Faster-CAN-FL: A Federated Convolutional Attention Network with AC-GAN for Privacy-Preserving Skin Cancer Classification

Ponnam Lalitha[1], Dr. V. Janardhan Babu[2]
[1]*Dept.of CSE Research Scholar JNTUA, Anatapuram Andhra Pradesh, India*
[2]*Asst.Professor Dept.of CSE, Supervisor, Sri Venkatesa Perumal College of Engineering & Technology Andhra Pradesh, India*

***Abstract*—Early diagnosis and treatment of skin cancer depend heavily on its classification; nevertheless, manual classifi- cation is frequently laborious and prone to human error. Deep learning approaches have demonstrated significant promise in automating and enhancing the precision of skin lesion classi- fication, particularly in light of the swift progress of artificial intelligence. However, issues like data imbalance and restricted sample availability can sometimes hinder the efficacy of deep learning models. In order to improve model training, this study uses Auxiliary Classifier Generative Adversarial Networks (AC- GAN) to balance class distributions and create synthetic data. For the purpose of accurately and efficiently classifying skin cancer, we suggest a new hybrid model called Faster-CAN, which is a faster convolutional neural network architecture. Results from experiments show that adding data augmentation based on AC-GANs greatly enhances model performance. With a testing accuracy of 98.42%, precision of 98.50%, recall of 98.31%, and F1-score of 98.40%, the suggested Faster-CAN model produces exceptional results. These findings show how well synthetic data augmentation and a quicker, more efficient hybrid architecture work together to provide a reliable and scalable method for diagnosing skin cancer in clinical settings.**

***Index Terms*—Skin Cancer; Faster-CNN, Attention Mechanism, Federated Learning, HAM1000**

I. INTRODUCTION

Millions of cases of skin cancer are diagnosed annually, making it one of the most prevalent types of cancer in the world. Multi-type cancer is possibly the most dangerous disease and a major threat to human health [1]. Melanoma skin cancer is among the deadliest types of cancer. Skin cancer is the second leading cause of death globally, after heart disease [2]. It is at significant danger because it is currently the 19th most frequent cancer worldwide and its incidence is increasing [3]. There are several varieties of skin cancer, each with unique traits, risk factors, and treatment choices. Skin cancer is a type of cancer that starts in the skin's cells [4]. The human body is primarily covered by the skin, which also provides protection from the elements, including heat, dust, UV radiation, and contaminated water [5]. Even with this protective role, the skin can still be susceptible to a number of illnesses that can affect people of all ages, such as eczema, rosacea, moles, and different types of cancer. Skin cancer has become a growing risk in addition to these diseases [6]. There are two sorts of tumors: benign and malignant. Whereas benign cells stay confined and noninvasive, malignant cells proliferate and spread quickly. Because cancer cells are aggressive, a variety of therapeutic modalities are used, including radiation, chemotherapy, and surgery.

The World Health Organization (WHO) reports that three out of every four cancer cases are caused by skin cancer, which is also notably on the rise in countries like the US, Canada, and Australia. The increasing prevalence of skin cancer, which is brought on by a 10% ozone layer depletion and increased UV radiation, is a major global health problem [7]. Protecting the skin from the sun is crucial by applying sunscreen, donning protective clothing, and avoiding direct sunlight. Early diagnosis of new or developing skin lesions depends on routine skin inspections and prompt medical care.

It is crucial to remain vigilant and take preventative action because aging, hormonal changes, medications, smoking, HPV infections, artificial UV exposure, and underlying medical disorders can also bring on skin lesions. As a result, cancer patients, particularly those with skin cancer, require early detection and treatment [8]. Thirty to fifty percent of skin cancer cases are treatable if detected early and treated adequately [9]. Previously, this diagnosis was done manually. But, in recent years, deep learning (DL) and machine learning (ML) techniques have greatly aided skin cancer detection and classification. But the limited number of images hinders the accurate classification of the model [10].

This work uses Auxiliary Classifier Generation Adversarial Networks (AC-GAN) for enhanced data augmentation to over- come the problem of limited sample availability in medical imaging [11]. By producing high-quality synthetic pictures conditioned on class labels, AC-GAN successfully balances class distributions and adds a variety of dermatoscopic patterns to the training dataset. Model generalization and robustness are greatly improved by this augmentation strategy, particu- larly in situations involving rare illness categories or class imbalance. The study presents Faster-CAN (Faster Convolu- tional Attention Network) [12], a revolutionary deep learning architecture for automated and privacy-preserving skin cancer classification, built on this enriched dataset. The HAM10000 dermatoscopic image dataset is used to train and assess the model, and Federated Learning (FL) protects patient privacy

in dispersed healthcare settings [13]. Faster-CAN enhances classification accuracy and makes feature selection more effi- cient and comprehensible by fusing lightweight convolutional encoders with both channel and spatial attention techniques. By using FL, it is possible to train the model cooperatively across several clients (such as clinics or hospitals) without exchanging sensitive raw data, protecting patient privacy. The overall framework of the research is shown in Figure 1.

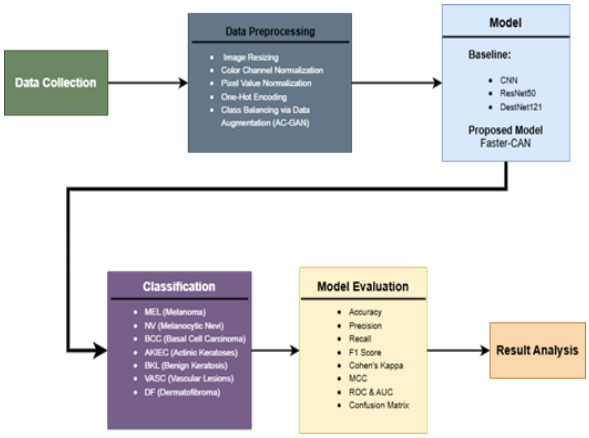

Fig. 1.

The major contributions of the research are as follows:

- Auxiliary Classifier Generative Adversarial Networks (AC-GAN) are used in this study to enrich the data.

This method successfully increases dataset diversity and balances class distributions by generating high-fidelity synthetic images conditioned on class labels.

- Faster-CAN (Faster Convolutional Attention Network) is a new deep learning model that is suggested. The architecture allows for focused and effective feature

extraction by combining channel and spatial attention techniques with lightweight convolutional encoders. The model's capacity to locate and identify pertinent characteristics linked to different skin lesions is im- proved by this attention-guided structure, which raises classification accuracy.

- The model is coupled with Federated Learning (FL) to guarantee the confidentiality of patient data. With this method, several clients—like clinics or hospitals—can

work together to train the model without exchanging raw picture data. FL preserves data locality, allowing for secure and privacy-preserving learning over dis- persed data sources.

- Standard classification metrics like accuracy, preci- sion, recall, F1-score, Cohen's Kappa, and Matthews

Correlation Coefficient (MCC) are used to completely assess the model's performance. The suggested frame- work's robustness, dependability, and appropriateness for clinical deployment are confirmed by the results, which demonstrate its

outstanding and balanced per- formance across all criteria. .

This paper has been organized into multiple sections. First, we will review the existing methods for identifying skin cancer
in Section II. Then, Section III will provide an overview of the dataset description and preprocessing steps. Section IV presented our proposed model. The outcomes of the research and analysis is discussed in Section V. Ultimately; we will showcase our research findings in Section VI.

II. LITERATURE REVIEW

Many studies have been conducted in the last ten years to increase the classification of skin cancer's precision, generalizability, and clinical usefulness. With differing emphasis on segmentation, feature extraction, controlling data imbalance, and real-time diagnosis, these attempts range from conventional deep learning methods to more complex hybrid frameworks.

Monica et al. [14] and Rajesh et al. [15] only used deep convolutional neural networks (DCNNs) enhanced by pre-trained models like Inception V3, DenseNet201, and GRU variants, whereas Naeem et al. [16], [17] combined handcrafted features with deep learning models (SNC_Net and DVFNet) to capture both low-level and high-level representations. Using a Swish-ReLU-activated GRU model, Monica et al. notably out- performed the majority of traditional CNN-based techniques, reaching near-perfect accuracies (99.98%) on HAM10000. However, the performance ceiling of traditional architectures without sophisticated augmentation or feature fusion was re- vealed by Rajesh et al.'s ensemble technique, which used Inception V3 and DenseNet201 and capped at 87.42%.

To illustrate the superiority of hybrid features over pure CNNs in capturing discriminative lesion characteristics, DVFNet [17] stands out in terms of feature engineering and hybridization by integrating Histogram of Oriented Gradients (HOG) with VGG19, resulting in a 98.32% accuracy. Similarly, SNC_Net demonstrated the efficacy of complementary feature integration by surpassing multiple state-of-the-art models on the ISIC 2019 dataset [16] thanks to its utilization of both handmade and deep features.

Handling data imbalances is another crucial element that has a big impact on the classification of uncommon skin lesion types. SMOTE-Tomek was used by Naeem et al. [17] to address this, and Kousis et al. [18] used targeted augmenta- tion and fine-tuning, which was particularly advantageous for underrepresented lesion types. In contrast to many previous models that were primarily tested in academic or server-based settings, their implementation of DenseNet169 in a mobile application not only exhibited realistic deployment possibil- ities but also achieved excellent accuracy (92.25%). Strategies for localization and segmentation also differed significantly. Sivakumar et al. [19] coupled ResNet50 with hybrid pool- ing and innovative preprocessing for spectral enhancement, resulting in a 94% classification accuracy, whereas Monica et al. [14] used Mask R-CNN for accurate lesion segmen- tation, enabling superior downstream feature extraction. This demonstrates how using efficient segmentation before classi- fication can greatly enhance model performance, particularly for intricate lesion boundaries.Campos et al. [20] investigated a non-computational treatment-based strategy that used MoS nanoparticles and photothermal therapy to specifically target malignant skin cells, in contrast to algorithmic advancements.

This work complements the diagnostic focus of machine learning research and highlights an alternative, therapeutic breakthrough, although having nothing to do with classification methods.

AlSadhan et al. [21] evaluated many YOLO structures for real-time skin lesion localization and classification from the standpoints of optimization and computing efficiency. With 86.3% IoU and quick inference (0.32s/image), YOLOv7 was found to be the most efficient, indicating that lightweight object detectors hold promise for latency-sensitive real-world applications. In a similar vein, Rahman et al. [22] modified NASNet designs by adding a layer to manage sparse and irregular data; this resulted in a respectable accuracy of 85.62 percent, but it did not equal hybrid or fine-tuned deep models. From simple CNN models to more hybrid, ensemble, and application-aware frameworks that emphasize feature fusion, data imbalance, segmentation, and

deployment efficiency, the field of skin cancer classification research is generally showing a transition. Some studies emphasize practicality and inter- pretability (e.g., Kousis et al. [18], Sivakumar et al. [19]), while others place more emphasis on raw classification performance (e.g., Monica et al. [14]). Combining these elements could help future methods produce models that are not only precise and generalizable but also quick, easy to understand, and clinically applicable.

Many deep learning approaches have been used to tackle the universal problem of classifying skin cancer. In comparison to current machine learning models, Sowmya and Balasub- ramanian [23] achieved greater accuracy on the HAM10000 dataset by merging convolutional neural networks (CNN) with the Discrete Wavelet Transform (DWT) for feature extraction. In a similar vein, Naz et al. [24] developed an ensemble CNN model based on VGG-16 and a hair removal preprocessing strategy to detect malignant melanoma. This improved classi- fication performance and showed increased accuracy and F1- score on the same dataset. Using sequential 2D CNN layers and data augmentation techniques to balance the dataset, Javed et al. [25] presented a stacked CNN model that performed better on the MINST-HAM10000 datasets than other pretrained CNN variations.To improve the accuracy of skin disease cat- egorization, Basha et al. [26] investigated multimodal data integration by integrating metadata and high-resolution images of Squamous Cell Carcinoma (SCC) using multi-model CNNs. Their strategy gave priority to the 92% accurate detection of SCC, demonstrating the value of integrating diverse data sources with cutting-edge machine learning techniques for early skin cancer detection. Although these studies agree that deep learning can be used to classify skin cancer, they take different approaches, ranging from feature engineering and ensemble learning to multimodal data integration. This shows that efforts are still being made to improve diagnostic accuracy and early prediction through a variety of methodological advancements.

III. DATASET DESCRIPTION & PREPROCESSING

A popular benchmark dataset for analyzing and classifying skin lesions in dermatology is the HAM10000 dataset, which stands for Human Against Machine with 10,000 training pho- tos. It was created to aid in diagnosing pigmented skin lesions and was explicitly designed to promote machine learning and deep learning research in dermatoscopic image recognition. The dataset, which is housed on the ISIC (International Skin Imaging Collaboration) archive, was selected and made public by the Medical University of Vienna in association with the Australian Skin Cancer Institute.10,015 dermatoscopic images in all, each identified by one of seven different types of skin lesions, are included in this collection. These include dermatofibroma (df), actinic keratoses and intraepithelial car- cinoma (akiec), vascular lesions (vasc), benign keratosis-like lesions (bkl), melanocytic nevi (nv), melanoma (mel), and basal cell carcinoma (bcc). Of these, dermatofibroma is the rarest, with just 115 samples, and melanocytic nevi is the most common class, with 6,705 photos. This dataset is perfect for testing out sophisticated classification methods that deal with imbalance, like data augmentation, class weighting, and focus loss, because of its unequal distribution, which poses a serious class imbalance problem. Because of their excellent quality and uniform resolution (600x450 pixels), the photos in HAM10000 provide reliable input for deep learning models. The dataset supports multimodal learning techniques by integrating patient metadata such as age, sex, and lesion site with the image data. The dataset samples are shown in Figure 2.

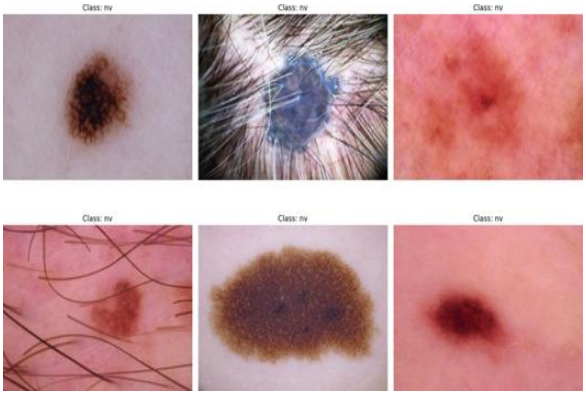

Fig. 2. Samples of the dataset

A. Preprocessing

Some preparation procedures were used on the HAM10000 dataset to guarantee consistency and

enhance the classification model's performance. To assign each image to its appropriate class, the metadata file comprising image identifiers and diag-nostic labels was first parsed. Since the original photos were saved in JPEG format using OpenCV standards, they were loaded from the source directory and converted from BGR to RGB format.

Every dermoscopic image was downsized to a fixed 224 by 224 pixel size to satisfy the input specifications of the models, including ResNet, VGG, MobileNet, and the proposed model. While preserving enough lesion detail for classification, this downsizing procedure helps standardise the data and lower computational overhead. Pixel intensity data were then adjusted by dividing by 255.0 and scaling them to the [0, 1] range. In deep learning models, this normalisation guarantees numerical stability and promotes quicker convergence during training. Normalisation was followed by label encoding. While classes were labelled as 0 to 6 for multi-class classification tasks, one-hot encoding was used to encode the seven diag- nostic categories in the multi-class classification instance. Data augmentation techniques were utilised to address the notable class imbalance present in the dataset. Below is a description of the data augmentation methods used in the study.

B. Data Balancing Using AC-GAN

An Auxiliary Classifier Generative Adversarial Network (AC-GAN) was used as a data balancing tool to address the notable class imbalance in the HAM10000 dataset. Conven- tional augmentation methods, such as rotation, flipping, and zooming, are often limited in their capacity to add meaningful variety, particularly when only a few examples are available from underrepresented classes. Consequently, we employed AC-GAN to produce class-conditional synthetic dermoscopic images, increasing sample diversity and preserving class-specific characteristics crucial for accurate classification.

The class balancing strategy in our study was based on equalizing the number of samples in each class to match the count of the majority class. Assume the dataset contains C total classes, and each class c has N_c original images. The class with the highest number of instances is denoted as $N_{\max}$, computed as:

$$N_{\max} = \max_{c \in \{1,2,\ldots,C\}} N_c, \tag{1}$$

For every class c where $N_c < N_{\max}$, the number of synthetic images to be generated is calculated as:

$$N_c^{\text{gen}} = N_{\max} - N_c. \tag{2}$$

In our dataset, the Melanocytic nevi (nv) class was the most populous, with Nmax = 6705 samples. Accordingly, N gen synthetic images were generated for each minority class using AC-GAN, resulting in a perfectly balanced dataset with Nmax samples per class.

The AC-GAN architecture comprises two neural networks trained adversarially: a generator G and a discriminator D. The generator is conditioned on a class label y and a random noise vector $z \sim \mathrm{N}(0, I)$, allowing it to generate synthetic images $G(z, y)$ that resemble dermoscopic images from the target class. The discriminator receives either real or generated images and produces two outputs: one predicting whether the input is real or fake, and another predicting the class label. This dual-task setting enables the discriminator to extract and learn rich, class-specific features while distinguishing between authentic and synthetic images.

The training objective of the AC-GAN incorporates both adversarial and classification loss components. The discrimi- nator loss L_D is defined as:

$$\begin{aligned} \mathrm{L}_D &= E\,[\log P(S = \text{real} \mid x_{\text{real}})] \\ &+ E\,[\log P(S = \text{fake} \mid G(z, y))] \\ &+ E\,[\log P(C = y \mid x_{\text{real}})] \end{aligned} \tag{3}$$

and the generator loss L_G is given by:

$$\begin{aligned} \mathrm{L}_G &= E\,[\log P(S = \text{real} \mid G(z, y))] \\ &+ E\,[\log P(C = y \mid G(z, y))] \end{aligned} \tag{4}$$

where $P(S \mid x)$ represents the probability that the discrim- inator assigns to the image being real or fake, and $P(C \mid x)$ is the probability distribution over the class labels. These loss functions are optimised using alternating gradient descent, encouraging the generator to produce semantically accurate and visually realistic images.

The generator begins with a fully connected input layer that accepts the concatenated noise vector and a

one-hot encoded class label. It then passes through a series of upsampling layers implemented using transposed convolutions, each followed by batch normalisation and ReLU activations. The final output layer uses a Tanh activation function to generate a 224×224
RGB image.

Conversely, the discriminator is a convolutional neural net- work that processes images using several convolutional layers, LeakyReLU activations, and dropout layers for regularisation. It has two output heads: one uses a sigmoid function for real/fake discrimination, and the other uses softmax for class label prediction.

Once the AC-GAN model was adequately trained and its outputs were visually verified, the required number of synthetic images for each minority class was generated and merged with the real training data. This resulted in a final, balanced dataset with all lesion categories equally represented with $N_{\max} = 6705$ samples per class. In addition to addressing the imbalance issue, the AC-GAN introduced complex, high- fidelity variations that enhanced the robustness and gener- alisation capability of the downstream classification model. The parameters of AC-GAN during training are shown in Table I. The number of samples of each class before and after augmentation is shown in Table II. Architecture of the proposed customised AC-GAN augmentation techniques is shown in Figure 3.

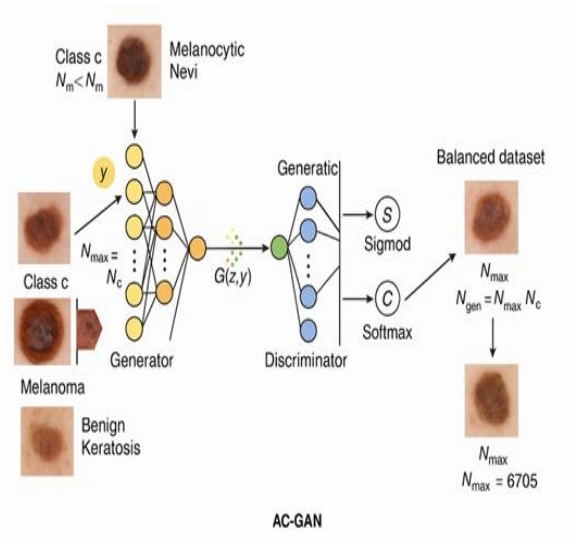

Fig. 3. Proposed AC-GAN schema

TABLE I. TRAINING PARAMETERS OF THE AC-GAN MODEL

Parameter	Value
Optimizer	Adam
Learning Rate	0.0002
Beta1 (Adam)	0.5
Beta2 (Adam)	0.999
Batch Size	64
Epochs	200
Latent Vector Dimension (z)	100
Image Resolution	$224 \times 224 \times 3$
Activation (Generator)	ReLU (intermediate), Tanh (output)
Activation (Discriminator)	LeakyReLU (slope = 0.2)
Loss Functions	Categorical Cross-Entropy (class)
Weight Initialization	Normal distribution ($\mu = 0$, $\sigma = 0.02$)
Dropout Rate (Discriminator)	0.3
Conditioning Method	Concatenation of z and one-hot class label

TABLE II. NUMBER OF SAMPLES PER CLASS BEFORE AND AFTER AC-GAN AUGMENTATION

Class	Before Augmentation	After Augmentation
Melanocytic Nevi (nv)	6,705	6,705
Melanoma (mel)	1,113	6,705
Benign Keratosis (bkl)	1,099	6,705
Basal Cell Carcinoma (bcc)	514	6,705
Actinic Keratoses (akiec)	327	6,705
Vascular Lesions (vasc)	142	6,705
Dermatofibroma (df)	115	6,705

C. Dataset Splitting and Experimental Setup

The fully balanced HAM10000 dataset, which consists of N images after AC-GAN augmentation,

was divided into training and test subsets in an 80:20 ratio to assess our classification system thoroughly. In particular, we calculated

$N_{\text{train}} = 0.8\,N$, $N_{\text{test}} = N - N_{\text{train}}$.

To ensure that the proportion of each class in the training and test sets matched that of the entire dataset, this split was implemented using the train_test_split function from the scikit-learn library with the parameters test_size=0.2, shuffle=True, and stratify=labels. A fixed random seed (random_state=42) ensured reproducibility between tests. We reduced sampling bias and made it possible to reliably evaluate the model's generalization performance on unseen dermoscopic images by maintaining class frequencies via stratification and rearranging the data before splitting.

IV. PROPOSED MODEL

This work uses the HAM10000 dermatoscopic image dataset to propose a novel model called Faster-CAN (Faster Convolutional Attention Network) integrated with Federated Learning (FL) for automated and privacy-preserving skin cancer classification. The suggested model improves classification accuracy and facilitates effective feature selection by combining lightweight convolutional encoders with channel and spatial attention methods. Federated learning is used to disperse the learning process without allowing customers (such as clinics or hospitals) to share raw data to protect patient data privacy. Figure 4 shows the proposed model architecture.

The suggested Faster-CAN model requires convolutional feature encoding to extract hierarchical visual patterns from dermatoscopic images. Low-level information like edges and textures is captured, and it gradually picks up more intricate

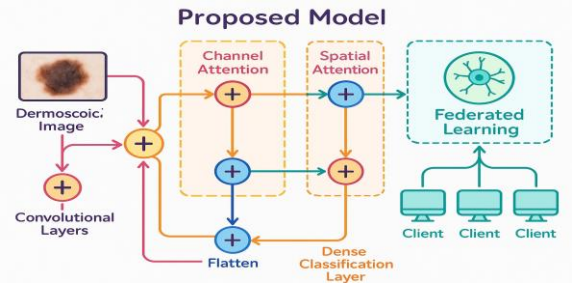

Fig. 4. Proposed model architecture

representations like the shape and structure of lesions. Con- volutional layers improve learning effectiveness and general- isation by maintaining spatial links and permitting parameter sharing. The most informative areas of skin lesions are the focus of later attention modules, which are built upon these stored traits. For a reliable and precise classification of skin cancer, this encoding procedure is essential.

With H, W, and C standing for the image height, width, and number of channels, respectively, let $X \in R^{H\times W\times C}$ rep- resent an input dermatoscopic image. A series of convolutional layers are applied to the input in order to extract hierarchical low- to high-level features.

The following transformation is carried out by each layer:

$$F^{(l)} = \text{ReLU}\,(\text{BN}\,(\text{Conv}\,(F^{(l-1)}; W^{(l)})))$$

where BN stands for batch normalization, ReLU is the rectified linear unit activation function, and $F^{(0)} = X$. Here, $W^{(l)}$ indicates the learnable weights of the l-th convolutional layer. Training stability and convergence are enhanced by this design.

After convolutional feature extraction, the Channel At- tention Module (CAM) is included to improve the model's capacity to concentrate on the most informative feature maps across channels [27]. Not every channel makes an equivalent contribution to the final classification task, even when convo- lutional layers encode a large set of data. By simulating inter- channel interactions, CAM dynamically reweights the channel- wise feature responses, enabling the network to highlight more pertinent features essential for skin cancer detection, such as colour variations, lesion borders, and pigmentation patterns.

We implement a channel attention module inspired by the CBAM framework to prioritize informative channels that contribute significantly to lesion classification.

Global max pooling (GMP) and global average pooling (GAP) are first carried out over the spatial dimensions:

$$F_{\text{avg}} = \frac{1}{H'W'}\sum_{i=1}^{H'}\sum_{j=1}^{W'} F_{i,j,c}, \quad F_{\max} = \max_{i=1,\ldots,H'}\max_{j=1,\ldots,W'} F_{i,j,c}$$

To determine the inter-channel relationship, these two feature descriptors are fed into a shared multi-layer perceptron

(MLP) with a bottleneck (reduction ratio r):

$$M_c = \sigma(W_2 \cdot \mathrm{ReLU}(W_1 \cdot F_{\mathrm{avg}}) + W_2 \cdot \mathrm{ReLU}(W_1 \cdot F_{\max}))$$

where $W_1 \in R^{C'/r \times C'}$, $W_2 \in R^{C' \times C'/r}$, and σ is the sigmoid function.

The refined feature map by channel is: $F' = Mc \cdot F$

The Spatial Attention Module (SAM) is intended to highlight" where" the most pertinent information is found within the feature map, whereas the CAM concentrates on "what" useful features to stress across channels. The lesion's location and shape are crucial in distinguishing between benign and ma-lignant instances in dermatoscopic pictures. SAM improves the feature representation by giving spatial regions with prominent lesion structures, like irregular borders, pigmentation clusters, or asymmetrical patterns, higher attention weights [28]. SAM improves classification accuracy and resilience by integrating spatial dependencies, which help the model concentrate on the most discriminative regions. This is especially useful when background noise or irrelevant skin texture could normally cause the model to become confused.

We compute two spatial descriptors using average and max over the channel dimension, pooling is applied as follows:

$$F^s_{\mathrm{avg}}(i,j) = \frac{1}{C'} \sum_{k=1}^{C'} F'_{i,j,k}, \quad F^s_{\max}(i,j) = \max_{k=1,\ldots,C'} F'_{i,j,k}$$

These two spatial descriptors are then concatenated and passed through a 7×7 convolutional layer:

$$M_s = \sigma\left(\mathrm{Conv}_{7\times7}\left(F^s_{\mathrm{avg}}; F^s_{\max}\right)\right)$$

The spatial attention map Ms is then applied to the feature map from the previous channel attention stage:

$$F'' = M_s \cdot F'$$

The model can filter background noise and flexibly focus on pertinent semantic regions by this dual attention technique. This dual attention mechanism allows the model to focus on relevant semantic regions and suppress background noise dynamically.

After being flattened, the attention-enhanced feature map F'' is passed through a dense classification layer:

$$\hat{y} = \mathrm{Softmax}(z), \qquad z = W_f \cdot \mathrm{Flatten}(F'') + b_f$$

In this case, the projected class probabilities over the C = 7 skin cancer categories in HAM10000 are represented by $\hat{y} \in RC$.

The model is trained using the categorical cross-entropy loss function:

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{C} y_{ij} \log(\hat{y}_{ij})$$

where $\hat{y}_{ij}$ is the predicted probability for class j of the i-th image, and $y_{ij} \in \{0, 1\}$ is the corresponding ground truth label.

Federated Learning (FL) is used to train the suggested architecture in order to safeguard private patient data. Let $n = \sum_{k=1}^{K} n_k$, where each of the K participating clients has a local dataset D_k of size n_k.

At communication round t, each client trains the model locally and transmits updated weights w^t to a central server.

The server uses the Federated Averaging (FedAvg) algorithm to aggregate the weights:

$$w^{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w^t_k$$

Stochastic gradient descent is used by clients to update their weights:

$$w^{t+1}_k = w^t_k - \eta L_k(w^t_k)$$

where Lk(wt) is the gradient of the local loss computed on dataset Dk, and η is the learning rate.

The use of convolutional blocks makes hierarchical feature extraction efficient. The network can prioritise critical lesion sites and reduce noise in the channel and spatial attention mod- ules, which also improves model interpretability. FL lessens overfitting to the data of any one institution by ensuring pri- vacy, equity, and collaborative learning across several clinical sites.

Algorithm 1 Federated Faster-CAN for Skin Cancer Classification

Input: HAM10000 dataset split across K clients as D_k, initial global model weights w^0, learning rate η, number of communication rounds T, number of local epochs E, batch size B

Output: Trained global model w^T

Initialize global model w^0 (Faster-CAN) on server
for each communication round $t = 0$ to $T - 1$ **do**
  Server: Broadcast current model weights w^t to all clients
  for all clients $k = 1$ to K **in parallel do**
    Receive w^t
    for each local epoch $e = 1$ to E **do**
      for each mini-batch $(x, y) \in D_k$ **do**
        Forward Pass:
        Apply convolutional layers to extract $F^{(l)}$
        Apply **Channel Attention Module (CAM)**:
        Compute F_{avg}, F_{max}
        Compute M_c via shared MLP
        Refine: $F' = M_c \cdot F$
        Apply **Spatial Attention Module (SAM)**:
        Compute F^s_{avg}, F^s_{max}
        Convolve to get M_s
        Refine: $F'' = M_s \cdot F'$
        Flatten F'' and apply dense layer
        Compute predicted probabilities $\hat{y}$
        Compute cross-entropy loss L_k
        Update weights: $w^{t+1}_k = w^t_k - \eta \nabla L_k (w^t_k)$
      end for
    end for
    Send updated weights w^{t+1}_k to server
  end for
  Server: Aggregate updates via Federated Averaging:

$$w^{t+1} = \sum_{k=1}^{K} \frac{n_k}{n} w^{t+1}_k$$

end for

All things considered, the suggested Faster-CAN with FL strikes a good balance between precision, interpretability, effectiveness, and data security, which qualifies it for practi- cal implementation in dermatology clinical decision support systems. The proposed model parameters are shown in Table III.

V. RESULT AND DISCUSSION

The study uses a multifaceted assessment methodology to fully examine the robustness and performance of the suggested model. The effect of AC-GAN data balancing approaches on model accuracy, recall, precision, and generalization is one overfitting or underfitting in order to evaluate the model's gen- eralization capabilities. Comprehensive classification reports are produced for the suggested hybrid model and a number of benchmark models after these diagnostics. These reports give a comprehensive picture of model performance by include specific measures like accuracy, precision, recall, F1-score, Cohen's Kappa, and Matthews Correlation Coefficient (MCC). In order to visually evaluate the model's strengths and flaws, a number of graphical evaluations are also provided, including metric comparison bar charts, ROC curves, and confusion matrices.Performance measures are calculated for each class to evaluate bias and fairness, enabling a thorough analysis of the model's treatment of different categories. Particularly with datasets that are unbalanced, this phase is essential for detecting any systemic bias.Using the TensorFlow and PyTorch deep learning frameworks for model generation and training, the complete model development and assessment pipeline is implemented in Python. Essential Python modules like NumPy, Pandas, and Matplotlib are used for data preprocessing, ma- nipulation, and visualization. Throughout the research process, this integrated development environment guarantees scalability, reproducibility, and adaptability. Data augmentation techniques of the research are shown in Table IV.

At first, the accuracy of the baseline CNN model was 87.12%. Its performance increased to 91.45% after adding AC-GAN, indicating a discernible improvement brought about by improved class representation in the training set. Similarly, after AC-GAN, the deep convolutional model ResNet50 increased from 89.34% to 94.21%, demonstrating its capacity to better utilize more balanced input.

The improvement from 90.58% to 95.63% for DenseNet121, renowned for its dense connections and feature reuse, suggests that models with deeper and more intricate architectures typically gain a great deal from the synthetic data augmentation offered by AC-GAN.

Above all, the suggested model showed a significant im- provement. Before using AC-GAN, it already outperformed the baseline models with an accuracy of 93.65%. Following the incorporation of samples generated by AC-GAN, the model's accuracy was an impressive 98.42%. This demonstrates the important factor examined. This guarantees that the model operates consistently across all categories and does not show bias toward the majority classes. Additionally, training and validation curves are monitored to look for indications of strength of the suggested architecture as well as how well AC-GAN addresses class imbalance and improves deep learning models' capacity for generalization.

Overall, the experimental findings show that using AC- GAN greatly improves the accuracy and fairness of more complex, custom-built models and the performance of con- ventional deep learning architectures. This emphasizes how crucial synthetic data creation methods are in fields where data imbalance is a significant problem.

A. Model generlization analysis

The performance of different models, such as their training, testing, and validation accuracy, is shown in Table V.

Following the application of AC-GAN for data augmentation, the performance of four distinct models—CNN, ResNet50, DenseNet121, and the suggested hybrid model—was assessed based on training, validation, and testing accuracy. The ultimate performance metric for unseen data is testing accuracy. The baseline models' greatest testing accuracy was 95.63% for DenseNet121, 94.21% for ResNet50, and 91.45% for CNN.With training accuracy of 98.84%, validation accuracy of 98.11%, and testing accuracy of 98.42%, the suggested Faster-CAN model performed noticeably better than the others. This demonstrates unequivocally how the proposed architecture is superior in generalization and learning capacity, making it ideal for the categorization challenge. The performance improvement seen in all models following the use of AC-GAN validates the advantages of class-balancing and the creation of synthetic data in enhancing neural networks' learning dynamics and decision boundaries. Insightful trends in model fitting behavior are revealed by closely examining the training, validation, and testing accuracies. The CNN model demonstrated 92.35% training accuracy, 90.14% validation accuracy, and 91.45% testing accuracy. Although the CNN model's lower overall performance reveals limits in catching deeper patterns in the data, the little difference between training and validation/testing suggests that the model is reasonably well-fitted.With ResNet50 at 94.20% (training) vs. 93.01% (validation) and DenseNet121 at 95.10% (training) vs. 94.28% (validation), both models demonstrated less variation between training and validation accuracy. Both models are well-fitted and not substantially overfitted, as these small gaps show. Their capacity to generalize to unknown data is further supported by improved test accuracy. With closely comparable validation (98.11%) and testing (98.42%) accuracies, the suggested model showed an extraordinarily high training accuracy (98.84%). The slight variation between these phases attests to the model's high generalization and avoidance of overfitting. This consistency throughout the evaluation process shows that, even with highly enriched synthetic data from AC-GAN, the suggested model is strong at learning from the training data and resistant to overfitting. The proposed model's training and validation accuracy and training and validation loss are shown in Figure 5,6.

TABLE III. PROPOSED FASTER-CAN MODEL PARAMETERS

Parameter	Value / Description
Input Image Size	$224 \times 224 \times 3$
Convolution Layers	Multiple layers with BatchNorm and ReLU
Channel Attention Module	Global Average Pooling (GAP), Global Max Pooling (GMP), MLP with bottleneck ratio $r = 16$
Spatial Attention Module	7×7 convolution over concatenated spatial descriptors
Attention-enhanced Feature Map	$F'' \in R^{H' \times W' \times C'}$
Pooling for Classification	Global Average Pooling (GAP)
Classifier	Dense Layer + Softmax
Number of Classes	$C = 7$ (HAM10000 categories)
Loss Function	Categorical Cross-Entropy
Learning Rate η	0.01 (can be adjusted)
Federated Learning Algorithm	FedAvg (Federated Averaging)
Number of Clients K	Variable, depending on deployment
Local Dataset Size n_k	Varies per client
Communication Rounds t	Typically 100–500

TABLE IV. COMPARISON OF MODEL ACCURACY BEFORE AND AFTER APPLYING AC-GAN

Model	Accuracy Before AC-GAN (%)	Accuracy After AC-GAN (%)
CNN	87.12	91.45
ResNet50	89.34	94.21
DenseNet121	90.58	95.63
Proposed Model	93.65	98.42

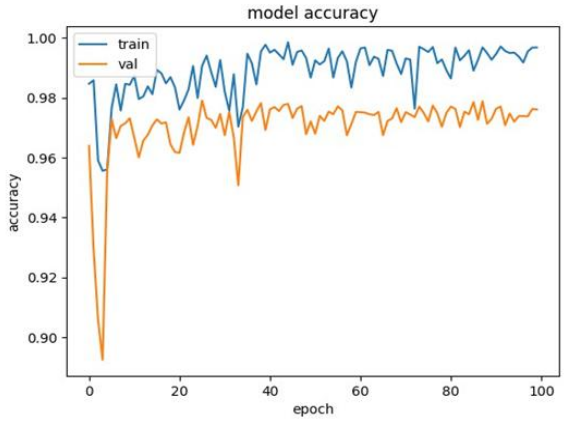

Fig. 5. Training and validation accuracy

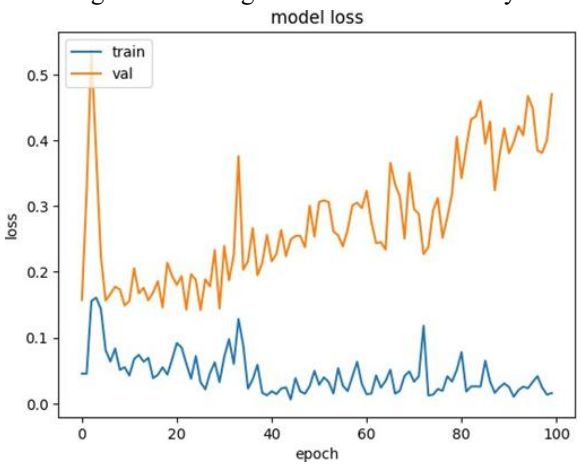

Fig. 6. Proposed model architecture

TABLE V. TRAINING, TESTING, AND VALIDATION ACCURACY OF THE PROPOSED MODELS AND OTHERS

Model	Training Accuracy (%)	Validation Accuracy (%)	Testing Accuracy (%)
CNN	92.35	90.14	91.45
ResNet50	94.20	93.01	94.21
DenseNet121	95.10	94.28	95.63
Proposed Model	98.84	98.11	98.42

TABLE VI. CLASSIFICATION REPORT OF DIFFERENT MODELS AFTER APPLYING AC-GAN

Model	Precision (%)	Recall (%)	F1-score (%)
CNN	90.76	91.21	90.98
ResNet50	93.85	94.05	93.95
DenseNet121	95.40	95.80	95.60
Proposed Model	98.50	98.31	98.40

B. Classification report and agreement analysis

Additionally, the suggested model is assessed using com- mon classification metrics such as accuracy, precision, recall, and F1-score. Table VI, which highlights the classification report received after applying AC-GAN for class balance and data augmentation, provides a detailed performance comparison among all models. F1-score of 90.98% and recall of 91.21%. These findings show that, despite AC-GAN support, the model occasionally had trouble with misclassifications, especially when processing cases of minority classes, even if it was able to capture broad patterns.

With a precision of 93.85%, recall of 94.05%, and F1-score of 93.95%, ResNet50 showed enhanced performance. Better feature extraction made possible by the deeper architecture resulted in more precise class predictions. A favorable balance between erroneous positives and false negatives is also indicated by the narrow precision-recall gap. DenseNet121's dense connectivity allowed for effective feature reuse, outperforming both CNN and ResNet50. Its F1-score was 95.60%, its preci- sion was 95.40%, and its recall was 95.80%. This consistency in metrics demonstrates that DenseNet121 maintained robust- ness across the dataset's different class distributions in addition to generalizing successfully. The suggested Faster-CAN model performed the best of all the assessed designs. It achieved a remarkable F1-score of 98.40%, recall of 98.31%, and precision of 98.50%. These findings demonstrate the model's exceptional efficacy in accurately determining class labels with few misclassifications. The model appears to be neither biased toward over-prediction nor under-prediction, based on the precision and recall metrics' extremely near agreement. The model's capacity to efficiently manage precision and recall is further validated by its high F1-score, which makes it the most dependable option for practical implementation.

TABLE VII. KAPPA AND MCCC PERFORMANCE

Model	Kappa	MCC
CNN	0.878	0.881
ResNet50	0.917	0.920
DenseNet121	0.940	0.943
Proposed Model	0.981	0.982

As indicated in Table VII, further analysis was carried out utilizing Cohen's Kappa and the

Matthews Correlation Coefficient (MCC) to supplement the conventional classifica- tion metrics. These metrics are especially useful for assessing performance on unbalanced datasets and offer greater insights into the agreement between anticipated and true labels.

Despite obtaining respectable classification metrics, the CNN model showed fair agreement and correlation with a Kappa score of 0.878 and an MCC of 0.881. These num- bers imply that while CNN has a modest level of gen- eralization, it may occasionally misclassify, particularly in minority classes. With a Kappa of 0.917 and an MCC of 0.920, ResNet50 demonstrated superior reliability and stronger predictive alignment. With a Kappa of 0.940 and an MCC of 0.943, DenseNet121 significantly enhanced these results, which is in line with its higher accuracy and F1-score. This demonstrates that DenseNet121 performs consistently across class distributions in addition to being accurate.

Notably, the suggested model obtained the greatest scores on both measures (MCC of 0.982 and Kappa of 0.981). With low chance agreement and a strong connection between the predicted and actual class labels, these nearly flawless scores demonstrate remarkable consistency and durability.

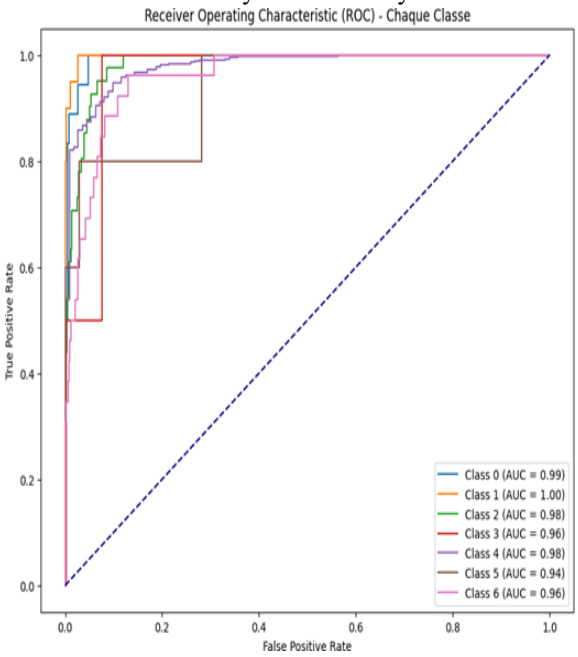

Fig. 7. ROC curve AUC values of proposed model

The Area Under the Receiver Operating Characteristic Curve (AUC), a reliable metric that measures the model's capacity to distinguish between positive and negative instances independent of the chosen threshold, was used to assess the model's discriminative performance to supplement the agreement analysis (Figure 7). The suggested model's strong discriminative capacity was further supported by the ROC- AUC analysis, which showed that it consistently obtained high AUC scores across all seven classes in the multi-class classification test. In particular, Class 1 obtained a perfect score of 1.00, suggesting perfect separability between positive and negative cases for that class, but Class 0 obtained an AUC of 0.99. The model's remarkable ability to differentiate between Class 2 and Class 4 is demonstrated by their respective AUCs of 0.98. Additionally, Class 3 and Class 6 both received AUC scores of 0.96, highlighting the model's strong predictive capabilities. Class 5 achieved a high AUC value of 0.94, which is well within the range regarded as excellent in medical imaging classification tasks, albeit being somewhat lower.

These AUC results offer further proof of the model's diag- nostic reliability and confirm the conclusions of the agreement- based evaluation.

C. Discussion

To improve the classification model's training, the sug- gested study's main goal was to generate synthetic data using the AC-GAN model. As demonstrated by the earlier findings, a significant improvement in model performance was noted after data augmentation approaches were applied. Strong general- isation capacity was indicated by the model's well-fitted be- haviour in both the training and validation stages. Furthermore, the categorisation performance held up well across all evalu- ation metrics. High Kappa and MCC scores from agreement analysis further supported the model's dependability. Figure 8 thoroughly summarises the model's assessment. Class-wise performance analysis was carried out to evaluate potential bias and guarantee balanced learning. Table VIII offers a comprehensive analysis of this class-wise assessment.

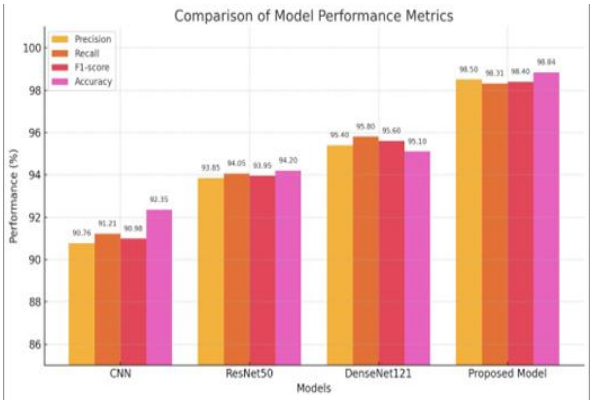

Fig. 8. Summary of all models performance

TABLE VIII. CLASSIFICATION REPORT OF THE PROPOSED MODEL (CLASS-WISE

Class	Precision	Recall	F1-Score
0	1.00	1.00	1.00
1	0.99	1.00	0.99
2	0.94	0.97	0.96
3	1.00	1.00	1.00
4	0.97	0.85	0.91
5	1.00	1.00	1.00
6	0.91	0.98	0.95

Class-wise precision, recall, and F1-score values were analysed to find any differences in the model's capacity to categorise particular classes in order to evaluate model bias. According to the analysis, the model produces highly accurate and consistent predictions for Classes 0, 3, and 5 with excellent precision, recall, and F1-scores (1.00). Class 1 performs almost flawlessly, with a recall of 1.00 and an accuracy of 0.99, indicating even fewer categorisation errors. Even though they are still doing well, Classes 2, 4, and 6 have somewhat lower scores, particularly in recall (0.85) and precision (0.91). With a recall of 0.85, Class 4 stands out in particular. This means that
15% of real Class 4 cases were incorrectly classified, which may indicate a slight underrepresentation or confusion with other classes. Class 6's somewhat lower precision of 0.91 also raises the possibility of some false positives. Despite these little variances, all F1-scores (≥ 0.91) show a good balance
between recall and precision in every class. This implies that
there is no discernible bias towards or against any specific class and that the model is generally well-calibrated. Overall, the model shows equitable performance across the class spectrum, while the relatively minor decline in scores for a few classes points to areas that could use improvement, such as targeted augmentation or class rebalancing.A sample output of the proposed model is shown in Figure 1, where the original class is melanoma (MEL) and the predicted class matches it. The model's confidence score during prediction is 100%.

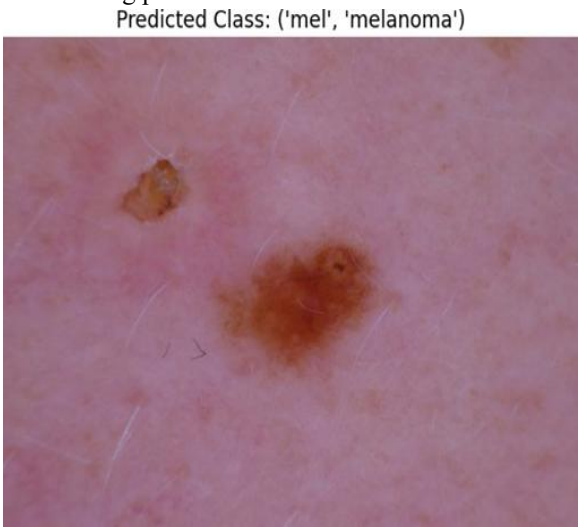

Fig. 9. Sample output of the model

VI. CONCLUSION

This study introduces a new method for classifying skin cancer by combining the suggested Faster-CAN model, a faster convolutional neural network architecture, with AC- GAN-based synthetic data creation. The training process and overall model performance are improved by the practical usage of AC-GAN, which tackles the problems of class imbalance and limited data. Experimental findings show that Faster-CAN outperforms baseline models in accuracy, precision, recall, and F1-score while achieving high agreement metrics (MCC and Cohen's Kappa). The model reduces the likelihood of bias and misclassification by maintaining balanced performance across all skin lesion categories, as confirmed by the class-wise study. The results demonstrate how cutting-edge data augmentation methods and well-designed hybrid architectures can create accurate and effective skin cancer diagnosis systems.In order to further enhance model generalizability, future studies can concentrate on growing the dataset by adding a variety of skin

lesion photos from various sources. Furthermore, inves- tigating the integration of multi-modal data—for example, dermoscopic pictures paired with patient metadata—may im- prove the accuracy of classification. Performance may also be improved by fine-tuning the AC-GAN design to produce even more varied and realistic synthetic examples, particularly for underrepresented classes. Lastly, evaluating the Faster-CAN model's practicality and directing advancements for scalable skin cancer diagnostic systems will require its deployment and validation in actual clinical settings.

REFERENCES

[1] C. Xia, X. Dong, H. Li, M. Cao, D. Sun, S. He, F. Yang, X. Yan, S. Zhang, N. Li, et al., "Cancer statistics in china and united states, 2022: profiles, trends, and determinants," Chinese medical journal, vol. 135, no. 05, pp. 584–590, 2022.

[2] F. Bray, M. Laversanne, B. Cao, C. Varghese, B. Mikkelsen, E. Weider- pass, and I. Soerjomataram, "Comparing cancer and cardiovascular dis- ease trends in 20 middle-or high-income countries 2000–19: A pointer to national trajectories towards achieving sustainable development goal target 3.4," Cancer Treatment Reviews, vol. 100, p. 102290, 2021.

[3] N. A. AlSadhan, S. A. Alamri, M. M. Ben Ismail, and O. Bchir, "Skin cancer recognition using unified deep convolutional neural networks," Cancers, vol. 16, no. 7, p. 1246, 2024.

[4] M. Sardar, M. M. Niazi, F. Nasim, et al., "Ensemble deep learning methods for detecting skin cancer," Bulletin of Business and Economics (BBE), vol. 13, no. 1, 2024.

[5] A. A. Adegun and S. Viriri, "Fcn-based densenet framework for automated detection and classification of skin lesions in dermoscopy images," IEEE Access, vol. 8, pp. 150377–150396, 2020.

[6] D. Adla, G. V. R. Reddy, P. Nayak, and G. Karuna, "A full-resolution convolutional network with a dynamic graph cut algorithm for skin can- cer classification and detection," Healthcare Analytics, vol. 3, p. 100154, 2023.

[7] R. R. Kumar, R. Varun, J. Sreeshwan, K. A. Kumar, U. Rana, and A. Rajyalakshmi, "Feasible skin lesion detection using cnn and rnn," in E3S Web of Conferences, vol. 430, p. 01050, EDP Sciences, 2023.

[8] M. A. Albahar, "Skin lesion classification using convolutional neural network with novel regularizer," IEEE Access, vol. 7, pp. 38306–38313, 2019.

[9] J. Dong and A. P. Thrift, "Alcohol, smoking and risk of oesophago- gastric cancer," Best practice & research Clinical gastroenterology, vol. 31, no. 5, pp. 509–517, 2017.

[10] M. Goyal, T. Knackstedt, S. Yan, and S. Hassanpour, "Artificial intelligence-based image classification methods for diagnosis of skin cancer: Challenges and opportunities," Computers in biology and medicine, vol. 127, p. 104065, 2020.

[11] M. Alsaidi, M. T. Jan, A. Altaher, H. Zhuang, and X. Zhu, "Tackling the class imbalanced dermoscopic image classification using data augmentation and gan," Multimedia Tools and Applications, vol. 83, no. 16, pp. 49121–49147, 2024.

[12] M. H. Li, Y. Yu, H. Wei, and T. O. Chan, "Classification of the qilou (arcade building) using a robust image processing framework based on the faster r-cnn with resnet50," Journal of Asian Architecture and Building Engineering, vol. 23, no. 2, pp. 595–612, 2024.

[13] H. Guan, P.-T. Yap, A. Bozoki, and M. Liu, "Federated learning for medical image analysis: A survey," Pattern Recognition, p. 110424, 2024.

[14] K. Monica, J. Shreeharsha, P. Falkowski-Gilski, B. Falkowska-Gilska, M. Awasthy, and R. Phadke, "Melanoma skin cancer detection using mask-rcnn with modified gru model," Frontiers in Physiology, vol. 14, p. 1324042, 2024.

[15] A. Rajesh, K. N. Rao, G. N. V. Sai, K. D. Kumar, and K. R. S. Karthik, "Skin cancer detection and intensity analysis using deep learning," in 2024 International Conference on Emerging Systems and Intelligent Computing (ESIC), pp. 376–381, IEEE, 2024.

[16] A. Naeem, T. Anees, M. Khalil, K. Zahra, R. A. Naqvi, and S.-W. Lee, "Snc net: Skin cancer detection by integrating handcrafted and deep learning-based features using dermoscopy

images," Mathematics, vol. 12, no. 7, p. 1030, 2024.

[17] A. Naeem and T. Anees, "Dvfnet: A deep feature fusion-based model for the multiclassification of skin cancer utilizing dermoscopy images," Plos one, vol. 19, no. 3, p. e0297667, 2024.

[18] I. Kousis, I. Perikos, I. Hatzilygeroudis, and M. Virvou, "Deep learning methods for accurate skin cancer recognition and mobile application," Electronics, vol. 11, no. 9, p. 1294, 2022.

[19] M. S. Sivakumar, L. M. Leo, T. Gurumekala, V. Sindhu, and A. S. Priyadharshini, "Deep learning in skin lesion analysis for malignant melanoma cancer identification," Multimedia Tools and Applications, vol. 83, no. 6, pp. 17833–17853, 2024.

[20] M. T. Campos, F. A. Silva, J. R. Fernandes, S. G. Santos, F. D. Magalha˜es, M. J. Oliveira, and A. M. Pinto, "New mos2/tegafur- containing pharmaceutical formulations for selective led-based skin cancer photo-chemotherapy," Pharmaceutics, vol. 16, no. 3, p. 360, 2024.

[21] N. A. AlSadhan, S. A. Alamri, M. M. Ben Ismail, and O. Bchir, "Skin cancer recognition using unified deep convolutional neural networks," Cancers, vol. 16, no. 7, p. 1246, 2024.

[22] M. A. Rahman, E. Bazgir, S. S. Hossain, M. Maniruzzaman, et al., "Skin cancer classification using nasnet," International Journal of Science and Research Archive, vol. 11, no. 1, pp. 775–785, 2024.

[23] M. Pandiyan, S. Sivakumar, and S. R. Nayak, "Skin cancer classifi- cation using convolutional neural network with dwt features," in 2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 309–313, IEEE, 2024.

[24] S. Mushtaq and O. Singh, "A deep learning based architecture for multi- class skin cancer classification," Multimedia Tools and Applications, vol. 83, no. 39, pp. 87105–87127, 2024.

[25] K. T. Ahmed, F. Rustam, A. Mehmood, I. Ashraf, and G. S. Choi, "Predicting skin cancer melanoma using stacked convolutional neural networks model," Multimedia Tools and Applications, vol. 83, no. 4, pp. 9503–9522, 2024.

[26] A. Khan, N. Chandran, and D. Gangodkar, "Integration of multimodal data sources for enhanced skin disease classification and c̄ancer predic- tion: A study leveraging pre-trained models on ham 10000 metadata and squamous cell carcinoma (scc) images.," 2024.

[27] N. N. I. Prova, "Enhancing agricultural research with an attention- based hybrid model for precise classification of rice varieties," Authorea Preprints, 2025.

[28] N. N. I. Prova, "Enhancing agricultural research with an attention- based hybrid model for precise classification of rice varieties," Authorea Preprints, 2025.