet - Biocybernetics and , 2022 – Elsevier

[4]Skin cancer classification using convolutional neural networks: systematic review TJ Brinker, A Hekler, JS Utikal, N Grabe - Journal of medical , 2018 - jmir.org

[5]Melanoma skin cancer detection using deep learning and advanced regularizer MA Hossin, FF Rupom, HR Mahi - 2020 International , 2020 - ieeexplore.ieee.org

[6]Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering M Nawaz, Z Mehmood, T Nazir - Microscopy , 2022 - Wiley Online Library

[7]Classification of melanoma and nevus in digital images for diagnosis of skin cancer MQ Khan, A Hussain, SU Rehman, U Khan - IEEE , 2019 - ieeexplore.ieee.org

[8]DUNEScan: a web server for uncertainty estimation in skin cancer detection with deep neural networks B Mazoure, A Mazoure, J Bédard, V Makarenkov - Scientific Reports, 2022 - nature.com

[9]Recent advances in deep learning applied to skin cancer detection AGC Pacheco, RA Krohling - arXiv preprint arXiv:1912.03280, 2019 - arxiv.org

[10]Dermatologist-level classification of skin cancer with deep neural networks A Esteva, B Kuprel, RA Novoa, J Ko, SM Swetter - nature, 2017 - nature.com

[11]Deep learning can improve early skin cancer detection AA Mohamed, WA Mohamed, AH Zekry - International JoDetection of Cancer in Medical Images using Deep Learning. E Shweikeh, J Lu, M Al-Rajab - International Journal of , 2021 - search.ebscohost.comurnal of Electronics, 2019 - ijet.pl

[12]Skin lesion analysis towards melanoma detection using deep learning network Y Li, L Shen - Sensors, 2018 - mdpi.com

[13]Designing deep learning studies in cancer diagnostics A Kleppe, OJ Skrede, S De Raedt, K Liestøl - Reviews Cancer, 2021 - nature.com

[14]Proposition of convolutional neural network based system for skin cancer detection EC Adjobo, ATS Mahama, P Gouton - 2019 15th International , 2019 - ieeexplore.ieee.org

[15]Effects of label noise on deep learning-based skin cancer classification A Hekler, JN Kather, E Krieghoff-Henning - Frontiers in, 2020 - frontiersin.org