Dermatological Cancer Classification

Pratish Khurdamoje
Information Technology
Shah & Anchor Kutchhi
Engineering College,Mumbai
pratish.khurdamoje16711@sakec.ac.in

Siddhesh Chavan
Information Technology
Shah & Anchor Kutchhi
Engineering College,Mumbai
siddhesh.chavan17686@sakec.ac.in

Piyush Parekh
Information Technology
Shah & Anchor Kutchhi
Engineering College,Mumbai
piyush.parekh16672@sakec.ac.in

Aditya Pathak
Information Technology
Shah & Anchor Kutchhi
Engineering College,Mumbai
adityapathak16687@sakec.ac.in

Jyoti Bansode
Assistant Professor
Information Technology
Shah & Anchor Kutchhi
Engineering College,Mumbai
Jyoti.bansode@sakec.ac.in

Vaishali Korade
Assistant Professor
Information Technology
Shah & Anchor Kutchhi
Engineering College,Mumbai
Vaishali.korade@sakec.ac.in

***Abstract*—Early and accurate detection of skin cancer is critical for effective treatment and improved patient outcomes. In this study, we develop and evaluate deep learning models, including DermaNet, ResNet50, EfficientNetB0, and a hybrid CNN, for dermatological cancer classification. The HAM10000 dataset was preprocessed using data augmentation and oversampling techniques to address class imbalances. Additionally, a "normal" class was introduced to filter out non-relevant images, enhancing the model's real-world applicability. DermaNet achieved the highest classification accuracy of 97.14%, followed by ResNet50 (95.74%) and EfficientNetB0 (88.50%), while the hybrid CNN effectively distinguished skin cancer from non-skin cancer cases with 91.87% accuracy. The study further explores model interpretability, hyperparameter tuning, and computational efficiency, making it suitable for real-time clinical deployment. These findings highlight the potential of deep learning for automated skin cancer diagnosis and emphasize the need for further validation on diverse datasets.**

***Index Terms*—Deep learning, dermatological cancer, CNN, DermaNet, ResNet50, EfficientNetB0, HAM10000 dataset, medical AI, skin cancer classification.**

I. Introduction

Skin cancer is the most common type of cancer worldwide, accounting for approximately one in three cancer diagnoses. According to the World Health Organization (WHO), approximately 2-3 million cases of non-melanoma skin cancer and 132,000 cases of melanoma are reported each year. Early diagnosis is of great importance, especially in melanoma, which accounts for 75 percent of skin cancers.[1]

Traditional diagnostic procedures such as eye examinations and blood tests are invasive, time-consuming, and resource-intensive. With the rise of artificial intelligence, deep learning models such as convolutional neural networks (CNNs) have shown great potential in automatic cancer detection. In this study, we built a CNN model using MNIST skin data that can classify seven types of skin diseases with over 90% accuracy, providing better diagnostic strategies and no intervention.[2]

II. Literature Review

Skin cancer is a significant public health concern worldwide, with increasing incidence rates in recent years. Early detection and diagnosis are crucial for successful treatment and improved patient outcomes. Advances in machine learning and computer vision have shown promising potential in aiding skin cancer identification.[1]

Numerous studies have explored the application of deep learning techniques for skin cancer classification. Convolutional Neural Networks (CNNs) have been widely adopted due to their ability to extract intricate features from images. CNN-based models have demonstrated impressive accuracy in differentiating between benign and malignant skin lesions, surpassing traditional methods.[2]

Researchers have also investigated the use of transfer learning, where pre-trained CNN models are fine-tuned on skin cancer datasets. This approach leverages the knowledge learned from large-scale image datasets, such as ImageNet, to improve performance on smaller, domain-specific datasets.[4]

Ensemble methods, combining multiple models, have been explored to enhance classification accuracy. Techniques like bagging and boosting have been employed to reduce overfitting and improve generalization.[6]

Despite notable advancements, there are still obstacles to overcome in the identification of skin cancer. Imbalance in data distribution, where one class (e. g. , benign lesions) is more common than others, can impact the accuracy of the model. Furthermore, the diversity in skin lesions, such as variations in size, hue, and texture, presents difficulties in accurately categorizing them. [4] Future research directions involve creating more robust models that can effectively handle data imbalance and enhance their ability to generalize to unseen data. By integrating deep learning with conventional image analysis methods, researchers can potentially improve the accuracy of image classification. Additionally, it is important to address privacy concerns and ensure ethical use of

artificial intelligence in the identification of skin cancer.

III. Methodology

A. Data Collection and Preprocessing

The MNIST: HAM10000 dataset comprises 10,015 skin lesion records, each with a resolution of 2351 pixels. It includes essential metadata such as patient age, gender, skin location, and cancer type. The majority of patients fall within the 40–70 age group, indicating a higher incidence of skin cancer in older individuals. The dataset contains seven types of skin cancer, with an imbalanced distribution—notably, Nevus (nv) has 6,705 images, significantly more than other categories: Melanoma (mel) - 1,113, Benign Keratosis-like lesions (bkl) - 1,099, Basal Cell Carcinoma (bcc) - 514, Actinic Keratoses (akiec) - 327, Vascular lesions (vasc) - 142, and Dermatofibroma (df) - 115.

To address this imbalance, random oversampling was applied, resulting in a balanced skin cancer dataset with 46,935 images in both (28,28,3) and (32,32,3) dimensions. Additionally, a ”normal” class was introduced to detect non-skin cancer images or cases where a patient submits an incorrect or unrelated image. This class was created by adding random images that do not correspond to any skin lesion, leading to a total of 53,640 non-skin cancer images in (28,28,3) dimensions. This enhancement ensures that the model can accurately distinguish between valid and invalid inputs.

The dataset was then split into 80% training and 20% testing data. This expanded dataset, including the ”normal” class, was exclusively used for training the Hybrid CNN model to improve its ability to identify incorrect or non-skin inputs. However, the ResNet50, EfficientNet-B0, and DermaNet models were trained only on skin cancer-related images without the ”normal” class.

B. Model Training

The bar chart titled “Model Accuracy Comparison” in Figure 1 shows the performance of our state-of-the-art deep learning model for classifying cancer cells on a piece of paper. skin: ResNet, EfficientNet, and DermaNet. Each bar represents the percentage accuracy achieved by a particular model. ResNet, shown in blue, demonstrated its ability to train deep networks by achieving 95.74% accuracy. The EfficientNet model has an accuracy of 88.55%. Although it falls slightly behind ResNet and DermaNet, it still shows a balance between performance and efficiency. DermaNet (shown in green) stands out as the best performing model in this comparison with an accuracy of 97.14%. The superior performance of DermaNet, a special design for skin cancer diagnosis, demonstrates its effectiveness in solving this difficult task. Among these models, DermaNet clearly stands out with its better results, structure and layers. These findings highlight the importance of selecting and optimizing designs to achieve high accuracy, especially in important applications such as cancer diagnosis.

The hybrid CNN model achieves an accuracy of 91.87% when tested on a dataset containing both skin cancer and non-skin cancer images. This indicates the model’s strong capability to distinguish between skin cancer and normal images. However, there remains room for improvement, particularly in further refining its ability to minimize false classifications and enhance overall reliability.

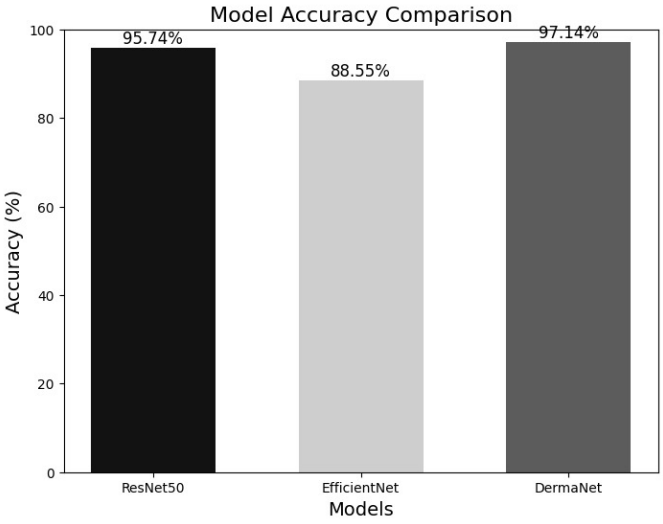

Fig. 1. Used Different Models for Skin Cancer MNIST Dataset

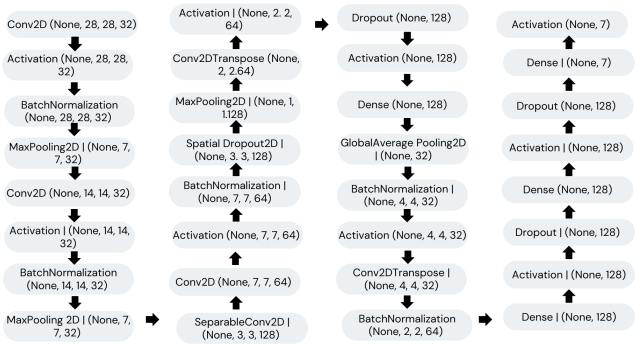

Fig. 2. Architecture of DermaNet

C. DermaNet Architecture for Classifying Skin Cancer

Fig. 2: The proposed dermanet architecture is a specialized convolutional neural network (cnn) specifically designed for accurately classifying skin cancer. The network starts with an input layer that receives images of dimensions 28x28x3, specifically designed for the dataset. The architectural design includes three layers of convolution, each consisting of 32, 32, and 64 filters of size 2×2, all employing the’swish’ activation function. After each convolutional layer, batch normalization and max-pooling are applied to improve feature extraction and reduce the spatial dimensions, resulting in a more efficient computational process.

One of the significant advancements in dermanet is the integration of a separable convolutional layer with 128 filters, which separates spatial and depth-wise convolutions to enhance parameter efficiency. After the first layer, spatial

Classification Report

	precision	recall	f1-score	support
0	0.99	1.00	1.00	1324
1	0.98	1.00	0.99	1344
2	0.92	1.00	0.96	1362
3	1.00	1.00	1.00	1353
4	0.97	0.84	0.90	1339
5	1.00	1.00	1.00	1340
6	0.94	0.96	0.95	1325
accuracy			0.97	9387
macro avg	0.97	0.97	0.97	9387
weighted avg	0.97	0.97	0.97	9387

Fig. 3. Classification Report of DermaNet

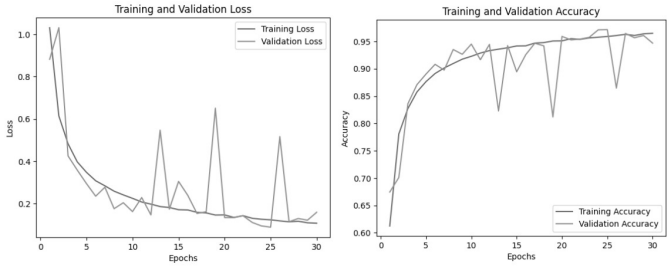

Fig. 4. Model Accuracy and Loss of DermaNet

dropout is applied to prevent overfitting, guaranteeing that the model can effectively handle new, unseen data. For feature reconstruction, the network incorporates two upsampling layers using Conv2DTranspose layers with 64 and 32 filters. These layers upscale the feature maps while maintaining the resolution of the input data. A global average pooling layer is then applied to reduce the feature dimensions, followed by a stack of three fully connected dense layers with 128 neurons each, using the 'swish' activation function. Dropout layers with a rate of 0.5 are added after each dense layer to further prevent overfitting.

The output layer comprises 7 neurons, each representing a different class, and utilizes a softmax activation function to generate class probabilities. The model is built using the nadam optimizer and the sparse categorical cross-entropy loss function, guaranteeing effective optimization and compatibility with sparse label encoding.

This hierarchical structure of convolutional, separable convolutional, and fully connected layers, combined with advanced techniques like batch normalization, spatial dropout, and swish activation, enables dermanet to achieve high accuracy and robust performance in skin cancer classification tasks. The Fig. 3 demonstrates outstanding performance on a test dataset of 9387 samples across 7 classes. It achieves a 97% accuracy, with precision ranging from 0.92 to 1.00 and recall between 0.84 and 1.00. The F1-scores are consistent, ranging from 0.90 to 1.00, except for class 4. Both macro and weighted averages are 0.97, reflecting strong and balanced performance. While the model performs excellently overall,

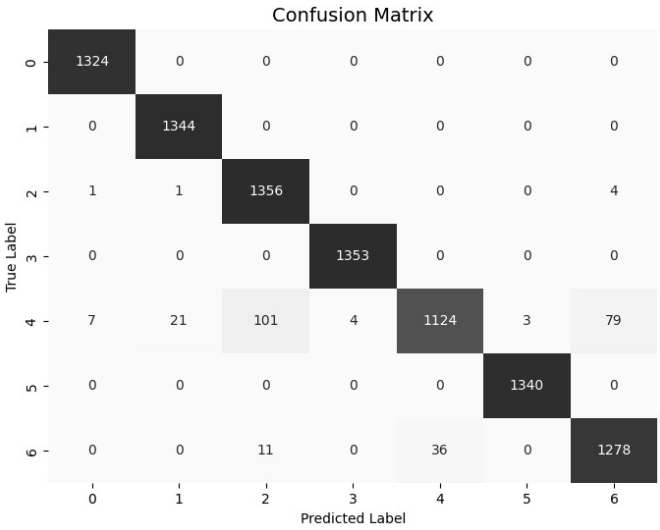

Fig. 5. confusion matrix of DerMaNet

slight improvements are suggested for class 4.

Fig 4.The process of training the skin cancer detection model took 30 epochs, with a consistent improvement in accuracy and decrease in loss. In the first epoch, the model was at 49.65% accuracy with a loss of 1.2851, but it rapidly improved in the following epochs. By Epoch 5, accuracy had increased to 88.01%, with a much lower loss of 0.3420. In spite of periodic dips in validation accuracy and loss, the model showed a strong trend of generalization. At Epoch 10, accuracy was at 92.15%, while validation accuracy was 93.12%, which meant substantial learning was achieved. Additional training further polished the model, with an accuracy of 96.61% and validation accuracy of 96.52% at Epoch 30. There were some epochs that dipped in terms of validation accuracy, which meant possible overfitting. Even with these variations, the final outcomes show that the model performed well in categorizing skin cancer images. The confusion matrix also shows some misclassifications, especially with Nevus (nv), Melanoma (mel), and Seborrheic Keratosis (bkl). The largest misclassification is with Nevus (nv) (Class 4), where 101 samples were falsely predicted as Seborrheic Keratosis (bkl) (Class 2), 79 as Melanoma (mel) (Class 6), and smaller misclassifications to other classes. This indicates that Nevus visually resembles these diseases, and the model finds it difficult to distinguish between them.

Furthermore, Melanoma (mel) (Class 6) is incorrectly classified 11 times as Seborrheic Keratosis (bkl) (Class 2) and 36 times as Nevus (nv) (Class 4). This suggests possible overlaps in features between these three entities, which could necessitate further refinement of the model or more distinguishing features in the dataset.

Overall, the model is good but is unable to discriminate between visually equivalent skin conditions. Feature extraction can be enhanced, training samples can be increased, or domain-specific preprocessing methods can be included to

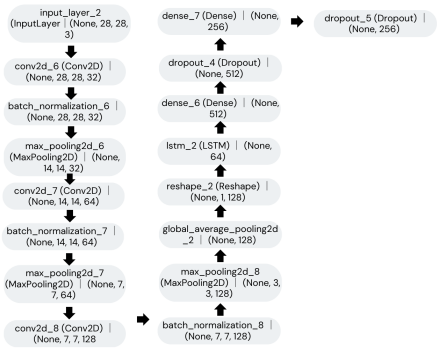

Fig. 6. Hybrid CNN architecture

decrease these misclassifications.

D. HybridCNN Architecture for Classifying Skin Cancer and Non-Skin Cancer Images

The hybrid convolutional neural network (CNN) model first processes the input images and then goes through a series of convolutional layers to extract hierarchical features from the data. Initially, the model uses multiple convolution operations and gradually increases the depth of the feature map. Each convolution is followed by batch normalization, which helps stabilize the training by normalizing the activations and reducing the auxiliary variable variations. The maximum performance is then used to gradually reduce the size of the feature map, minimizing the amount of data while preserving the important features. Abstract features transform the data into higher-level and higher-level representations. To ensure that the model remains robust and prevents overfitting, a release operation is introduced after the bound operation to randomly set a portion of the input data to zero during the training period. This technique helps the overall model perform better when faced with unknown data. The model also incorporates a global pooling layer, which reduces the feature maps' spatial dimensions to a fixed-size vector, regardless of the input image size. This vector is then passed through a reshaping operation to adjust the data format for subsequent processing. A Long Short-Term Memory (LSTM) layer is utilized to capture sequential dependencies within the data, enabling the model to learn temporal relationships between features.

Following the LSTM layer, fully connected dense layers are employed to perform high-level decision-making. These dense layers progressively reduce the number of features, while dropout layers continue to ensure regularization. The final output layer produces a set of predictions corresponding to the target classes, allowing the model to classify the input data into one of several categories.

This hybrid CNN architecture effectively combines the power of convolutional layers for spatial feature extraction with the sequence-learning capabilities of LSTM networks, making it well-suited for tasks that involve both spatial and

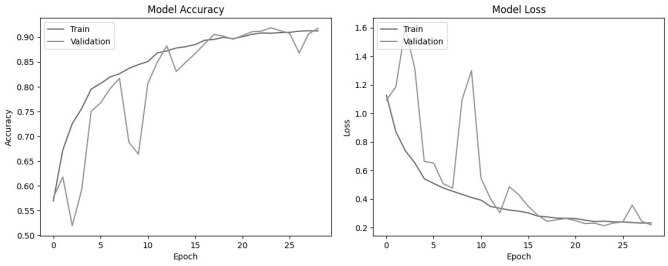

Fig. 7. Model Accuracy and Loss

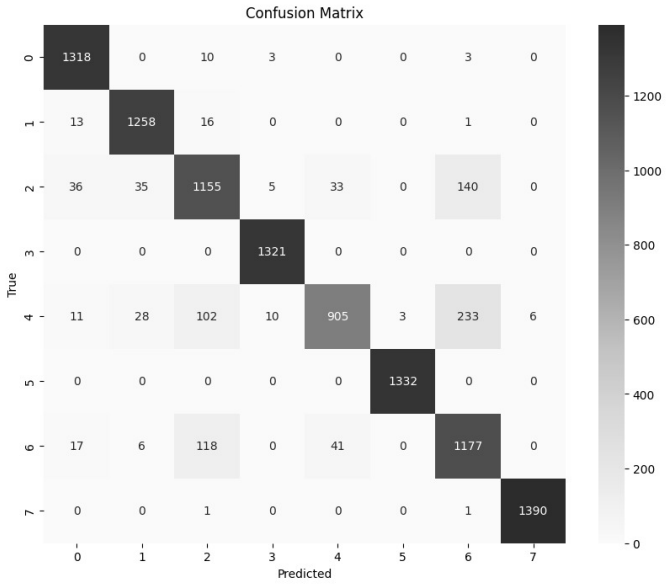

Fig. 8. Confusion matrix of Hybrid CNN

temporal patterns in the input data. The use of batch normalization and dropout layers throughout the network ensures robust training and generalization, ultimately leading to a highly efficient and accurate model.

Fig 7. The training process shows a steady improvement in accuracy and a decrease in loss, with fluctuations in validation performance. Initially, the model starts with around 50% accuracy, improving to 75% by the fourth epoch, while the loss decreases but with some instability in validation loss, indicating potential overfitting. The learning rate is adjusted using ReduceLROnPlateau to stabilize training. As training progresses, accuracy exceeds 80%, and validation accuracy peaks at 81.6%, though occasional drops suggest further fine-tuning is needed. By the final epochs, training accuracy reaches 93.3%, and validation accuracy stabilizes at 91.8%, with low loss values indicating good convergence. The learning rate adjustments played a crucial role in stabilizing performance, and additional regularization techniques like dropout or batch normalization could further improve generalization. Early stopping around epoch 30-40 might be beneficial, as performance appears to plateau.

The confusion matrix indicates that the model performs

well in classifying most skin lesion types, with high accuracy in detecting Actinic Keratosis (akiec), Basal Cell Carcinoma (bcc), Dermatofibroma (df), Nevus (nv), and Vascular Lesion (vasc) as seen from the strong diagonal values. However, Seborrheic Keratosis (bkl) and Melanoma (mel) show notable misclassifications, particularly with 140 cases of Seborrheic Keratosis predicted as Melanoma, and 118 cases of Melanoma classified as Nevus, suggesting overlapping visual features. Additionally, 36 cases of Actinic Keratosis were confused with Seborrheic Keratosis, likely due to similarities in texture or color. The normal class (7) is well-classified, with minimal errors. To improve performance, techniques such as data augmentation, class balancing, and advanced feature extraction could help differentiate similar lesion types. Transfer learning with deeper models like EfficientNet or ResNet may further enhance accuracy, while fine-tuning hyperparameters and incorporating domain-specific knowledge could address classification challenges.

IV. Results

Models	Accuracy
ResNet50	95.74%
EfficientB0	88.50%
DermaNet	97.14%
Hybrid CNN	91.87%

In this study, we evaluated multiple deep learning models for dermatological cancer classification using the Skin Cancer MNIST: HAM10000 dataset. The models were assessed based on their accuracy, loss, and classification effectiveness. The results are summarized as follows:

Model Performance Comparison DermaNet achieved the highest accuracy of 97.14ResNet50 followed closely with an accuracy of 95.74EfficientNetB0 attained an accuracy of 88.50The Hybrid CNN model, which was specifically designed to distinguish between skin cancer and non-skin cancer images, achieved 91.87Confusion Matrix and Model Insights The confusion matrix analysis highlighted that DermaNet exhibited superior classification ability, but some misclassifications were observed, particularly between Nevus (nv) and Melanoma (mel) due to their visual similarities. The Hybrid CNN model successfully identified non-skin cancer images, reducing false positives in medical diagnosis. Training and Validation Performance The training process showed a steady improvement in accuracy with decreasing loss across all models. DermaNet's training curve exhibited a consistent upward trend, reaching optimal accuracy around 30 epochs. The Hybrid CNN model, although effective, showed occasional fluctuations in validation accuracy, suggesting possible overfitting, which can be addressed with further regularization techniques. Conclusion from the Results The study demonstrates the effectiveness of deep learning models in classifying skin cancer. DermaNet is the most suitable model for medical applications, while the Hybrid CNN model enhances classification by correctly distinguishing between skin cancer and non-skin cancer images. Future improvements in feature extraction and dataset balancing can further enhance accuracy and reliability.

V. Conclusion

The provided conclusion compares the performance of three models (DermaNet, EfficientNet, and ResNet50) based on their accuracy and loss for skin cancer classification. It suggests that **DermaNet** is the best-performing model, achieving the highest accuracy (97.14%) and the lowest loss (0.0879). **ResNet50** is a strong contender with an accuracy of 95.74% and a loss of 0.16, while **EfficientNet** has an accuracy of 88.55% and a loss of 0.3421, indicating it performs well but is not as effective as the other two models.

Furthermore conclusion discusses a **Hybrid CNN model**. It classifies skin cancer and non-skin cancer. It achieves 90% accuracy. Yet it's critical to note something. Other models were utilized specifically on skin cancer dataset. This may have influenced their performance.

To sum up DermaNet is seen as the leading model for skin cancer classification. Yet the Hybrid CNN model 90% accuracy is notable. This is for both skin cancer and non-skin cancer classification.

VI. Future Scope

In the future, the combination of Hybrid-CNN models can greatly improve the classification accuracy of skin cancer detection systems. By integrating multiple architectures like ResNet50, DenseNet, and DermaNet, the model can utilize the strengths of various deep learning methods to enhance feature extraction and generalization. This mixed strategy may be especially effective for the classification of the seven skin cancer categories from the mini-ISIC HAM10000 dataset and a normal class of randomly gathered benign skin images. Transfer learning using pre-trained models such as ResNet50 can assist the system in learning generalizable representations, while the domain-specific adjustments from DermaNet may enhance the system. Furthermore, ensemble learning methods can be investigated to blend predictions from different models to improve overall classification reliability. Real-time deployment in clinical applications can also be addressed in future work, incorporating the model into a web-based or mobile diagnostic tool for early detection and decision support. In addition, active learning and semi-supervised learning can be used to decrease the reliance on big annotated datasets to make the system more applicable in real-world medical scenarios.

References

[1] N. Abuared, A. Panthakkan, M. Al-Saad, S. A. Amin and W. Mansoor, "Skin Cancer Classification Model Based on VGG 19 and Transfer Learning," 2020 3rd International Conference on Signal Processing and Information Security (ICSPIS), DUBAI, United Arab Emirates, 2020, pp. 1-4, doi: 10.1109/ICSPIS51252.2020.9340143.

[2] Azadeh Noori Hoshyar, A. Al-Jumaily and R. Sulaiman, "Review on automatic early skin cancer detection," 2011 International Conference on Computer Science and Service System (CSSS), Nanjing, China, 2011, pp. 4036-4039, doi: 10.1109/CSSS.2011.5974581.

[3] H.Patel,D. Thakore and N. Patel, "Skin Lesion Classification Using Deep Learning Techniques "2024 Parul International Conference on Engineering and Technology (PICET),Vadodara, India,2024,pp. 1-6,doi:10.1109/PICET60765.2024.10716067.

[4] N. Darapaneni et al., "Early Detection of Skin Cancer - Solution for Identifying and Defining Skin Cancers using AI," 2022 Interdisciplinary Research in Technology and Management (IRTM), Kolkata, India, 2022, pp. 1-6, doi: 10.1109/IRTM54583.2022.9791676.

[5] H. Alquran et al., "The melanoma skin cancer detection and classification using support vector machine," 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Aqaba, Jordan, 2017, pp. 1-5, doi: 10.1109/AEECT.2017.8257738.

[6] N. P. Behera, A. Pratap Singh, J. K. Rout, B. Kumar Balabantaray and A. Swetapadma, "Skin cancer detection using image features and ensemble learning," 2023 OITS International Conference on Information Technology (OCIT), Raipur, India, 2023, pp. 559-563, doi: 10.1109/OCIT59427.2023.10431321.

[7] S. -N. Tang, Y. -H. Chen, Y. -W. Chang, Y. -T. Chen and S. -H. Chou, "Hybrid CNN-LSTM Network for ECG Classification and Its Software-Hardware Co-Design Approach," 2023 20th International SoC Design Conference (ISOCC), Jeju, Korea, Republic of, 2023, pp. 173-174, doi: 10.1109/ISOCC59558.2023.10396448.

[8] N. Abuared, A. Panthakkan, M. Al-Saad, S. A. Amin and W. Mansoor, "Skin Cancer Classification Model Based on VGG 19 and Transfer Learning," 2020 3rd International Conference on Signal Processing and Information Security (ICSPIS), DUBAI, United Arab Emirates, 2020, pp. 1-4, doi: 10.1109/ICSPIS51252.2020.9340143.

[9] H. K. Kondaveeti and P. Edupuganti, "Skin Cancer Classification using Transfer Learning," 2020 IEEE International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI), Buldhana, India, 2020, pp. 1-4, doi: 10.1109/ICATMRI51801.2020.9398388. keywords: Biological system modeling;Transfer learning;Skin;Data models;Lesions;Task analysis;Skin cancer;Convolutional Neural Networks (CNN);Deep Learning;Transfer Learning;Skin Cancer classification,

[10] R. Bhuvaneswari and P. C. V. S. Raman, "Enhancing Skin Cancer Diagnosis with AlexNet Convolutional Neural Networks," 2023 International Conference on Integrated Intelligence and Communication Systems (ICIICS), Kalaburagi, India, 2023, pp. 1-5, doi: 10.1109/ICIICS59993.2023.10421461. keywords: Squamous cell carcinoma;Image recognition;Basal cell carcinoma;Melanoma;Skin;Lesions;Medical diagnostic imaging;Skin cancer;Melanoma;Benign;Malignant;dermatoscopic images.

\