e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:03/March-2022 Impact Factor- 6.752 www.irjmets.com

SKIN CANCER CLASSIFICATION USING DEEP LEARNING TECHNIQUES

Tummalapalli Kaushik[*1], Tati Siva Neeraj[*2], Tumukunta Kishan Rahul[*3], Mr S Anjanayya[*4]

[*1,2,3]Student, Department Of Computer Science And Engineering, B V Raju Institute Of Technology, Medak, Telangana State, India.

[*4]Associate Professor, Department Of Computer Science And Engineering, B V Raju Institute Of Technology, Medak, Telangana State, India.

ABSTRACT

Skin cancer is one the most challenging and difficult type of cancer to predict it with a human eye. Stanford certified dermatologists tried to predict the cancer with their experience and can classify the cancer type accurately with an accuracy rate of 66-69% respectively. Skin cancer if identified at the earlier stages, can be cured to the max limit but if it is postponed, it can turn hazardous in less Two(2) years duration. The most dangerous skin cancer which takes 9 million lives every year around the world is the Melanoma type. The main aim of the paper is to use more computational methods which can help in detecting the skin cancer more easily as the resources for the detection and treatment for the skin cancer is very scarce. Some of the computational methods that we can use are deep learning techniques and machine learning techniques to better understand the patterns in the data and to efficiently classify the data into major 3 types of skin cancer types without using much highly used resources which are very less in general.

Keywords: Deep Learning, Machine Learning, Skin Cancer Classification, Computational Methods.

I. INTRODUCTION

For our research problem statement to understand more about the patterns of the skin cancer classification, we will be using the below AI based techniques to better understand patterns.

MACHINE LEARNING: Machine learning is a branch of Artificial Intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. With the help of Machine Learning, we try to find the best fit for our model.

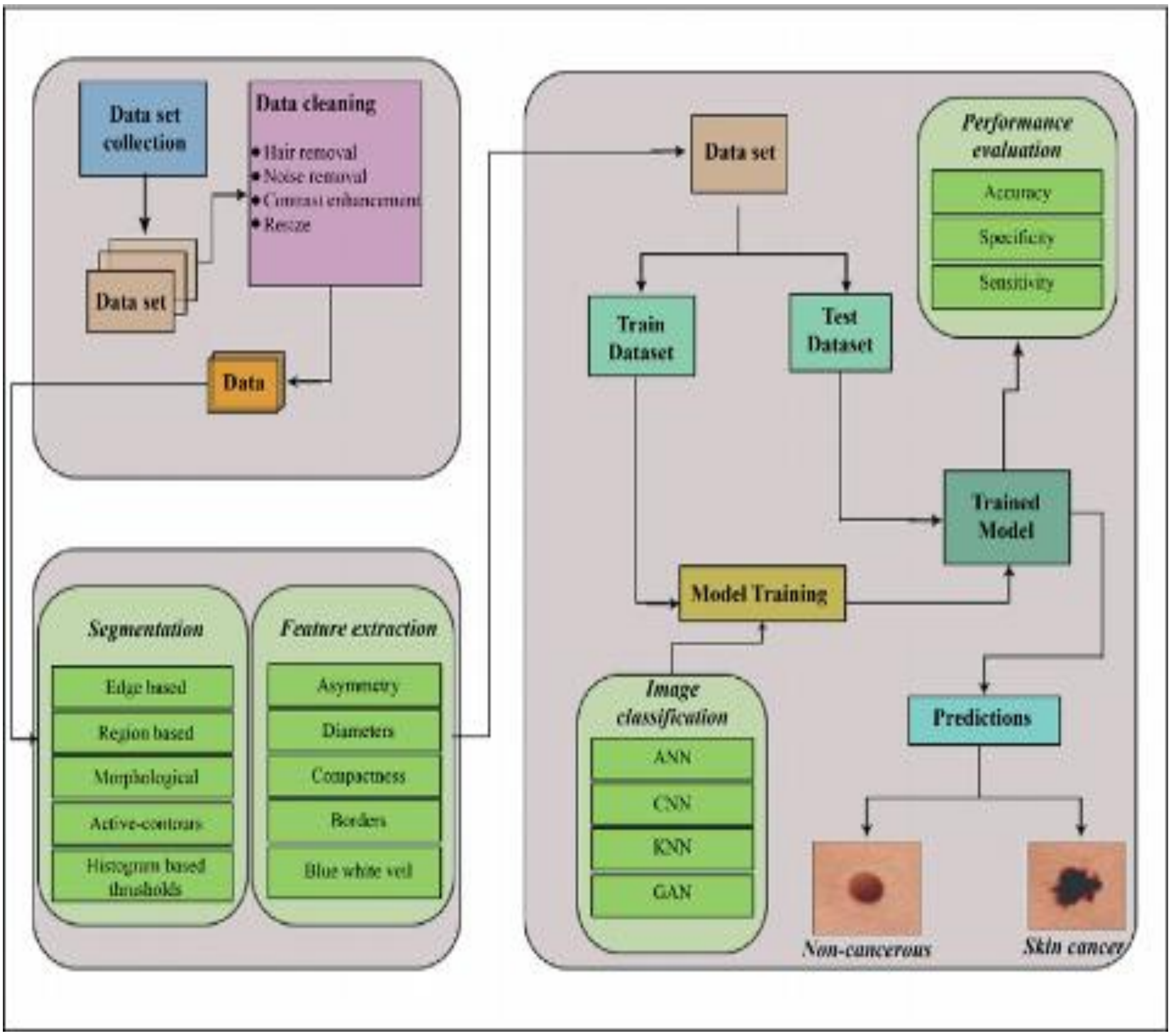
DEEP LEARNING: Using neural network architectures, we will be able to extract some of the in-depth features that leads to better results. Skin cancers, both non-melanoma and melanoma, have become more common in recent decades.

- Each year, around 2 to 3 million non-melanoma skin cancers and 132,000 melanoma skin cancers occur worldwide.

- According to the Skin Cancer Foundation, one out of every three cancers diagnosed is skin cancer, and one out of every five Americans will acquire skin cancer during their lifetime.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:03/March-2022 Impact Factor- 6.752 www.irjmets.com

II. METHODOLOGY



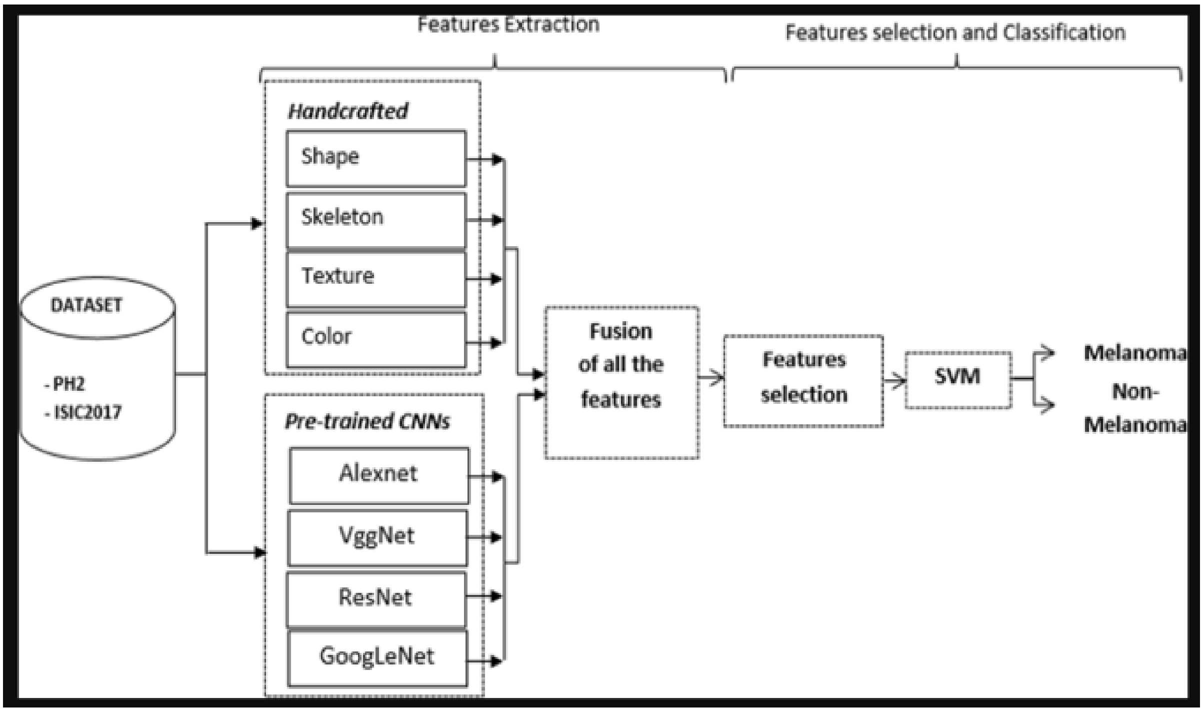
We get the data from the ISIC medical conference. The entire dataset which contains numerous images of images representing spots and marks on the skin. The Dataset is 11GB in total where we spilt the data into two or three types which is training dataset, testing dataset, and validation dataset which can be done in the ratio of 6:3:1 or 6.5:2:1.5 in general or can also be done without the usage of the validation dataset.

Explanation of key functions and modules used for this research project:

- Collecting the dataset
- Data Cleaning and Processing of the collected data.
- Data Segmentation
- Feature Extraction
- Splitting the Data

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:03/March-2022 Impact Factor- 6.752 www.irjmets.com

UML Diagram for the research Project:



- **Collecting the Dataset** : From ISIC-More than 10GB]
- **Data cleaning and processing of the data set:** In Data cleaning the system detect and correct corrupt or inaccurate records from database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing , modifying or detecting the dirty or coarse data. In Data processing the system convert data from a given form to a much more usable and desired form i.e. make it more meaningful and informative.
- **Data Segmentation:** In the images for various types of cancers in the data set, we will be doing Data Segmentation, as only need the important details of the image which is the spots identification on the skin, so in the images, it is mentioned with a red collar mark which tells that it is important for our project to capture and we will be training on that respective source.
- **Feature Extraction:** In the images, the algorithms in the training phase, identifies some of the important features like borders, asymmetry and diameter of the respective spots on the skin and the same will be used as parameters for our pattern understanding by the algorithms.
- **Training and Testing**: After the splitting of the data that we collected from the conference, we will be using this for training and testing purposes for the algorithms to identify various patterns in our data.

OTHER MODULES USED:

• **StandardScaler:** It follows Standard Normal Distribution (SND). Therefore, it makes mean = 0 and scales the data to unit variance. It standardize the data values into a standard format.

• **Numpy:** NumPy aims to provide an array object that is up to 50x faster than traditional Python lists. The array object in NumPy is called ndarray, it provides a lot of supporting functions that make working with ndarray very easy.

III. MODELING AND ANALYSIS

Using Resnet 50 model and standard ML models like Linear SVM and Non-linear SVM to train the model, and for testing it.[Predictions could be submitted on validation and test data sets.] In existing systems, the current systems which are out in the market are very costly and not very efficient as we have seen that the accuracy for the skin cancer classification is way less when the researchers got the predictions done by the Stanford dermatologists.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:03/March-2022 Impact Factor- 6.752 www.irjmets.com

```
model_transfer = models.resnet152(pretrained=True)

for param in model_transfer.parameters():
    param.requires_grad = True

model_transfer.fc = nn.Linear(2048, 3)
fc_parameters = model_transfer.fc.parameters()

for param in fc_parameters:
    param.requires_grad = True

use_cuda = torch.cuda.is_available()

print(model_transfer)
if use_cuda:
    model_transfer = model_transfer.cuda()

model_transfer.load_state_dict(torch.load(
    'final.pth', map_location=torch.device('cpu')))
```

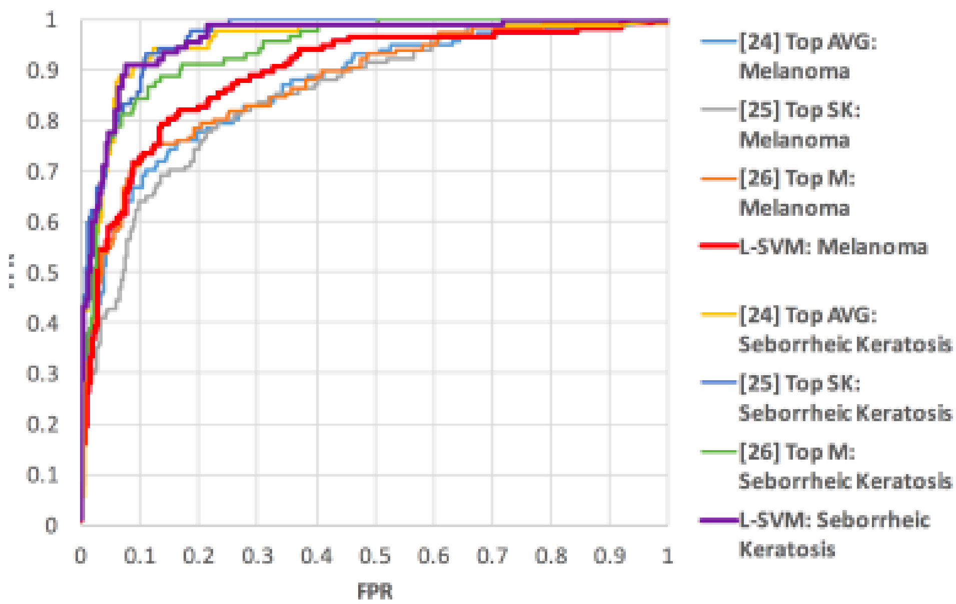
• **Resnet 50 Model:** ResNet-50 is **a convolutional neural network that is 50 layers deep**. You can load a pre-trained version of the network trained on more than a million images from the ImageNet database [1]. The pre-trained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals.

* We use techniques like Backpropagation in our training phase which will help us to reduce the overall error rate and can improve our model compoundly based on the number of epochs or the iterations that we travel in the training loop.

IV. RESULTS AND DISCUSSION

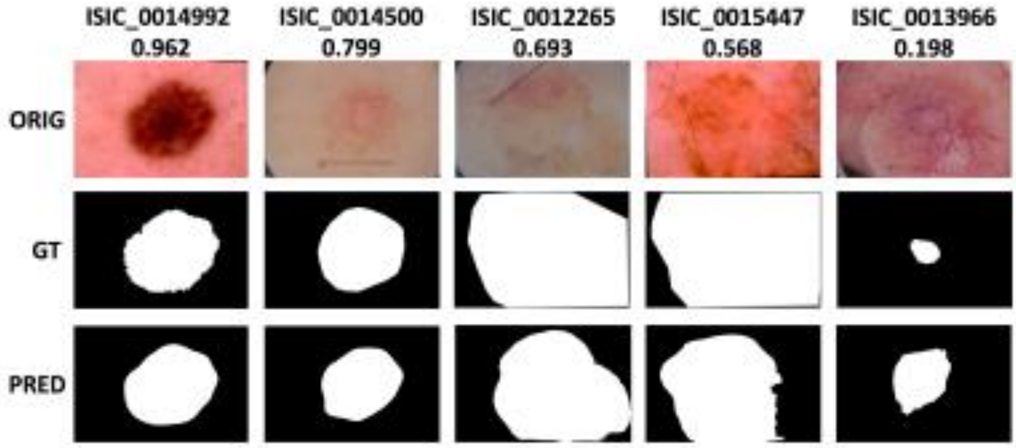
MEASURES OF EVALUATION:

The area under the curve (AUC) measurement calculated from the receiver operating characteristic (ROC) curve was used to evaluate categorization judgments. Any confidence level above 0.5 was seen as positive for a category when making classification decisions. Pixel values greater than 128 were deemed positive, while those less than 128 were regarded negative for segmentation tasks. Furthermore, specificity was tested on the operational curve for melanoma classification, with sensitivity of 82 percent, 89 percent, and 95 percent, respectively, conforming to dermatologist classification and management performance levels, and theoretically desirable sensitivity values. The Jaccard Index, Dice coefficient, and p values were used to compare segmentation submissions.



e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:04/Issue:03/March-2022 Impact Factor- 6.752 www.irjmets.com

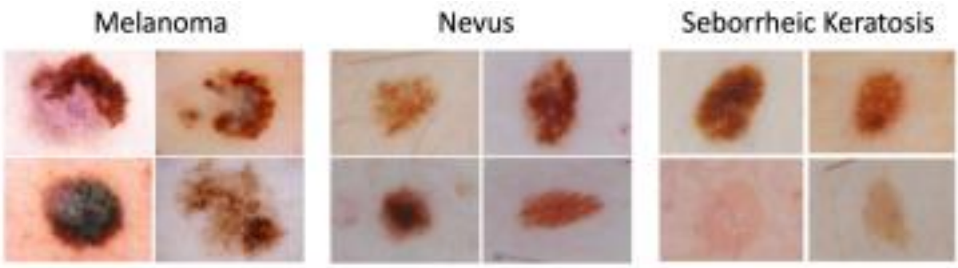
Lesion Segmentation task: We perform the Lesion segmentation task for all of the images which represents the skin images with various spots and after segmentation, we can move ahead with feature extraction and then using the Resnet pre-trained models and various machine learning algorithms.



Lesion Segmentation

The **accuracy** of RESNET Architecture is nearly **75%.** The accuracy can be improved overall by using techniques which reduce the imbalanced proportion of the images, so that the bias can be reduced to a great extent. Other CNN based architectures can perform well but RESNET bing trained on 1M+ images with some fine-tuning can bring wonders when it comes to pattern recognition and RESNET is one of the best performing CNN architectures.

IMAGE RESULT:



V. CONCLUSION

The research project can be used to perform real time analysis which will help in detecting the skin cancer early and can save many lives across the world as its really efficient, easy to use and also not all expensive to use. The classification experiment revealed that ensembles of deep learning algorithms combined with additional data produced the best results. Furthermore, collaborative fusions of all participant systems outperformed any single system by a significant margin. Except for the approaches offered produce minimal human-interpretable evidence of disease diagnosis. Future research or difficulties may place a greater emphasis on the importance of proper integration into healthcare workflows. Dataset bias (not all diseases, ages, devices, or races were represented equally across categories) and insufficient dermoscopic feature annotations were among the study's limitations. It's also possible that relying on single evaluation metrics rather than combinations is a drawback. Future challenges will work together with the community to address these issues. This Project Work Can Be Extended To Higher Level In Future. There are Many ML algorithms and Deep learning architectures that can be incorporated to increase the accuracy. This project has the scope to be embedded in an end-end application like a web application or an android application.

VI. REFERENCES

[1] Noel C. F. Codella , David Gutman, Brain Helba, Aadi kaloo, Micheal A. Marchetti, Stephen, Nabin Mishra, Allan Halpern: “Skin Lesion Analysis Towards Melanoma Detection”, International Symposium on Bio Medical Imaging, Hosted by the International Skin Imaging Collaboration(ISIC) pp.1–5.

[2] Skin Cancer Detection: A review using Deep learning techniques pp 1-20.

[3] RogersHW,WeinstockMA,FeldmanSR,ColdironBM.: “Incidence estimate of nonmelanoma skin cancer (ker- atinocyte carcinomas) in the US population, 2012” JAMA Dermatol vol. 151, no. 10, pp. 1081-1086. 2015

[4] Daz IG. “Incorporating the Knowledge of Dermatologists to Convolutional Neural Networks for the Diagnosis of Skin Lesions”. International Skin Imaging Collaboration (ISIC) 2017 Challenge at the International Symposium on Biomedical Imaging (ISBI). Available: https://arxiv.org/abs/1703.01976

[5] Menegola A, Tavares J, Fornaciali M, Li LT, Avila S, Valle E. “RECOD Titans at ISIC Challenge 2017”. International Skin Imaging Collaboration (ISIC) 2017 Challenge at the International Symposium on Biomedical Imaging (ISBI). Available: https://arxiv.org/pdf/1703.04819.pdf

[6] Fishbaugh, J, et al. ”Data-Driven Rank Aggregation with Application to Grand Challenges.” International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Springer, pp. 754-762. 2017

[7] Argenziano, G. et al.: “Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet” J. American Academy of Dermatology. vol. 48, no. 5, 2003.

[8] Gachon, J., et. al.:“First Prospective Study of the Recognition Process of Melanoma in Dermatological Practice”. Arch Dermatol. vol. 141, no. 4, pp. 434-438, 2005.

[9] Kimball, A.B., Resneck, J.S. Jr.: “The US dermatology workforce: a specialty remains in shortage.” J Am Acad Dermatol. vol. 59, no. 5, pp. 741-5. 2008.

[10] Mishra, N.K., Celebi, M.E.: “An Overview of Melanoma Detection in Dermoscopy Images Using Image Processing and Machine Learning” arxiv.org: 1601.07843. Available: http://arxiv.org/abs/1601.07843

[11] Ali, A.A., Deserno, T.M.: “A Systematic Review of Automated Melanoma Detection in Dermatoscopic Images and its Group.