Skin Cancer Detection and Classification Using Deep Learning Approach

Dhanush S, Bhargava Reddy D, Mahesh Kumar G,
R.M.K College of Engineering and Technology
Puduvoyal, Tamilnadu, India

Email: sdha20cs117@rmkcet.ac.in, dont20cs028@rmkcet.ac.in, gorl20cs175@rmkcet.ac.in

Abstract: Skin cancer is one of the most threatening and rapidly spreading illnesses in today's world. Detecting skin cancer was a challenging task for dermatologists, but the rapid development of technologies has made skin cancer detection easier. Using these advancements, the early stages of skin cancer can be identified, and better treatments can be provided. In recent years, with the help of deep learning, Convolutional Neural Networks (CNNs) have emerged as powerful tools for diagnosing skin cancer. In this paper, the Human Against Machine (HAM) 10000 dataset is used to demonstrate skin cancer classification, consisting of seven different types of skin lesions with a sample size of 10,015 used for experimentation. Along with the CNN network, the MobileNet model is used for detection, following the Deep Learning methodology. The dataset preprocessing steps are illustrated, and the training process is explained. The performance of the models is displayed.

Index Terms - Skin Cancer, Deep Learning, CNN, MobileNet, Image Processing, Classification and Detection.

I. INTRODUCTION :

Skin cancer is becoming more common, so finding it early is very important for treating it well. Doctors used to mostly look at skin with their eyes to find cancer signs, but this method wasn't always accurate. Now, with new technology like machine learning and deep learning, we have better ways to find skin cancer. These technologies can quickly and accurately analyze lots of medical information. As skin cancer cases increase worldwide, there's a bigger need for quick and reliable ways to diagnose it. Old methods, while somewhat helpful, could take a long time and weren't always accurate. But by using advanced algorithms like Convolutional Neural Networks (CNNs) and models like MobileNet, we can analyze large amounts of data to create more accurate and automatic diagnosis tools. These new methods could change how we find and treat skin cancer, possibly leading to earlier detection, better treatment, and saving lives. This paper looks at how we can use CNNs and the MobileNet model to understand different types of skin problems using the HAM 10000 dataset. By using advanced technology, we hope to improve how we find and treat skin cancer early. Our goal is to make skin cancer diagnosis easier and more accurate for doctors and patients. We're working hard to help fight skin cancer and make a big difference in saving lives. With continued research and innovation, we aim to push the boundaries of skin cancer diagnosis and treatment to ensure better outcomes for patients worldwide.

II. LITERATURE REVIEW :

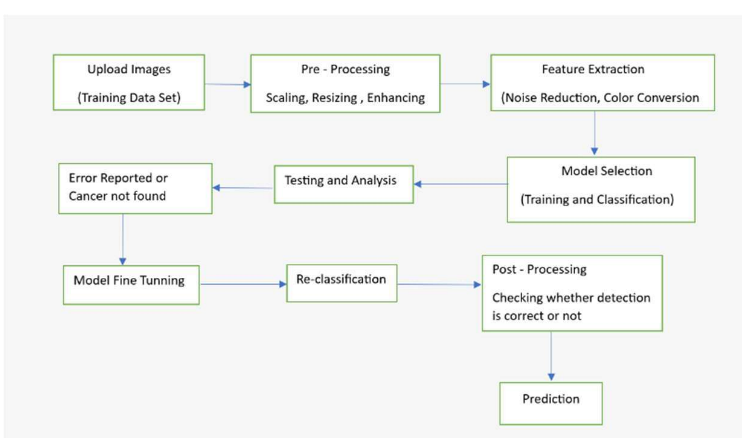
In the field of skin cancer research, there have been big improvements in how we classify the disease. One exciting development is the use of deep learning, especially Convolutional Neural Networks (CNNs). These methods are really good at telling different types of skin cancer apart, but they can be tricky to train when we don't have a lot of examples to learn from. To tackle this challenge, researchers have come up with partial transferable CNNs, which can adapt to new datasets with different qualities. Another helpful method is transfer learning, which can handle images with different levels of detail, making our classifications more accurate.

In the past, most efforts focused on simply telling benign (harmless) and malignant (dangerous) lesions apart. Some techniques, like K-Means and Support Vector Machines, were quite accurate but had limitations, like biases in the data and incomplete information. Now, with datasets like HAM10000, which include many different types of skin lesions, we can build more advanced models. Custom CNNs and transfer learning have become powerful tools for identifying different types of skin lesions accurately. Looking ahead, researchers are working hard to improve classification accuracy even more. They're balancing datasets better and trying out different transfer learning models to see which ones work best. At the same time, they're exploring new ways to prepare images and separate lesions from surrounding skin. As skin cancer research progresses, we're getting closer to better ways of detecting and treating the disease early, which is really promising for patients everywhere. In addition to making skin cancer classification better, researchers are working on ways to diagnose the disease more easily and quickly. They're using technology like telemedicine and digital health platforms to create tools that can help doctors and patients spot potential skin cancer spots from afar. These tools, like phone apps and websites, can analyze pictures of skin and tell if a spot needs a doctor's attention. This could be especially helpful in places where there aren't many skin doctors around. Also, scientists are studying the genes and molecules involved in skin cancer to find new ways to treat it. By understanding these tiny details, they hope to create treatments that work better for each person and have fewer side effects.

III. METHODOLOGY :

Detecting skin cancer early is hard for doctors. Using deep learning methods like Convolutional Neural Networks (CNNs) and MobileNet can help classify the different types of skin cancer. These methods are good at finding objects and sorting them. Here's how we used CNNs and MobileNet in our study:

Figure 1. Flow Diagram



1. Data Preprocessing :

Compile a diverse dataset containing images depicting various types of skin cancer lesions. Standardize image formats, resolutions, and color spaces to ensure consistency across the dataset. Apply preprocessing methods to enhance image quality and remove noise for improved model performance.

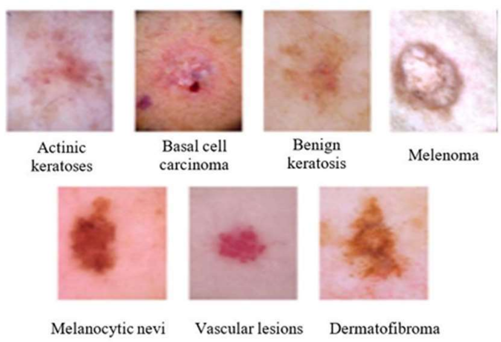


Figure 2. Skin lesions

Table 1. Hyper parameters used in data augmentation

Hyper parameter	Value
Rescale	1/255
Rotation range	40
Width shift range	0.2
Height shift range	0.2
Shear range	0.2
Zoom range	0.2
Horizontal flip	True
Fill mode	Nearest
Validation split	0.4

2. Model Selection and Design :

Choose between MobileNet and CNN architectures based on factors like computational efficiency and performance. Design the selected model architecture with appropriate layers and parameters optimized for skin cancer classification tasks.

3. Transfer Learning and Training :

Utilize transfer learning to initialize the MobileNet and CNN models with pre-trained weights from large-scale image datasets. Train the models using the skin cancer dataset, fine-tuning the parameters to adapt to specific lesion characteristics.Validate the trained models using a separate validation set to assess their performance and prevent overfitting.

Table 2. Transfer learning

Layer (type)	Output Shape	Param #
keras_layer (KerasLayer)	(None, 1280)	2257984
flatten (Flatten)	(None, 1280)	0
dense (Dense)	(None, 512)	655872
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 2)	1026

Total params: 2914882 (11.12 MB)
Trainable params: 656898 (2.51 MB)
Non-trainable params: 2257984 (8.61 MB)

4. Evaluation and Performance Analysis :

Evaluate the performance of the MobileNet and CNN models using standard evaluation metrics such as accuracy, precision, recall, and F1-score. Analyze the models' performance on a testing dataset to assess their ability to generalize and detect skin cancer lesions accurately.

Table 3. Performance metrics

Performance metric	Acronym	Equation
Positive Predictive Value	PPV	TP / TP + FP
Negative Predictive Value	NPV	TP / TP + FN
True Positive Rate	TPR	TP / TP + FN
True Negative Rate	TNR	TP / TP + FP
Accuracy	ACC	TP / TP + FN

5. Post-processing :

Implement post-processing techniques to refine model predictions and enhance classification accuracy. Optimize model parameters and architectures based on the evaluation results to improve overall performance and robustness.

Table 4. Training Model for 15 Epochs

Epochs	Train_loss	Val_loss	Accuracy
1	1.2462	0.8517	0.5000
2	0.5312	0.8384	0.7500
3	0.1674	0.8674	0.9167
4	0.1511	1.0748	0.9167
5	0.0523	0.9116	1.0000
6	0.0800	0.9188	0.9583
7	0.0378	0.8621	1.0000
8	0.0492	0.7957	1.0000
9	0.0051	1.2771	1.0000
10	0.0571	1.3072	0.9583
11	0.0048	1.0587	1.0000
12	0.0180	1.9372	1.0000
13	0.1211	1.6062	0.9583
14	0.0636	1.0675	0.9583
15	0.0647	1.2212	0.9583

Softmax is a function that takes the output values from the model and turns it into a probability function. The softmax function is given by the formula in equation,

$$e^{zi} / \sum_{j=1}^{k} e^{z_j}$$

Learning rate = 0.001

Batch Size = 16

Figure 3. Training and Validation accuracy

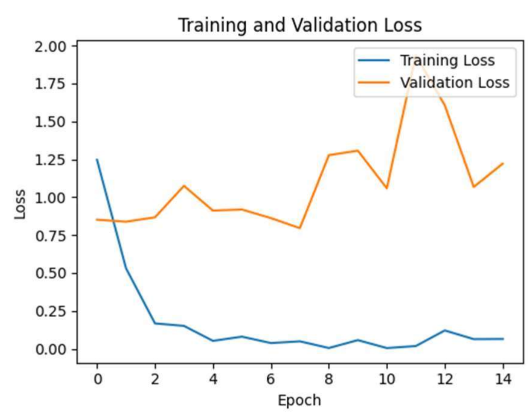


Figure 4. Training and Validation loss

Table 5. Training Model for 20 Epochs

Epochs	Train_loss	Val_loss	Accuracy
1	0.0182	1.5293	1.0000
2	0.0756	1.4915	0.9583
3	0.0140	1.1138	1.0000
4	0.0133	1.0600	1.0000
5	0.0060	1.1284	1.0000
6	0.0014	1.3557	1.0000
7	0.0160	1.3077	1.0000
8	0.0802	1.8347	0.9583
9	0.0016	1.0599	1.0000
10	0.0028	1.3072	1.0000
11	0.0053	1.4885	1.0000
12	0.0017	1.3288	1.0000
13	0.0163	1.5221	1.0000
14	0.1331	0.9494	0.9583
15	0.4000	1.5189	1.0000
16	0.0261	1.1816	1.0000
17	0.0065	1.4341	1.0000
18	0.0312	1.7379	0.9583
19	0.1484	1.7767	0.9583
20	0.0073	1.6506	1.0000

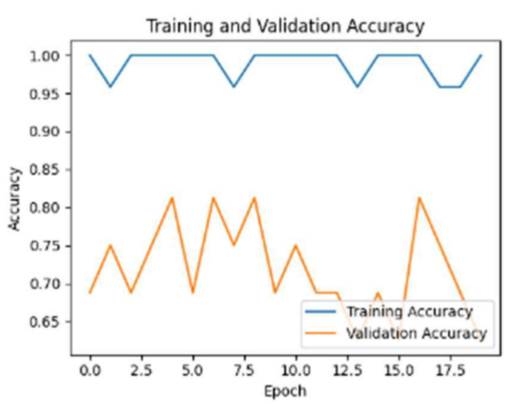


Figure 5. Training and Validation accuracy

Figure 6. Training and Validation loss

IV. RESULT AND DISCUSSION :

The confusion matrix is a crucial tool used to evaluate the performance of a classification model, such as our skin cancer classification system. It provides a detailed breakdown of the model's predictions compared to the actual ground truth across different classes.

Evaluation Loss: 1.2895033359527588

Evaluation Accuracy: 0.6875

1/1

[==============================]

- 1s 1s/step

Total Percentage for Benign: 60.64562563167419

Total Percentage for Malignant: 39.35437496355735

Thus, Our skin cancer classification system yielded promising results.

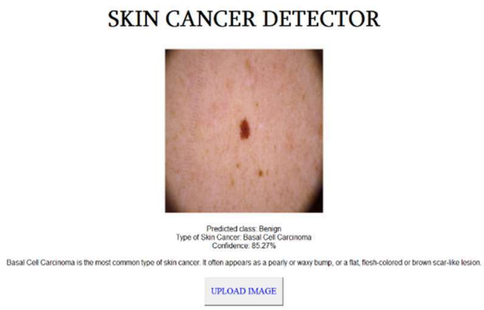


Figure 7. MobileNet model Result

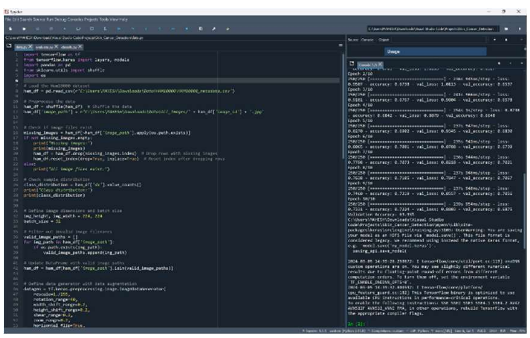


Figure 8. CNN model Result

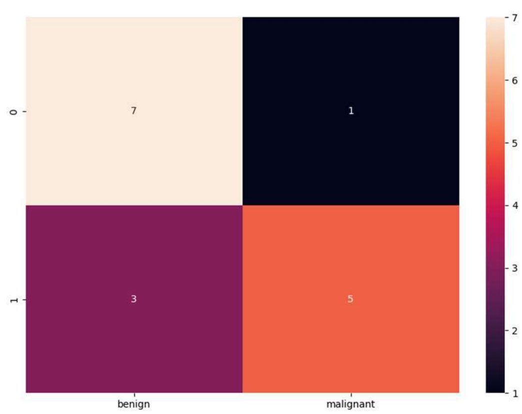


Figure 9. Confusion Matrix

V. CONCLUSION :

We conclude that the use of deep learning including the CNN algorithm and MobileNet model, has revolutionized skin cancer detection. The CNN algorithm, with its ability to recognize patterns in images, has been instrumental in analyzing dermatological images and distinguishing between cancerous and benign lesions. Additionally, the MobileNet model's lightweight architecture and efficient computation have further enhanced the accuracy and speed of detection.The combination of CNN and MobileNet has not only improved diagnostic precision but also holds promise for broader accessibility to effective detection methods. With ongoing innovation and implementation, we are moving closer to a future where skin cancer can be detected and treated more effectively, ultimately improving outcomes for patients worldwide.

VI. REFERENCES :

[1]https://www.wcrf.org/dietandcancer/cancer-trends/skin-cancer

Skin cancer statistics, date: Mar 30,2020.

[2] Geoffrey E. Hinton Alex Krizhevsky, Ilya Sutskever. 2012. Imagenet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems (2012).

[3] Xino Yao. 1999. Evolving artificial neural networks. Proc. IEEE 87, 9 (1999), 1423—1447

[4] Mirreza Mahbod, Gerald Schaefer, Chunliang Wang, Rupert Ecker, Isabella

Ellinger, “Skin Lesion Classification Using Hybrid Deep Neural Networks”, IEEE, International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.1229-1233,2019.

[5] Balazs Harangi, Agnes Baran , Andras Hajdu, “Classification of Skin Lesions Using An Ensemble of Deep Neural Networks”, IEEE, 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp.2575-2578, 2018.

[6] Signs and Symptoms of Skin Cancer, https://www.cancer.org/cancer/melanoma-skin-cancer/detection-diagnosis-staging/signsand-symptoms.html, accessed date: Mar 30,2020.

[7] Tests for Melanoma Skin Cancer, https://www.cancer.org/cancer/melanoma-skincancer/detection-diagnosis-staging/how-diagnosed.html,accessed date: Mar 30

[8] Esteva, Andre, et al. "A Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks." Nature 542.7639 (2017): 115-118.

[9] Yu, Lei, et al. "Automated melanoma recognition in dermoscopy images using very deep residual networks." IEEE transactions on medical imaging 37.12 (2018): 2821-2830.

[10] Ismael-Cima, Francisco, et al. "Skin lesion classification using deep learning and semantic segmentation." IEEE Access 6 (2018): 46450-46460.