International Scientific Journal of Engineering and Management
Volume: 03 Issue: 04 | April – 2024
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata
ISSN: 2583-6129
DOI: 10.55041/ISJEM01573

Enhancing Skin Cancer Detection Through Deep Learning Techniques

Prof. Sofiya Mujawar, Kumar Mayank, Aisha Singh, Kamiya Mujawar, Kalyani Hankare
sofiya.mujawar@dypatiluniversitypune.edu.in, mayankjaiswal1411@gmail.com, aishasingh17672@gmail.com, kamiyahm19@gmail.com, kalyanihankare1@gmail.com
B.Tech, Department of Computer Engineering, SOET
D Y Patil University, Ambi

Abstract– **Skin cancer is a common and possibly deadly disease that affects the skin's outer layers. Promoting knowledge of skin cancer, its risk factors, and the need for early detection can help combat the illness and reduce its impact on individuals and communities around the world. In this project, we describe a novel use of deep learning techniques to detect skin cancer in its early stages using dermatoscopic images. The primary goal of this study is to create an accurate and reliable model for predicting various types of skin cancer, including actinic keratoses and intraepithelial carcinoma (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (bkl), dermatofibroma (df), melanoma (mel), melanocytic nevi (nv), and pyogenic granulomas and hemorrhage (vasc). The project is written in Python, and the primary algorithm/model used is the Convolutional Neural Network (CNN) architecture. CNNs are ideal for image classification jobs because they can automatically learn relevant characteristics from input. Using CNNs, our model is trained on the HAM10000 (Human Against Machine with 10,000 training images) dataset, which contains 10015 high-resolution dermatoscopic images obtained from distinct populations and acquired via multiple modalities. The achieved findings illustrate the effectiveness of our proposed method. The model had an amazing training accuracy of 96.00% and a validation accuracy of 97.00%. These high accuracy rates demonstrate the promise of our deep learning-based skin cancer prediction system as a trustworthy tool for early diagnosis, allowing healthcare practitioners to make more informed decisions and improve patient outcomes. Our effort contributes to the field of dermatological research and machine learning, providing valuable insights into the use of deep learning algorithms for skin cancer prediction. Furthermore, the publicly available HAM10000 dataset, which includes a wide range of dermatoscopic pictures, can be a great resource for academic study and future advances in the field of medical image processing and classification.**

Key Words– **benign, CNN, early detection, skin cancer, machine learning**

I. INTRODUCTION

Skin cancer is one of the most prevalent types of cancer in the current decade [2]. Skin cancer is the most common type of cancer in humans, which is understandable given that it is the body's largest organ [3]. Skin cancer is categorized into two types: melanoma and nonmelanoma [4]. Melanoma is a dangerous, rare, and fatal form of skin cancer. Melanoma skin cancer accounts for only 1% of total incidences, but is associated with a higher death rate [5]. Melanoma develops in melanocytes. It begins when healthy melanocytes proliferate out of control, resulting in a malignant tumor. It can affect any part of the human body. It commonly arises on sun-exposed areas like hands, face, neck, and lips. Melanoma malignancies are only curable if detected early; otherwise, they spread to other body parts and cause the victim's terrible death [6]. There are several kinds of melanoma skin cancer, including nodular melanoma, superficial spreading melanoma, acral lentiginous, and lentigo maligna[4]. The majority of cancer cases are classified as nonmelanoma, which includes basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and sebaceous gland carcinoma. BCC, SGC, and SCC develop in the middle and top layers of the epidermis, respectively. These cancer cells have a low dissemination rate to other body areas. Non-melanoma cancers are easier to treat than melanoma malignancies.

As a result, early detection is crucial for skin cancer treatment [7]. Doctors typically utilize the biopsy procedure to identify skin cancer. This treatment involves removing a sample of a suspected skin lesion for medical testing to determine if it is malignant or not. This procedure is uncomfortable, slow, and time-consuming. Computer-based technology allows for a more comfortable, cost-effective, and timely identification of skin cancer symptoms. Several noninvasive approaches are offered for examining skin cancer symptoms, whether they reflect melanoma or not. The usual technique for skin cancer diagnosis is to take an image, preprocess it, segment it, extract the desired feature, and categorize it.

Deep learning has transformed the entire field of machine learning in recent decades. It is regarded as the most advanced machine learning subfield focused with artificial neural network techniques. These algorithms are modelled after the function and structure of the human brain. Deep learning approaches are applied in a variety of fields, including speech recognition [8], pattern recognition [9], and bioinformatics [10]. Deep learning systems have outperformed traditional machine learning approaches in various areas. In recent years, a variety of deep learning algorithms have been applied to identify skin cancer using computers. This study discusses and analyzes deep learning-based skin cancer detection approaches. This paper presents a comprehensive, systematic review of classical deep learning approaches for skin cancer detection, including artificial neural networks (ANN), convolutional neural networks (CNN), Kohonen self-

International Scientific Journal of Engineering and Management ISSN: 2583-6129
Volume: 03 Issue: 04 | April – 2024 DOI: 10.55041/ISJEM01573
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

organizing neural networks (KNN), and generative adversarial neural networks (GAN).

II. LITERATURE SURVEY

In the year 2021, Y. Jusman built a deep neural network to detect skin cancer using the HAM10000 data set, resulting in Multilayer Perceptron and Deep Neural Networks. By this, they have trained Transfer Learning. For additional reading, refer [11]

In the year 2021, R. Raja Subramanian used Convolutional Neural Network (CNN) for skin cancer detection and used the HAM10000 data set, achieving 83.11% accuracy, F score (0.82797%), precision (0.818642%), and recall. They have trained original images.For any further reading, refer to [12].

In the year 2020, Hari Krishan Kondaveeti employed Transfer Learning for skin cancer detection and used the HAM10000 data set to increase recall. They trained a Neural Network over 30 epochs and achieved category accuracy. For detailed information, see [13].

In the year 2020, Nourabuared employed VGG19 and Transfer Learning for skin cancer detection, and they used the HAM10000 data set to obtain the proposed strategy. By this, they have trained Relaced last layer of the Deep CNN with a Soft max layer. For further reading, refer [14]

Emara used a modified Inception_V4 for skin cancer detection in 2019 and used the HAM10000 data set, resulting in model classification. They used this to train the ImageNet Data Set. For more information, see [15].

Ahmed Demir employed Deep Learning Architectures: Resnet_101 and Inception-V3 for skin cancer detection in 2019 and obtained the ResNet_101 Model from a biomedical data set. By this, they have trained two separate Deep Learning Methods. For any more queries, refer [16].

In the year 2021, a Javaid employed Machine Learning for skin cancer detection, and they used the ISIC-ISBI data set to obtain a trained algorithm as part of the evaluation function. This resulted in Benign data being trained; for more information, see [17].

Yann Lecun introduced CNN in the 1980s. CNN (Convolution Neural Network) is a sort of artificial neural network used to recognize images. CNN is a design that uses numerous building pieces as a convolution network to automatically learn new features for hierarchies. It comprises of two layers: input layer and output layer. When you submit input to the CNN, it performs a number of tasks, which are then passed on to subsequent layers. It requires a vast amount of data for both training and testing. It works on the basis of a confusion matrix, which has three levels known as RGB. The output is transmitted to the next layer after creating several features.

The first layer extracts basic information such as horizontal and diagonal edges. Where the output is transferred to the next layer. The next layer detects more complex features such as corners or combinational edges, and as we progress further into the network, it can detect more complex features such as objects, faces, and so on. This is how the CNN works; based on the three layers, it recognizes the image and outputs it. Table 2.1 shows the details of the survey.

TABLE 2.1: RELATED WORKS

S. No	Method	Data set	Performance
1	Multilayer perceptronand Deep Neural Network	HAM10000	Accuracy (74.75%) UGC (81.46%) Sensitivity(90.09%)
2	Convolutional Neural Network (CNN)	HAM10000	Accuracy (83.11%) , F1 score (0.827) precision (0.818),
3	Transfer Learning	HAM10000	Accuracy (90%) Recall (0.90%)
4	VGG19 and Transfer Learning	HAM10000	Training Accuracy (0.985) Testing Accuracy (0.975)
5	A Modified Inception_ V4	HAM10000	Accuracy(82.89%)
6	Artificial Neural Network,SVM, Back-Propagation Neural Network	DermIS, DermQuest	BNN Accuracy(89.9%) AANN Accuracy(80.8%)
7	SVM, KNN	ISIC-ISBI	KNN Accuracy(86%) SVM Accuracy (87.5) SVM+KN N Accuracy(94%)
8	Diagnostic Algorithm	Image dataset	Accuracy (93.3%)
9	CNN	ISIC	Accuracy (71%), sensitivity (0.68) Specificity (0.74) F1 score (0.7)
10	Multilayer Decomposing aided	SCD	Accuracy (Melanoma [97.1%]), Accuracy (non melanoma [95.6%])
11	SVM and KNN	1.Dermoscopic 2.Histoapthalog ical	Accuracy (94%)
12	KNN &Decision Tree & SVM	ISIC	Accuracy (KNN [76.4%]) (Decision Tree[76.4%]),(SVM[78.2%])
13	Deep Learning	HAM10000	Accuracy (94%)
14	Neural Network	463 images	Accuracy (76.9%)
15	SVM and RANDOM FOREST	ISIC-ISBI	Accuracy (93.89%)

III. EXISTING SYSTEM

The existing skin cancer prediction system used the RESNET (Residual Neural Network) architecture, a deep learning model noted for its high performance in image recognition tasks. The major goal of this technique was to reliably categorize dermatoscopic images into three types: melanoma, basal cell carcinoma, and squamous cell skin cancer.

RESNET is a form of Convolutional Neural Network (CNN) that makes use of skip connections or shortcuts to help the model learn from both shallow and deep layers efficiently. This feature enables RESNET to overcome the vanishing gradient problem, which is common in very deep neural networks, making it ideal for complex image processing jobs such as skin cancer classification.

After lengthy training on a broad collection of dermatoscopic pictures, the system obtained an impressive accuracy of 82.87%. This degree of accuracy reveals its capacity to distinguish between the three forms of skin cancer with an acceptable success rate.

The system's classification procedure consisted of several stages. First, the dermatoscopic images were preprocessed to improve quality and standardise the data for the model. The RESNET architecture was then trained on the preprocessed pictures, learning to extract relevant characteristics and patterns that distinguish between melanoma, basal cell carcinoma, and squamous cell skin cancer.

To assess the system's performance, a separate test dataset was employed that was not part of the training procedure. The trained RESNET model then predicted on this test dataset, with an accuracy of 82.27% indicating the proportion of right predictions compared to the total number of samples in the test set.

The earlier system's effectiveness in reaching a reasonably high accuracy rate demonstrates the promise of deep learning techniques to help medical practitioners detect skin cancer early. However, it is critical to note that no diagnostic system is perfect, and human skill is still required in the ultimate diagnosis and treatment decisions.

Despite the encouraging results, future research and development in the field of medical image analysis will focus on refining existing models, investigating ensemble approaches, and combining other cutting-edge deep learning architectures to improve accuracy and resilience. By constantly developing skin cancer prediction tools, the medical community may make considerable progress in improving patient outcomes and lowering the worldwide burden of skin cancer.

IV. DISADVANTAGES OF EXISTING SYSTEM

1) Limited accuracy for key instances: While an accuracy of 82.27% is commendable, it indicates that the system may still misclassify a considerable proportion of cases. In critical cases where rapid and precise diagnosis is vital, this degree of accuracy may result in misdiagnoses, delaying appropriate medical interventions for skin cancer patients.

2) Sensitivity to image quality: Deep learning models, including RESNET, can be sensitive to fluctuations in image quality, such as lighting, resolution, and noise. If the dermatoscopic images in the dataset are of poor quality or significantly different from the training data, the system's performance may decline, resulting in lower accuracy.

3) Limited interpretability: Deep learning models, such as RESNET, are frequently referred to as "black boxes" since they do not provide insight into how they make their predictions. The absence of interpretability makes it difficult to explain the rationale for a certain categorization decision, which can be problematic in vital medical applications where interpretability is critical for confidence and acceptance.

4) Data imbalance issues: Skin cancer databases may have class distribution imbalances, with some types of skin cancer underrepresented in comparison to others. This imbalance might result in biased predictions because the model may favour the dominant class, thus diminishing accuracy for less common cancer forms.

5) Overfitting: Deep Learning models, such as RESNET, are prone to overfitting, which occurs when they memorize specific patterns in training data instead of learning generalizable characteristics. Overfitting can produce good training accuracy but poor performance on unknown data, limiting the system's real-world applicability.

6) Computationally intensive: Deep learning models like RESNET frequently require large computing resources, such as powerful GPUs or TPUs. This might lead to expensive hardware and infrastructure expenditures while implementing and maintaining the system.

7) Generalization to different populations: The qualities of the training data may have an impact on the existing system's performance, especially if it is skewed toward specific demographics or groups. This raises questions about the system's ability to efficiently generalize to various populations with different skin tones and cultural origins.
Ethical considerations: Automated skin cancer prediction systems may pose ethical questions about patient privacy, data security, and the possibility of algorithmic bias. Ensuring patient permission, data anonymization, and fairness in algorithmic decision-making are critical issues to overcome.

To address these drawbacks, ongoing research and development in medical image analysis and machine learning should prioritize data augmentation techniques, model interpretability methods, class imbalance resolution, and rigorous testing on diverse and representative datasets. Furthermore, integrating the strengths of deep learning models with clinical expertise from healthcare experts can result in more accurate and trustworthy skin cancer prediction systems.

International Scientific Journal of Engineering and Management ISSN: 2583-6129
Volume: 03 Issue: 04 | April – 2024 DOI: 10.55041/ISJEM01573
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

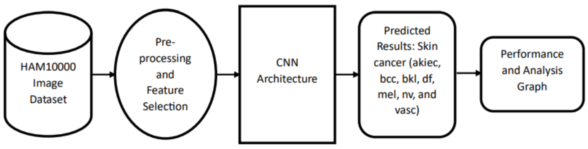


Fig 4.1 System Architecture

V. PROPOSED SYSTEM

The proposed method uses sophisticated deep learning techniques to improve skin cancer prediction, with a specific focus on dermatoscopic pictures. The system aims to improve accuracy, efficiency, and overall performance in skin cancer classification by using the capabilities of convolutional neural networks (CNNS) and implementing specific upgrades.

The suggested system is built around a cutting-edge deep learning architecture, such as an ensemble of CNNs or a more complicated variant of the standard CNN, which is intended to effectively learn and extract subtle patterns from dermatoscopic images. This architecture allows the model to discover key characteristics linked to various forms of skin cancer.

Enhance the model's generalization capabilities. The suggested system uses data-augmentation techniques. By applying changes like rotation, scaling, flipping, and cropping to the training dataset, the system can successfully improve data variety, lowering the danger of overfitting and increasing the model's performance on previously unseen images.

The suggested approach handles the possible issue of class imbalance in the dataset using strategies such as oversampling, undersampling, and class weighting. These strategies ensure that the model does not favour the majority class, hence enhancing its capacity to reliably forecast rare types of skin cancer.

The proposed system is optimized for real-time inference. This makes it acceptable for use in therapeutic environments. Minimizing inference time is critical for smooth integration into healthcare operations and timely diagnosis. The suggested solution meets ethical requirements by protecting patient data privacy, anonymity, and secure storage. It also works to eliminate algorithmic biases in order to reduce potential inequalities in predictions based on race, gender, or ethnicity.

By embracing these developments, the proposed system hopes to overcome the limits of earlier systems, resulting in a more accurate, efficient, and interpretable skin cancer prediction solution. Rigorous study and benchmarking against published datasets and real-world clinical data will be required to evaluate its efficacy and potential for widespread clinical application.

VI. ADVANTAGES OF PROPOSED SYSTEM

1) Increased Accuracy: By employing complex deep learning architectures and data augmentation approaches, the suggested system can reach greater levels of skin cancer classification accuracy. The capacity to acquire and learn detailed patterns from dermatoscopic images allows for more accurate and trustworthy predictions.

2) Robust Generalization: The suggested system's use of transfer learning and data augmentation improves generalization to new data. This ensures that the model works well on a variety of datasets and can handle differences in image quality, lighting conditions, and patient demographics.

3) Efficient Training: Using hyperparameter tuning and validation methodologies such as k-fold cross-validation, the proposed system optimizes training. Efficient training allows for speedier convergence and deployment of the model in real-world applications.

4) Class Imbalance Handling: The suggested approach effectively handles class imbalances in the dataset. Oversampling, undersampling, and class weighting approaches are used to reduce bias toward the majority class, resulting in better forecasts for rare skin cancer types.

5) Real-Time Inference: The suggested system is optimized for real-time inference and makes rapid predictions, making it appropriate for inclusion in clinical workflows. Fast inference times guarantee quick outcomes. Assisting healthcare practitioners in making timely, educated decisions.

6) It can handle larger datasets and adapt to future advances in medical picture data collection, meeting the growing demand for reliable skin cancer prediction.

7) Ethical Considerations: The system follows ethical guidelines, protecting patient data privacy and anonymity. By tackling algorithmic biases, it strives to give fair and impartial forecasts for a heterogeneous patient group.

8) Clinical Impact: Ultimately, the suggested system's increased accuracy and efficiency improve patient care. Early and accurate skin cancer predictions can help healthcare providers diagnose the disease at an early stage, potentially improving treatment outcomes and patient survival rates.

9) Research Advancements: The suggested system's creative application of deep learning techniques advances the fields of medical image analysis and skin cancer research. It lays the door for future advances in the use of artificial intelligence for dermatological diagnostics and may motivate additional research in this vital field of medicine.

By combining these characteristics, the suggested method provides a strong, accurate, and dependable tool for skin cancer prediction, aiding both healthcare professionals and patients in the fight against skin cancer.

VII. IMPLEMENTATION AND RESULT DISCUSSION

International Scientific Journal of Engineering and Management ISSN: 2583-6129
Volume: 03 Issue: 04 | April – 2024 DOI: 10.55041/ISJEM01573
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

The photos from Fig 7.1 shown below were obtained for both testing and training. It offers sample photos for various forms of skin cancer. With this image, we will train the system, which will then be tested with another sample of photos. We examine 10,000 photos for training and testing. We assess a bigger number of photos to be melanoma skin cancer kind. Because it is more hazardous than other types of skin cancer.

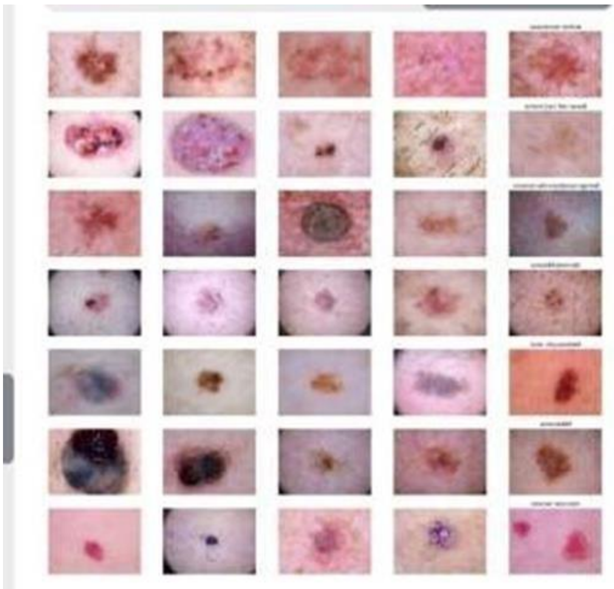


Fig. 7.1 Different types of image samples of skin cancer

From the below Fig 7.2 you can easily understand that Melanocytic nevi is the most frequent disease, with a prevalence of 6000 to 7000 cases. When compared to other conditions, dermatofibroma is less common. So, melanoma is the second most common and the most hazardous; it affects 10% of people out of every 10,000.

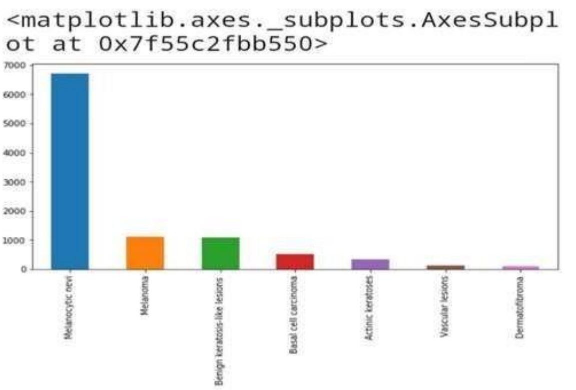


Fig. 7.2 Visualization of different types of skin cancer

The following Fig 7.3 bar-graph depicts various types of body parts and the amount of distribution of skin cancer in each type of body part. It is apparent that skin cancer in the back of the body is more widely disseminated than in other areas of the body. It rarely spreads to the acral body portion.

Figure 7.4 depicts the many types of people who suffer from skin cancer, with the majority of cases falling between the ages of 30 and 60.

Figure 7.5 illustrates the confusion matrix as well as the calculated accuracy, pression, true positive rate, and true negative rate.

Figure 7.6 shows the accuracy plotted versus the number of epochs. The blue graphic depicts the training accuracy curve, while orange reflects testing accuracy.

In addition, Figure 7.6 depicts the accuracy measured with the RESNET -50 model. Figure 7.7 depicts the training and testing losses acquired for the developed model.

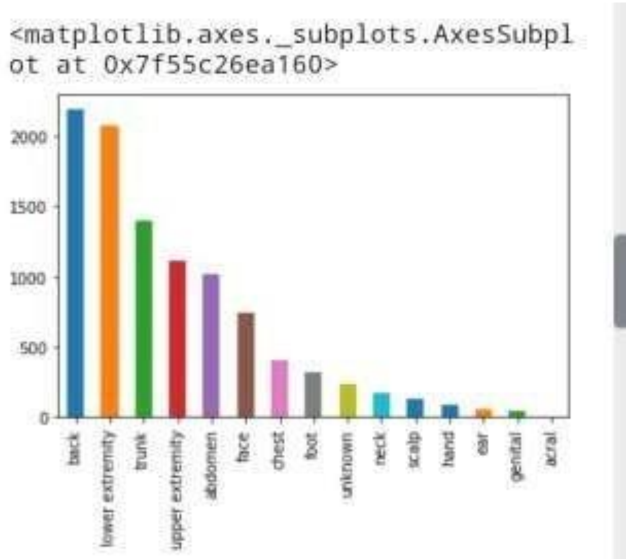


Fig. 7.3 Visualization of skin cancer in different parts of body

International Scientific Journal of Engineering and Management ISSN: 2583-6129
Volume: 03 Issue: 04 | April – 2024 DOI: 10.55041/ISJEM01573
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

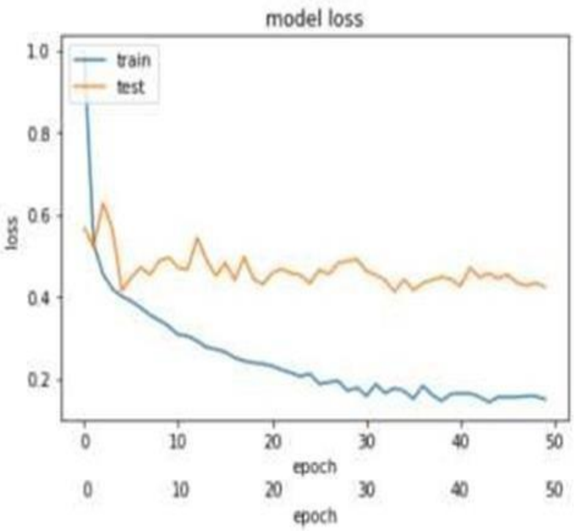
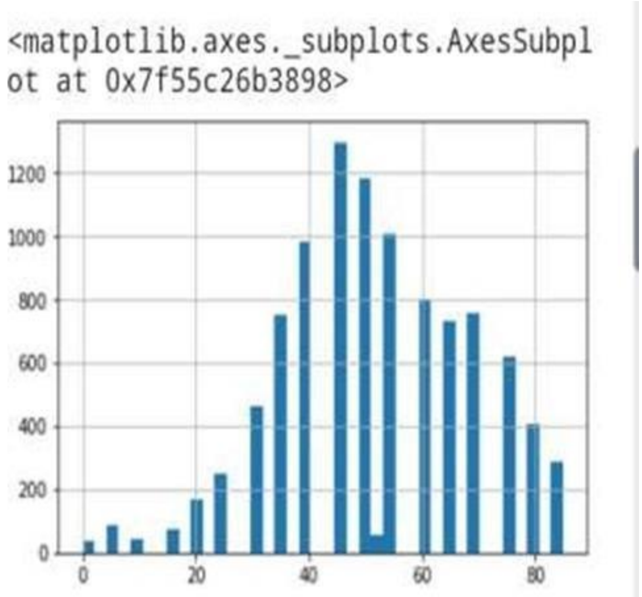


Fig. 7.4 Visualization of skin cancer among different ages



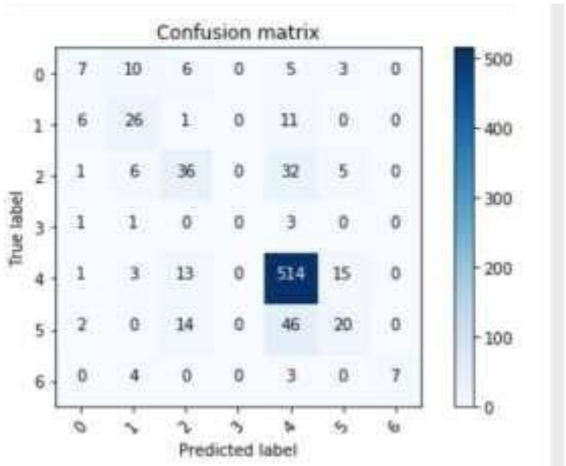


Fig. 7.5 Confusion Matrix

Fig. 7.6 Model Accuracy Using RESNET 50

Fig. 7.7 Model Loss Using RESNET 50

VIII. CONCLUSION

Based on the study presented above, it is obvious that RESNET - 50 outperforms the detection of skin cancer with an average training accuracy of 96.00%. Furthermore, the model's loss decreases dramatically from 0.7 to 0.4, as demonstrated in Figure 7.7. We find that future hypertuning of the model can increase accuracy and loss.

REFERENCES

[1] R. R. Sekar, Y. J. M. Reddy, K. Nani, C. S. P. Reddy, K. Chiranjeevi and K. Vikram, "Skin Cancer Prediction Using Deep Learning Techniques," 2023 4th International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India, 2023, pp. 28-31, doi: 10.1109/ICSPC57692.2023.10126035.

[2] Y Jusman," Performance of Multi-Layer Perceptron and Deep Neural Networks in Skin Cancer Classification", IEEE Transactions, vol 10, pp. 118198 -118212, 2022.

[3] R Raja Subramanian, "Skin Cancer Classification using CNN", IEEE Conference, 2021 11th International Conference on Cloud Computing, Data Science & Engineering 10.1109/Confluence51648.2021.9377155, 2021.

[4] Harikrishna, "Skin Cancer Classification using Transfer Learning IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI-2020)

[5] Nourabuared, "Skin Cancer Based on VGG19 and Transfer Learning", In 3rd International Conference on Signal Processing and Information Security IEEE Conference, pp 1-4, DOI- https://doi.org/10.1109/ICSPIS51252.2020.93401432021.

[6] Emara, "A Modifier Inception_v4 for imbalanced skin Cancer Classification", 2019 14th International Conference on Computer Engineering and Systems , pp. 28-33, December 2019.

[7] Enakshi Jana, "Research on Skin cancer Cell Detection using Image Processing", 2017 IEEE International Conference on Computational Intelligence and Computing Research, DOI: 10.1109/ICCIC.2017.8524554. pp. 1-8, 2017.

[8] A. Javaid, M. Sadiq, and F. Akram, "Skin cancer classification using image processing and machine learning," in Proceedings of the 2021 International Bhurban Conference on Applied Sciences and Technologies, pp. 439–444, IEEE, Islamabad, Pakistan, 12-16 Jan 2021.

[9] N. J. Dhinagar and M. Celenk, "Analysis of regularity in skin pigmentation and vascularity by an optimized feature space for early cancer classification," 2014 7th International Conference on Biomedical Engineering and Informatics, Dalian, China, 2014, pp. 709-713, DOI: 10.1109/BMEI.2014.7002865.

[10] P. Sedigh, R. Sadeghian and M. T. Masouleh, "Generating Synthetic Medical Images by Using GAN to Improve CNN Performance in Skin Cancer Classification," 2019 7th International Conference on Robotics and Mechatronics , Tehran, Iran, 2019, pp. 497-502, DOI: 10.1109/ICRoM48714.2019.9071823.

[11] Y Jusman," Performance of Multi-Layer Perceptron and Deep Neural Networks in Skin Cancer Classification", IEEE Transactions, vol 10, pp. 118198 -118212, 2022.

[12] R Raja Subramanian, "Skin Cancer Classification using CNN", IEEE Conference, 2021 11th International Conference on Cloud Computing, Data Science & Engineering 10.1109/Confluence51648.2021.9377155, 2021.

[13] Harikrishna, "Skin Cancer Classification using Transfer Learning IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI-2020)

[14] Nourabuared, "Skin Cancer Based on VGG19 and Transfer Learning", In 3rd International Conference on Signal Processing and Information Security IEEE Conference, pp 1-4, DOI- https://doi.org/10.1109/ICSPIS51252.2020.93401432021.

[15] Emara, "A Modifier Inception_v4 for imbalanced skin Cancer Classification", 2019 14th International Conference on Computer Engineering and Systems , pp. 28-33, December 2019.

[16] Enakshi Jana, "Research on Skin cancer Cell Detection using Image Processing", 2017 IEEE International Conference on Computational Intelligence and Computing Research, DOI: 10.1109/ICCIC.2017.8524554. pp. 1-8, 2017.

[17] A. Javaid, M. Sadiq, and F. Akram, "Skin cancer classification using image processing and machine learning," in Proceedings of the 2021 International Bhurban Conference on Applied Sciences and Technologies, pp. 439–444, IEEE, Islamabad, Pakistan, 12-16 Jan 2021.