

Melanoma Skin Cancer Detection

Table of Contents

- [General Info](#)
- [Algorithms Used](#)
- [Dataset Information](#)
- [Technologies Used](#)
- [Steps Involved](#)
- [Prapred Modal](#)
- [Conclusions](#)
- [Acknowledgements](#)
- [Contact](#)

General Information

Problem Summary

In cancer, there are over 200 different forms. Out of 200, melanoma is the deadliest form of skin cancer. The diagnostic procedure for melanoma starts with clinical screening, followed by dermoscopic analysis and histopathological examination. Melanoma skin cancer is highly curable if it gets identified at the early stages. The first step of Melanoma skin cancer diagnosis is to conduct a visual examination of the skin's affected area. Dermatologists take the dermoscopic images of the skin lesions by the high-speed camera, which have an accuracy of 65-80% in the melanoma diagnosis without any additional technical support. With further visual examination by cancer treatment specialists and dermoscopic images, the overall prediction rate of melanoma diagnosis raised to 75-84% accuracy. The project aims to build an automated classification system based on image processing techniques to classify skin cancer using skin lesions images.

Problem Statement

To build a CNN based model which can accurately detect melanoma. Melanoma is a type of cancer that can be deadly if not detected early. It accounts for 75% of skin cancer deaths. A solution that can evaluate images and alert dermatologists about the presence of melanoma has the potential to reduce a lot of manual effort needed in diagnosis.

Algorithms Used

Convolutional Neural Network

Dataset Information

The dataset consists of 2357 images of malignant and benign oncological diseases, which were formed from the International Skin Imaging Collaboration (ISIC). All images were sorted according to the classification taken with ISIC, and all subsets were divided into the same number of images, with the exception of melanomas and moles, whose images are slightly dominant.

The data set contains the following diseases:

- Actinic keratosis
- Basal cell carcinoma
- Dermatofibroma
- Melanoma
- Nevus
- Pigmented benign keratosis
- Seborrheic keratosis
- Squamous cell carcinoma
- Vascular lesion

Technologies Used

- Python
- Tensorflow
- Keras
- Augmentor
- Matplotlib
- NumPy

Steps Involved

- Data Loading
- Baseline Model Building
- Training the Model and testing the model
- Building an augmented model
- Training the augmented model and testing the model
- Countering Class Imbalance with augmentor
- Building the final model
- Training the final model and testing the model
- Verifying the model
- Using modal against test data

Prepared Modal

To classify skin cancer using skin lesions images. To achieve higher accuracy and results on the classification task, I have built custom CNN model.

- Rescalling Layer - To rescale an input in the $[0, 255]$ range to be in the $[0, 1]$ range.
- Convolutional Layer - Convolutional layers apply a convolution operation to the input, passing the result to the next layer. A convolution converts all the pixels in its receptive field into a single value. For example, if you would apply a convolution to an image, you will be decreasing the image size as well as bringing all the information in the field together into a single pixel.
- Pooling Layer - Pooling layers are used to reduce the dimensions of the feature maps. Thus, it reduces the number of parameters to learn and the amount of computation performed in the network. The pooling layer summarises the features present in a region of the feature map generated by a convolution layer.
- Dropout Layer - The Dropout layer randomly sets input units to 0 with a frequency of rate at each step during training time, which helps prevent overfitting.
- Flatten Layer - Flattening is converting the data into a 1-dimensional array for inputting it to the next layer. We flatten the output of the convolutional

layers to create a single long feature vector. And it is connected to the final classification model, which is called a fully-connected layer.

- Dense Layer - The dense layer is a neural network layer that is connected deeply, which means each neuron in the dense layer receives input from all neurons of its previous layer.
- Activation Function(ReLU) - The rectified linear activation function or ReLU for short is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. The rectified linear activation function overcomes the vanishing gradient problem, allowing models to learn faster and perform better.
- Activation Function(Softmax) - The softmax function is used as the activation function in the output layer of neural network models that predict a multinomial probability distribution. The main advantage of using Softmax is the output probabilities range. The range will 0 to 1, and the sum of all the probabilities will be equal to one.

Conclusions

The final model has an accuracy of 83% and a loss of 0.45. The model is able to predict the class of the lesion with a high accuracy. Augmenting the data and countering class imbalance helped in improving the accuracy of the model.

Acknowledgements

Melanoma Skin Cancer from <https://www.cancer.org/cancer/melanoma-skin-cancer/about/what-is-melanoma.html>

CNN Explained from <https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939?gi=85d3a69d4225>

Image classification using CNN from <https://www.analyticsvidhya.com/blog/2020/02/learn-image-classification-cnn-convolutional-neural-networks-3-datasets/>

Efficient way to build CNN architecture from <https://towardsdatascience.com/a-guide-to-an-efficient-way-to-build-neural-network-architectures-part-ii-hyper-parameter-42efca01e5d7>

Contact

This solution is created by @pragya2611. You can post me your questions on the above modal and dataset.