IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 3, May-June 2025

Impact Factor: 8.152

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203084

Skin Cancer Detection using Transformer-Based Deep Neural Networks on Enhanced Image Datasets

D. Srinivas[1], M.Vaishali Gopal[2], J. Devarani[3], M. Akshaya[4]

Assistant Professor, Department of Computer Science & Engineering, Guru Nanak Institute of Technology, Hyderabad, India[1]

Student, Department of Computer Science & Engineering, Guru Nanak Institute of Technology, Hyderabad, India[2]

Student, Department of Computer Science & Engineering, Guru Nanak Institute of Technology, Hyderabad, India[3]

Student, Department of Computer Science & Engineering, Guru Nanak Institute of Technology, Hyderabad, India[4]

ABSTRACT: Skin Cancer Detection Using Transformer-Based Deep Neural Networks on Enhanced Image Datasets reflects a research approach aimed at improving the accuracy of skin cancer diagnosis by utilizing cutting-edge deep learning techniques, specifically Transformer models, on a large dataset. Using a substantial dataset related to skin cancer plays a crucial role in this research, as larger datasets provide more diverse examples that help the model generalize better to unseen data. This enhances the model's ability to recognize patterns across various types of skin conditions, including Melanoma, Melanocytic Nevi, Basal Cell Carcinoma, Actinic Keratoses, Benign Keratosis-like Lesions, Dermatofibroma, and Vascular Lesions. Transformer-based deep neural networks are applied here, leveraging self-attention mechanisms that allow the model to analyse and capture complex relationships within image data. This self-attention mechanism works by weighing the importance of each part of an image in relation to others, effectively capturing spatial dependencies across image regions. This allows the model to focus on relevant features in diverse and detailed skin images, making Transformer-based architectures a promising technique for medical image analysis.

I. INTRODUCTION

According to estimates from the World Health Organization, there are between two to three million new cases of skin cancer annually. Despite advancements in treatment, the 5-year su`rvival rate for skin cancer patients starts from almost 00% survivability in the early stages and drops to approximately 25-35% in the latest stages of all skin cancer-related deaths [] [2]. When BCC, SCC, or other skin cancers metastasize, the prognosis is generally bleak, significantly threatening patients' lives and often severely impacting their appearance [3]–[5]. Machine Learning (ML) technologies have emerged as a cornerstone for developing innovative and efficient solutions to support clinicians. The advent of ML in dermatology has revolutionized diagnostics, enhancing accuracy, speed, and scalability. The International Skin Imaging Collaboration (ISIC) [6] provides a repository of diverse datasets often accompanied by challenges [7] to advance nevus image classification techniques through ML and Deep Learning (DL) solutions. Despite considerable progress, developing models that generalize across large and heterogeneous datasets remains a challenge. Many datasets in the ISIC repository contain duplicate images, both within and between datasets, introducing undesirable biases in ML models [8]. In addition, some datasets, such as ISIC2020, offer many images with binary labels (malignant or benign). Although such datasets provide a substantial number of images, their limited labeling constrains the classification potential of DL models. More granular labeling would allow the development of more specific tools to support clinicians.

The integrity of test datasets is crucial for unbiased model performance evaluation. Issues such as the internal partitioning of datasets into training and testing sets can introduce variability, complicating objective model assessment. A standardized test dataset with ground-truth labels would facilitate fairer performance comparisons across different models. Among these challenges, DL and Transformer Models (TMs) have shown exceptional promise due to their ability to discern intricate patterns in complex datasets, surpassing the capabilities of traditional analytical methods. However, these models still face deficiencies in objective assessments of test performance. This paper delves into the investigation and application of TMs within dermatology, focusing particularly on the classification of skin lesions a critical aspect of early skin cancer detection and diagnosis.

Our work introduces a comprehensive approach that minimizes data augmentation techniques and leverages specific neural network architectures, starting with pre-trained networks to adapt different training and testing experiments for skin disease classification. With the advent of TMs, a shift has been observed in the methodologies employed for skin lesion analysis. Recently, the exploitation performances of Transformer networks with multiclass medical imaging were investigated by Matsoukas et al. [23], demonstrating their ability to handle sequential data within images and providing an ternative to the spatial hierarchies of DNNs. Moreover, Pedro and Oliveira [24] investigated how selfattention mechanisms can be integrated into TMs to focus more precisely on relevant image features, thereby potentially increasing model interpretability and performance. By using both Transformer and DNNs, which are based on end-to-end mapping and do not require previous information, an ST model for multiclass skin lesion classification is suggested by Ayas [30]. Moreover, a weighted cross-entropy loss was used to solve the class imbalance issue Hao et al. [29] proposed ConvNetXT, a model with high multiclass classification capabilities. The proposed model uses pre-trained ConvNeXt and ST networks to extract local and global features from pictures. Attentional Feature Fusion (AFF) submodules are then utilized to fuse the extracted features. Furthermore, an Efficient Channel Attention (ECA) module is included in the ConvNeXt network to improve the model's focus on the skin lesion locations during training. Transformer-based deep neural networks, originally developed for natural language processing, have been adapted for image-based tasks like classification, object detection, and segmentation due to their powerful ability to model complex spatial relationships across an image. When applied to images, the Transformer first divides the image into smaller patches (for instance, 6x6 pixels each). Each patch is then flattened into a D vector, essentially transforming each one into an embedding similar to a "token" in text processing. These patch embeddings are enriched with positional information to indicate each patch's location within the overall image.

At the core of the Transformer is the self-attention mechanism, where each image patch can assess and weigh the importance of every other patch. This is achieved by calculating three vectors for each embedding: a query (indicating relevance), a key (providing the patch's feature information), and a value (containing the feature's actual data). By comparing these vectors, the model allows each patch to consider the relationships and dependencies with other patches, thus capturing both local and global spatial features. To improve its capacity to detect complex patterns, Transformers use multi-head attention, which enables multiple attention "heads" to process different parts of the image simultaneously. Each head identifies unique spatial patterns, adding depth to the model's understanding. After attention processing, the embeddings pass through feed-forward neural networks, which introduce non-linearity and further enhance feature extraction. Additionally, a special classification token is introduced at the start of the sequence to gather information from all patches, which is then used for the final classification

II. LITURATURE SURVEY

Title: Vision Transformers for Skin Cancer Detection on Large Dermoscopic Image Datasets
Author: Wang, T., Zhao, L., & Smith, R.
Year: 2023.
Description: Use of artificial intelligence to identify dermoscopic images has brought major breakthroughs in recent years to the early diagnosis and early treatment of skin cancer, the incidence of which is increasing year by year worldwide and poses a great threat to human health. Achievements have been made in the research of skin cancer image classification by using the deep backbone of the convolutional neural network (CNN). This approach, however, only extracts the features of small objects in the image, and cannot locate the important parts.

Title: Hybrid CNN-Transformer Models for Enhanced Skin Lesion Classification on Combined Datasets
Author: Singh, P., Kumar, A., & Verma,
Year: 2024.
Description: Skin cancers are the most cancers diagnosed worldwide, with an estimated > .5 million new cases in 2020. Use of computer-aided diagnosis (CAD) systems for early detection and classification of skin lesions helps reduce skin cancer mortality rates. Inspired by the success of the transformer network in natural language processing (NLP) and the deep convolutional neural network (DCNN) in computer vision, we propose an end-to-end CNN transformer hybrid model with a focal loss (FL) function to classify skin lesion images. First, the CNN extracts low-level, local feature maps from the dermoscopic images. In the second stage, the vision transformer (ViT) globally models these features, then extracts abstract and high-level semantic information, and finally sends this to the multi-layer perceptron (MLP) head for classification. Based on an evaluation of three different loss functions, the FL-based algorithm is aimed to improve the extreme class imbalance that exists in the International Skin Imaging Collaboration (ISIC) 208 dataset.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203084

III. METHODOLOGY

MODULES NAME:

- **Data collection**
- **Dataset preprocessing and augmentation**
- **Transformer models**
- **Classification optimization**
- **Model evaluation and fine-tuning**
- **Predictions**

MODULES EXPLANATION

3.1 DATA COLLECTION:

We have seen that Transformer-based models offer great performance scalability depending on the cardinality of the training dataset and the number of model parameters. In this section, we propose an approach that investigates the potential of the ST model [9] on skin disease classification. We have performed several experiments where we investigate the behavior of the model's performance when increasing the number of training images, and we evaluate its classification accuracy on the external test dataset. Figure provides a comprehensive understanding of our methodology. In the first group of experiments, we considered the HAM dataset as training data. In the second group of experiments, we use the LD proposal as training. In both cases, the external test dataset was used for a fair comparison of the seven-class classification accuracies. For both groups of experiments, data augmentation techniques are applied. In subsection IV-A, we explain how the datasets are partitioned into training and validation sets and the relative augmentation techniques. Section III already discussed the methods used to choose datasets, their compositions, and the LD creation. Most of the experiments were conducted using the pre-trained weights of the chosen TMs, while some experiments involved training from scratch and TL (see subsection IV-B and IV-C). A.

3.2 DATASET PREPROCESSING AND AUGMENTATION:

An important aspect of this work concerns the preprocessing of data. Splitting the datasets in training and validation was the first key aspect. From Table 4, we notice that the datasets are highly imbalanced, and this led us to perform different types of experiments to find the optimal split choice. Regarding the training datasets, we have considered HAM, BCN, and DL datasets, and each time, 80% images were selected, class by class and randomly, to create the training dataset and the remaining 20% to create the validation dataset. Data Augmentation (DA) allows. us to increase the data by producing different images using the original ones as a base. DA techniques were investigated to mitigate the problem of unbalanced data. The primary purpose is to increase the amount of data by leading to randomness and increasing the model's generalization. To implement this, we decided to implement dynamic DA by randomly selecting the techniques to provide the model with new information all the time.

3.3 TRANSFORMER MODELS:

Classic deep learning methods, such as DNNs, excel at capturing local patterns through hierarchical feature extraction but struggle with long-range dependencies. In contrast, Transformers leverage self attention mechanisms to capture global context and relationships between all input elements, allowing for more flexible and powerful modeling of complex data structures. One of the best models proposed in recent literature is the ST [9]. It enhances the representation power while maintaining efficiency by incorporating locally computed self-attention in nonoverlapping windows and shifting these windows between successive Transformer layers. The main innovation of ST is its ability to adjust the processing scale dynamically, seamlessly moving from local to more global representations. The model focuses on fine-grained details within small windows in the initial layers. As information progresses through the stages, the feature maps are down sampled, and the model begins to attend to broader areas, integrating more context into its representations. This progression from local to global processing is crucial for capturing complex visual patterns and relationships in images [9]. The primary components of the ST architecture are as follows: Hierarchical Structure: Similar to CNNs, ST hierarchically processes inputs. B is the bias term, which can be a matrix for relative positional biases in the context of Transformers; • The entire expression inside the soft max function is then multiplied by V (value matrices). Cyclic Shift and Window Partitioning: The window partitioning is shifted by half the window size in horizontal and vertical directions after every ST block, enabling cross-window connections and enhancing modeling power. To perform self attention within local windows and allow for cross.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203084

3.4 CLASSIFICATION OPTIMIZATION:

A twofold evaluation strategy was adopted to optimize model classification performance on the test dataset and ensure accurate image classification. The first involved testing the experiment with higher accuracy for each group of experiments in the table 7. The training was repeated five times so that a statistical evaluation of the mean and standard deviation over the five training cycles could be made. Conversely, the second involved the random rotation of images during the inference process for each train to check the incidence of images in the classification. The rotation procedure consists of the following steps: • Evaluation through n rotation cycles are iteratively performed. In each cycle, all images in the test dataset are randomly rotated before being submitted to the model for classification; • After each rotation cycle, the model predictions for each image are recorded. At the end of all cycles, the list of predictions obtained for each image is analyzed, and the most frequent classification using the mode is calculated; • For each image, the classification that obtained the highest number of occurrences during the model run over all rotation cycles is selected. This approach aims to improve classification accuracy by considering the variability introduced by image rotation. Since it is impossible to decide a priori the optimal number of rotation cycles for this experiment, a random number of 500 was set to monitor the rotation behavior. After a series of tests, it was set at 00. This choice was motivated by the observation that the accuracy value tends to decrease beyond this threshold, as shown in Figure 5. This value represents an optimal compromise between the variation introduced by the rotations and the model's overall accuracy. Therefore, using evaluation cycles with random rotations aims to improve the robustness and accuracy of the classification model, allowing better generalization to images not seen during training.

3.5 MODEL EVALUATION AND FINE-TUNING:

We evaluate the performance of the classification tasks in terms of accuracy, precision, recall, and F- Score: Accuracy = TP + TN TP + FP + TN + FN (2) Precision = TP TP + FP (3) Recall = TP TP + FN (4) F-Score = 2 · Precision × Recall Precision + Recall (5) Where TP indicates the True Positives, TN the True Negatives and FP, FN indicates the False Positives and the False Negatives, respectively. Accurate model evaluation and fine-tuning are paramount for enhancing ML models' predictive performance and reliability. Model evaluation takes place in two different stages. The first occurs during model training. In each experiment, a large number of training epochs is set. At each epoch, the images in the training dataset are used to train the model, and the images in the validation dataset are used to evaluate the model's learning on images never seen during the training. At the end of the process, accuracy values for training and validation and loss values for training and validation are printed. Each time a training epoch ends, the loss function value is checked. If this value is lower than the previous epoch, the model is saved; otherwise, a new epoch starts immediately. This process is repeated until a higher loss value is found for 20 or 30 consecutive times (depending on the type of the experiment) than the best saved one. At the end of the training process, the confusion matrix and the accuracy, precision, recall, and f-score metrics of the best epochs were selected through the process described before being printed.

3.6 PREDECTIONS:

In a skin cancer classification project, the prediction is the model's best guess about whether a given skin lesion image represents a benign or malignant tumor, or it might even predict specific types of skin cancer if the model is set up for multi-class classification. This prediction is based on the features learned by the model during training, where it identifies patterns and relationships in the input data to map to the correct label or value.

3.3.1 EXISTING SYSTEM

Convolutional Neural Networks (CNNs) for skin disease classification come with several limitations. While CNNs excel at recognizing visual patterns, they are highly dependent on large datasets and extensive computational resources, which can make training costly and inaccessible for smaller clinics or institutions. Additionally, CNNs often struggle to capture global context in images, focusing primarily on localized features rather than the holistic view, which can lead to misclassifications, especially with subtle skin conditions that require broader context. Moreover, CNNs are prone to overfitting when the dataset is limited or lacks diversity, which can result in reduced accuracy on unseen images or varied skin types populations or subtle skin conditions.

3.3.2 EXISTING SYSTEM DISADVANTAGES

Limited Global Context Problems struggle to capture relationships across distant regions in an image.
CNNs need significant processing power and memory.
They can easily overfit on small datasets without regularization.

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203084

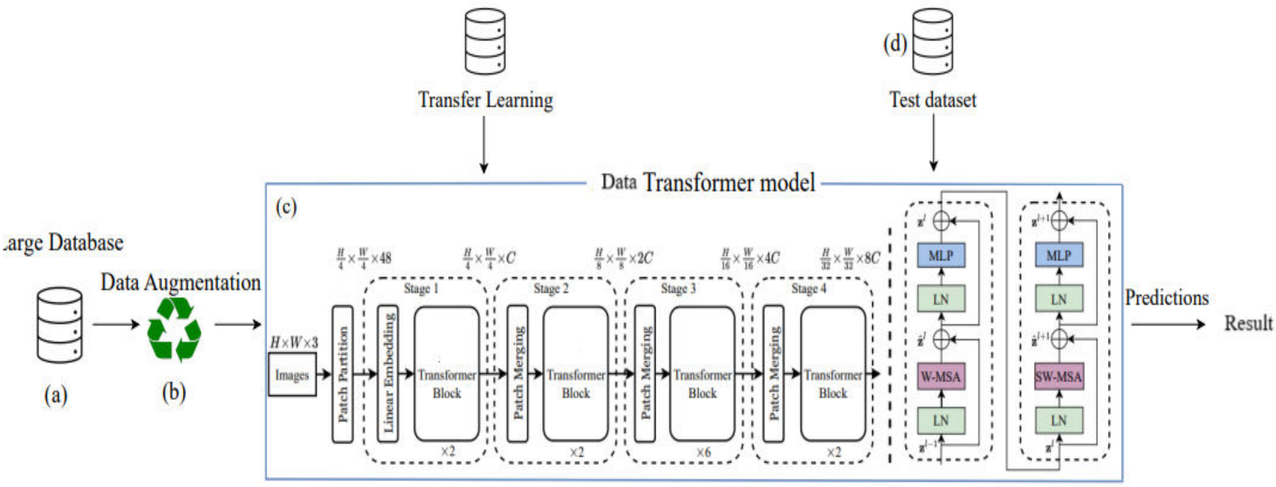
IV. PROPOSED SYSTEM

Classic deep learning methods, such as CNNs, excel at capturing local patterns through hierarchical feature extraction but struggle with long-range dependencies. In contrast, Transformers leverage self-attention mechanisms to capture global context and relationships between all input elements, allowing for more flexible and powerful modeling of complex data structures. One of the best models proposed in recent literature is the ST [9]. It enhances the representation power while maintaining efficiency by incorporating locally computed self-attention in nonoverlapping windows and shifting these windows between successive Transformer layers. The main innovation of ST is its ability to adjust the processing scale dynamically, seamlessly moving from local to more global representations. The model focuses on fine-grained details within small windows in the initial layers.

PROPOSED SYSTEM ADVANTAGES:

- Transformers capture relationships across the entire image.
- Improved Performance on Challenging Tasks.
- Flexibility with Large Datasets
- Focus on multiple important areas of the image simultaneously.

V. SYSTEM ARCHITECTURE



EXPLANATION:

The system architecture consists of multiple modules that work together to detect skin cancer from images. It starts with data collection, followed by preprocessing and augmentation to enhance image quality. The images are then passed to transformer-based deep learning models for classification. The system includes evaluation and fine-tuning stages to improve accuracy, and finally, it provides predictions to the user through a user-friendly interface.

IMPLEMENTATION:

User Interface Design

Input : Enter Login name and Password

Output : If valid user name and password then directly open the home page otherwise show error message and redirect to the registration page.

Input : Upload Image

Output : Gives the skin cancer prediction according to the given image.

IJARETY | ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

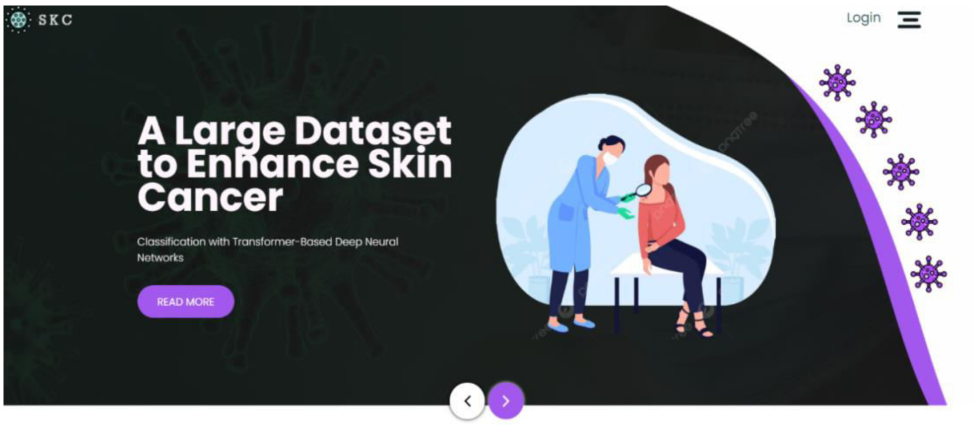
|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203084

VI. EXPERIMENTAL RESULTS

RESULTS

Figure 3.8.2.1 Home page



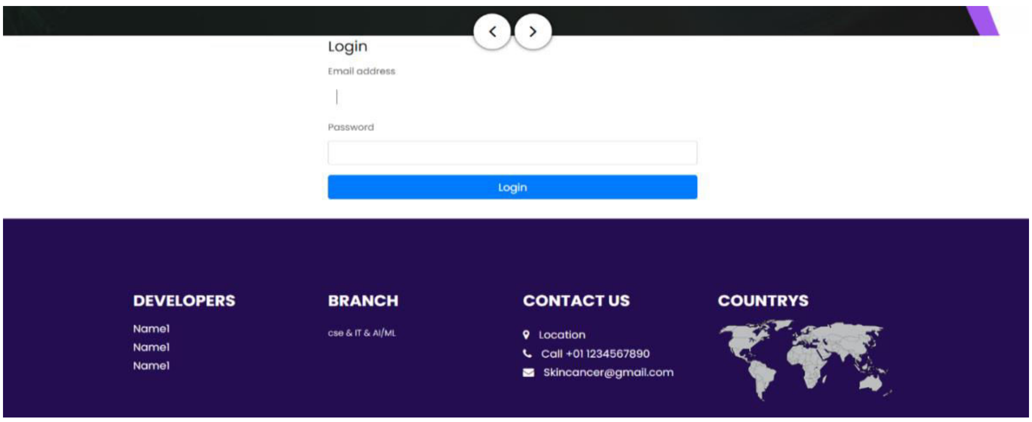


Figure 3.8.2.2 Login page

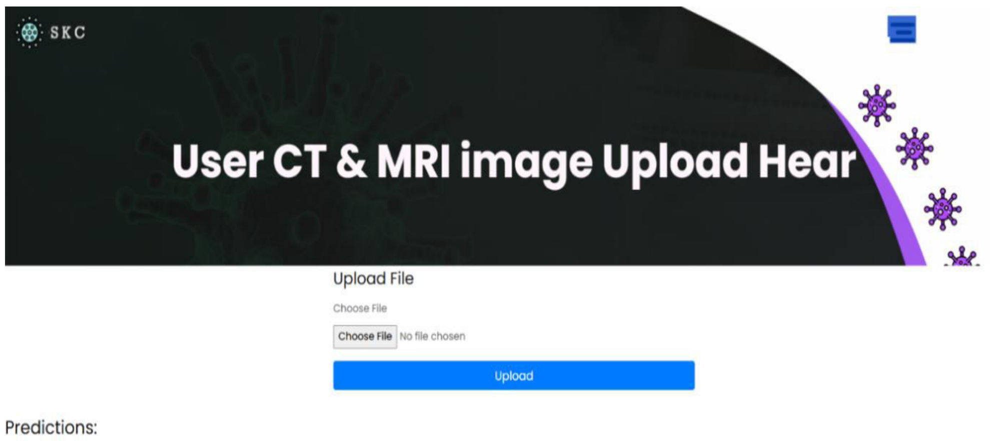


Figure 3.8.2.3 File Upload page

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203084

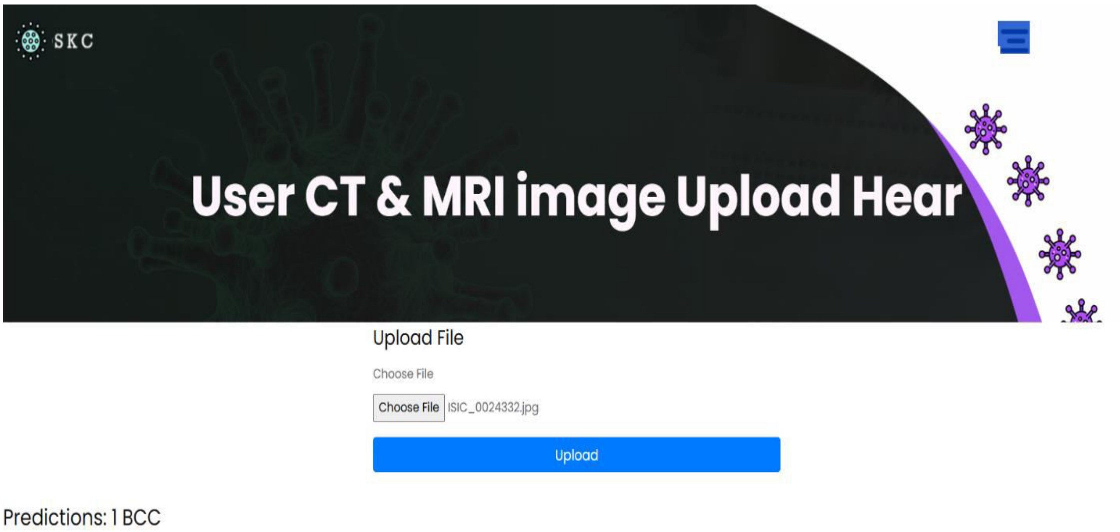


Figure 3.8.2.4 Skin Cancer Classification Result

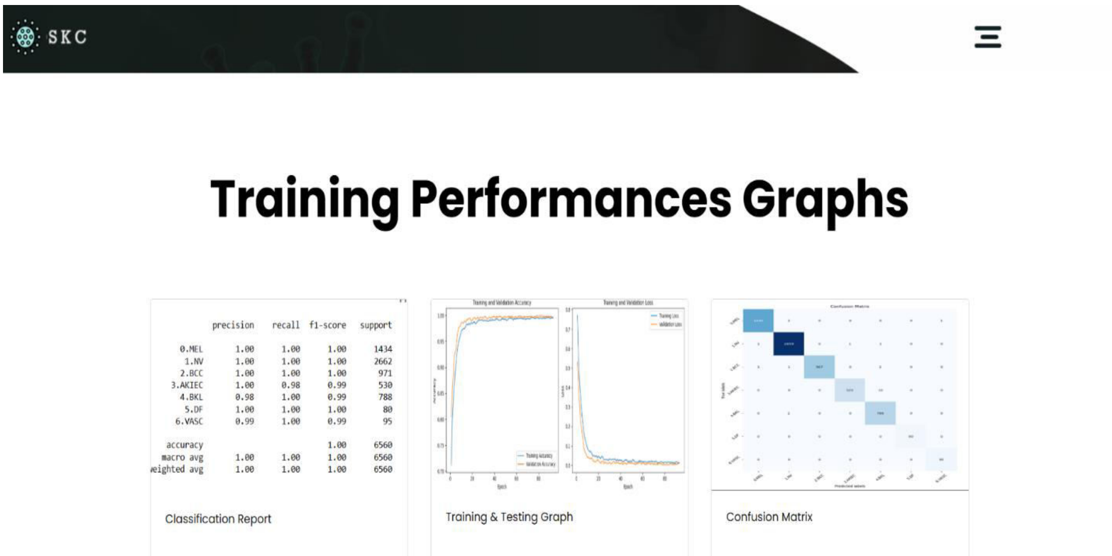


Figure 3.8.2.5 Training Performance Graphs

VII. CONCLUSION

This paper explored using Transformer-based deep neural networks, specifically the DNN model, for multiclass skin lesion classification. Our approach aimed to harness the self-attention mechanism intrinsic to TMs to capture intricate spatial dependencies within skin lesion images, bypassing the need for handcrafted features and extensive preprocessing steps. The performance of our proposed model was rigorously evaluated using a benchmark test set which includes ground truth labels for various types of skin lesions, including melanoma. Our experimental results demonstrate that the Transformer-based architecture achieves state-of the-art performance in skin lesion classification, outperforming traditional DNNs and other DL models previously employed for similar tasks. The superior performance can be attributed to the model's ability to effectively manage long-range dependencies and spatial relationships in the image data, which are crucial for accurate medical image analysis. Our work underscores the potential of Transformer based deep neural networks in advancing skin lesion classification, highlighting their clinical utility in aiding early and accurate skin cancer diagnosis. This research opens avenues for future studies to delve deeper into integrating advanced DL techniques in medical image analysis, ultimately contributing to developing powerful diagnostic tools for clinicians.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203084

VIII. FUTURE ENHANCEMENT

In Feature selection and dimensionality reduction enhance computational efficiency and simplify models by focusing on the most relevant data features. Techniques like Principal Component Analysis (PCA) help reduce the dimensionality of various extracted features, such as texture or color, by retaining only the components that contain the most valuable information.

For multi-class classification tasks, Linear Discriminant Analysis (LDA) can be applied to prioritize features that best separate classes, improving model performance. Both PCA and LDA can be implemented with scikit-learn to streamline data before feeding it into the model.

REFERENCES

[1] S. Waseh and J. B. Lee, ‘‘Advances in melanoma: Epidemiology, diagnosis, and prognosis,’’ Frontiers in Medicine, vol. 0, p. 268479, 2023.

[2] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, ‘‘Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 85 countries,’’ CA: A Cancer Journal for Clinicians, vol. 7, no. 3, pp. 209–249.

[3] Y. R. Woo, S. H. Cho, J. D. Lee, and H. S. Kim, ‘‘The human microbiota and skin cancer,’’ International Journal of Molecular Sciences, vol. 23, no. 3, p. 83, 2022.

[4] S. Chen, C. Han, X. Miao, X. Li, C. Yin, J. Zou, M. Liu, S. Li, L. Stawski, B. Zhu et al., ‘‘Targeting mcr depalmitoylation to prevent melanomagenesis in redheads,’’ Nature communications, vol. 0, no. , p. 877, 209.

[5] D. Schadendorf, A. C. Van Akkooi, C. Berking, K. G. Griewank, R. Gutzmer, A. Hauschild, A. Stang, A. Roesch, and S. Ugurel, ‘‘Melanoma,’’ The Lancet, vol. 392, no. 05, pp. 97–984, 208.

[6] ‘‘The international skin imaging collaboration,’’ https://www. isic-archive.com/, Last accessed on 2024-05-20. [7] ‘‘Isic challenge webpage,’’ https://challenge.isic-archive.com, Last accessed on 2024-07-04.

[7] B. Cassidy, C. Kendrick, A. Brodzicki, J. JaworekKorjakowska, and M. H. Yap, ‘‘Analysis of the isic image datasets: Usage, benchmarks and recommendations,’’ Medical image analysis, vol. 75, p. 02305, 2022.

[8] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, ‘‘Swin transformer: Hierarchical vision transformer using shifted windows,’’ in Proceedings of the IEEE/CVF international conference on computer vision, 202, pp. 0 02– 0 022.

[9] https://challenge.isic-archive.com/leaderboards/208/, Last accessed on 2024-05-20.

[10] M. S. Ali, M. S. Miah, J. Haque, M. M. Rahman, and M. K. Islam, ‘‘An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models,’’ Machine Learning with Applications, vol. 5, p. 00036, 202.

[11] P. Tschandl, C. Rosendahl, and H. Kittler, ‘‘The ham0000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions,’’ Scientific data, vol. 5, no. , pp. –9, 208.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ‘‘Imagenet classification with deep convolutional neural networks,’’ Advances in neural information processing systems, vol. 25, 202.

[13] K. Simonyan and A. Zisserman, ‘‘Very deep convolutional networks for large-scale image recognition,’’ arXiv preprint arXiv:409.556, 204.

[14] S. Gulati and R. K. Bhogal, ‘‘Detection of malignant melanoma using deep learning,’’ in International conference on advances in computing and data sciences. Springer, 209, pp. 32–325.

[15] M. Naqvi, S. Q. Gilani, T. Syed, O. Marques, and H.-C. Kim, ‘‘Skin cancer detection using deep learning—a review,’’ Diagnostics, vol. 3, no. , p. 9, 2023.

[16] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Deep residual learning for image recognition,’’ in Proceedings of the IEEE conference on computer vision and pattern recognition, 206, pp. 770– 778.

[17] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, ‘‘Densely connected convolutional networks,’’ in Proceedings of the IEEE conference on computer vision and pattern recognition, 207, pp. 4700–4708.

[18] S. K. Datta, M. A. Shaikh, S. N. Srihari, and M. Gao, ‘‘Soft attention improves skin cancer classification performance,’’ in Interpretability of Machine Intelligence in Medical Image Computing, and Topological Data Analysis and Its Applications for Medical Data: 4th International Workshop, iMIMIC 202, and st International Workshop, TDA4MedicalData 202, Held in Conjunction with MICCAI 202, Strasbourg, France, September 27, 202, Proceedings 4. Springer, 202, pp. 3–23.

[19] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, ‘‘Inceptionv4, inception-resnet and the impact of residual connections on learning,’’ in Proceedings of the AAAI conference on artificial intelligence, vol. 3, no. , 207.

[20] N. C. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler et al., ‘‘Skin lesion analysis toward melanoma detection: A challenge at the 207 international

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 3, May-June 2025 ||

DOI:10.15680/IJARETY.2025.1203084

symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic),'' in 208 IEEE 5th international symposium on biomedical imaging (ISBI 208). IEEE, 208, pp. 68–72.

[21] A. Yilmaz, M. Kalebasi, Y. Samoylenko, M. E. Guvenilir, and H. Uvet, ''Benchmarking of lightweight deep learning architectures for skin cancer classification using isic 207 dataset,'' arXiv preprint arXiv:20.2270, 202.

[22] C. Matsoukas, J. F. Haslum, M. Söderberg, and K. Smith, ''Is it time to replace cnns with transformers for medical images?'' arXiv preprint arXiv:208.09038, 202.

[23] R. Pedro and A. L. Oliveira, ''Assessing the impact of attention and self-attention mechanisms on the classification of skin lesions,'' in 2022 International Joint Conference on Neural Networks (IJCNN). IEEE, 2022, pp. –8.

[24] Y. Gulzar and S. A. Khan, ''Skin lesion segmentation based on vision transformers and convolutional neural networks—a comparative study,'' Applied Sciences, vol. 2, no. 2, p. 5990, 2022.

IJARETY

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com