[1]Ashish Tripathi
[2]Rajnesh Singh
[3]Rajani Singh
[4]Anubhi Bansal
[5]Amit Kumar
[6]Shivam Agarwal

Melanoma Classification Using Deep Learning

Abstract: - Melanoma is a life-threatening and serious skin cancer with an increasing number of fatal cases worldwide. Early detection of Melanoma is of vital importance because the survival of the patient is based on the stage of the disease. In the first stage survival rate is more than compared to further stages. Melanoma is a life-threatening and serious skin cancer type with an increasing number of fatal cases worldwide, early detection of Melanoma is of vital importance because of its survival rates at different stages that is starts from 1st stage and 98.4% at 5th stage. With the advancement in technology, early detection is possible through several applications which uses the concepts of machine learning and deep learning. In this paper, a convolution deep learning model named MelaNet is designed from scratch to classify benign and malignant type of melanoma skin cancer. MelaNet achieved the high accuracy of 92.67% to classify the skin lesion accurately. One of the main differences between MelaNet and other existing models is that MelaNet's parameters are optimally specified and tuned, which aids in avoiding over/under-fitting.

Keywords: Melanoma, Machine Learning, Deep Learning, Cancer Detection and Classification

I. INTRODUCTION

Melanoma is a form of widely spread skin cancer that arises in cells responsible for producing skin pigment [1] [2]. It can manifest on any body part but is most frequently observed on sun-exposed areas like the face, legs, chest, and back [3]. The disease can result from genetic factors, as well as exposure to UV radiation via tanning beds or sunlight. Due to the ability to spread to lymph nodes and other regions of the body, melanoma is the most lethal type of skin cancer. There are five stages of melanoma skin cancer given below which includes different layers of skin that is Epidermis, Dermis and Subcutaneous Tissue [4].

Stage 1: Melanoma is limited to epidermal layer of the body skin.
Stage 2 is the thin and localized illness solely affects the skin.
Stage 3 is a thicker, and more localized illness than stage 2.
Stage 4 is the cancer has spread to the lymph nodes.
Stage 5 is spread to other organs.

Early detection and treatment are important for improving the prognosis of melanoma. Melanoma can often be detected early by performing regular skin self-exams and having any suspicious moles or lesions evaluated by a healthcare professional. Some warning signs of melanoma include a mole or spot that changes in size, shape, or color, or one that bleeds or becomes itchy. If melanoma is detected early, it can often be treated with surgery to remove the cancerous cells. In some cases, additional treatments such as chemotherapy, radiation therapy, or immunotherapy may also be recommended. The earlier melanoma is detected and treated, the greater the chances of a successful outcome.

Melanoma can be either benign or malignant. Benign melanomas are not cancerous and typically do not require treatment. On the other hand, malignant melanomas are carcinogenic and can spread to other parts of the body if they are not treated quickly. The most important factor in determining whether a melanoma is malignant is the depth of the melanoma, which is measured by a pathologist during a biopsy [5]. Other factors that can affect the prognosis of melanoma include the size and location of the melanoma, as well as whether the lymph nodes or other organs nearby that the melanoma has migrated to. Treatment for malignant melanoma typically involves surgical removal of the melanoma, along with any nearby lymph nodes that may be affected. In some instances, other therapies like chemotherapy, radiation therapy, or immunotherapy could also be suggested to help eradicate any cancer cells that might still be present and stop the melanoma from returning [6].

Diagnosis of skin disease can be time-consuming and costly for certain patients. This can be particularly challenging in developing countries, where access to healthcare may be limited and many people may not have the resources to seek medical attention for minor health issues [7]. As a result, some people may ignore the primary symptoms of a skin disease and delay seeking treatment until the disease has progressed. This can make treatment more complex and may lead to poorer outcomes for the patient.

[1] *Corresponding author: Ashish Tripathi, SCSE, Galgotias University, Greater Noida, India
[2] Rajnesh Singh, [5] Amit Kumar, SCSE, Galgotias University, Greater Noida, India,
[3] Rajani Singh, [4] Anubhi Bansal, [6] Shivam Agarwal, Department of IT, G. L. Bajaj Institute of Technology and Management, Greater Noida, India

The application of deep learning for melanoma classification involves training a deep neural network on a dataset comprising skin lesion images to distinguish between benign and malignant melanomas. To achieve this, various techniques are employed, including pre-processing the images, employing data augmentation methods, and utilizing different Convolutional Neural Network (CNN) architectures like VGG, ResNet, or Inception. These architectures enable feature extraction and image classification [8]. The model's performance can be assessed using metrics like accuracy, precision, recall, and F1-score. Improvements can be made by fine-tuning hyperparameters, implementing regularization techniques, or leveraging transfer learning.

Deep learning techniques, such as CNNs, prove to be highly effective in analyzing microscopic images of infected skin and aiding in the diagnosis of skin diseases. CNNs are a type of artificial neural network specifically designed for image analysis tasks. Through extensive training on a large dataset, CNNs can learn to recognize and identify patterns in images. Once trained, they can accurately classify new images based on the patterns they have acquired [9].

This paper aims to build an automatic classification system that classifies the stages of melanoma from images of skin lesions based on image processing technology. Our overall goal is to support efforts to reduce deaths from melanoma skin cancer. The main motivation driving the proposed work is to use advanced image classification techniques for the benefit of the people. Computer vision has greatly advanced scalable machines and deep learning across multiple domains.

A. Objectives of the Proposed Work

The following are the objectives of the proposed work.

- To overcome CNN's constraints.
- To accurately classify melanoma skin diseases.
- To increase the model's acceptance for faster and smoother categorization of melanoma.

B. Overview of the Proposed Work

The proposed model (MelaNet) is made up of a succession of convolutional layers with various filter sizes, followed by pooling layers that extract significant characteristics from the input images. Dropout strategies are used to avoid overfitting and increase model generalization. The model's last layers are dense layers for classification, with a sigmoid activation function used to generate a probability output for each class.

The model architecture initiates with an input layer that receives 100 x 100 x 3 images. Sequentially, two convolutional layers with 16 filters of size 3 x 3 are applied, utilizing the ReLU activation function to introduce non-linearity. Subsequently, a max pooling layer with a pool size of 2 x 2 is employed to reduce the feature map dimensions.

To address overfitting concerns, a dropout layer with a rate of 0.2 follows the first pooling layer. This is succeeded by two additional convolutional layers, each consisting of 32 filters of size 3 x 3, and ReLU activation functions. Another max pooling layer with a pool size of 2 x 2 is implemented to further decrease dimensionality.

Continuing the process, two more convolutional layers are introduced, employing 64 filters of size 3 x 3 and ReLU activation functions. To further reduce dimensionality, a second max pooling layer with a pool size of 2 x 2 is incorporated.

Finally, two additional convolutional layers are applied, utilizing 128 filters of size 3 x 3 and ReLU activation functions. Another max pooling layer with a pool size of 2 x 2 is added for additional dimensionality reduction. To tackle overfitting concerns, a second dropout layer with a rate of 0.4 is included.

After flattening the output, five fully connected dense layers with 128, 64, 32, 16, and 2 neurons respectively are employed. The last dense layer utilizes a sigmoid activation function to generate probability outputs for each class, while the preceding four dense layers utilize ReLU activation functions.

This model architecture leverages a series of convolutional and pooling layers to extract features from the input images. Dropout layers are integrated to address overfitting, and dense layers with varying neuron configurations are utilized for classification, culminating in probability outputs for the respective classes.

The proposed work includes the MelaNet deep learning model which classifies the type of Melanoma Skin Cancer into benign or malignant.

C. Novelty in the Proposed Work Over CNN

The proposed MelaNet is a deep learning model based on the CNN architecture, with a high accuracy of 92.67%. One of the main differences between MelaNet and CNN is that MelaNet's parameters are optimally specified, and seem to help the model in avoiding over- or under-fitting, resulting in a good curve.

Six conv2D layers in MelaNet are followed by the ReLU activation functions and max-pooling layers, which preserve the image's dimension. In contrast to CNN, which uses either max-pooling or average pooling layer, this design has a predetermined upper limit for the number of conv2D layers.

MelaNet includes three fully connected dense layers, each of which is passed a certain set of parameters values in a particular order. In contrast to MelaNet, there is no limit on the number of dense layers that can be utilized in CNN, each with a separate set of parameters' values. While the dropout layer in CNN is optional, MelaNet employs the dropout layer very efficiently to keep the model away from over-fitting.

D. Summary of the Proposed Work

The model starts with an input layer that takes images of size 100x100 with 3 color channels (RGB), two consecutive convolutional layers with 16 filters of size 3x3 are applied, followed by a rectified linear unit (ReLU) activation function. Padding is set to 'same' to ensure the spatial dimensions remain the same, a max pooling layer with a pool size of 2x2 is added to reduce the spatial dimensions by half, and a dropout layer is introduced with a rate of 0.2, which randomly sets 20% of the previous layer's outputs to zero during training to reduce overfitting. Two more sets of convolutional layers with 64 filters each and a size of 3x3 are applied, followed by ReLU activation and max pooling. Another set of convolutional layers with 128 filters of size 3x3 is added, along with ReLU activation and max pooling, a dropout layer with a rate `of 0.2 is inserted after the previous convolutional layers, the feature maps are then flattened to a 1D vector to be fed into the dense layers, four dense layers with an increasing number of units are added:128 unit, 64 unit, 32 unit and 16 unit each activated by ReLU, dropout layers with decreasing rates of 0.2, 0.1, 0.1, 0.1, and 0.1 are added after each dense layer, finally, a dense output layer with 2 units and a sigmoid activation function is added to classify the input images into two classes.

II. BACKGROUND DETAILS

According to three decades of cancer registry data (1982-2011) from six populations (US, UK, Sweden, Norway, Australia, and New Zealand) with moderate to high melanoma cases, age period cohort models were used to describe the current trends and projection of future incidence rates [2]. Among Caucasians people in the United States, the annual number of new cases will increase from about 70,000 in 2007-2011 to 116,000 in 2026-2031. 79% of the increase is due to an increase in age proportions, and 21% is due to population growth and an aging population. The American Academy of Dermatology (AAD) reports that melanoma is the second most common cancer in women aged 20 to 29, with 68,000 new cases diagnosed in the United States in 2010 [2].

Thus, it is now more important than ever to avoid skin disorders. In order to ensure that more people may live healthy lives, effective data-driven methods for predicting melanoma problems can enhance the overall research and preventive process.

In this study, CNN & binary classification issue has been used to calculate the likelihood that a participant will get melanoma skin cancer. Healthy conditions receive a score of 0, whereas all conditions involving the cancer symptom receive a score of 1. We discuss how to score the data gathered in this study in order to evaluate a variety of research hypotheses. The study that has recently been conducted in numerous domains to learn how to forecast melanoma skin cancer is examined.

Convolutional Neural Networks, or CNNs, do exceptionally well in Deep Learning while dealing with real data. These networks can execute pretty complex jobs with photos, text files, audio files, video files, and so forth. Professor Yann LeCunn created the first successful complex networks for Bell Labs in the late 1990s [10].

Convolutional Neural Networks (CNNs) have been proved to be extremely effective in picture recognition and classification. Which includes the ABCDE rule that serves as a helpful for deep learning model in identifying potential indications of skin cancer, specifically melanoma, which is known to be the most dangerous type of skin cancer [11]. The given proposed model is design using the CNN, that is, MelaNet. The scope of a melanoma classification project would typically involve the development of a system or model that can predict the likelihood of disease based on certain risk factors or symptoms.

III. LITERATURE SURVEY

Melanoma classification is a common task in medical image analysis and has been the focus of extensive research. There are several algorithms that can be used for melanoma classification, and the choice of algorithm depends on several elements include the dataset's size and complexity, available computing power, and performance standards. The accuracy of each algorithm for melanoma classification can vary depending on several factors such as the quality and size of the dataset, pre-processing techniques, hyperparameter tuning, and evaluation metrics used. Here, Table 1 provides a brief overview of the reported accuracy of some commonly used algorithms for melanoma classification:

Table 1. Literature Survey

Author/Reference	Algorithms/ Models	Application/Work
Esteva et al. (2017) [8]	CNNs	CNNs have shown high accuracy for melanoma classification in several studies. For instance, a study by Esteva et al. (2017) reported an accuracy of 91% for a CNN model trained on a dataset of over 129,000 clinical images. Other studies have reported similar or slightly lower accuracy, ranging from 80% to 95%. The deep learning model's performance is not generalize well to classify the actual clinical scenarios.

Han et al.(2018)[9]	Random Forests	A study by Han et al. (2018) reported an accuracy of 85.7% for a Random Forest model trained on a dataset of over 2,000 dermoscopic images. The study's dataset only contains a small number of images, which prevents the model from fully capturing the features of melanoma classification.
Giotis et al. (2019) [10]	KNN and K-means	KNN has shown good accuracy for melanoma classification in some studies, but its performance may depend on the choice of k and distance metric. A study by Giotis et al. (2019) reported an accuracy of 81.8% for a KNN model trained on a dataset of over 2,500 dermoscopic images. The performance of KNN model varies based on the dataset, the data preprocessing methods, and the unique features of the melanoma skin images.
Dhawan et al. (2019) [11]	SVM	SVMs have shown good accuracy for melanoma classification in some studies, but their performance may depend on the choice of kernel and hyperparameters. A study by Dhawan et al. (2019) reported an accuracy of 87.5% for an SVM model trained on a dataset of over 2,500 dermoscopic images. The Support Vector Machine (SVM) model's performance is dependent on the precise collection of features used, which may introduce biases and restrict its generalizability to other feature sets or datasets.
Tsantoulis et al. (2020) [12]	ANN	In a recent study conducted by Tsantoulis et al. (2020), an Artificial Neural Network (ANN) model was trained on a dataset consisting of 3,986 dermoscopic images. The results of the study demonstrated an accuracy of 88.75% on the test set. ANN models are renowned for their complicated architecture. While they can do a variety of tasks with great accuracy, it becomes difficult to understand the model's fundamental decision-making process.
Haque et al. (2020) [13]	ResNet	Haque et al. (2020) conducted a study utilizing ResNet for melanoma classification based on dermoscopic images. The dataset consisted of 10,015 dermoscopic images, with 7,025 images allocated for training, 1,245 images for validation, and 1,745 images for testing. The findings of the study demonstrated that ResNet achieved an accuracy of 88.6% on the test set, surpassing the performance of other deep learning models like VGG-16 and Inception-v3. To successfully train and analyze deep learning models like ResNet, a lot of computing power is needed, requiring powerful GPUs and RAM.
Liang et al.(2021)[14]	MobileNetV2	Liang et al. (2021) conducted a study where they proposed a melanoma classification system based on deep learning. They utilized MobileNetV2, a widely used convolutional neural network architecture, as the backbone network for their system. The training and testing of the system were performed on a dataset containing over 20,000 dermoscopic images. The results showed an impressive overall accuracy of 90.5% for the proposed system.
Zhang et al. (2021) [15]	Channel burst LSTM (CB-LSTM)	Zhang et al. (2021) conducted a study on melanoma classification utilizing CB-LSTM, that incorporates channel-wise choosing features to enhance classification accuracy. Dermoscopic images were used in the study, with a dataset comprising 2,000 images, of which 1,000 were allocated for training and 1,000 for testing. Prior to training, the images underwent preprocessing, including normalization and augmentation techniques. The results of the study demonstrated that CB-LSTM achieved an impressive accuracy of 93.5%, surpassing the performance of other

		deep learning models such as CNN and ResNet. This indicates the effectiveness of CB-LSTM for melanoma classification tasks. The study uses the less number of training datatset which lacks the generalizability of the model for the classification of melanoma ski cancer.
Kaur et al. (2021) [16]	ResNet50	Kaur et al. (2021) conducted a study utilizing ResNet50 for melanoma classification. The dataset consisted of 3,740 dermoscopic images, with 2,992 images allocated for training and 748 images for testing. The study reported an impressive accuracy of 91.4% for ResNet50, surpassing the performance of other deep learning models such as InceptionV3 and VGG16. These findings highlight the effectiveness of ResNet50 for melanoma classification tasks.
Lu et al.(2021)[17]	XceptionNet	Lu et al. (2022) conducted a study using the Skin Cancer MNIST: HAM10000 datasets. The study employed the Xception approach on the MNIST skin cancer dataset and compared the results with several state-of-the-art methodologies. It achieved 100% accuracy, 94.05% sensitivity, 97.07% precision, and 95.53% F1 score.
Priyadharshini et al. (2023)[18]	ELM-TLBO	Priyadharshini et al. (2023) conducted a study using the ELM-TLBO algorithm for image classification. The dataset used in the study comprised 300 sample images, with 200 images designated for training and 100 images for testing. The study reported an impressive accuracy of 93.18% for ELM-TLBO. This indicates the effectiveness of the algorithm in accurately classifying the images. The study utilized relatively small dataset comparing with other studies which lacks the comprehensive evaluation of algorithm performance.

IV. MATERIAS AND METHODS

A. Dataset Details

The images in the dataset were gathered from several sources, including Kaggle, IEEEDataPort, and BioGSP. The given dataset was generated by Kaggle-melanoma-skin-cancer-dataset-of-10000-images data set for Classification of Malignant and Benign. The dataset contains demographic Melanoma Skin cancer images of the melanoma unique benign and malignant skin lesions from over approx. 2000 patients. Most of the Malignant images were confirmed via histopathology and benign images were confirmed using either expert agreement or histopathology. The given dataset contains 10,000 images [22]. Melanoma skin cancer is deadly cancer, early detection and cure can save many lives. This dataset will be useful for developing the deep learning models for accurate classification of melanoma. Sample image dataset is shown in Figure 1.

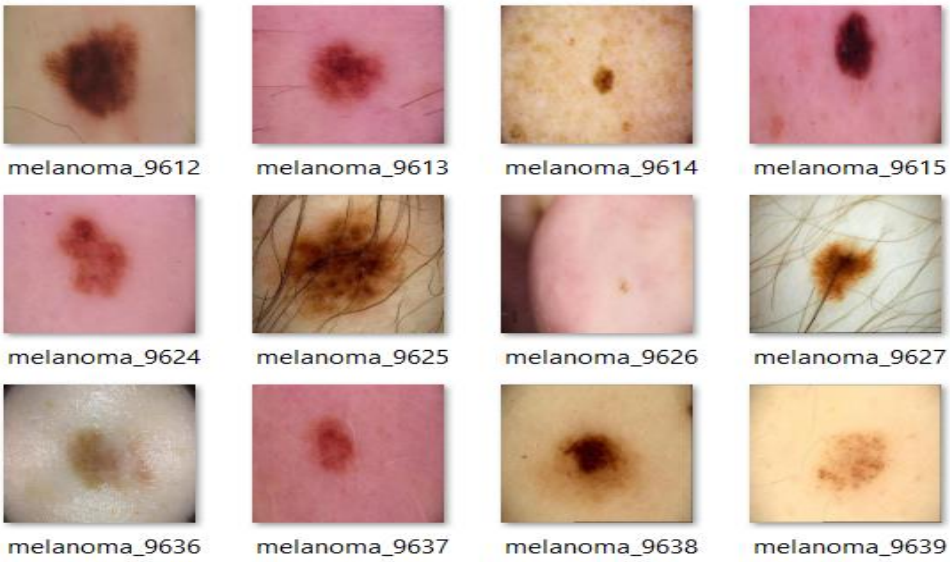

Figure 1. Sample Image Dataset

B. Analysis of Dataset

The dataset is classified as:

- Training Dataset
- Test Dataset

The melanoma dataset contains 9600 images for training the melanoma model and 1000 images for model evaluation [22]. The melanoma dataset contains 5000 images for benign and 4600 for malignant images for classification, given below is the graphical Analysis of the binary melanoma dataset [22]. Graphical Analysis for Benign and Malignant data is shown in Figure 2.

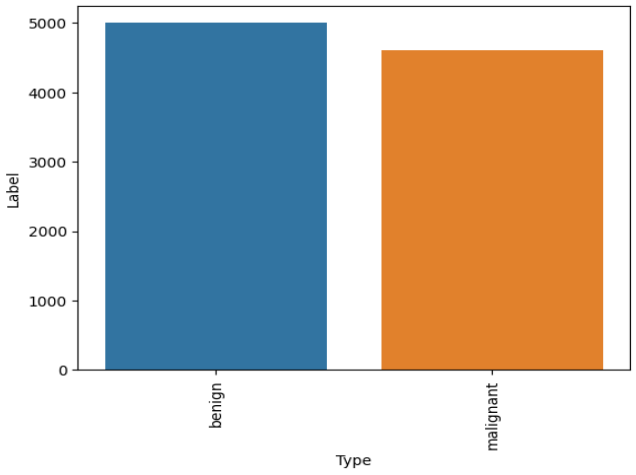

Figure 2. Graphical Analysis for Benign and Malignant data

C. Proposed Methodology

1) Layers used in Proposed Model (MelaNet): Following layers have been used for creating the Proposed Model (MelaNet). Architecture of the model is shown in Figure 3. Table 2 shows the layers and parameters of the model.

a) Convolution Layer: The Convolution Layer is the most important layer in the CNN and is used to extract the important features from images. The output of the Convolution layer is then supplied as input to next layer to continue the process.

b) Pooling Layer: The pooling layer often comes after the convolution layer and is used to lower the size of convoluted feature map of the melanoma image in order to save on computation cost.

c) *Fully Connected Layer:* The neurons connecting the two layers are connected via this layer, which also contains biases and weights. Two fully connected layers perform better than one layer, which is the main justification for linking the two layers.

d) Dropout: This layer is employed in neural networks when the data are overfitted, resulting in a model that performs well on the training data but has poor performance on real data, which leads to erroneous prediction. The Dropout layer, which involves the removal or dropping out of a few neurons during training operations to reduce the model size, is used to overcome this issue. The machine learning model's performance improves as a result.

e) *Activation Function:* Finally, most crucial function is this one. This process determines which information should be forwarded and which should not at the very end. Proposed Model(MelaNet) uses a variety of activation functions, including ReLU, SoftMax, Sigmoid Function, and tanH. Each of these operations has a specific purpose and area of expertise when creating models. Simply said, the Activation Function determines whether or not to activate the neuron [14-17].

2) Pseudo-code of the Proposed Model (MelaNet):

```
# Importing the python libraries
import OS, glob, seaborn as sns, numpy as np, matplotlib, pandas

# Loading the dataset
filepaths □ list(glob.glob(file_path+'/**/*.*'))
labels □ list(map(lambda x: os.path.split(os.path.split(x)[0]), filepaths))
data □ pd.concat([filepath, labels], axis=1)


# Preprocessing the defined datasets
train_gen □ ImageDataGenerator (
   preprocessing □ preprocess_input,
   rotation □ 30
   zoom □ 0.2
   width_shift □ 0.1
   height_shift □ 0.1
          horizontal_flip □ True
          vertical_flip □ True )

# Train the model with custom layers
model □ Sequential()
conv_filters □ [16, 32, 64, 128]
conv_kernel_sizes □ [3]
conv_paddings □ ['same']
conv_activations □ ['relu']
pool_sizes □ [2, 2]
input_shape □ (100, 100, 3)
for filters, kernel_size, padding, activation in zip (conv_filters, conv_kernel_sizes, conv_paddings,
conv_activations):
   model.add(Conv2D(filters, kernel_size, padding, activation, input_shape))
   input_shape □ None
    if padding □ 'same':
       model.add(Conv2D(filters, kernel_size, padding, activation))
       model.add(MaxPooling2D(pool_size=pool_sizes[0]))
   else:
      model.add(Conv2D(filters, kernel_size, padding, activation))
      model.add(MaxPooling2D(pool_size=pool_sizes[0]))
model.add (Flatten())
dense_units □ [128, 64, 32, 16]
dropout_rates □ [0.2, 0.1]
   for units, dropout_rate in zip (dense_units, dropout_rates):
   model.add(Dense(units, activation='relu'))
   model.add(Dropout(dropout_rate))

model.add (Dense (2, activation='sigmoid'))

          # Make predictions on test data
      predictions □ model.predict(test_gen)
```

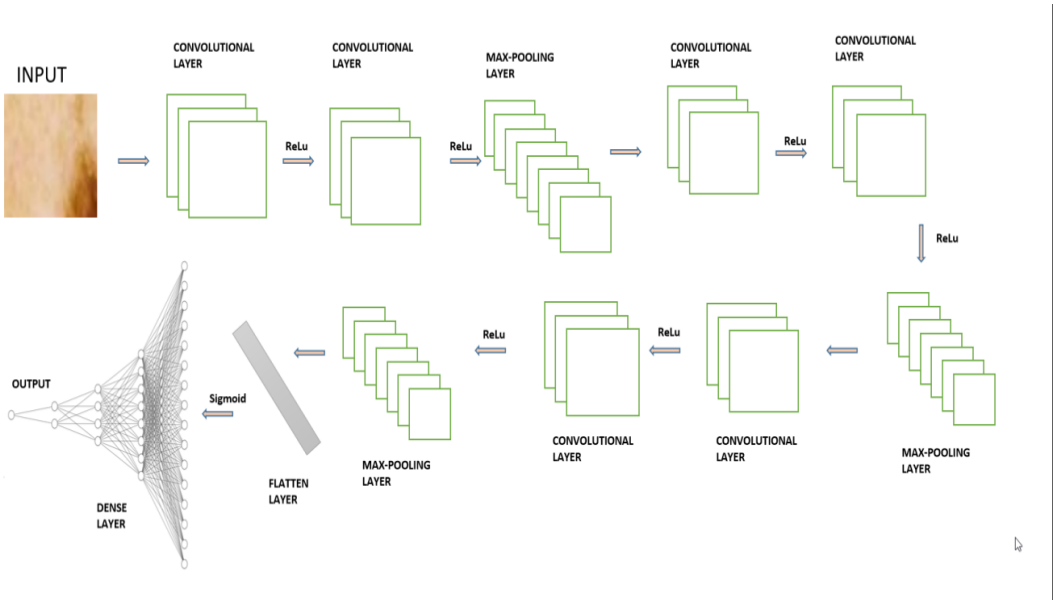

Figure 3. Architecture of the Proposed Model (MelaNet)

Table 2. MelaNet Layers and Parameters

Layer (Type)	Output Shape	Trainable Parameters
conv2d_0 (Convolutional)	100,100,16	448
conv2d_1 (Convolutional)	100,100,16	2320
conv2d_2 (Convolutional)	50,50,16	4640
conv2d_3 (Convolutional)	50,50,32	9248
conv2d_4 (Convolutional)	25,25,64	18496
conv2d_5 (Convolutional)	25,25,64	36928
conv2d_6 (Convolutional)	12,12,128	73856
conv2d_7 (Convolutional)	12,12,128	147584
Dense	128	2359424
Dense	64	8256
Dense	32	2080
Dense	16	528
Dense	2	34

3) Flowchart of the Proposed Model: The dataset is loaded initially in the suggested technique, followed by feature extraction from the dataset. The weight file is saved and later utilized to classify the melanoma disease when a picture is provided as input to the model. The flowchart of the suggested model with augmentation and without augmentation is shown in Figure 4 and Figure 5 respectively.

V. Experimental Work and Result Analysis

A. *Proposed Model - MelaNet*

MelaNet is a convolutional neural network model architecture that pulls out pertinent characteristics from input images by using different filter sizes and pooling layers. Additionally, dropout methods are used to lessen overfitting and boost model generalization. The final layers of the model consist of dense layers for classification, with a sigmoid activation function used to produce a probability output for each class. The model starts with an input layer that accepts 100 x 100 x 3 images as input.

The output is processed through three thick layers after the final pooling layer in order to produce the probability output for each class. ReLU activation is used in the first two dense layers, whereas sigmoid activation is used in the third and final dense layers. The architecture of the model is made to give accurate predictions while minimizing overfitting by extracting pertinent features from the input images.

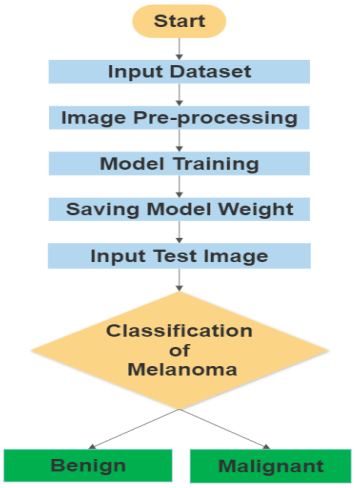

Figure 4. Flowchart of the proposed model without augmentation

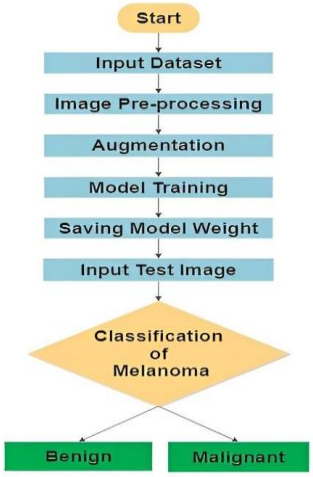

Figure 5. Flowchart of the proposed model with augmentation

B. Analysis of MelaNet without Augmentation

Initially, the model starts with loss of 0.5565 and accuracy of 0.7470 on the training dataset. As the training progresses, both the loss and accuracy improve. By the end of the 50 epochs, the model achieves loss of 0.0338 and an accuracy of 0.9299 on the training set.

On the validation set, the model starts with a loss of 0.3799 and an accuracy of 0.8676. Similar to the training set, the loss decreases, and the accuracy increases as the training proceeds. After the end of 40 epochs, the model achieves a validation loss of 0.5878 (Figure 6) and a validation accuracy of 0.9222 (Figure 7).

Overall, the model demonstrates good performance, with the accuracy steadily increasing and the loss decreasing over the training period. These results indicate that the model has learned to make accurate predictions on both the training and validation data, suggesting it may be effective for the given task.

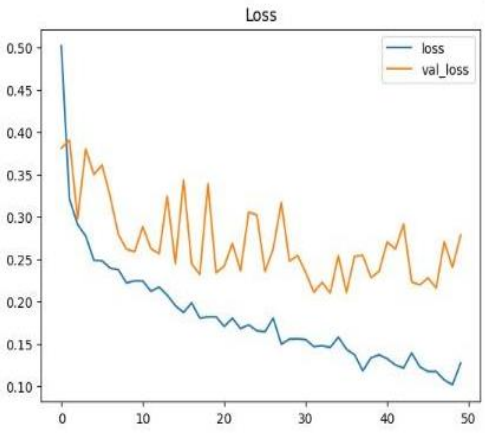

Figure 6. MelaNet Loss

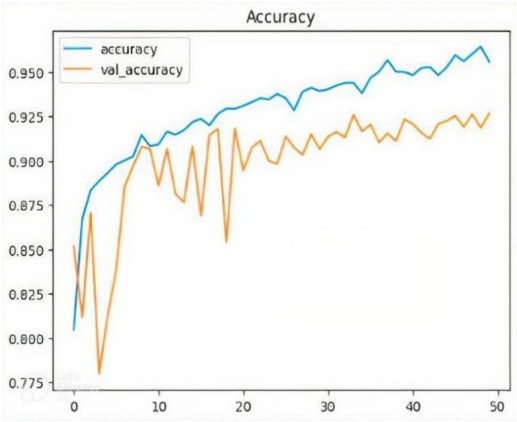

Figure 7. MelaNet Accuracy

1) Classification Report for MelaNet without Augmentation: The classification of MelaNet without augmentation is shown in Table 3. The value of precision, recall, F1-score, and support is shown in the table. The graphical representation for the same is shown in Figure 9. The detail of these parameters is as follows:

a) Precision: The precision for the "benign" class is 0.93, indicating that 93% of the samples predicted as benign are actually benign. Similarly, the precision for the "malignant" class is 0.84, meaning 84% of the samples predicted as malignant are truly malignant.
b) Recall: The recall for the "benign" class is 0.84, indicating that 84% of the true benign samples were correctly identified. Likewise, the recall for the "malignant" class is 0.93, suggesting that 93% of the true malignant samples were correctly identified.
c) F1-Score: The F1-score for "benign" class is 0.88, indicating a balanced performance between both precision and recall for this classification. Similarly, the F1-score for the "malignant" class is 0.88, suggesting a balanced performance as well.
d) Support: There are 1,281 samples labelled as "benign" and 1,121 samples labelled as "malignant."

Table 3. Classification Report of MelaNet

	Precision	Recall	F1-Score	Support
benign	0.93	0.84	0.88	1281
malignant	0.84	0.93	0.88	1121
accuracy			0.88	2402
macro avg	0.88	0.89	0.88	2402
weighted avg	0.89	0.88	0.88	2402

Overall, the classification report suggests that the classifier performs well for both classes, with high precision, recall, and F1-scores. The classifier demonstrates a good level of accuracy in identifying both "benign" and "malignant" samples, as indicated by the precision and recall values. The F1-scores also indicate a balanced performance between precision and recall for both classes, with similar values for both classes. The similar precision and recall values suggest that the classifier is performing consistently for both classes. The confusion matrix for both classes is shown in Figure 8.

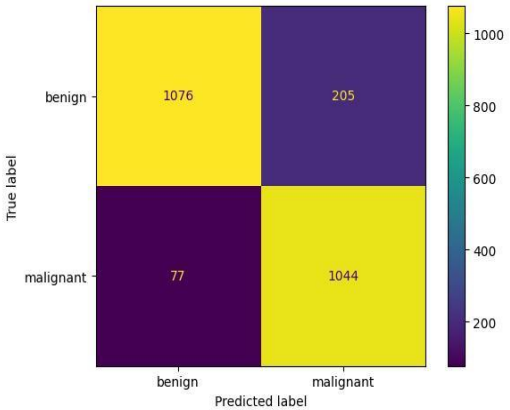

Figure 8. Confusion Matrix of MelaNet

The confusion matrix represents the performance of a melanoma skin classification model that aims to distinguish between benign and malignant cases. The matrix consists of four cells, each capturing a different outcome based on the model's predictions and the actual labels.

Starting with the true positives (TP), we have a value of 1076. This indicates the number of cases where the model correctly identified a sample as malignant and indeed it was malignant. These are the instances where the model successfully detected malignant melanomas.

Moving on to the false positives (FP), we find a value of 205. This represents the number of cases where the model incorrectly predicted a sample as malignant, while in reality, it was benign. These are the instances where the model produced a false alarm, classifying a benign case as malignant.

Next, we have the false negatives (FN) with a value of 77. This signifies the number of cases where the model mistakenly predicted a sample as benign when it was actually malignant. In other words, the model failed to identify 77 malignant melanomas, leading to a false classification as benign.

Lastly, the true negatives (TN) are represented by a value of 1044. This indicates the number of cases where the model correctly classified a sample as benign and it was indeed benign. These are the instances where the model accurately recognized benign skin conditions.

C. Analysis of MelaNet with Augmentation

The initial epoch started with a loss of 0.8061 and an accuracy of 0.5945. As the training progressed, the loss decreased, and the accuracy increased. By the end of the 50th epoch, the loss was reduced to 0.3714, and the accuracy reached 0.8397. The validation results also showed improvements over the epochs. The validation loss has been decreased from 0.6039 in the first epoch to 0.3669 in the last epoch (Figure 10), while the model validation accuracy increased from 0.6420 to 0.8347.

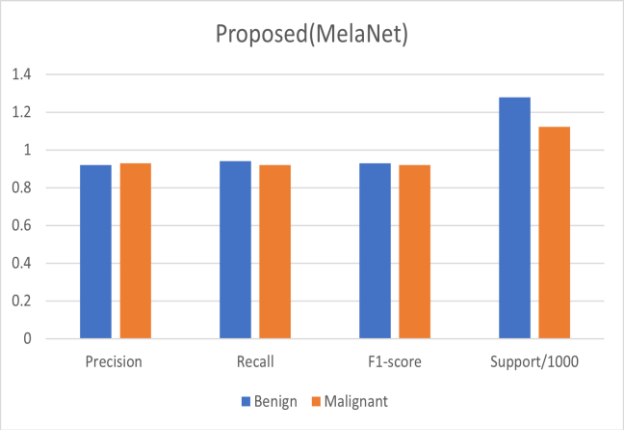

Figure 9. Graphical Analysis of MelaNet without Augmentation

The model's performance showed significant improvement, particularly in terms of accuracy, which increased from 59.45% to 83.97% (Figure 11). This suggests that the model learned to make better predictions as it was exposed to more training data.

The accuracy showed consistent improvement. Starting from an initial accuracy of 59.45%, the model's performance gradually increased, reaching 83.97% at the end of the training. This suggests that the model effectively learned to recognize and classify patterns within the given dataset.

The validation results provide further evidence of the model's progress. The validation loss, which measures the model's performance on unseen data, decreased from 0.6039 to 0.3669. Similarly, the validation accuracy improved from 64.20% to 83.47%. These results indicate that the model performance extends beyond the training data and can generalize well to new, unseen examples.

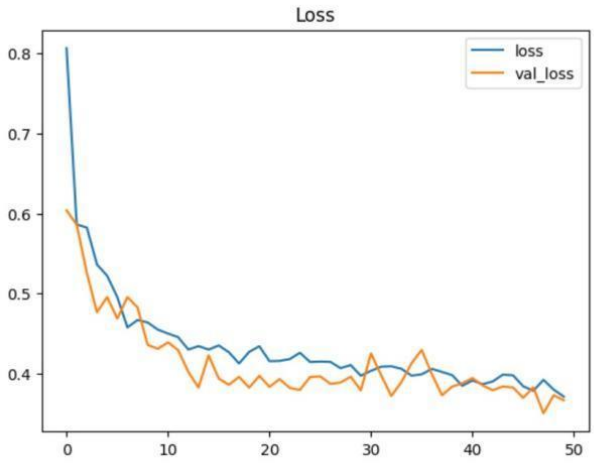

Figure 10. MelaNet Loss with Augmentation

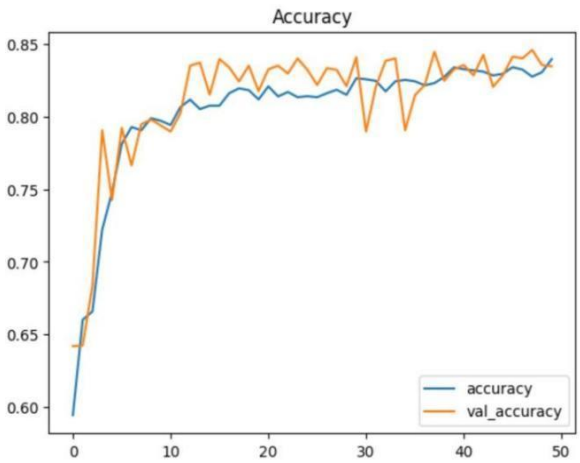

Figure 11. MelaNet Accuracy with Augmentation

1) Classification Report for MelaNet with Augmentation: Table 4 shows the classification report of MelaNet with augmentation. This table displays the values of precision, recall, F1-score, and support for binary classes i.e., benign and malignant. Also, the Graphical Analysis of MelaNet with Augmentation is shown in Figure 13. A detailed explanation of all these parameters is as follows:

a) Precision: The precision for the "benign" class is 0.86, indicating that 86% of the samples predicted as benign are actually benign. Similarly, the precision for the "malignant" class is 0.82, meaning 82% of the samples predicted as malignant are truly malignant.

b) Recall: The recall for benign class is 0.83, indicating that 83% of the true benign samples were correctly identified. Likewise, the recall for the malignant class is 0.85, suggesting that 85% of the true malignant samples were correctly identified.

c) F1-Score: The F1-score for the "benign" class is 0.84, indicating a balanced performance between precision and recall for this class. Similarly, the F1-score for the "malignant" class is also 0.84, suggesting a balanced performance as well.

d) Support: There are 1,247 samples labelled as "benign" and 1,155 samples labelled as "malignant."

Table 4. Classification Report of MelaNet with Augmentation

	Precision	Recall	F1-Score	Support
benign	0.86	0.83	0.84	1247
malignant	0.82	0.85	0.84	1155
accuracy			0.84	2402
macro avg	0.84	0.84	0.84	2402
weighted avg	0.84	0.84	0.84	2402

Overall, the classification report suggests that the classifier performs moderately well for both classes, with precision, recall, and F1-scores around 0.8-0.9. The classifier shows a reasonably good level of accuracy in identifying both "benign" and "malignant" samples, as indicated by the precision and recall values. The F1-scores also indicate a balanced performance between precision and recall for both classes, with similar values for both classes. The similar precision and recall values suggest that the classifier is performing consistently for both classes. The confusion matrix for the same is shown in Figure 12.

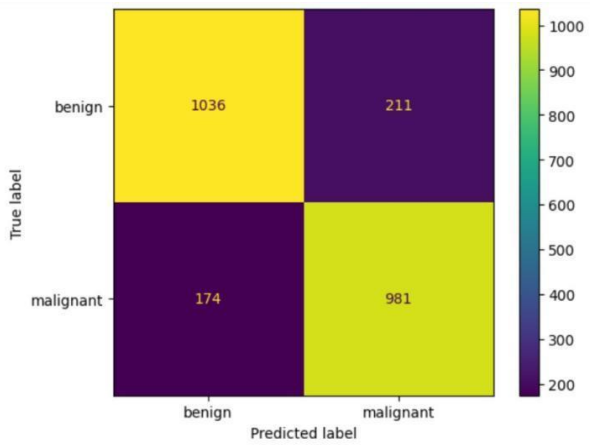

Figure 12. Confusion Matrix of MelaNet with Augmentation

In this context, true positive (TP) are the instances where the model correctly identified a sample as malignant when it is actually malignant. The matrix shows that there are 1036 true positives, indicating that the model accurately detected malignant melanomas.

False positive (FP) represent cases where the model incorrectly labeled a sample as malignant when it is actually benign. Here, we see that there are 211 false positives, implying that the model misclassified some benign cases as malignant.

False negative (FN) occur when the model mistakenly predicted a sample as benign when it is actually malignant. The confusion matrix displays 174 false negatives, meaning that the model failed to identify some malignant melanomas.

True negative (TN) indicate the instances where the model correctly identified a sample as benign when it is indeed benign. The matrix indicates that there are 981 true negatives, demonstrating that the model accurately recognized benign skin conditions.

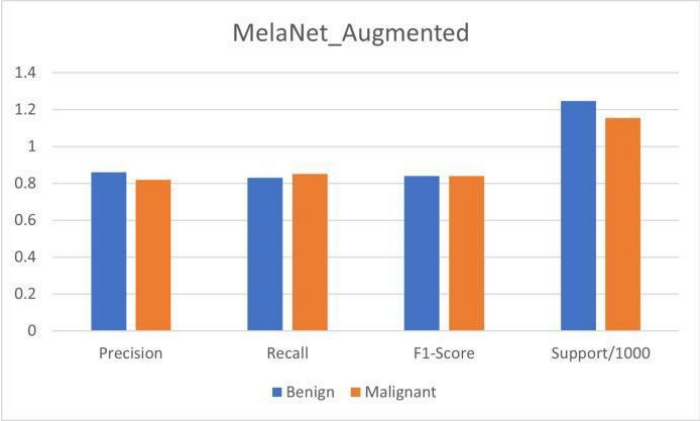

Figure 13. Graphical Analysis of MelaNet with Augmentation

VI. Comparative Analysis of MelaNet with other Models

A. Comparative Analysis of Accuracy and Loss without Augmentation

The models are evaluated based on their accuracy and loss values, as shown in the Table 5 below, which presents performance metrics for several classification models.

Based on this table, the MelaNet, achieved the highest accuracy of 92.67% and the lowest loss value of 0.27811. It outperformed the other models, including ResNet50, VGG-16, MobileNet, InceptionV3, and MobileNetV2, in terms of both accuracy and loss.

Table 5. Comparative Analysis before Performing Augmentation

Models	Accuracy	Loss
MelaNet (Proposed Model)	92.67	0.27811
ResNet50	92.13	0.54655

VGG-16	89.97	0.71918
MobileNet	86.01	0.89285
InceptionV3	84.18	0.65748
MobileNetV2	86.84	0.70583

B. Comparative Analysis of Accuracy and Loss with Augmentation

The Table 6 given below shows the performance metrics of various models on Classification, the models are evaluated based on their accuracy and loss values.

The accuracy and loss values of the MelaNet model are 84.18% and 0.28973, respectively. Although the accuracy of the MelaNet model is not the highest among the models, it is important to note that it has a smaller loss value when compared to models like ResNet50, VGG-16, MobileNet, Inception V3, and MobileNetV2. The smaller loss value means that MelaNet is less likely to overfit since it is better at minimizing the difference between its projected and actual outputs during training.

Table 6. Comparative Analysis after Performing Augmentation

Models	Accuracy	Loss
MelaNet (Proposed Model)	84.18	0.28973
ResNet50	87.80	0.29055
VGG-16	87.39	0.29573
MobileNet	81.31	0.39953
InceptionV3	77.44	0.49206
MobileNetV2	86.84	0.70583

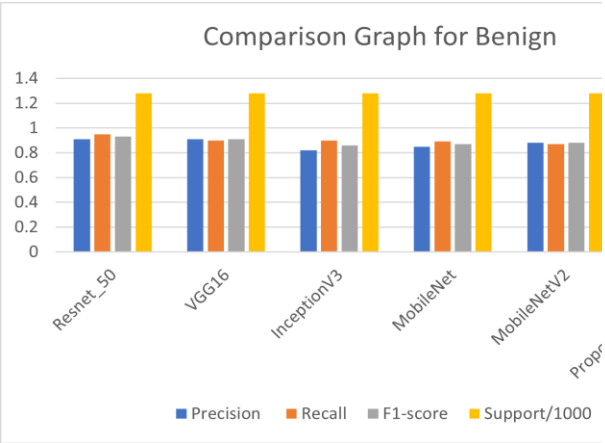

Figure 14. Comparison Graph for Benign

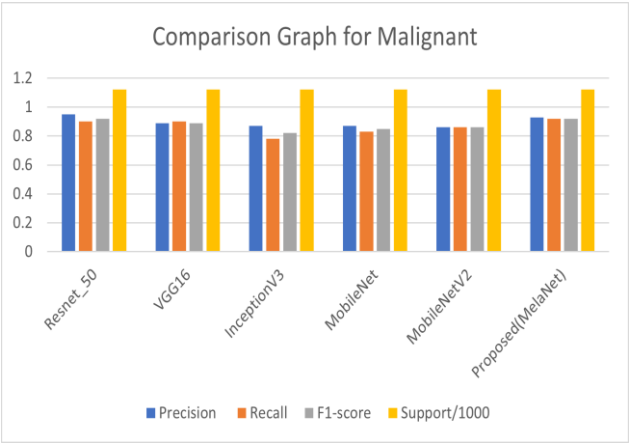

Figure 15. Comparison Graph for Malignant

VII. CONCLUSION

In conclusion, the use of deep learning and image processing techniques for the classification and categorization of melanoma has shown more effective and economical strategy compared to conventional procedures. The proposed model, MelaNet, utilizes a Convolutional Neural Network specifically designed for multi-classification tasks. It demonstrates impressive accuracy, achieving results of 74.05%, 85.36%, 88.12%, 90.94%, and 92.67% at different epochs.

When compared to other conventional and recently developed models, MelaNet exhibits superior accuracy. While this model cannot replace laboratory-based diagnostic techniques, it can provide valuable decision-making support. It is important to note that laboratory-based diagnosis remains more reliable than visual symptom-based diagnoses. Overall, the development of such a model utilizes deep learning for the classification of potential melanoma in skin lesion images and holds the potential to improve the accuracy and efficiency of skin disease diagnosis. These

advancements have the potential to aid in the prompt detection and treatment of melanoma, enhancing healthcare outcomes for patients.

VIII. FUTURE SCOPE

To improve the proposed model, further increasing the number of classes and adding a large dataset with high quality resolution skin lesion images can be beneficial. With more data, the model can learn to generalize better and detect melanoma more accurately. Additionally, more enhanced structured deep neural network can help achieve better accuracy. Through fine-tuning the model's parameters and layers, the performance can be further optimized. However, it is important to consider the computational resources required for training a more complex model. Balancing the model's complexity with computational efficiency is crucial to ensure practicality and scalability. When a model performs well on the training data but struggles to generalize to fresh, untried data, overfitting has taken place. MelaNet proves it is less prone to overfitting and may be able to generalize better to unseen data by attaining a lower loss value.

REFERENCES

[1] Jain, S., & Pise, N. (2015). Computer aided melanoma skin cancer detection using image processing. Procedia Computer Science, 48, 735-740.

[2] Davis, L. E., Shalin, S. C., & Tackett, A. J. (2019). Current state of melanoma diagnosis and treatment. Cancer biology & therapy, 20(11), 1366-1379.

[3] Carr, S., Smith, C., & Wernberg, J. (2020). Epidemiology and risk factors of melanoma. Surgical Clinics, 100(1), 1-12.

[4] Lv, H., Liu, X., Zeng, X., Liu, Y., Zhang, C., Zhang, Q., & Xu, J. (2022). Comprehensive analysis of cuproptosis-related genes in immune infiltration and prognosis in melanoma. Frontiers in pharmacology, 13, 2157.

[5] Ahmed, B., Qadir, M. I., & Ghafoor, S. (2020). Malignant Melanoma: Skin Cancer− Diagnosis, Prevention, and Treatment. Critical Reviews™ in Eukaryotic Gene Expression, 30(4).

[6] Kato, J., Horimoto, K., Sato, S., Minowa, T., & Uhara, H. (2019). Dermoscopy of melanoma and non-melanoma skin cancers. Frontiers in medicine, 6, 180.

[7] Islam, T., Sah, M., Baral, S., & Choudhury, R. R. (2018, April). A faster technique on rice disease detection using image processing of affected area in agro-field. In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 62-66). IEEE.

[8] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. nature, 542(7639), 115-118.

[9] Mallick, P. K., Mishra, S., Mohanty, B. P., & Satapathy, S. K. (2021). A deep neural network model for effective diagnosis of melanoma disorder. In Cognitive Informatics and Soft Computing: Proceeding of CISC 2020 (pp. 43-51). Springer Singapore.

[10] Pillay, V., & Viriri, S. (2019, March). Skin cancer detection from macroscopic images. In 2019 Conference on Information Communications Technology and Society (ICTAS) (pp. 1-9). IEEE.

[11] Abbas, Z., Rehman, M. U., Najam, S., & Rizvi, S. D. (2019, February). An efficient gray-level co-occurrence matrix (GLCM) based approach towards classification of skin lesion. In 2019 amity international conference on artificial intelligence (AICAI) (pp. 317-320). IEEE.

[12] Tsantoulis, P., Delorenzi, M., Bièche, I., Vacher, S., Mariani, P., Cassoux, N., ... & Cacheux, W. (2019). Prospective validation in epithelial tumors of a gene expression predictor of liver metastasis derived from uveal melanoma. Scientific reports, 9(1), 17178.

[13] Guha, S. R., & Rafizul Haque, S. M. (2020). Performance comparison of machine learning-based classification of skin diseases from skin lesion images. In International Conference on Communication, Computing and Electronics Systems: Proceedings of ICCCES 2019 (pp. 15-25). Springer Singapore.

[14] Liu, Y. P., Wang, Z., Li, Z., Li, J., Li, T., Chen, P., & Liang, R. (2021). Multiscale ensemble of convolutional neural networks for skin lesion classification. IET Image Processing, 15(10), 2309-2318.

[15] Li, H., Pan, Y., Zhao, J., & Zhang, L. (2021). Skin disease diagnosis with deep learning: a review. Neurocomputing, 464, 364-393.

[16] Kaur, R., & Kaur, N. (2021, December). Improved Skin Cancer Detection Classification Residual Network Feature Engineering. In 2021 International Conference on Computational Performance Evaluation (ComPE) (pp. 671-675). IEEE.

[17] Lu, X., & Firoozeh Abolhasani Zadeh, Y. A. (2022). Deep learning-based classification for melanoma detection using XceptionNet. Journal of Healthcare Engineering, 2022.

[18] Priyadharshini, N., Selvanathan, N., Hemalatha, B., & Sureshkumar, C. (2023). A noble hybrid Extreme Learning Machine and Teaching–Learning-Based Optimization algorithm for skin cancer detection. Healthcare Analytics, 3, 100161.

[19] Sheha, M. A., Mabrouk, M. S., & Sharawy, A. (2012). Automatic detection of melanoma skin cancer using texture analysis. International Journal of Computer Applications, 42(20), 22-26.

[20] "Kaggle", kaggle.com [Online]. Available: https://kaggle.com

[21] Islam, T., Sah, M., Baral, S., & Choudhury, R. R. (2018, April). A faster technique on rice disease detection using image processing of affected area in agro-field. In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 62-66). IEEE.

[22] Melanoma Skin Cancer Dataset of 10000 Images, Source:https://www.kaggle.com/datasets/hasnainjaved/melanoma-skin-cancer-dataset-of-10000-images