MELANOMA CLASSIFICATION

Project By – Manish Sri Sai Surya Routhu

Data Description

The project utilizes a dataset comprising 300 images of melanoma skin lesions. The dataset is divided into a training set of 300 images (150 benign and 150 malignant) and a test set of 50 images (25 benign and 25 malignant). Each image is labelled with the corresponding class (benign or malignant) for supervised learning.

The images in the dataset are of varying resolutions and may contain different skin tones, lighting conditions, and backgrounds. Preprocessing techniques such as resizing and normalization will be done to enhance the dataset and improve model generalization.

Our Dataset consists of 2 classes Benign and Malignant. Let's know about them a bit:

i. Benign

Benign refers to a medical condition or tumor that is non-cancerous and does not pose a significant threat to one's health. In the context of melanoma disease, a benign skin lesion refers to a non-cancerous growth or abnormality on the skin that is typically harmless and unlikely to spread or cause severe health problems.

In dermatology, benign lesions are often characterized by their regular shape, uniform color, and well-defined borders. They do not invade nearby tissues or organs and are generally considered low-risk or non-threatening. While benign skin lesions may still require medical attention or monitoring, they are not considered cancerous and do not have the potential to metastasize or spread to other parts of the body.

It's important to note that even though a lesion is initially determined to be benign, regular monitoring and follow-up with a healthcare professional are essential to ensure any changes or potential risks are identified promptly. Dermatologists are trained to evaluate and differentiate between benign and malignant skin lesions through various diagnostic methods, including visual inspection, dermoscopy, and in some cases, biopsy.

ii. Malignant

Malignant refers to a medical condition or tumor that is cancerous and has the potential to invade nearby tissues, spread to other parts of the body, and pose a significant threat to one's health. In the context of melanoma disease, a malignant skin lesion refers to a cancerous growth or abnormality on the skin that can be potentially life-threatening if left untreated or not detected at an early stage.

Malignant melanoma is a type of skin cancer that arises from the pigment-producing cells called melanocytes. It is characterized by the uncontrolled growth and division of these cells, leading to the formation of tumors or lesions on the skin. Malignant melanoma has the ability to invade surrounding tissues and can metastasize or spread to other organs in the body through the lymphatic system or bloodstream.

The identification of malignant skin lesions is crucial as early detection and treatment significantly improve prognosis and outcomes. Dermatologists and healthcare professionals assess various factors, including the lesion's size, shape, color, border irregularities, and changes over time, to determine the likelihood of malignancy. Diagnostic techniques such as dermoscopy, biopsy, and histopathological examination are often utilized to confirm the presence of malignancy.

Note: So, as Malignant is cancerous, it becomes necessary for us to identify it and accuracy of identifying Malignant will be considered as a key metric. How ever as the data is balanced, It can be easily interpretable from the confusion matrices and classification reports

Solution Approach

To tackle the melanoma classification problem, transfer learning will be employed. Transfer learning leverages the knowledge and learned representations from pre-trained models trained on large-scale datasets. It allows us to benefit from the feature extraction capabilities of these models and adapt them to our specific task with limited data.

The following steps will be taken to develop the solution:

i. Preprocessing

The images will undergo preprocessing steps to ensure uniformity and enhance model performance. These steps may include resizing the images to a fixed dimension i.e. (224, 224, 3), and normalizing pixel values between [0,1].

ii. Transfer Learning Model Selection

Several state-of-the-art pre-trained models will be considered for transfer learning, such as VGG16, ResNet50, InceptionV3, Xception, MobileNet, DenseNet. These models have been trained on large-scale image datasets and have demonstrated strong performance on various computer vision tasks.

iii. Model Customization

The selected pre-trained model will be customized by adding a few additional layers on top to adapt it to our specific classification task. These layers will enable the model to learn discriminative features from the melanoma skin lesion images.

iv. Fine-tuning

To further improve model performance, fine-tuning will be performed. Fine-tuning involves training the added layers along with a portion of the pre-trained model's layers. By updating the weights of these layers during training, the model can adapt to the specific characteristics of the melanoma dataset.

v. Model Evaluation

The trained model will be evaluated using the test set, consisting of 50 labelled images. Performance metrics such as accuracy and recall, will be calculated to assess the model's classification performance. The model will also undergo a Receiver Operating Characteristics curve and Precision-Recall Curve . Finally, a classification report is also given.

Implementation

I used 7 transfer learning models and calculated metrics for all of those

i. Resnet

ResNet (Residual Neural Network) is a deep neural network architecture introduced to address the problem of vanishing gradients during training of very deep neural networks. ResNet introduces the concept of residual connections, allowing the network to learn residual mappings instead of directly learning the desired mappings. These residual connections enable the network to effectively propagate gradients and learn deeper and more accurate representations. ResNet architectures come in different depths, such as ResNet-18, ResNet-34, ResNet-50, and so on, with the number indicating the number of layers. ResNet has achieved remarkable performance in various computer vision tasks, including image classification, object detection, and image segmentation.

In this case, Resnet is not at all training the model and it is pushing whole data to one class. Hence, It is not taken into consideration.

ii. Inception

Inception is a convolutional neural network (CNN) architecture that was designed to address the challenge of capturing multi-scale features efficiently. The Inception architecture introduced the concept of "inception modules," which consist of parallel convolutional layers with different filter sizes. These parallel layers capture features at different scales and concatenate their outputs. This allows the network to learn both local and global features simultaneously, leading to improved performance. Inception models, such as InceptionV3 and InceptionResNet, have been highly successful in image classification tasks and have significantly contributed to the advancement of deep learning in computer vision.

Even Inception is not training the model and pushing all the outputs to one class. Hence, it is also not considered

iii. VGG16

VGG16 is a popular convolutional neural network (CNN) architecture that was introduced by the Visual Geometry Group (VGG) at the University of Oxford. It is known for its simplicity and uniform structure. VGG16 consists of 16 layers, including 13 convolutional layers and 3 fully connected layers. It uses small-sized convolutional filters (3x3) throughout the network, which helps in capturing fine-grained features. VGG16 has achieved impressive performance on various computer vision tasks, particularly in image classification. Its deep architecture allows it to learn complex representations, and it serves as a benchmark for evaluating and comparing other CNN models.

Implementation and Observation:

The architecture of the Model is below

```
Model: "sequential_11"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 vgg16 (Functional)          (None, 7, 7, 512)         14714688

 flatten_11 (Flatten)        (None, 25088)             0

 dense_33 (Dense)            (None, 1024)              25691136

 dense_34 (Dense)            (None, 512)               524800

 dense_35 (Dense)            (None, 32)                16416

 dense_36 (Dense)            (None, 2)                 66

=================================================================
Total params: 40,947,106
Trainable params: 26,232,418
Non-trainable params: 14,714,688
```

From the above – we have 26, 232, 418 parameters to train and I used 10 epochs and batch size 32 to fit the model.

After fitting the data following metrics are evaluated

Classification Report:

```
TEST DATA EVALUATION
              precision    recall  f1-score   support

           0       0.92      0.48      0.63        25
           1       0.65      0.96      0.77        25

    accuracy                           0.72        50
   macro avg       0.79      0.72      0.70        50
weighted avg       0.79      0.72      0.70        50
```

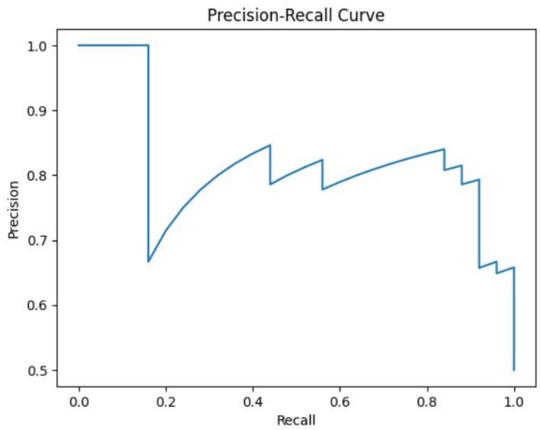
Confusion_matrix:

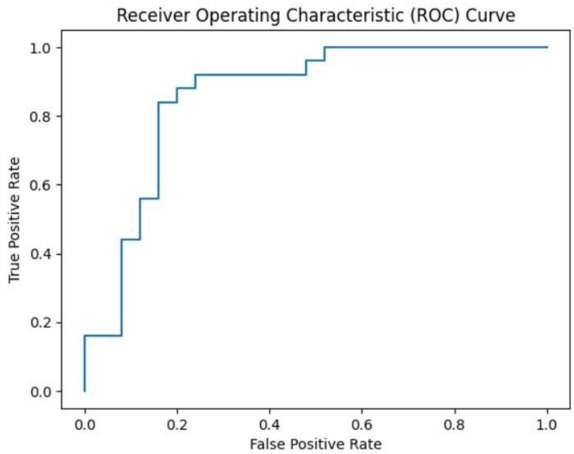
	Benign	Malignant
Benign	12	13
Malignant	1	24

Accuracy: 0.72

Recall: 0.96

ROC and PRC:





iv. Mobilenet

MobileNet is a lightweight convolutional neural network (CNN) architecture specifically designed for mobile and embedded devices with limited computational resources. It focuses on achieving a good balance between accuracy and model size. MobileNet employs depthwise separable convolutions, which split the standard convolution operation into a depthwise convolution followed by a pointwise convolution. This reduces the computational complexity and model size significantly while maintaining reasonable accuracy. MobileNet models come in different versions such as MobileNetV1, MobileNetV2, and MobileNetV3, with each version introducing improvements in efficiency and performance. MobileNet has been widely used for various applications on resource-constrained devices, such as image classification, object detection, and semantic segmentation.

Implementation and observation:

The architecture of the model is as below

```
Model: "sequential_6"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 mobilenet_1.00_224 (Functio  (None, 7, 7, 1024)       3228864
 nal)

 flatten_6 (Flatten)         (None, 50176)             0

 dense_19 (Dense)            (None, 512)               25690624

 dense_20 (Dense)            (None, 2)                 1026

=================================================================
Total params: 28,920,514
Trainable params: 25,691,650
Non-trainable params: 3,228,864
```

Classification Report:

```
TEST DATA EVALUATION
              precision    recall  f1-score   support

           0       0.83      0.76      0.79        25
           1       0.78      0.84      0.81        25

    accuracy                           0.80        50
   macro avg       0.80      0.80      0.80        50
weighted avg       0.80      0.80      0.80        50
```

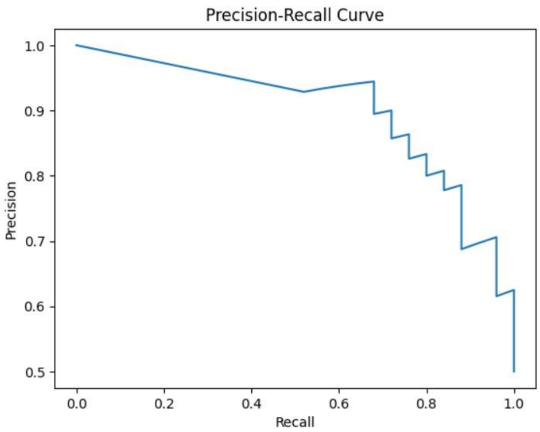
Confusion Matrix:

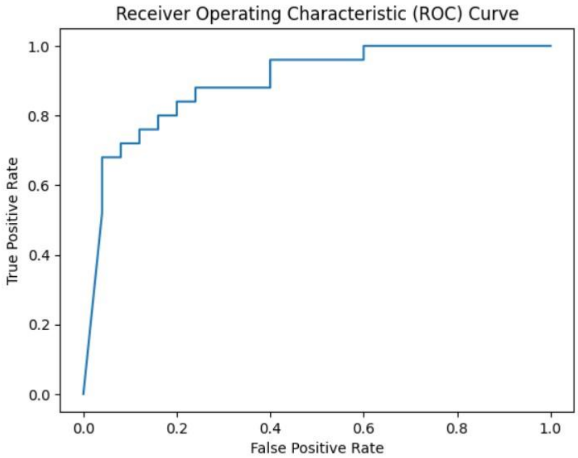
	Benign	Malignant
Benign	19	6
Malignant	4	21

Accuracy: 0.8

Recall: 0.84

ROC and PRC:





v. Xception

Xception is a deep convolutional neural network (CNN) architecture that aims to improve both accuracy and efficiency by enhancing the feature extraction process. It is based on the idea of an "extreme inception" module, where it replaces the standard convolutional layers with depthwise separable convolutions. Xception uses separable convolutions to capture spatial and channel-wise dependencies separately, resulting in a more efficient and expressive network. This architecture reduces the number of parameters and computations while maintaining or even improving performance compared to traditional CNNs. Xception has been successful in various computer vision tasks, such as image classification, object detection, and image segmentation, and it has demonstrated state-of-the-art results on benchmark datasets.

Implementaton and observation:

The architecture of the model is given below

```
Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 xception (Functional)       (None, 7, 7, 2048)        20861480

 flatten_5 (Flatten)         (None, 100352)            0

 dense_16 (Dense)            (None, 512)               51380736

 dense_17 (Dense)            (None, 32)                16416

 dense_18 (Dense)            (None, 2)                 66

=================================================================
Total params: 72,258,698
Trainable params: 51,397,218
Non-trainable params: 20,861,480
```

The model has 51, 397, 218 trainable parameters and the metrics taken is accuracy. The evaluation part of the model is as below.

Classification Report:

```
TEST DATA EVALUATION
              precision    recall  f1-score   support

           0       0.79      0.76      0.78        25
           1       0.77      0.80      0.78        25

    accuracy                           0.78        50
   macro avg       0.78      0.78      0.78        50
weighted avg       0.78      0.78      0.78        50
```

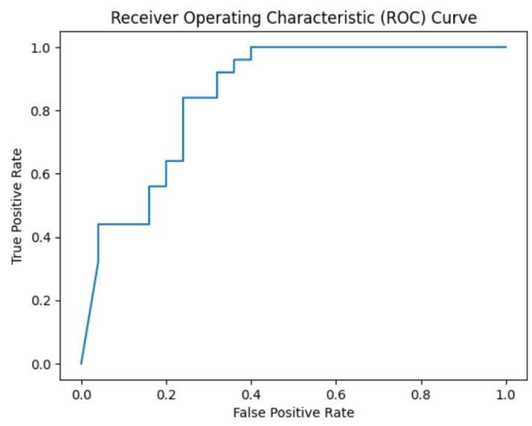
Confusion matrix:

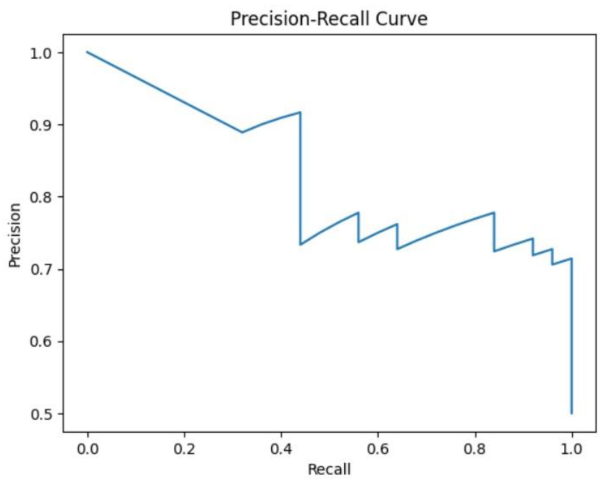
	Benign	Malignant
Benign	19	6
Malignant	5	20

Accuracy: 0.78

Recall: 0.8

ROC and PRC:





vi. Densenet

DenseNet is a deep convolutional neural network (CNN) architecture known for its dense connectivity pattern. Unlike traditional CNNs where each layer is connected only to the subsequent layer, DenseNet introduces dense connections, where each layer is connected to every other layer in a feed-forward fashion. This dense connectivity enables direct information flow between layers, facilitating feature reuse and gradient propagation. DenseNet's compact architecture leads to parameter efficiency and alleviates the vanishing gradient problem. It encourages feature reuse and enables better feature propagation throughout the network. DenseNet has demonstrated impressive performance on various computer vision tasks, including image classification, object detection, and semantic segmentation, with relatively fewer parameters.

Implementation and observation:

The architecture of the model is as given below

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 densenet121 (Functional)    (None, 1024)              7037504

 flatten_2 (Flatten)         (None, 1024)              0

 dense_4 (Dense)             (None, 1024)              1049600

 dense_5 (Dense)             (None, 512)               524800

 dense_6 (Dense)             (None, 32)                16416

 dense_7 (Dense)             (None, 2)                 66

=================================================================
Total params: 8,628,386
Trainable params: 1,590,882
Non-trainable params: 7,037,504
```

There are 1,590,882 trainable parameters and even with such less trainable parameters it is giving a phenomenal accuracy. Rest of the observations are shown below.

Classification report:

```
TEST DATA EVALUATION
              precision    recall  f1-score   support

           0       0.80      0.80      0.80        25
           1       0.80      0.80      0.80        25

    accuracy                           0.80        50
   macro avg       0.80      0.80      0.80        50
weighted avg       0.80      0.80      0.80        50
```

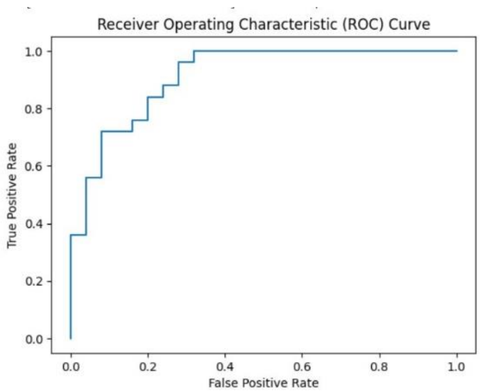
Confusion matrix:

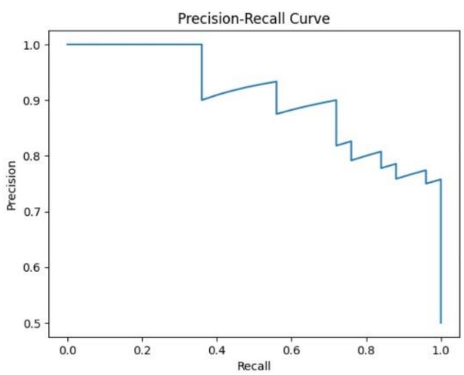
	Benign	Malignant
Benign	20	5
Malignant	5	20

Accuracy: 0.8

Recall: 0.8

ROC and PRC:





vii. EfficientNet

EfficientNet is a state-of-the-art deep convolutional neural network (CNN) architecture that achieves remarkable performance while maintaining efficiency in terms of model size and computational resources. It introduces a compound scaling method that uniformly scales the network's depth, width, and resolution in a principled manner. This approach allows EfficientNet to achieve superior accuracy by finding an optimal balance between model size and computational efficiency. EfficientNet models are typically named with a scaling coefficient, such as EfficientNet-B0, EfficientNet-B1, and so on, with increasing coefficients indicating larger and more accurate models. EfficientNet has shown outstanding results across various computer vision tasks and has become a go to choice for efficient yet powerful deep learning models.

Implementation and observation:

The architecture of the models is so complex that it cannot be shown on a single page. Basically, EfficientNet is a convolutional neural network architecture and scaling method that uniformly scales all dimensions of depth/width/resolution using a compound coefficient.

The evaluation part of the model is shown below

Classification report:

```
TEST DATA EVALUATION
              precision    recall  f1-score   support

           0       0.54      0.80      0.65        25
           1       0.62      0.32      0.42        25

    accuracy                           0.56        50
   macro avg       0.58      0.56      0.53        50
weighted avg       0.58      0.56      0.53        50
```

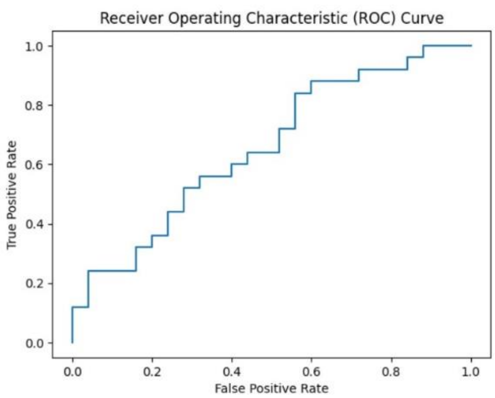
Confusion matrix:

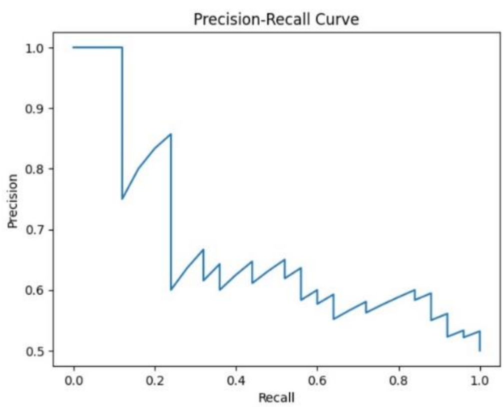
	Benign	Malignant
Benign	20	5
Malignant	17	8

Accuracy: 0.56

Recall: 0.32

ROC and PRC:





Voting on Transfer Learning Models

The 5 models we arer using for voting are VGG16, Xception, Inception, Mobilenet, and Densenet.

Doing so has improved the accuracy by 4 percent to 84.

Classification Report:

```
              precision    recall  f1-score   support

           0       0.88      0.81      0.85        27
           1       0.80      0.87      0.83        23

    accuracy                           0.84        50
   macro avg       0.84      0.84      0.84        50
weighted avg       0.84      0.84      0.84        50
```

Accuracy score: 0.84

Recall score: 0.8695652173913043

precision score: 0.8

f1 score: 0.8333333333333333

Confusion Matrix:

	Benign	**Malignant**
Benign	22	5
Malignant	3	20

Conclusion

Out of 7 models Densenet and Mobilenet were giving 80% accuracy and the densenet is more of a balanced kind. Where as Mobilenet is predicting one class better[malignant] than other. So, depending on the business case if we want to detect the malignant then Mobilenet is good model.

Malignant is a dangerous hence in some cases we may go with VGG 16 also which was predicting malignant with 96% accuracy.

Xception also does better job with 78% accuracy on test data.

Finally, in this case considering the stability aspect I consider Densenet as best model with accuracy of 0.8 on test data

Overall, if we employ voting of the transfer models there is an increase of 4% in overall accuracy.

Hence the best model is – Voting on Transfer Learning models