A CNN-based Melanoma Skin Cancer Detection and Classification Approach

BY

FARZANA AKTER
ID: 221-25-121

This report is presented in partial compliance with the Qualifications Requirements for Computer Science and Engineering.

Supervised By

Professor Dr. Md. Ismail Jabiullah
Professor
Department of CSE
Daffodil International University

Co-Supervised By

Mr. Md. Sadekur Rahman
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2023

APPROVAL

This Project/Thesis titled **"A CNN-based Melanoma Skin Cancer Detection and Classification Approach"**, submitted by Farzana Akter, ID No: 221-25-121 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 17-01-2023.

BOARD OF EXAMINERS

Chairman

Dr. Touhid Bhuiyan, PhD
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Ms. Nazmun Nessa Moon
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Fizar Ahmed
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Md. Safaet Hossain
Associate Professor & Head
Department of Computer Science and Engineering
City University

DECLARATION

We therefore make declaration that this work has been done by us under the watchful eye of, **Professor Dr. Md. Ismail Jabiullah, Professor** in the Department of CSE Daffodil International University. We also announce that neither this project nor any part of this project has been relocated to be awarded any degree or diploma.

Supervised by:

Professor Dr. Md. Ismail Jabiullah
Professor
Department of CSE
Daffodil International University

Co-Supervised by:

Mr. Md. Sadekur Rahman
Assistant Professor
Department of CSE
Daffodil International University

Submitted by:

Farzana Akter
ID: 221-25-121
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

At first, I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes me possible to complete the thesis successfully.

I really very grateful and wish my profound my indebtedness to **Professor Dr. Md. Ismail Jabiullah**, Professor, Department of CSE Daffodil International University, Dhaka. Deep knowledge and keen interest of our supervisor in the field of "Deep Learning and Machine Learning" to carry out this thesis. His scholarly guidance, patience, constructive criticism, motivation, constant and energetic supervision, continual encouragement, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this thesis.

I would also like to thank our co-supervisor **Md. Sadekur Rahman**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. When I face any problem, he helped me with valuable ideas and suggestions. He motivated me and help me to complete this work.

I would like to express our heartiest gratitude to **Professor Dr. Touhid Bhuiyan, Professor & Head,** Department of CSE, for his motivation and appreciation. I am also very thankful to other faculty members and the staff of CSE department of Daffodil International University.

Finally, I am very thankful to our parents and friends who were always motivate and criticize my work in a manner to improve my work. At least I thank all of them from the core of my heart.

ABSTRACT

Among various classes of skin cancer, Melanoma is a perilous pattern of skin cancer. Melanoma, widely familiar as malignant melanoma begins in cells which are called melanocytes. From ancient times, people are affected more by that right now. To overcome the complementary problem easier, need to accomplish Melanoma detection as soon as earlier. According to the keen observance and larger analysis of melanoma, CNN achieves better performance both for detection and classification efficiently, specifically deep learning feature-based Convolutional Neural Network, which has the automatic proficiency of skin cancer detection. The proposed method classifies melanoma into two classes, namely Malignant and Benign Melanoma, based on multitasking python libraries. In this research work, the process is come to an end by using the novel CNN model, working with both the training dataset at first and the testing dataset later which has been taken from kaggle platform and is publically available for 10000 images. In the report of the skin cancer image dataset, the experimental results demonstrate a higher level of accuracy rate from the image classifier.

TABLE OF CONTENTS

LIST OF FIGURES

LISTS OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Skin is one of the extensive organs that is surrounded by the entire body part. It consists of three layers each layer contains Epidermis, Dermis, and hypodermis. It works as a defensive shelter as opposed to infection, heat, and light. In other words, it is also known as a sensory organ with performing among the organism and its environment. But nowadays, the skin is at risk because deadly skin cancer like Melanoma is responsible for malignant tumors. Melanocytes are the foundation of melanin which produces melanin reason for providing skin its color. It may be in various colors like red, blue, pink, and purple. Being pigmented, it is hard to diagnose sometimes. The disease displays non-healing ulcer. The reason for occurring melanoma is (UV) radiation reflects directly from the sun. A mole is another reason for appearing melanoma which generally looks different from other moles that remain in the skin. For white skin persons, it remains to be ten times the chance to get melanoma [1]. It needs to be urgently examined for any suspicious lesions that can reduce into a less range. With the advancement of skin injury the disease reflected and changed in shape, size, color, and texture [2]. Cancer can be handled suitably if it is detected as before, and also will help the patient to get recovery soon. Today, CNN also well known as Convent has the automatic capability of disease detection together with classification. To work with large-scale datasets, the much higher accuracy rate and performance are shown in the model. The study is divided into various sections whereas section 2 discusses the literature review of this entire study, the proposed methodology is displayed in section 3, and experimental details and the deep learning feature-based model discusses in section 4. Finally, the results and discussion and comparative study and advantages were carried out in sections 5 and 6 including section 7 the conclusion of the research work. The original goal of this work is to detect melanoma as fast as an efficient procedure.

1.2 Motivation

- The proposed system is reliable.
- We can implement the system for the analysis of different datasets.
- To classify skin lesions we can assist the physician.
- The systems represent the classification performance the most greatly.

1.3 Objectives

A pre-trained CNN analyzes the SPL in the image. Determine how suspicious these lesions are, label them, and classify them using the label. CNN achieves comparable performance to all experts tested on both tasks (detection of melanoma and other types of cancer). Demonstrates artificial intelligence capable of classifying skin cancers at a skill level comparable to that of a dermatologist.

1.4 Research Questions

- To detect Melanoma why CNN approach is so important?
- What is the outcome of our experiment?

1.5 Expected Outcome

The expected outcome of our experiment is shown below.

- Use pre-trained CNN to classify images
- Detect and recognize an object
- Explore train and test image set
- Establish a more desirable accuracy rate from the suggested algorithm

CHAPTER 2

LITERATURE REVIEW

2.1 Literature Review

In this section, there are a lot of numerous observation is reviewed. For improvement and correction in a few preprocessing steps, useless raw image features are remaining [3]. To recognize the physical attributes of melanoma, it is commonly called ABCD skin cancer [4]. ABCD represents the classification performance of Melanoma. Both color feature and texture are executed for melanoma detection. Our target distinguish from previous work. The study shows a publically available dataset taken from ISIC (International Skin Imaging Collaboration). They used a dataset containing 900 images. The whole dataset exists from different sources. Regarding classification accuracy with sensitivity and specificity, our method is one of the most powerful approaches [5]. Allows light to penetrate deep into the skin, this imaging technique reveals most pigment structures together [2]. The classification tree brings out the calculated outcome. Segmentation techniques, feature computation, and finally recognition are included in the presented method [6]. In this work, the ISIC dataset has achieved 91.2% accuracy. As feature generators, ANN, VGG, AlexNet, and ResNet18 are commonly used. The work demonstrates preprocessing steps. The training dataset images are 2800. Due to the efficient use of labeled data rather that unlabeled data it deals with the classification of skin scratch problem [7]. The occurring stage of melanoma has made the highest record in the entire world. Melanoma displays 2% mortality for each annum. The symptoms of melanoma that happen in European countries is 3% among all categories of new cancer. Mostly young patients, also with a middle-aged diagnosis impacting melanoma incidence is rising to 57-64 in the whole world [8]. In 2020, there were over 150,000 new cases of skin melanoma [9]. Choose clothing with a UV protection factor (UPF) on the label for sun protection [10].

CHAPTER 3

PROPOSED METHODOLOGY

3.1 Block Diagram

The proposed system methodology was analysed with the following major phases:

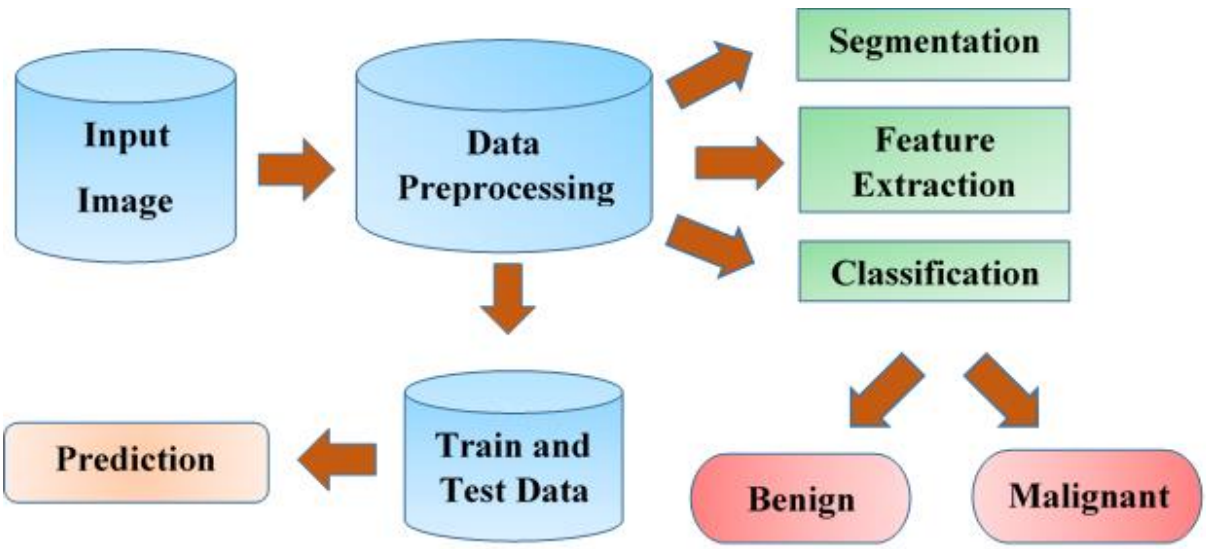

Figure 3.1.1: Block Diagram of the Proposed Model

3.2 Input Image

Input image denotes input as an image.

3.3 Data Preprocessing

In our strategy, to make sure a better outcome data preprocessing step is the precondition before training the dataset. This becomes the initial phase of our experiment, performing genuine results. It needs to be mandatory and clear that the image has been taken. Cleaning, instance selection, and transformation include preprocessing stage. There are many raw data around us in the present world and the following steps convert them into a building process. The data preprocessing steps are-

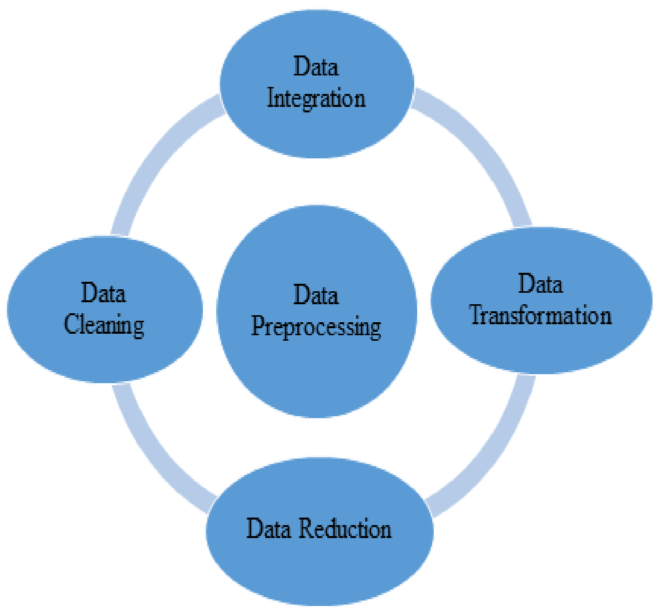

Figure 3.3.1: Data Preprocessing

3.4 Segmentation

Segmentation is referred to two tasks involving similarities and discontinuous. Similarities through the divided images whereas discontinuous performs through the fixed criteria based on divided images.

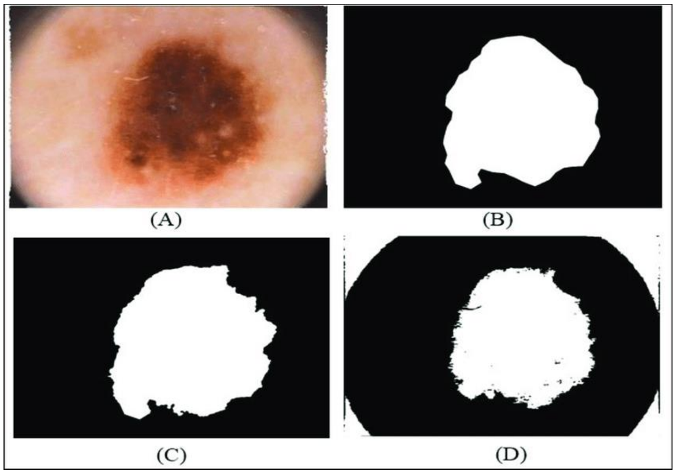

Figure 3.4.1: Segmented Images of Melanoma Skin Cancer

3.5 Feature Extraction

As ABCDE helps to detect Melanoma are stands with given phases:

Figure 3.5.1: ABCDEs of Melanoma Skin Cancer

1. **Asymmetry:** One half differs from the presence of another half.
2. **Border irregularity:** The border becomes notched, blurred, or ragged.
3. **Color:** The color may be brown or black. Gray, pink, white, or blue areas may also be present.
4. **Diameter:** Mole size is 6mm or more exists as an eraser. Mole growth should be assessed.
5. **Evolving:** The mole turns into another shape and size. It causes the skin area to break down.

Strategies both for Feature Extraction and Classification:

Training Period:

- Read the dataset as the training set.
- Preprocess the dataset
- Segment the wound from the given dataset images
- Implement with CNN model
- Identify it as Melanoma Skin Cancer

Testing Period:

- Read the dataset as a testing set
- Put in both training and testing dataset
- Classify Melanoma Skin Cancer

3.6 Train and Test data:

We enquire about both training and testing data using an epoch. epoch derives all training datasets at once.

CHAPTER 4

EXPERIMENTAL DETAILS

4.1 Dataset Description

The experimental Melanoma Skin Cancer dataset is publically available for 10000 images. We have taken the dataset from Kaggle platform. As it is a fatal disease, it may cure with early detection. The dataset is helpful for accurate melanoma skin cancer classification. The images belong to both Benign and Malignant skin cancer. The entire dataset divided into two categories: 9000 for the training dataset and 1000 for the testing dataset.

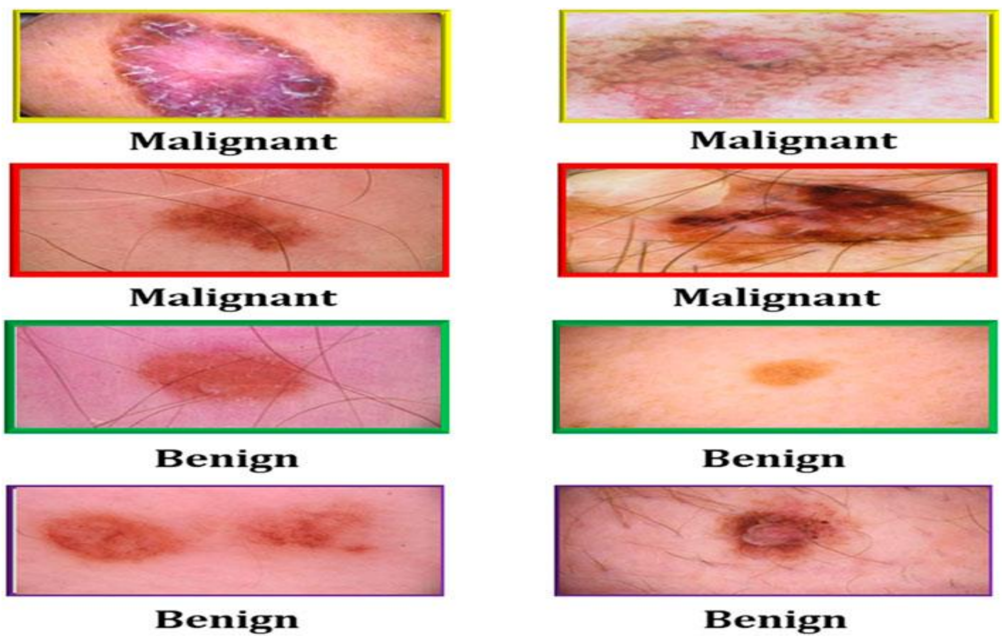

Figure 4.1.1: Sample Image of Melanoma Skin Cancer

To obtain greater accuracy we have used the CNN model. The model classifies Benign and Malignant skin cancer.

We clarify the necessary hardware and software, advanced libraries, and required tools to fulfill our work.

Hardware and Software:

- Web IDE: Google Colab
- Hard Disk
- High Speed Internet Connections

Advanced Libraries:

- Python
- Pandas
- NumPy
- Matplotlib

Required Tools:

- Windows 10
- Dataset: From Kaggle Platform
- Algorithm: CNN (Convolutional Neural Network)

4.2 CNN Model

The following figure shows the melanoma Skin Cancer Detection Model:

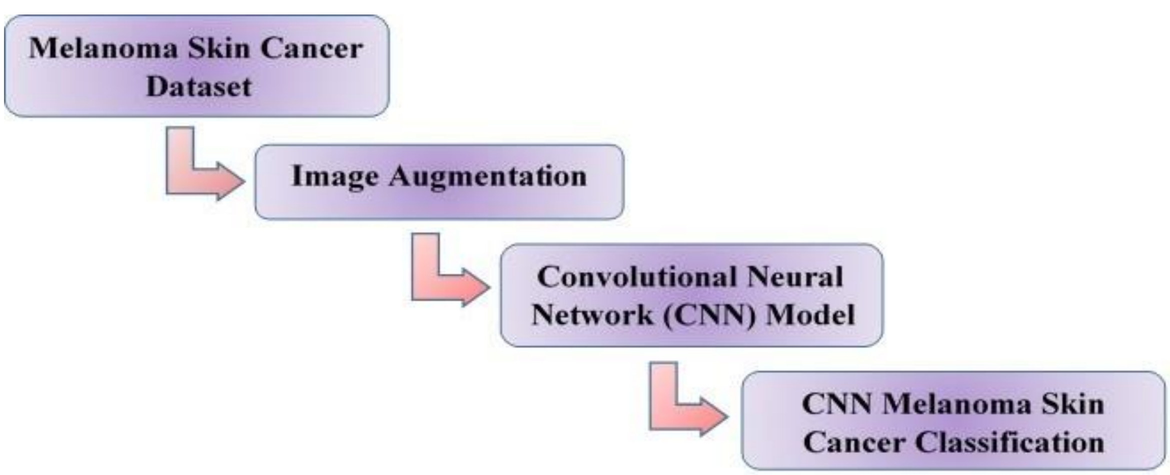

Figure 4.2.1: Melanoma Skin Cancer Detection Model

4.3 CNN Layers

As we implement our experiment with the CNN model that is associated with some layers involving-

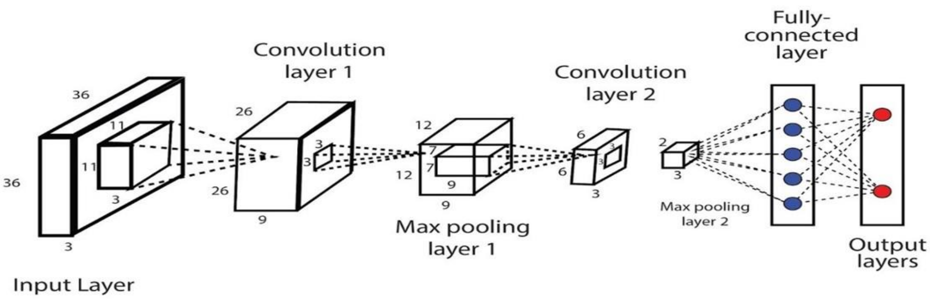

Figure 4.3.1: Layers of CNN

- **Input layer:** The input layer contains image data as input. Image data belong to a 3-dimensional matrix.
- **Convolution layer:** It is called the original blockchain of CNN while performing kernels is widely known as a set of filters. The filters act a little size. A lot of patches around this layer.
- **Max pooling layer:** To select the broadest element the layer is being utilized. The Max pooling layer is better if we compared it with the average max pooling layer.

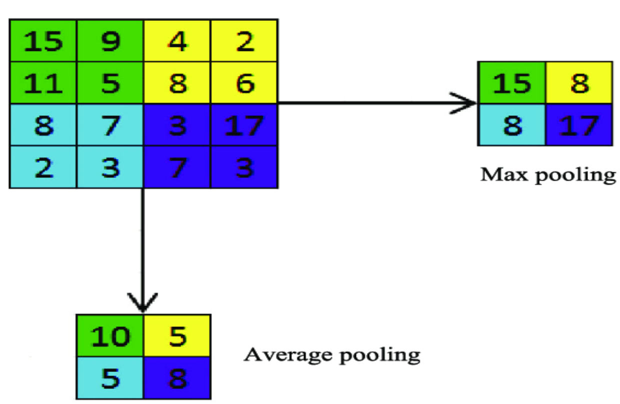

Figure 4.3.2: Comparison of Max Pooling and Average Pooling

- **Fully Connected layer:** The last layer of CNN is also called the dense layer.

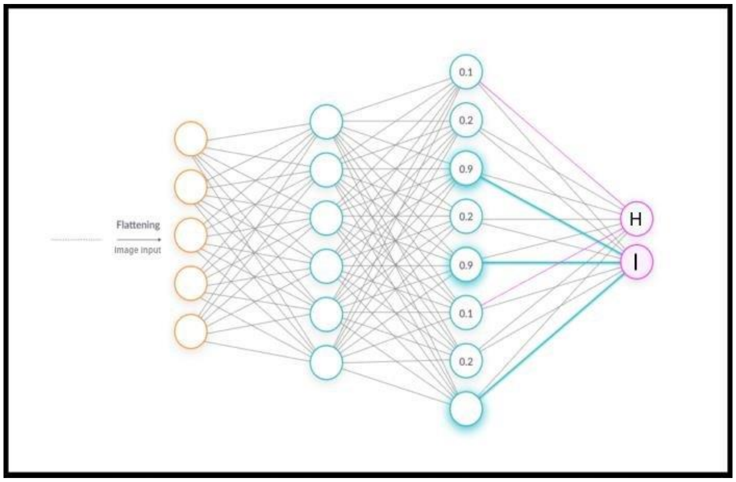

Figure 4.3.3: Fully Connected Layer

- **Output Layer:** Holds the final classification results.

4.4 CNN Architecture

CNN architecture is an architecture of deep learning which learns directly from the data. Visual imagery has been used more.

Model Content	Details
First Convolution Layer	128 filters of size 3x3, ReLU
First Max Pooling Layer	Pooling Size 2x2
First Dropout Layer	0.2
Second Convolution Layer	64 filters of size 3x3, ReLU
Second Max Pooling Layer	Pooling size 2x2
Second Dropout Layer	0.2
Fully Connected Layer	128 nodes, ReLU, Dropout=0.2
Output Layer	50 nodes for 50 classes, SoftMax

Table 4.4.1: CNN Architecture

CHAPTER 5

RESULT AND DISCUSSION

5.1 Confusion Matrix

An special kind of contingency table which is produced from a simple CNN application. The matrix expresses true positive, true negative, false positive and false negative.

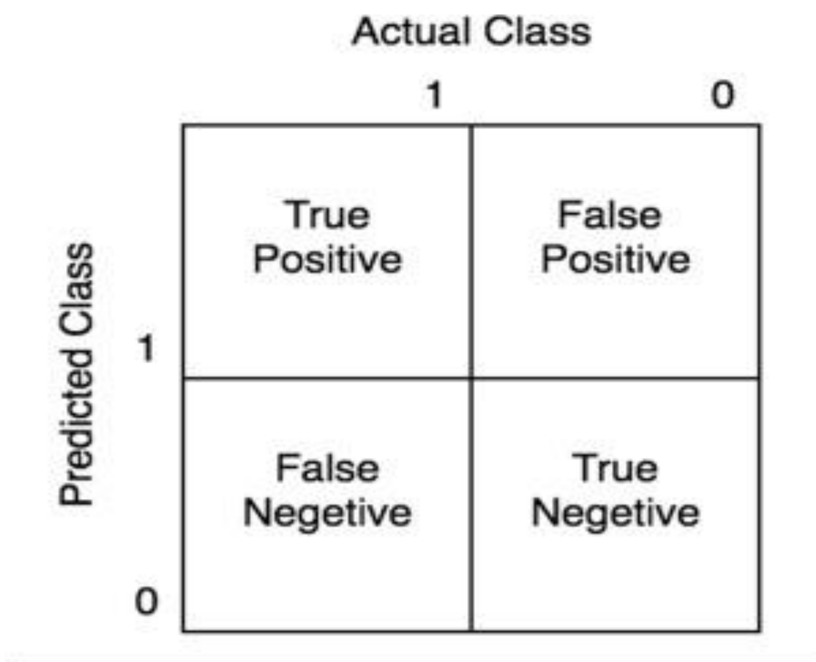

Figure 5.1.1: Confusion Matrix

Other accuracy matrices are precision, f1 score, and recall present in the model. Precision and recall instruct true positive and true negative while the f-1 scoreis also known as the f1 measure generating the same weightage based on precision and recall.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\text{-}score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Figure 5.1.2: Accuracy Matrices

5.2 Model Accuracy

We use Matplotlib.Pyplot function for measuring accuracy. The graph indicates the model accuracy of our experiment.

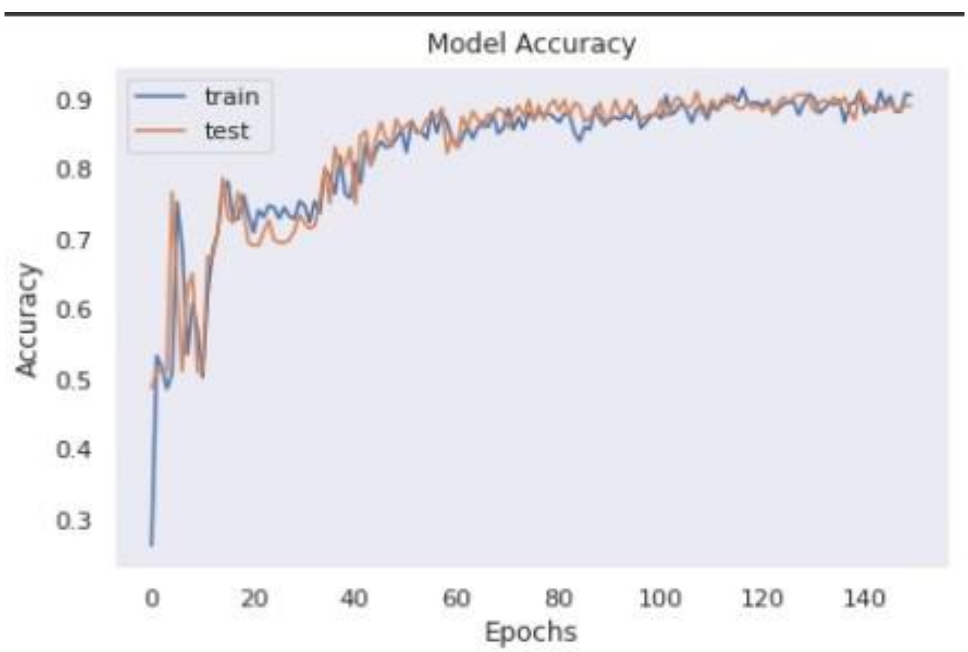

Figure 5.2.1: Model Accuracy

5.3 Model Loss

Another graph points out the testing performance of CNN.

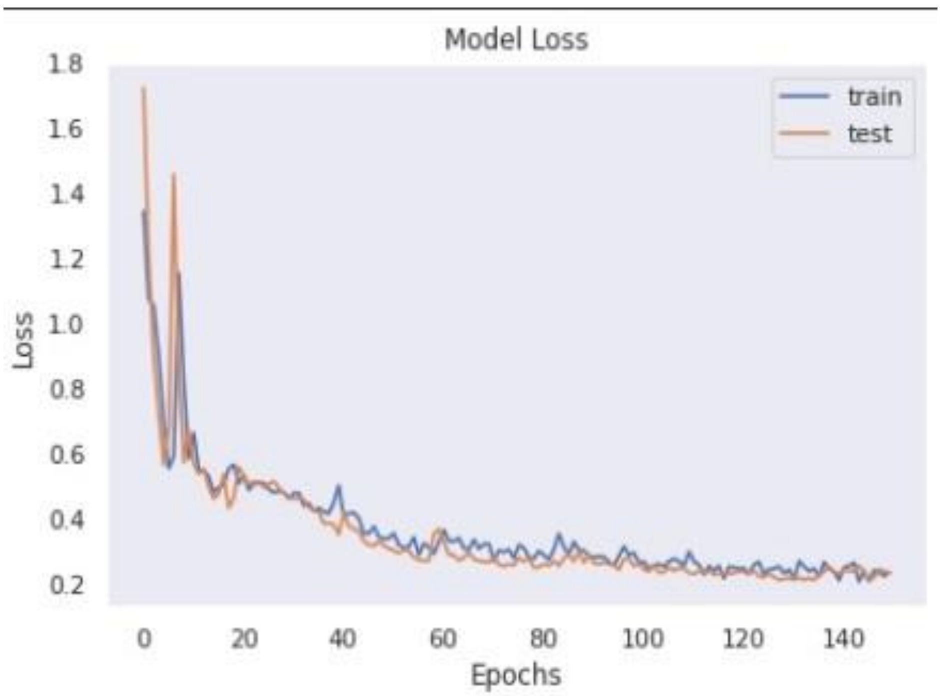

Figure 5.3.1: Testing Performance of CNN

Some of code screen shots:

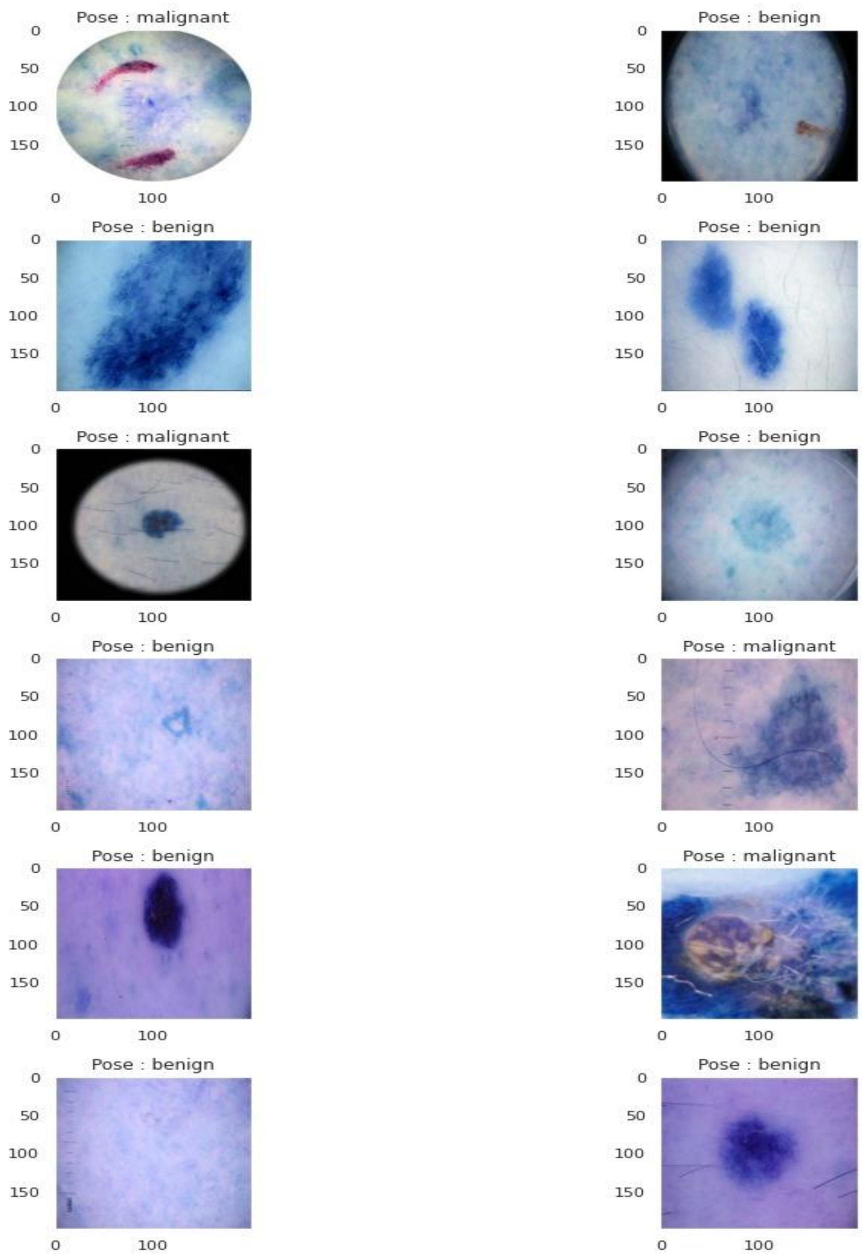

Figure 5.3.2: Some Pictures of Code

```
[7]         path = os.path.join(DIR,img)
            img = cv2.imread(path,cv2.IMREAD_COLOR)
            img = cv2.resize(img, (IMG_SIZE,IMG_SIZE))

            X.append(np.array(img))
            Z.append(str(label))
```

```
[8]  make_train_data('benign',benign)
     print(len(X))

     100%|██████████| 500/500 [00:09<00:00, 51.03it/s] 500
```

```
[9]  make_train_data('malignant',malignant)
     print(len(X))

     100%|██████████| 500/500 [00:11<00:00, 44.48it/s] 1000
```

Figure 5.3.3: Some Pictures of Code

```
[19] model.add(Dense(6, activation = "softmax"))
```

```
[20] batch_size=280
     epochs=80
     from keras.callbacks import ReduceLROnPlateau
     red_lr= ReduceLROnPlateau(monitor='val_acc',patience=3,verbose=1,factor=0.1)
```

```
[21] datagen = ImageDataGenerator(
             featurewise_center=False,
             samplewise_center=False,
             featurewise_std_normalization=False,
             samplewise_std_normalization=False,
             zca_whitening=False,
             rotation_range=10,
             zoom_range = 0.1,
             width_shift_range=0.2,
             height_shift_range=0.2,
             horizontal_flip=True,
             vertical_flip=False)


     datagen.fit(x_train)
```

Figure 5.3.4: Some Pictures of Code

```
[24] model.summary()

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 200, 200, 32)      2432

 max_pooling2d (MaxPooling2D  (None, 100, 100, 32)     0
 )

 conv2d_1 (Conv2D)           (None, 100, 100, 64)      18496

 max_pooling2d_1 (MaxPooling  (None, 50, 50, 64)       0
 2D)

 conv2d_2 (Conv2D)           (None, 50, 50, 96)        55392

 max_pooling2d_2 (MaxPooling  (None, 25, 25, 96)       0
 2D)

 conv2d_3 (Conv2D)           (None, 25, 25, 96)        83040

 max_pooling2d_3 (MaxPooling  (None, 12, 12, 96)       0
 2D)

 flatten (Flatten)           (None, 13824)             0
```

Figure 5.3.5: Some Pictures of Code

```
[29] Y_pred = model.predict(x_test)
     print(Y_pred)

8/8 [==============================] - 7s 817ms/step
[[8.3141178e-01 1.6858374e-01 8.1453095e-07 7.2238629e-07 1.9261872e-06
  1.0818666e-06]
 [8.0945498e-01 1.9054440e-01 7.3622076e-08 1.2019029e-07 2.2737461e-07
  1.9313252e-07]
 [8.1116876e-14 9.9999994e-01 0.0000000e+00 0.0000000e+00 3.5925497e-37
  3.0248476e-38]
 ...
 [9.8132741e-01 1.8672507e-02 1.3311847e-11 2.7104884e-11 7.7503240e-11
  5.9872038e-11]
 [9.8025399e-01 1.9746006e-02 2.7642232e-11 5.3724483e-11 1.5186072e-10
  1.1514609e-10]
 [7.5367814e-01 2.4632192e-01 2.5178206e-09 2.0444999e-09 1.7182307e-08
  4.2015440e-09]]
```

Figure 5.3.6: Some Pictures of Code

CHAPTER 6

COMPARATIVE STUDY AND ADVANTAGES

6.1 Comparative Study

No.	Existing Model	Our Model
1	Existing model discovers only investigate mistakes.	Our model discovers preparing exactness, and exactness.
2	Maximum research based on ANN algorithm which is considered to be less powerful than CNN.	Our model is based on CNN algorithm which is considered to be more powerful than ANN.
3	The accuracy of current model is 86-90%	Our accuracy is 97%.

Table 6.1.1 Comparative Study

6.2 Advantages:

1. Our model is unexpectedly intuitive
2. Easy to use
3. Illustrates complex solution easier
4. Extremely productive for dataset
5. Accuracy is so high

CHAPTER 7

CONCLUTION, LIMITATIONS AND FUTURE SCOPE

7.1 Conclusion and Future Scope

From the above study, we can detect earlier melanoma skin cancer accurately by using the CNN model which detects benign and malignant melanoma. We obtained 97% accuracy. Being absent of required hardware, we used a few numbers of an epoch. We aim to build a clear relationship between melanoma and skin cancer so that we can generate a strong network to detect more categories of skin cancer in the future, Meanwhile, this model ensure to make that technology-based application in order that anyone can easily identify skin cancer symptoms without moving to another place.

7.2 Limitations

We have some drawbacks on our work. We use deep learning algorithms in our work. We can't improve the accuracy of a deep learning model without increasing the amount of data.

REFERENCES

[1] Diwakar Gautam and Mushtaq Ahmed. Melanoma detection and classification using svm based decision support system. In 2015 Annual IEEE India Conference (INDICON), pages 1–6. IEEE, 2015.

[2] JC Kavitha, A Suruliandi, D Nagarajan, and T Nadu. Melanoma detection in dermoscopic images using global and local feature extraction. International Journal of Multimedia and Ubiquitous Engineering, 12(5):19–28, 2017.

[3] Ashwin R Jadhav, Arun G Ghontale, and Vimal K Shrivastava. Segmentation and border detection of melanoma lesions using convolutional neural network and svm. In Computational Intelligence: Theories, Applications and Future Directions-Volume I, pages 97–108. Springer, 2019.

[4] Samy Bakheet. An svm framework for malignant melanoma detection based on optimized hog features. Computation, 5(1):4, 2017.

[5] Majtner T, Yildirim-Yayilgan S, Hardeberg J Y. Combining deep learning and hand-crafted features for skin lesion classification[C]//2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA). IEEE, 2016: 1-6.

[6] Khan, Muhammad Attique, Muhammad Sharif, Tallha Akram, Syed Ahmad Chan Bukhari, and Ramesh Sunder Nayak. "Developed Newton-Raphson based deep features selection framework for skin lesion recognition." Pattern Recognition Letters 129 (2020): 293-303.

[7] Masood A, Al-Jumaily A, Anam K. Self-supervised learning model for skin cancer diagnosis[C]//2015 7th International IEEE/EMBS Conference on Neural Engineering (NER). IEEE, 2015: 1012-1015.

[8] Florentia Dimitriou, Regina Krattinger, Egle Ramelyte, Marjam JBarysch, Sara Micaletto, Reinhard Dummer, and Simone M Goldinger. The world of melanoma: epidemiologic, genetic, and anatomic differences of melanoma across the globe. Current oncology reports, 20(11):87, 2018.

[9] https://www.wcrf.org/dietandcancer/cancer-trends/skin-cancer-statistics

[10] https://www.aad.org/media/stats-skin-cancer

A CNN-based Melanoma Skin Cancer Detection and Classification Approach

ORIGINALITY REPORT

3%	2%	2%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	xajzkjdx.cn Internet Source	1%
2	intubatedangel.tumblr.com Internet Source	1%
3	research.ijcaonline.org Internet Source	1%
4	mdpi-res.com Internet Source	<1%
5	Priti Bansal, Ritik Garg, Priyank Soni. "Detection of melanoma in dermoscopic images by integrating features extracted using handcrafted and deep learning models", Computers & Industrial Engineering, 2022 Publication	<1%
6	JaneY. Nancy, Sk Charanya, M. Amsaprabhaa, Preetiha Jayashanker, Khanna Nehemiah H. "2-HDCNN: A two-tier hybrid dual convolution neural network feature fusion approach for	<1%