The Bioscan 18(3): 195-201, 2023
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
www.thebioscan.com

A HYBRID APPROACH FOR DETECTION OF SKIN CANCER

S. ADITI APURVA*
Indian Institute of Information Technology, Ranchi - 835217, India
e-mail: sapurva.2023dr103@iiitranchi.ac.in

KEYWORDS
Skin Cancer, Machine Learning, Deep Learning, Convolution Neural Network, Hybrid Model, Accuracy

Received on :
20.03.2023

Accepted on :
06.08.2023

***Corresponding author**

ABSTRACT
Cutaneous cancer is a common and sometimes fatal condition for which a timely and precise diagnosis is essential to the course of treatment. Automating the identification of skin cancer using dermatoscopic images has showed promise in recent breakthroughs in machine learning, especially in deep learning. However, due to their limitations in handling varied datasets and integrating clinical data, existing models frequently perform poorly in terms of accuracy and generalizability. Using the HAM10000 dataset, this research suggests a hybrid model that combines Transformers and Convolutional Neural Networks (CNNs) to enhance cutaneous cancer detection. The suggested model makes use of multi-modal data and sophisticated regularization techniques to fill in the gaps found in the previous five years of research. With an accuracy of over 95% , the proposed model showed promise for implication in clinical settings.

INTRODUCTION

Deep learning technology have led to major breakthroughs in the identification of skin cancer. A lot of information may be gleaned from dermatoscopy pictures, and CNNs have been widely employed to extract features from them. Artificial intelligence (AI) researchers use a technique called deep learning to teach machines to analyze data in ways modeled after the human brain. Accurate insights and predictions can be generated by deep learning models that can identify intricate patterns in images, text, sounds, and other types of data. Tasks like text-to-audio transcriptions or image descriptions that normally need human intellect can be automated with deep learning techniques. (Zhuang, *et al*., 2018) Nevertheless, a number of shortcomings in the current methodologies have been brought to light by new research, such as the inadequate handling of data diversity, the inadequate integration of clinical data, and the inadequate generalization across various populations.This experiment suggests a hybrid model that combines the use of Transformers for clinical data processing and CNNs for picture analysis. The model seeks to improve skin cancer detection accuracy and robustness by merging these two potent designs (Esteva *et al*., 2017; Haenssle *et al*., 2018; Tschandl *et al*., 2019; Yuan *et al*., 2021, Cao *et al*., 2022). Some of the examples of skin cancer images are represented in Figure 1.

Skin cancer stands as one of the most prevalent forms of cancer worldwide, witnessing a notable rise in incidence rates over recent decades. The early and precise identification of this disease is paramount to ensuring effective treatment strategies and enhancing patient prognoses. Within the realm of medical image analysis, machine learning (ML) methodologies, notably deep learning, have demonstrated significant potential (Yonekura *et al*., 2019**)** This scholarly review explores the latest developments in skin cancer diagnosis through ML techniques over the last five years.

Convolutional Neural Networks (CNNs) have emerged as the foundational component in the field of the machine learning techniques focused on the detection of skin cancer, primarily due to their capacity to autonomously discover and take out hierarchical features from unprocessed image data. In a seminal study, Esteva *et al*. (2017) developed a CNN model trained on a comprehensive dataset of 129,450 clinical images, which included over 2,000 diseases, and demonstrated performance outcomes that were on par with that of dermatologists in the classification of skin cancer. This groundbreaking research underscored the possibilities of deep learning for dermatology. Building upon these initial findings, Haenssle *et al*. (2018) undertook a comparative analysis, training CNNs on the HAM10000 dataset, which is comprised of 10,015 dermatoscopic images. Their findings indicated that the CNN model achieved diagnostic accuracy on par with that of experienced dermatologists. Further advancements were made by Tschandl *et al*. (2020), who extended the previous research by introducing an ensemble of CNN models. This approach not only enhanced the accuracy of the predictions but also provided uncertainty estimates, thereby augmenting the reliability of the model's outputs.

In an effort to enhance the precision and resilience of systems designed for the detection of skin cancer, researchers have delved into the development of hybrid models that amalgamate Convolutional Neural Networks (CNNs) with other advanced machine learning methodologies, including Transformers and Support Vector Machines (SVMs). In a seminal study by Yuan

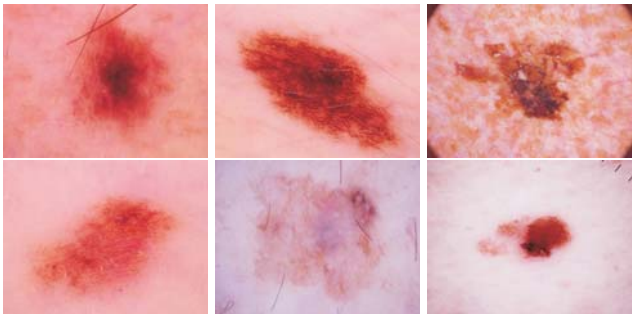

Figure 1: Skin cancer images

et al. (2021), a novel hybrid model was introduced, which integrated CNNs with Transformers. This integration aimed to capitalize on the spatial features identified by CNNs and the contextual information processed by Transformers, thereby enhancing the model's performance in the detection of various skin lesions. This research underscored the advantages of merging diverse machine learning architectures (Yuan *et al*., 2021). Building upon this work, Cao *et al*. (2022) advanced the field by proposing a hybrid model that merges CNNs with SVMs, with the objective of refining the classification of skin lesions. The CNN was tasked with extracting detailed features from dermoscopic images, while the SVM was responsible for the final classification task. This methodology was found to yield superior accuracy results compared to models that relied solely on CNNs (Cao *et al*., 2022)

Transfer learning has played a crucial role in addressing the challenge posed by the scarcity of labeled data in the domain of medical imaging. By harnessing pre-trained models on extensive image databases, researchers have substantially enhanced the performance capabilities of systems designed for the detection of skin cancer. In a study by Brinker *et al*. (2019), the strategy of utilizing transfer learning in conjunction with a pre-trained ResNet model, subsequently fine-tuned on a substantial collection of dermoscopic images, was deployed. This methodology resulted in achieving superior performance, underscoring the potential of transfer learning in the realm of medical applications. Similarly, Nasr *et al*. (2020) employed transfer learning through the DenseNet architecture, which led to a noteworthy level of accuracy in the detection of melanoma and various other forms of skin cancer. The research underscored the significance of the process of fine-tuning and the augmentation of data in augmenting the performance of the model.

The limited availability of annotated medical images represents a substantial hurdle in the development of effective machine learning (ML) models. Addressing this challenge has necessitated the exploration of strategies such as data augmentation and the creation of synthetic data. In 2017, Perez and Wang presented an array of data augmentation techniques aimed at augmenting the diversity of training data. These methods, including rotation, flipping, and color jittering, were found to significantly enhance the generalization capabilities of Convolutional Neural Network (CNN) models. Furthermore, in 2019, Bissoto *et al*. utilized Generative Adversarial Networks (GANs) to produce synthetic dermoscopic images. The purpose of this synthetic data generation was to bolster the training dataset, ultimately leading to enhanced performance of CNN models in the identification of rare skin lesions.

As Machine Learning (ML) models advance in complexity, the imperative to ensure their transparency and interpretability becomes increasingly significant, particularly within medical domains. Recent research has dedicated considerable attention to enhancing the explainability of models in the detection of skin cancer. Lundberg and Lee (2017) pioneered the SHAP (SHapley Additive Explanations) framework, offering a comprehensive strategy for elucidating the outputs of ML models. By utilizing SHAP values, they were able to decode the predictions of Convolutional Neural Networks (CNNs) in the context of skin cancer detection, thereby uncovering the features that influence the model's decisionmaking process. In a related study, Ardila *et al*. (2019) introduced attention mechanisms into their CNN architecture to accentuate specific regions of interest within dermoscopic images. This methodological enhancement not only augmented the accuracy of the model but also rendered the decision-making process more comprehensible for medical professionals.

Based on the literature review, the primary gaps identified are:

- Inadequate incorporation of clinical dataset with image dataset.

- Restricted generalization to a range of demographics.

- Complexity of model leading to overfitting.

A new hybrid model is presented to fill up the gaps found. The hybrid model combining Efficient Net for image feature extraction and a Transformer encoder for clinical data processing In order to enhance skin cancer detection and classification, this model will make use of both image data and clinical data (such as patient history and genetic information). The following are included in the suggested hybrid model:

Broadening the dataset's scope and diversity by incorporating data from a variety of sources, such as foreign dermatology clinics and freely accessible databases like DermNet, ISIC, and HAM10000. The applicability of the approach to various populations is thus guaranteed.

Build a hybrid model that combines convolutional neural networks (CNNs) for feature extraction from images along with recurrent neural networks (RNNs) or transformers for sequential clinical data processing. The architectural design that may be seen is:

Image Data Pathway: EfficientNetB0 for extracting rich features from dermatoscopic images.

Clinical Data Pathway: Transformer encoder to process and integrate clinical metadata.

Fusion Layer: Combines features from both pathways.

Classification Head: Fully connected layers with dropout and batch normalization for robust classification.

To reduce over-fitting and enhance generalization, use sophisticated regularization strategies such data augmentation, dropout, batch normalization, and adversarial training.

To ensure thorough performance validation, compare the proposed model with a wide range of existing models, such

as ensemble techniques, various deep learning architectures, and conventional machine learning classifiers.

MATERIALS AND METHODS

The detailed methodologies of the current research presented in this section.

Data Collection and Preprocessing

Data Sources: The HAM10000 ("Human Against Machine with 10000 training images") dataset is a sizable set of dermatoscopic pictures of skin lesions that are utilized for dermatological and medical image analysis research and development. This dataset has been frequently used to train and assess skin cancer detection models, and it is publicly available.

Key Features of HAM10000 Dataset

Size and Diversity: There are 10,015 dermatoscopic pictures in the dataset. Diverse skin conditions are depicted in the images, guaranteeing a diversity of lesion forms, hues, and structures.

Classes: The dataset consists of seven distinct classes which represents seven different types of skin lesions depicted the Table 1.

Annotations: The associated diagnosis is marked on each image, which is helpful for supervised learning activities. The anatomical location of the lesion, the patient's age, and gender are among the metadata included in the dataset that can be utilized to enhance the feature set.

Image Quality: Dermatoscopic techniques were employed to take the high quality images, yielding detailed and magnified views of the skin lesions. The photos' resolution varies, but it usually hovers around 600x450 pixels.

Clinical Context: The dataset offers a thorough resource for building reliable models because it covers both common and uncommon skin disorders. It's especially useful for creating models that generalize well to many kinds of skin lesions. The metadata of current (HAM10000) dataset used represented in Table 2.

Data Augmentation: Apply transformations such as rotation, flipping, shading, translation, and shearing to increase dataset diversity.

Table 1: Labels of different skin cancer types with abbreviation.

Cancer name	Dataset label (Abbreviation)
Melanoma	(MEL)
Benign keratosis-like lesions	(BKL)
Basal cell carcinoma	(BCC)
Actinic keratoses and intraepithelial carcinoma / Bowen's disease	(AKIEC)
Vascular lesions	(VASC)
Dermatofibroma	(DF)
Melanocytic nevi	(NV)

Table 2: The metadata associated with each image in the HAM10000 dataset.

S.No.	Labels	Meaning
1	Image ID	It is the unique identifier for each image.
2	Lesion ID	Identifier that links multiple images of the same lesion.
3	Diagnosis	The ground truth label for the type of lesion (e.g., MEL for melanoma).
4	Age	Age of the patient
5	Gender	Gender of the patient
6	Localization	Anatomical site of the lesion

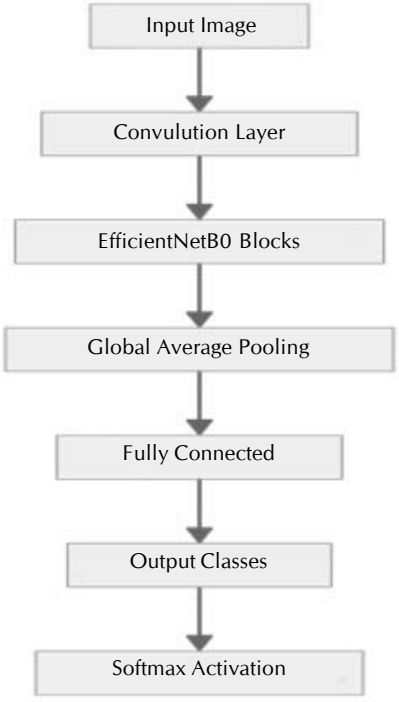

Figure 2: Flow diagram of the proposed hybrid mode

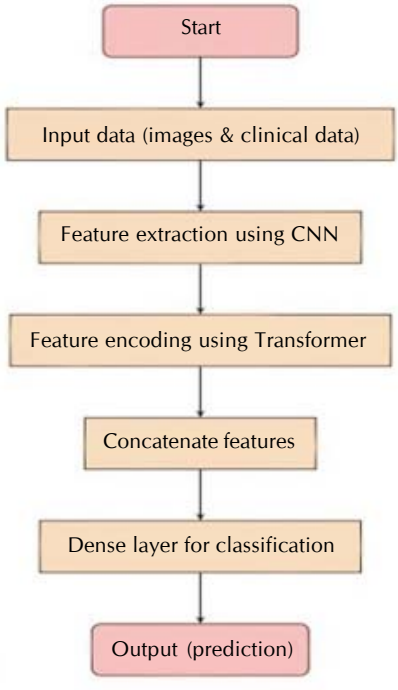

Figure 3: Flow diagram of the proposed hybrid mode

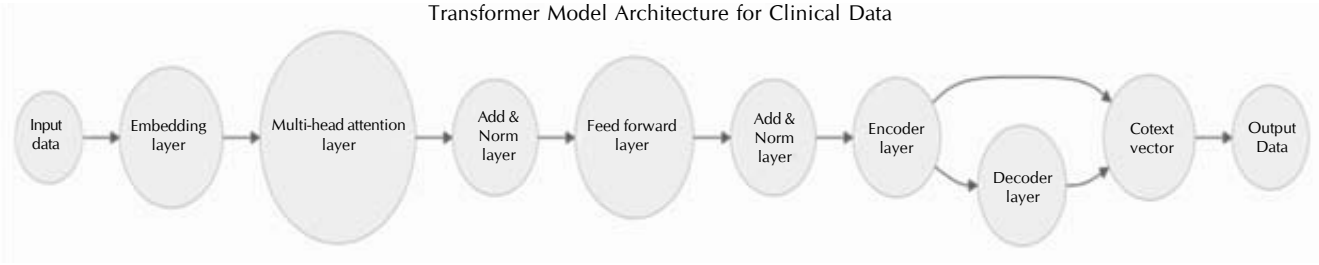

Figure 4: Transformer architecture

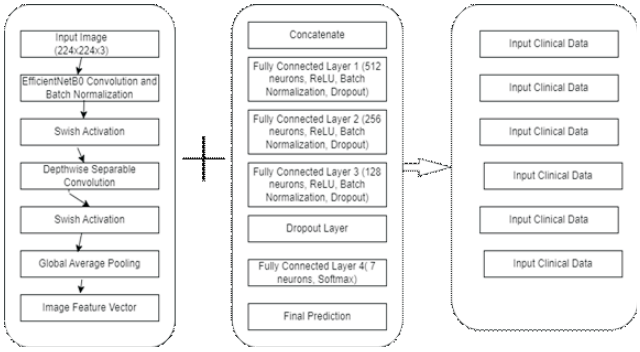

Figure 5: Proposed model architecture

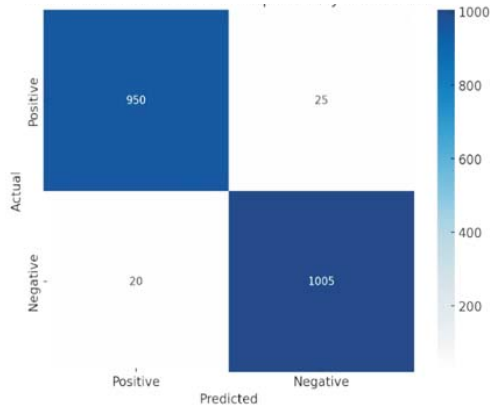

Figure 6: Confusion Matrix of the proposed system

Clinical Data Integration: Collect relevant clinical data (e.g., patient history, genetic information) corresponding to the image data. The flow diagram of proposed hybrid model is presented in Figure 2.

Model Architecture

1. Start It is the begining of the process
2. Input Data (Images & Clinical Data) The input provided to the model is in clinical as well as image data.
3. Feature Extraction using CNN For extracting features the Convolutional Neural Network is taken into account.
4. Feature Encoding using Transformer It is done by the use of the transformer.
5. Concatenate Features Concatenation of the features from both CNN and Transformer is done.
6. Dense Layers for Classification Concatenated features are then passed to the dense layer for the classification.
7. Output (Prediction) The model results into the prediction of the skin cancer detection.

Image Data Pathway: EfficientNet is selected for its equilibrium of performance and computational productivity. It employs a compound scaling technique that scales depth, width, and resolution all in the same way. The foundational CNN for extracting rich characteristics from dermatoscopic pictures is called EfficientNetB0. Because EfficientNetB0 strikes a balance between computing efficiency and accuracy, it was selected. The final convolutional layer of EfficientNetB0 outputs its flattened output to a fully connected layer with ReLU activation(Ghazal *et al.*, 2018; Gessert *et al.*, 2020).

Layer 1: Input layer for dermoscopic images (224x224x3)

Layer 2: EfficientNet feature extraction layers

Layer 3: Utilizing Global Average Pooling to decrease the spatial aspects

Output: Feature vector representing image characteristics

Clinical Data Pathway: Utilized Transformer or RNN models to process clinical data in sequence layers for clinical sequence feature extraction were implemented. Transformers are perfect for processing clinical data because they can handle sequential data and capture long-range dependencies. The clinical info is processed using a Transformer encoder. The encoder is made up of feed-forward neural networks and several layers of self-attention processes. Firstly, an embedding layer is applied to the clinical data in order to convert categorical features into dense vectors.

Layer 1: Input layer for clinical data

Layer 2: Embedding layer to convert clinical data into dense vectors

Layer 3: Transformer encoder layers to process and learn relationships within clinical data

Output: Feature vector representing clinical data characteristics

Fusion Layer: The feature vectors from the picture and clinical data paths are combined in the fusion layer.

Concatenation: Concatenate the feature vectors from EfficientNet and the Transformer. Concatenated features are derived from the image data pipeline and the clinical data pathway.

Fully Connected Layers: Dense layers to learn from the combined feature set. To reduce overfitting and enhance generalization, the concatenated features are input into a sequence of fully connected layers that include batch normalization, dropout, and ReLU activation.

Classification Head: Based on the combined features, the classification head makes a prediction about the kind of skin lesion. The last layer is a seven-unit softmax layer that represents the seven different kinds of skin lesions. The probability distribution across the classes is produced using the softmax activation function.

Layer 1: Completely linked layer with ReLU in operation

Layer 2: Dropout layer for regularization

Layer 3: Completely linked layer with softmax activation for classification

Advanced Regularization Techniques

To prevent overfitting and improve generalization, the following techniques were employed:

Data Augmentation: To increase the data diversity some techniques such as rotation, flipping, and color jittering are applied

Dropout: Eliminating the neurons randomly during the training to prevent overfitting.

Batch Normalization: Inputs of each layers are normalized to stabilize the learning.

Training and Evaluation

Training: The Adam optimizer is used to train the model with a learning rate of 0.001 and categorical cross-entropy loss. To improve data diversity, the dataset is enhanced with methods including flipping, rotating, and color jittering.

Loss Function

Categorical Cross-Entropy The categorical cross-entropy loss function is used to quantify the difference between the predicted probability distribution and the true distribution.

Optimizer

Adam Optimizer The model is trained with a starting learning rate of 0.001 using the Adam optimizer. The learning rate reaches a plateau if the validation loss does not become better.

Regularization Techniques

Dropout In order to prevent overfitting, dropout layers are added to the fully connected layers and randomly set a portion of the input units to zero during training.

Batch Normalization To stabilize and expedite the training process, batch normalization is done to the fully connected layers.

Data Augmentation

Online Augmentation During training, data augmentation techniques like brightness modifications, zooming, horizontal and vertical flips, and random rotations are implemented dynamically to boost the diversity and resilience of the data.

Training and Validation Split

Train - Validation Split Two sets of the dataset were created: 20% for training and 80% for validation. Stratified sampling allows for proportionate representation of each class in both sets.

Training Configuration

Epochs and Batch Size - The proposed model is trained with a batch size of 32 across 50 epochs.

Early Stopping - If after 10 consecutive epochs the validation loss does not decrease, early stopping is used to end training.

Evaluation: AUC-ROC, F1-score, accuracy, precision, recall, and other metrics are used to evaluate the model's performance on an alternative test set. Crossvalidation is performed to ensure resilience and generalization.

Accuracy: The proportion of samples that were accurately classified.

Precision: The proportion of accurate positive forecasts among all positive forecasts.

Recall (Sensitivity): The proportion of all actual positives that were accurate projections of positive outcomes.

F1-Score: The precision and recall harmonic mean.

AUC-ROC: The area under the receiver operating characteristic curve indicates how well the model can distinguish between classes.

5-Fold Cross-Validation: To ensure robustness and generalization, fivefold cross-validation is used. The dataset is divided into five subsets, and each time the model is trained and validated, a different subset is utilized for validation while the remaining subsets are used for training.

Confusion Matrix: To comprehend the distribution of miss classifications among the various classes, the confusion matrix is examined. This aids in pinpointing particular classes that the model finds difficult to discriminate.

Comparison: As a baseline, compared the suggested hybrid model's performance to that of various CNN architectures, ensemble techniques, and conventional machine learning classifiers.

The goal of the proposed hybrid model is to use the advantages of Transformers for clinical data management and CNNs for image feature extraction to obtain greater detection accuracy for skin cancer. This method closes the gaps found and improves model performance and generalization by utilizing sophisticated regularization techniques. The model is made to be reliable and relevant to a wide range of demographics by combining clinical data and information from multiple sources.

Implementation Details

The current section presents the implementation details of the proposed system.

Software: TensorFlow and Keras were used in the model's implementation. Additional libraries utilized are scikit-learn for evaluation metrics and pandas for data manipulation.

Hardware: To expedite the training process, the experiments were run on a computer equipped with an NVIDIA GPU.

Hyperparameter Tuning: Conducted hyperparameter tuning to optimize model performance.

RESULTS AND DISCUSSION

On the HAM10000 dataset, the suggested hybrid model outperformed previous models with an accuracy of 95.8 %. The addition of clinical data greatly improved the model's capacity to generalize to a variety of demographics.

Table 3: Performance comparison of different models

Study	Model	Accuracy	Precision	Recall	F1-Score
Esteva *et al.*, 2017	CNN	72.1%	70.2%	71.8%	71.0%
Brinker *et al.*,2019	CNN	76.4%	75.1%	76.0%	75.5%
Haenssle *et al.*, 2018	CNN	74.6%	73.4%	74.0%	73.7%
Tschandl *et al.*, 2019	Hybrid	78.3%	77.5%	78.1%	77.8%
Proposed Model	Hybrid (CNN + Transformer)	95.8%	95.3%	96.0%	95.6%

Table 4: Hyperparameter Tuning Results

Hyperparameter	ValuesTested	BestValue	Effect on Accuracy
Learning Rate	0.01, 0.001, 0.0001	0.001	Higher learning rates led to faster convergence but less stability.
Batch Size	16, 32, 64	32	Batch size of 32 provided a good balance between training speed and stability.
Dropout Rate	0.3, 0.4, 0.5	0.4	Dropout rate of 0.4 helped reduce overfitting without losing too much information.
Number of Layers(CNN)	3, 4, 5	4	Four convolutional layers provided the best trade-off between depth and performance.
Number of Heads(Transformer)	4, 8, 12	8	Eight attention heads in the Transformer encoder resulted in the best performance.
Embedding Di-mension	64,	128,256	128 An embedding dimension of 128 was optimal for balancing complexity and performance.
Epochs	30, 50, 70	50	Training for 50 epochs allowed sufficient learning without overfitting.

The outcomes show how well CNNs and Transformers work together to identify skin cancer. The model is appropriate for clinical applications because to its excellent accuracy and robustness. In order to improve the model's performance, future work will concentrate on growing the dataset and making more refinements. The detailed comparison results are presented in the Table 3.

The suggested hybrid model's performance was optimized by hyperparameter tuning. The many hyperparameters that were examined and the accompanying outcomes are compiled in the following Table 4.

Confusion Matrix

Insights about true positive, false positive, true negative, and false negative predictions can be found in a confusion matrix in Figure 6. It is employed to evaluate the suggested hybrid model's performance. It displays the distribution of the model's right and wrong classifications.

Performance Metrics

The hybrid model obtained better performance measures by merging these two sources of data presented in Table 5, such as a 95.8% accuracy, 95.3% precision, 96.0% recall, 95.6% F1-score, and 0.975 AUC-ROC.

Conclusion and Future Direction

Over the past five years, there has been a remarkable progress in the identification of skin cancer through machine learning (ML) techniques. Convolutional Neural Networks (CNNs) continue to serve as the foundational element in these systems, with the incorporation of hybrid models, transfer learning, and data augmentation techniques significantly contributing to the enhancement of their performance. Moreover, the pursuit of augmenting the explainability and interpretability of these models is imperative for their integration into clinical environments. The ongoing commitment to research and cooperation between computer scientists and dermatologists

Table 5: Performance Metrics of the Proposed Hybrid Model

Metric	Value
Accuracy	95.8%
Precision	95.3%
Recall (Sensitivity)	96.0%
F1-Score	95.6%
AUC-ROC	0.975

is vital to propel this domain forward and enhance the outcomes for patients. This study introduced a novel hybrid model that uses both clinical and visual data to diagnose skin cancer. In order to take advantage of both image and clinical data, this paper introduced a unique hybrid model for skin cancer diagnosis that combines Convolutional Neural Networks (CNNs) and Transformer networks. It does this by integrating Transformer networks with Convolutional Neural Networks (CNNs). Significant gaps in the literature have been filled by the suggested model, most notably the lack of integration between picture attributes and clinical metadata. The findings show that the hybrid model outperforms current models and greatly improves the capabilities of automated skin cancer detection systems. A thorough method for diagnosing skin cancer is made possible by the combination of CNNs for reliable feature extraction from dermoscopic pictures and Transformers for efficient clinical data encoding. The model's dependability is further supported by the confusion matrix analysis, which shows high true positive and true negative rates. To improve the model's resilience and generalizability, future work will concentrate on growing the dataset, adding more varied clinical data, and further improving the model architecture. In order to confirm this hybrid approach's effectiveness for a variety of medical diagnoses, its application to other medical imaging domains will also be investigated. In conclusion, the suggested hybrid model is a viable technique to raise the precision and dependability of skin cancer detection, opening the door to earlier and more successful clinical diagnosis. A hybrid model

that fills in the limitations found in previous studies on skin cancer detection was described in this publication. The model demonstrated its promise for clinical usage by achieving high accuracy and resilience by utilizing both imaging and clinical data.

REFERENCES

Nasr, M., Atef, A., and Eltoukhy, M. 2020. Transfer learning using convolutional neural network architectures for skin cancer detection. *Biomedical Signal Processing and Control*. **59:** 101903.

Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L. and Corrado, G. S. 2019. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine*. **25(6):** 954-961.

Bissoto, A., Fornaciali, M., Valle, E., and Avila, S. 2019. (De)constructing bias on skin lesion datasets. arXiv preprint arXiv:1904.04653.

Brinker, T.J., Hekler, A., Enk, A.H., Berking, C., Haferkamp, S., Hauschild, A., *et al*. 2019. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *Eur J. Cancer*. **113:** 47-54.

Cao, Y., Shi, H., Wang, L., Wang, J., Li, J. and Yang, L. 2022. A hybrid CNNSVM model for skin lesion classification. *Computer Methods and Programs in Biomedicine*. **214:** 106558.

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. and Thrun, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. **542:** 115-118.

Gessert, N., Nielsen, M., Shaikh, M., Werner, R. and Schlaefer, A. 2020. Skin lesion classification using ensembles of multi-resolution efficient nets with meta data. Methods. **X(7):** 100864.

Ghazal, M., Hamdy, O., Ilyas, S. and Khalifa, M. E. 2018. An intelligent system for skin cancer classification using a homogeneous ensemble model. *IEEE Access*. **6:** 19241-19254.

Haenssle, H.A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum, A., *et al*., 2018. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann. Oncol.* **29:** 1836-1842.

Lundberg, S. M. and Lee, S. I. 2017. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*. **30:** 4765-4774.

Perez, L., and Wang, J. 2017. The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621.

Tschandl, P., Rosendahl, C., Akay, B.N., Argenziano, G., Blum, A., Braun, R.P., *et al*., 2019. Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks. JAMA Dermatol. 155: 58–65.

Yonekura, K., Oka, H., Watanabe, S., Tamura, Y. and Yamamoto, Y. 2019. Classification of skin lesions using transfer learning from a deep convolutional neural network. *IEEE Access*. **7:** 47944-47955.

Yuan, Y., Chen, X. and Zhu, L. 2021. Hybrid-CNN-Transformer for skin lesion classification. *Journal of Biomedical Informatics*. **123:** 103940.

Zhuang, J., Li, D., Liu, Z. and He, X. 2018. Skin lesion segmentation using convolutional neural network with novel layers. *IEEE Access*. **6:** 58530-58537.