Skin Lesion Analysis Towards Melanoma Detection and Classification Using Machine Learning

Yash Mudgal
Indian Institute of Information Technology
Sonepat - 131021
yash.11912080@iiitsonepat.ac.in

Abstract

Skin cancer is a typical common cancer. Melanoma is also known as malignant melanoma which is the most lethal form of skin cancer and responsible for 75% of skin cancer deaths, despite being the least common skin cancer. The best way to combat from this - is trying to identify it as early as possible and treat it with minor surgery.

During this paper, I will systematically study about melanoma and how can we detect melanoma in early stages using Machine learning and deep learning techniques and try to optimize previous works based in these areas. Previous researches are done this detection near accuracy of 80% and I will try to optimize this accuracy and try to increase to near 90%. I will also compare different machine learning and deep learning techniques during this research.

Keywords—skin cancer, melanoma, convolutional neural network, classification, machine learning algorithms, transfer learning.

I. Introduction

Skin cancer is the most prevalent cancer worldwide. It turns out that survival is very high if you can diagnose it at an early stage and choose the ideal treatment. Therefore, it is essential to find out as soon as possible whether a patients symptoms correspond to cancer. Traditionally, doctors detected skin cancer with the naked eye. However, this often leads to results that are not so accurate that humans make mistakes. Even experts have a hard time saying that, especially when the cancer is in its very early stages. This is where machine learning can help automate the entire process or pipeline. Deep neural networks can process thousands of images in both categories - Benign and malignant.

By learning the non-linear interactions, the model can tell whether a new image corresponds to a malignant or benign class. This automation could not only have a higher efficiency in avoiding both false positives and false negatives but also reduce time and manual spent on more productive work. Our model can be deployed on locations which lack professional docs. This work could dramatically change the healthcare and well-being of humans in developing countries of Africa and Asia.

II. Project Objectives

- Analysis of previous implemented research papers related to project.
- Gathering Data Set and Feature Engineering.
- Training different Machine Learning and deep learning Models on Dataset.

- Comparison of different methods to implement Melanoma Detection.

III. Related Work

Lot of researches has been published in the area of skin cancer classification using deep learning and computer vision techniques. Various approaches have been used in these works, including classification only, segmentation and detection, and image processing with different kinds of filters.

(Mahbod A, Schaefer G, Wang C, et al.,2019) used CNN but they have included popular AlxNet, VGG16, ResNet18. I will use the novel models for ResNet50, Inception V3, MobileNet and EfficientNet. EfficientNet uses a baseline network created by neural architecture search to scale all dimensions of depth, width, image resolution using a simple yet highly effective method called compound coefficient. This greatly enhances the ability to capture richer and much more complex features for melanoma recognition.

(Jinen Daghrir, Lotfi Tlig, Moez Bouchouicha, Mounir Sayadi, et al., 2020) used hybrid approach to achieve this classification. They used deep learning and classical machine learning algorithms and at last fuse them. But they have not very efficient ml model. Even after fusion of 3 models – KNN, SVM and CNN. They have got the accuracy of 88.4%.

(Esteva et al., 2017) separately used AdaBoost to classify the skin lesions. (Xu et al., 2014) used various sets of features including type of texture, lesions, colour etc and neural networks for the making of a robust diagnosis system.

The examples so far have only have only shown algorithms using traditional machine learning techniques, but recently deep learning has proven to be more accurate. The reason for that it automates the feature extraction process completely. It is upto the algorithms to find the better features and train the model accordingly. (Lopez et al., 2017) used a pre-trained GoogleNet Inception v3 CNN model to classify 129,450 clinical skin cancer images for skin cancer classification. This was the breakthrough that has been achieved. (Dorj et al., 2018) developed a convolutional neural network with over 50 layers on ISBI 2016 challenge dataset for the classification of malignant melanoma. In 2018, (Brinker et al., 2018) utilized a deep convolutional neural network to classify 2 a binary class problem of dermoscopy images. (Rezvantalab et al., 2018) developed an algorithm that uses a support vector machine combined with a deep convolutional neural network approach to classify four diagnostic categories of clinical skin cancer images. (Codella et al., 2017) used a deep convolutional neural network to classify the clinical images of 12 skin diseases.

In this paper I will tackling skin cancer classification which is of binary type i.e., there are 2 classes present in the dataset benign and malignant.

IV. Dataset

I am using Kaggle dataset – Skin Cancer- Malignant or benign dataset which contains total of 3297 images. This dataset includes balanced data of images of benign skin moles and malignant skin moles. I am using 660 images for testing the model and for training I am using 2637 images for training the model. Data is split into 2 folders – train and test. Both folders contain 2 separate folders for benign and malignant(cancerous) images. This dataset is pre-processed from ISIC. The data consists of two folders with each 1800 pictures (224x244) of the two types of moles.

The pictures have all been resized to low resolution (224x224x3) RGB.

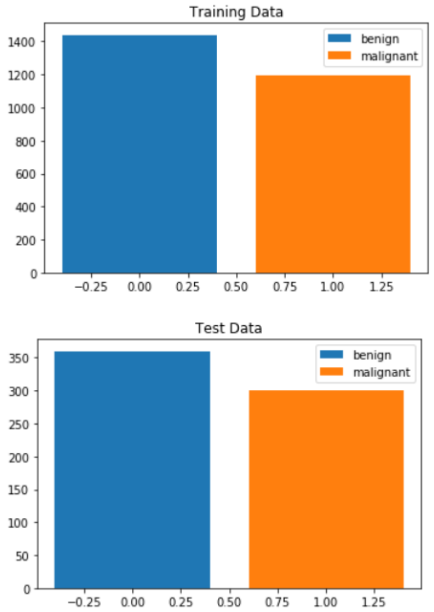


Fig 1. Dataset Visualization

V. Proposed Solution

My proposed is that I should try all the best classifiers models and pick best one of them or compare between them and suggest best model for this problem. So, I am able to find best model – DenseNet-121 with 90.90%.

I have implemented these model which are as follows:

1. Support Vector Classifier (SVC)
2. InceptionV3
3. MobileNet.
4. Logistic Regression
5. DenseNet-121
6. ResNet-50
7. VGG16
8. VGG19
9. ResNet-152
10. DenseNet-201
11. Efficient-B6

VI. Support Vector Classifier

Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

VII. InceptionV3

Google's Inception v3 architecture works by fine-tuning all layers, replacing the upper layer with an average pooling layer, two fully connected layers and finally a softmax layer that allows classification of two diagnostic categories was retrained on the dataset. The input image has been resized to (224, 224) to be compatible with this model. The learning rate was set to 0.0001 and Adam was used as the optimizer.

VIII. MobileNetV2

A fundamental part of MobileNet is a depthwise separable filter called depthwise separable convolution. These convolution layers are a variant of factored convolution, which decomposes standard convolutions into depth-wise convolutions called point-wise convolutions. In MobileNet, depth convolution applies one filter to each input channel. Convolution by points applies a convolution operation to combine the outputs of convolution by depth.

IX. Logistic Regression

Logistic regression is a statistical model used for binary classification tasks. It estimates the probability of an event occurring based on input features and assigns a class label. It's a popular and interpretable algorithm widely used in various fields for its simplicity and effectiveness.

X. ResNet-50/152

ResNet-50 and ResNet-152 are deep convolutional neural network architectures that have significantly advanced image recognition tasks. They are part of the

ResNet family, known for their skip connections, which alleviate the vanishing gradient problem. ResNet-50 consists of 50 layers, while ResNet-152 has 152 layers, making it deeper and more powerful. These models employ residual learning, where each layer learns the residual mapping instead of directly estimating the desired output. This enables them to handle more complex features and achieve higher accuracy on large-scale image classification benchmarks, making them popular choices for deep learning-based computer vision applications.

XI. VGG16/19

The VGG (Visual Geometry Group) family includes VGG16 and VGG19, which are popular convolutional neural network architectures for image classification tasks. They are known for their simplicity and uniformity, comprising stacked convolutional layers with small filter sizes (3x3) and max pooling layers. VGG16 has 16 layers, while VGG19 has 19 layers. These models demonstrate impressive performance by capturing intricate patterns through deep feature extraction. However, their increased depth makes them computationally expensive. Despite being surpassed by newer architectures like ResNet, the VGG family remains influential in research and serves as a benchmark for evaluating the performance and generalization of newer deep learning models.

XII. DenseNet-121/201

DenseNet201 and DenseNet121 are deep convolutional neural network architectures that belong to the DenseNet family. These models introduce the concept of dense connections, where each layer receives direct input from all preceding layers. DenseNet201 has 201 layers, while DenseNet121 has 121 layers. This dense connectivity fosters feature reuse, alleviating the vanishing gradient problem and promoting information flow throughout the network. By combining feature maps from different depths, DenseNets can effectively extract complex features, enhance feature propagation, and achieve impressive performance on image classification tasks. They have demonstrated state-of-the-art results and are widely used in various computer vision applications.

XIII. Efficient-B6

Efficient-B6 is a variant of the EfficientNet family, a series of neural network architectures that aim to achieve excellent performance while minimizing computational resources. Efficient-B6 is specifically designed to balance accuracy and efficiency, making it suitable for a wide range of computer vision tasks. It has a moderate number of layers and parameters, striking a balance between accuracy and computational complexity. By employing advanced techniques like efficient channel and spatial attention modules, Efficient-B6 effectively captures intricate features in images. It has shown competitive performance on various benchmarks, making it a popular choice for applications where a good trade-off between accuracy and efficiency is required.

XIV. Transfer Learning

Transfer learning is a common technique in computer vision that can produce accurate models more quickly. Transfer learning does not start the learning process from scratch, but from patterns learned while solving another problem. The advantages of using transfer learning are: Very easy to integrate. Achieves the same or better performance quickly (depending on dataset model performance). I don't need so much labelled data. Diverse use cases such as transfer learning, prediction, and feature extraction. The proposed method can be summarized in the following three points.

- Split the data set into two parts. The training and test sets have 80% and 20% images respectively.
- I tried pre-trained models such as Inception v3, MobileNet by fine-tuning the last layer of the network. We also used traditional machine learning approaches such as SVC and logistic regression.
- I trained a model with stack size 32 and 20 epochs by varying hyperparameters such as learning rate, stack size, optimizer, and pretrained weights.

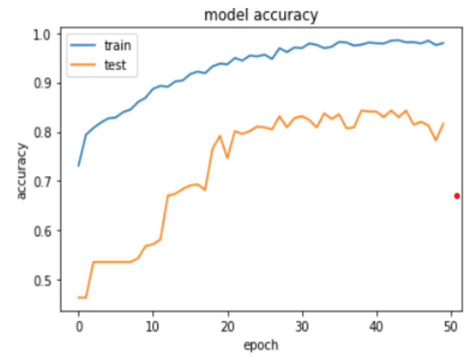
XV. Experimental Results

In this section, I shown my findings. I plotted the accuracy vs epochs, loss vs epochs and table for comparison of different used models. And I am able to achieve greater accuracy then my base papers. I have got DenseNet-121 as best performing model. This is transfer based learning model. I am able to achieve 90.90% accuracy during testing.

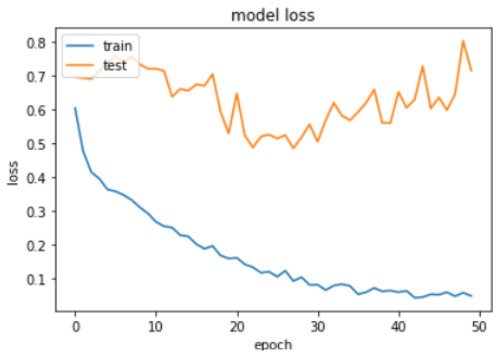
Model Name	**Accuracy Achieved During Testing in %**
Support Vector Classifier	83.48
InceptionV3	81.36
MobileNet	53.60
Logistic Regression	78
ResNet-50	81.32
VGG16	88.94
DenseNet-121	90.90
ResNet-152	88.18
Efficient-B6	86.36
DenseNet-201	89.24
VGG-19	88.40

Table 2. Experimental Results

- **For InceptionV3**



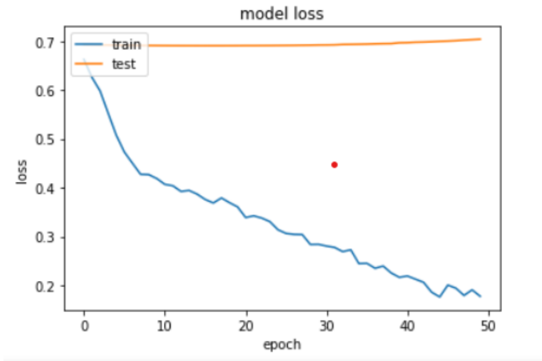
a.



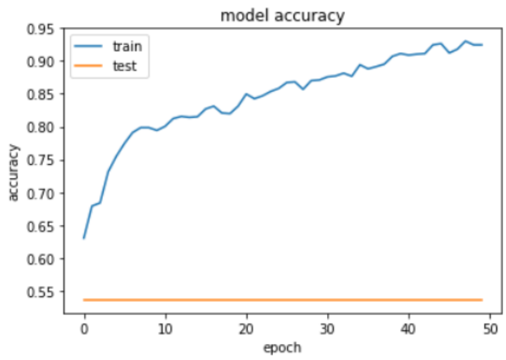
b.

Fig 2. (a) accuracy vs epoch
(b) loss vs epoch for InceptionV3

- **For MobileNetV2**



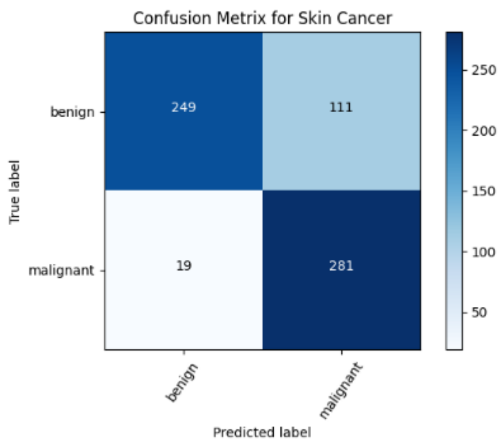
a.



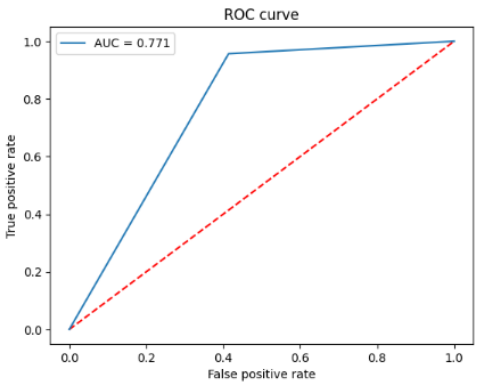
b.

Fig 3. (a) accuracy vs epoch
(b) loss vs epoch for MobileNet

- **For ResNet-50:**



a.



b.

Fig 5. (a) Confusion Metric for ResNet-50 model
(b) ROC Curve for ResNet-50 model

- **For DenseNet-121 (Best Model):**

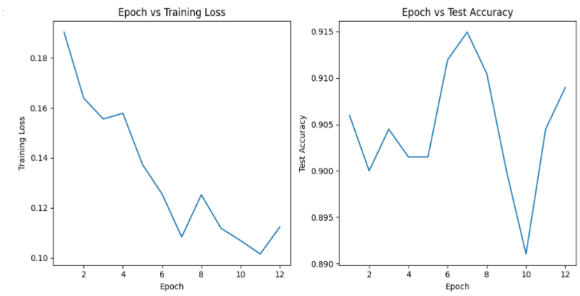


Fig 4. (a) Epoch vs Training Loss & Test Accuracy for DenseNet-121

XVI. Conclusion

In summary, this study investigated the ability of deep convolutional neural networks in classifying benign and malignant skin cancers. My results show that state-of-the-art Deep learning architecture trained on dermoscopy images (a total of 3297, consisting of 2637 training sessions and 660 tests) outperforms base papers models. I have shown that by using very deep convolutional neural networks with transfer learning and fine-tuning them with dermoscopy images, we can achieve better diagnostic accuracy compared to experienced physicians and clinicians. These models can be easily implemented in dermatoscopy systems and smartphones to support dermatologists. Further improvements require a more diverse dataset (different categories, different ages) with more dermoscopy images and balanced samples per class.

I am able to surpass base papers accuracy. I am getting 90.90% accuracy on testing. I have compared different classical ML algorithms with newly transfer based learning algorithms. The best performed model in this comparison is DenseNet-121.

XVII. Future Scope

In the upcoming months, I will try to implement frontend part of detection of melanoma skin detection using this highly accurate DenseNet-121 model.

XVIII. References

[1]. Mahbod A, Schaefer G, Wang C, et al. Skin lesion classification using hybrid deep neural networks[C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019: 1229-1233.

[2]. Jinen Daghrir, Lotfi Tlig, Moez Bouchouicha, Mounir Sayadi. Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach//5th International Conference on Advanced Technologies For Signal and Image Processing, ATSIP' 2020 September 02-05, 2020, Sfax, Tunisia.

[3]. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologistlevel classification of skin cancer with deep neural networks. nature, 542(7639):115–118, 2017.

[4]. Xu, J. S. Ren, C. Liu, and J. Jia. Deep convolutional neural network for image deconvolution. In Advances in neural information processing systems, pages 1790–1798, 2014.

[5]. Lopez, X. Giro-i Nieto, J. Burdick, and O. Marques. Skin lesion classification from dermoscopic

images using deep learning techniques. In 2017 13th IASTED international conference on biomedical engineering (BioMed), pages 49–54. IEEE, 2017.
[6]. Brinker, A. Hekler, J. S. Utikal, N. Grabe, D. Schadendorf, J. Klode, C. Berking, T. Steeb, A. H. Enk, and C. von Kalle. Skin cancer classification using convolutional neural networks: systematic review. Journal of medical Internet research, 20(10):e11936, 2018.
[7]. Codella, Q.-B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, A. C. Halpern, and J. R. Smith. Deep learning ensembles for melanoma recognition in dermoscopy images. IBM Journal of Research and Development, 61(4/5):5–1, 2017.
[8]. A. R. Lopez, X. Giro-i Nieto, J. Burdick, and O. Marques. Skin lesion classification from dermoscopic images using deep learning techniques. In 2017 13th IASTED international conference on biomedical engineering (BioMed), pages 49–54. IEEE, 2017.
[9]. Kawahara and G. Hamarneh. Multi-resolution-tract cnn with hybrid pretrained and skin-lesion trained layers. In International workshop on machine learning in medical imaging, pages 164–171. Springer, 2016.
[10]. Nasr-Esfahani, S. Samavi, N. Karimi, S. M. R. Soroushmehr, M. H. Jafari, K. Ward, and K. Najarian. Melanoma detection by analysis of clinical images using convolutional neural network. In 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 1373–1376. IEEE, 2016.
[11]. Nazia Hameed, Antesar M. Shabut, M. A. Hossain, "Multi-Class Skin Diseases Classification Using Deep Convolutional Neural Network and Support Vector Machine", International Conference on Software, Knowledge Information, Industrial Management and Applications, SKIMA, ISBN-9781538691410, ISSN-25733214, 2019.
[12]. Hasan, Mahamudul & Barman, Surajit & Islam, Samia & Reza, Ahmed Wasif. (2019). Skin Cancer Detection Using Convolutional Neural Network. 254-258. 10.1145/3330482.3330525. Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M. F., & Petkov, N. (2015). MED-NODE: a computer-assisted melanoma diagnosis system using nondermoscopic images. Expert systems with applications, 42(19), 6578-6585
[13]. U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee. The skin cancer classification using deep convolutional neural network. Multimedia Tools and Applications, 77(8):9909–9924, 2018.