International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(4s) 2026 pp. 19-36
https://doi.org/10.70917/ijcisim-2026-2118
Article

An Efficient Attention-Guided Xception Framework for Robust Multiclass Skin Lesion Classification

Shamira Arshad Shaikh[1], Yogita Mistry[2]

[1] Department of Electronics and Telecommunication Engineering, Ramrao Adik Institute of Technology (RAIT), Nerul, Navi Mumbai, Maharashtra, India.
Email: syed.sameera222@gmail.com
[2] Department of Information Technology Engineering, Ramrao Adik Institute of Technology (RAIT), Nerul, Navi Mumbai, Maharashtra, India.
Email: yogita.mistry@rait.ac.in
Corresponding Author: Shamira Arshad Shaikh, syed.sameera222@gmail.com

Abstract: - Skin cancer is one of the most common and potentially deadly diseases worldwide. Early and accurate diagnosis of skin cancer is very important for better outcomes for the patient and a reduction in mortality rates. However, the visual similarity of different skin lesion categories and the class imbalance are major challenges for automated skin lesion classification. This study proposes an Attention-Guided Xception framework for robust multiclass skin lesion classification. The proposed framework leverages transfer learning based on the Xception architecture for discriminative dermoscopic feature extraction, and an attention mechanism to enhance the representation of clinically relevant lesion regions and suppress irrelevant background information. The model was trained on the ISIC skin lesion dataset, with melanoma samples removed during dataset refinement to create an eight-class non-melanoma classification framework. Data preprocessing and class balancing techniques were used in order to improve the stability of training and the classification performance. Accuracy, precision, recall, F1-score and confusion matrix analysis were used to evaluate the proposed framework. The experimental results show that the proposed Attention-Guided Xception model achieved a classification accuracy of 93 % and outperformed CNN, VGG16, EfficientNet-B2 and Xception architectures. The results demonstrate the benefit of attention-guided feature learning in improving lesion discrimination and reducing inter-class ambiguity. The proposed framework provides an accurate and computationally efficient way to automate skin lesion classification, and it demonstrates its capability to support computer-aided dermatological diagnosis and early disease detection.
Keywords: - Skin Lesion Classification; Attention-Guided Xception; Deep Learning; Multiclass Classification; Dermoscopic Images.

1. Introduction

Skin cancer is one of the most common malignancies globally and remains a serious public health problem with rising incidence and mortality rates. It is mainly composed of basal cell carcinoma (BCC), squamous cell carcinoma (SCC) and melanoma, with distinct clinical features and prognosis [1]. Among these, melanoma is considered the most aggressive with a high metastatic potential and mortality [2]. Contemporary research on global cancer statistics has shown that the incidence of skin cancer has been steadily increasing in the last decade, thus emphasizing the need for early and accurate diagnosis to improve patient survival rates [3].

Early detection of skin lesions is important to slow down the progression of the disease and improve treatment outcomes. Studies have shown that early diagnosis improves survival rates and reduces the healthcare burden [4]. However, interpretation of dermoscopic images is still difficult due to the visual similarity between lesion categories, variability in color and texture, and inter-observer variability among dermatologists. Therefore, the development of reliable computer-aided diagnostic systems has become an active research area in medical image analysis.

Emerging developments in artificial intelligence and deep learning have shown remarkable success in the automated analysis of skin lesions. Convolutional Neural Networks (CNNs) have demonstrated a better ability to learn hierarchical image features directly from dermoscopic images. However, conventional deep learning models suffer from class imbalance, limited interpretability, and a lack of focus on clinically relevant lesion regions. Such drawbacks encourage the creation of robust and explainable frameworks that can provide accurate multiclass classification while facilitating clinical decision making [5].

State-of-the-art methods have achieved important advances for skin lesion classification using CNN-based models, transformer architectures, explainable AI frameworks, and attention-guided networks. These approaches have shown promising classification results and better feature extraction ability. Nevertheless, the literature still has several obvious limitations. Many studies are mainly focused on binary classification tasks, and others are limited to a small number of lesion categories. Besides, problems such as class imbalance, complexity of calculation, poor interpretability and generalization on different datasets still influence the reliability of models. These observations highlight the need for a robust and computationally efficient framework that can accurately classify multiple skin lesion categories, while also capturing clinically relevant lesion characteristics. Motivated by the above challenges, in this study, an attention-guided Xception framework is proposed for multiclass skin lesion classification. [6-20]

Hence, this study proposes an Attention-Guided Xception framework for robust multiclass skin lesion classification. The proposed framework combines efficient deep feature extraction, attention-guided feature refinement, and class balancing techniques to improve classification performance across multiple skin lesion categories. The experimental evaluation on the ISIC dataset shows that the proposed framework is effective for improving the classification performance and assisting the computer-aided dermatological diagnosis. The goals of the current study are as follows:

• To design an Attention-Guided Xception model for automated classification of skin lesions in multiple classes with the help of dermoscopic images.

• To promote feature discrimination with the help of attention mechanisms that highlight meaningful clinical areas within the skin lesion images.

• To utilize the technique of transfer learning with the help of the Xception framework for efficient extraction of hierarchical features from skin lesion images.

• To investigate the effectiveness of SMOTE-based data balancing in reducing class imbalance and improving classification performance across minority lesion categories.

• To measure the efficacy of the designed Attention-Guided Xception model based on various classification metrics, such as accuracy, precision, recall, F1-score, etc.

• To benchmark the designed attention-gated Xception model against traditional deep learning models and prove its efficacy in skin lesion classification.

• To contribute to designing efficient computer-aided diagnosis systems for skin cancer detection.

The rest of this paper is organized as follows. Section 2 presents a detailed review of previous works on skin lesion classification using machine learning and deep learning methods. It discusses convolutional neural networks, attention-based architectures, and recent advancements in automated skin lesion analysis and points out the research gaps addressed in this study. Section 3 presents the dataset, preprocessing techniques, data balancing strategy and the proposed Attention-Guided Xception framework for multiclass skin lesion classification. Section 4 presents the experimental results, comparative performance analysis, and evaluation using standard classification metrics, including accuracy, precision, recall, F1-score, and confusion matrix analysis. Section 5 concludes with a summary of the main findings and contributions of the study, as well as future research directions to improve model generalization, clinical applicability, and real-world deployment of intelligent systems for dermatological diagnosis.

2. Related Work

Riaz et al. proposed a holistic joint learning framework that integrates Convolutional Neural Networks (CNN) and Local Binary Patterns (LBP) for multiclass skin cancer classification from the HAM10000 dataset. The hybrid approach successfully fused deep and handcrafted features to enhance the representation of the lesion and achieve an accuracy of 98.60% and a validation accuracy of 97.32%. The study showed the potential of hybrid feature extraction techniques for improving the diagnosis of skin lesions. However, the lack of attention mechanisms and explainable learning strategies hindered the model's ability to capture fine-grained lesion characteristics and generate clinically interpretable predictions. [6]

Karthik et al. proposed a hybrid deep learning framework based on a Swin Transformer and a Dense Group Shuffle Non-Local Attention (DGSNLA) network for skin cancer classification using the HAM10000 dataset. The dual-track architecture was designed to fuse the global contextual information and local lesion characteristics for feature fusion. The experimental results show that it performs better, has high precision and recall, and achieves an accuracy of 94.21%. However, the framework was mainly evaluated on binary classification and did not include clinical metadata and explainability mechanisms, which limits its applicability in real-world clinical environments. [7]

Marie et al. introduced a Melano Hybrid Model for enhanced melanoma detection by integrating YOLOv9 and Faster R-CNN with an adaptive feature fusion mechanism. The model successfully integrated YOLOv9's real-time detection ability and Faster R-CNN's precise localization ability, and it achieved 96.2% accuracy and 95.1% F1-score on the benchmark datasets. Additionally, the framework showed clinical utility with robust cross-dataset generalization and 31.3 FPS inference speed. However, the study was mainly limited to the binary melanoma classification and limited its applicability to the multiclass skin lesion diagnosis. [8]

Mridha et al. proposed an explainable skin cancer classification framework using an optimized convolutional neural network (CNN) with the integration of explainable artificial intelligence (XAI) techniques such as Grad-CAM and Grad-CAM++. The model was trained on the HAM10000 dataset that includes seven classes of skin lesions and utilized multiple optimization functions like Adam and RMSprop, and activation functions like ReLU, Swish and Tanh. The experimental results showed that the ReLU-Adam configuration achieved the highest classification accuracy, i.e., around 82% with a low loss of 0.47, whereas Grad-CAM++ provided enhanced visualization of lesion regions for better interpretability. However, the study utilized a relatively shallow CNN architecture that restricted its capacity to learn complex lesion representations for multiclass classification tasks. [9]

Hosny et al. provided a thorough overview of the application of deep learning and optimization techniques to skin lesion segmentation. The importance of segmentation in skin lesion diagnosis was also underlined in terms of its usefulness in computer-aided systems in detecting early signs of skin cancer. Various approaches were examined in the study concerning traditional segmentation methods, deep learning approaches, optimization techniques, as well as different datasets like ISIC, HAM10000, PH2, and Dermofit. It was stressed in the article that deep learning and optimized techniques offer better segmentation results, as well as high computation performance and overcome problems such as artefacts, irregular lesion borders, and illumination differences. Nonetheless, the study covered mainly skin lesion segmentation but not attention-based multiclass classification schemes or interpretability of models [10].

Noronha et al. conducted an extensive systematic review on the latest deep learning methodologies for diagnosing skin diseases using dermoscopy images. Their research covered 22 cutting-edge methods, several public datasets, various metrics, and clinical diagnostic approaches used to diagnose the skin disease. They discussed important issues such as the availability of small labelled datasets, the imbalance between classes, non-interpretability of AI models, and inadequacy of patient metadata, among others, as well as possible research directions for creating reliable AI diagnostic systems in dermatology. Even though Noronha et al.'s work is highly informative and helpful, they have not come up with a new classification approach or attention mechanism for skin disease diagnosis. [11]

Gururaj et al proposed Deep Skin, a deep learning framework for multiclass skin cancer classification on the HAM10000 dataset, which has 10,015 dermoscopic images in seven categories of lesions. The study employed pre-processing techniques such as balancing the dataset through over- and under-sampling, hair removal using the Dull Razor algorithm, and lesion segmentation with an encoder–decoder architecture. Classification was done using transfer learning models, namely DenseNet169 and ResNet50. Experimental results show that DenseNet169 with under-sampling achieved the best performance with a classification accuracy of 91.2% with an F1-score of 91.7%, which outperformed the ResNet50 model. However, this study was limited to a seven-class classification without adding an attention mechanism for better feature representation and interpretability. Attention-guided hybrid framework is suggested to overcome these limitations by providing adaptive attention-based feature extraction for the betterment of multiclass skin lesion classification.[12]

Mittal et al. have proposed DermCDSM, an intelligent clinical decision support framework for skin disease diagnosis based on ICSO-based segmentation, MSSO feature optimization and CD-SNN classification. The model was tested on the ISIC 2017 dataset and achieved a high testing accuracy of 95.07%. It also included Explainable AI

(XAI) to improve interpretability in clinical settings. The framework performed better than the traditional methods like ABCD rule, SVM and DGT-NB classifiers. However, its applicability was limited to six skin disease categories, leaving scope for scalable multiclass attention-based architectures for complex dermatological datasets. [13]

Magdy et al. proposed two computer vision-based frameworks named KNN-PDNN and AlexGWO for automatic skin cancer diagnosis using dermoscopic image data from the ISIC database. Different machine learning and deep learning models such as AlexNet, VGG, ResNet, DenseNet, EfficientNet and MobileNet were used in the study, and the results indicated that the two proposed methods achieved classification accuracy higher than 99%. It is worth to mention that the use of Grey Wolf Optimization (GWO) algorithm to optimize the hyperparameters of these models leads to significant improvement in the performance of the models. However, both approaches were limited to a binary classification task only, which means that there is no use of an attention mechanism or any kind of explainable artificial intelligence methods for multiclass lesion detection. [14]

Musthafa et al. (2024) proposed an improved Convolutional Neural Network (CNN) model for an automated detection of skin lesions using the HAM10000 dataset. To overcome the class imbalance and achieve better generalization capacity of the model, the authors used many data augmentation techniques like rotations, zooming, and flips. CNN architecture included several layers of convolution, pooling and other operations such as checkpointing callbacks, early stopping and dropout to avoid overfitting during training. The experimental results showed that the model has 97.78% accuracy in classification, precision and recall of 97.9%. Thus, CNN models are very helpful in skin cancer diagnostics. However, the CNN was not able to include attention-based feature extraction techniques and explainable AI methods such as Grad-CAM visualization. [15]

Khalaf et al. have conducted a comprehensive review on deep learning algorithms for skin cancer segmentation and classification. The paper focused on the latest developments of CNN, attention models, transformer methods and XAI approaches. The researchers discuss several major problems, such as class imbalance, lack of skin type diversity, interpretability and computational burden that threaten the practical deployment of such systems for clinical purposes. The researchers also emphasized the need to design explainable models that are able to generalize.[16]

Hosny et al. proposed an Efficient Dual Attention ResNet50 (EDA-ResNet50) framework with a Multi-scale Feature Representation (MFR) block and channel-spatial attention mechanisms to improve lesion feature extraction and classification performance. The model is able to capture subtle morphological variations in dermoscopic images and enhances interpretability through Grad-CAM visualizations. The experimental results show an accuracy of 93.18% for the classification of benign and malignant skin lesions, better than several conventional CNN architectures with relatively low computational complexity. Furthermore, the proposed framework performed well on multiple skin cancer datasets, showing its promise for real-world clinical applications in early skin cancer detection. [17]

Saeed et al. proposed a dual-branch ensemble framework, EG-VAN, for multiclass skin cancer classification by combining EfficientNetV2S and attention-enhanced ResNet50. The model used a Spatial-Context Group Attention (SCGA), Non-Local Blocks, and Multi-Scale Feature Fusion to learn local and global lesion features. In addition, a new color-balancing method based on Gray World and Retinex theory was developed to improve the quality of dermoscopic images. The experimental results on the 9-class and 7-class datasets achieved the accuracies of 98.20% and 97.80%, respectively. However, the framework is more complex from an architectural viewpoint and requires validation on external datasets to evaluate generalizability. [18]

Bibi et al. introduced FoMoSkinNet, a dual-stream deep learning framework for classifying skin cancer using non-dermoscopic images. The architecture combines a transformer-based FocalNet branch that captures the global contextual information and long-range dependencies, with a customized Local Feature Network (LFNet) for extracting fine-grained spatial features. The combination of local and global representation enhanced the lesion discrimination and classification performance. On the ISIC 2024 SLICE-3D dataset, the model achieved 98.85% accuracy, 99.46% precision, 98.23% sensitivity (recall), 98.84% F1-score, and 99.47% specificity, which is higher than several state-of-the-art CNN and transformer models. Cross-dataset evaluation on PAD-UFES-20 showed that it can generalize well. However, the framework is limited to binary classification (benign vs. malignant) and was evaluated mainly on non-dermoscopic images, suggesting the need for multiclass evaluation and explainable AI integration in future studies. [19].

In their study, Kar et al. proposed the concept of LeSegGAN as an attention-driven GAN architecture that enables accurate segmentation of skin lesions. The generation process included convolution and inception layers with channel attention, and the ViT was employed as the discriminator for lesion segmentation. The results on various benchmark datasets such as ISIC-2016 and MED-NODE, showed a higher accuracy of 99.43% and IoU scores. The limitation of the model was that it only handled the segmentation of lesions, without considering the classification of multiclass skin cancers and techniques of explainable AI. [20].

Table 1. Review of State-of-the-Art Deep Learning Models for Skin Cancer Diagnosis

Ref.	Author (Year)	Method	Strengths	Limitations
[6]	Riaz et al. (2023)	CNN-LBP Joint Learning Framework	Combined deep and handcrafted features; achieved 98.60% accuracy on HAM10000.	Lacked attention mechanisms and explainable AI for clinical interpretation.
[7]	Karthik et al. (2024)	Swin Transformer + DGSNLA Network	Effectively fused local and global features; achieved 94.21% accuracy.	Evaluated mainly for binary classification without explainability.
[8]	Marie et al. (2024)	YOLOv9 + Faster R-CNN Hybrid Model	Real-time detection with accurate localization; 96.2% accuracy.	Limited to melanoma and binary classification.
[9]	Mridha et al. (2024)	Explainable CNN with Grad-CAM/Grad-CAM++	Improved interpretability through XAI techniques.	Shallow CNN restricted complex multiclass feature learning.
[10]	Hosny et al. (2024)	Review of Segmentation Methods	Comprehensive survey of segmentation approaches and datasets.	Did not address attention-based multiclass classification.
[11]	Noronha et al. (2024)	Systematic Review of DL for Skin Diseases	Identified challenges including class imbalance and lack of metadata.	Proposed no novel classification framework.
[12]	Gururaj et al. (2024)	Deep Skin using DenseNet169 and ResNet50	Achieved 91.2% accuracy with robust preprocessing.	Limited to seven classes without attention mechanisms.
[13]	Mittal et al. (2024)	DermCDSM Framework	Integrated XAI and achieved 95.07% accuracy.	Evaluated on only six skin disease categories.
[14]	Magdy et al. (2024)	KNN-PDNN and AlexGWO	Achieved >99% accuracy using GWO optimization.	Restricted to binary classification and lacked XAI.
[15]	Musthafa et al. (2024)	Improved CNN with Data Augmentation	Achieved 97.78% accuracy and strong generalization.	Did not incorporate attention or Grad-CAM.
[16]	Khalaf et al. (2025)	Review of DL, Transformers and XAI	Highlighted challenges in	Presented no new classification model.

			interpretability and skin diversity.	
[17]	Hosny et al. (2025)	EDA-ResNet50 with Dual Attention	Multi-scale attention improved feature extraction; achieved 93.18% accuracy.	Focused mainly on binary lesion classification.
[18]	Saeed et al. (2025)	EG-VAN Dual-Branch Attention Network	Achieved 98.20% accuracy for multiclass classification using SCGA and feature fusion.	Architecturally complex and requires external validation.
[19]	Bibi et al. (2026)	FoMoSkinNet (FocalNet + LFNet)	Captured local and global features; achieved 98.85% accuracy.	Limited to binary classification and non-dermoscopic images.
[20]		LeSegGAN: Hybrid Attention-Based GAN with ViT Discriminator		Focused on lesion segmentation rather than multiclass classification; lacked explainable AI techniques and involved high computational complexity.
[21]	Proposed Methodology	Transfer learning + attention mechanism + data balancing	Provides an efficient attention-guided multiclass classification framework capable of distinguishing eight skin lesion categories.	External validation remains future work

Table 1. Comparison of conventional deep learning methods for the classification of skin lesions. Previous studies have shown promising results in employing convolutional neural networks, transformer architectures, hybrid attention mechanisms, and explainable artificial intelligence techniques. Many approaches were limited to binary classification, restricted lesion categories or a lack of model interpretability, though several models achieved high classification accuracies of more than 95%. The attention-based frameworks enhanced feature extraction, but they were also often associated with high computational complexity and low clinical explainability. Moreover, there are still unresolved challenges such as class imbalance, poor generalization across datasets and lack of robust multiclass classification frameworks. Such limitations motivate the development of the proposed Attention-Guided Xception framework for the automated diagnosis of skin lesions, which requires accurate, robust, and computationally efficient models capable of effectively distinguishing multiple lesion categories.

2.1. Research Gap

Deep learning has significantly enhanced the automated diagnosis of skin cancer using convolutional neural networks, transformers, hybrid frameworks, and explainable artificial intelligence techniques. Some studies, such as CNN-LBP joint learning, Swin Transformer-based network, attention-enhanced CNNs and dual branch architectures have shown potential classification performance. But there are still a number of important issues that have not been settled. The majority of the surveyed studies are mainly designed for binary classification or limited multi-class settings with six or seven types of lesions. Most of the reported studies are mainly designed for binary classification or restricted multi-class settings with six or seven types of lesions. Although multiclass frameworks have achieved high accuracies, their scalability and generalization to different datasets have not been well investigated.

Many state-of-the-art solutions rely on computationally expensive architectures, such as transformers, GANs, or ensemble networks, which lead to increased model complexity and preclude their deployment in resource-constrained clinical settings. Hence, practical applications require efficient and accurate deep learning frameworks. Although deep learning has achieved great success, many studies lack robust attention mechanisms that can highlight the lesion regions that are relevant for diagnosis. As such, the subtle morphological differences between visually similar skin lesions may not be sufficiently captured, resulting in misclassification.

Another important limitation is the absence of explainability and clinical interpretability. Some studies have integrated explainable AI techniques, but many deep learning models remain black-box systems, which lowers clinician trust and restricts real-world adoption in healthcare settings.

Moreover, class imbalance and underrepresented minority lesion categories in the dataset continue to impact the classification performance, particularly sensitivity and recall for the rare skin cancer categories. Current approaches usually obtain high overall accuracy, but demonstrate inconsistent performance across various lesion types.

A framework based on an Attention-Guided Xception architecture for multiclass skin lesion classification using dermoscopic images from the ISIC Kaggle dataset is proposed. The original dataset comprised nine different categories of lesions, but the images of melanoma lesions were removed in the refinement process to centre the study on eight classes of non-melanoma skin lesions of clinical relevance and to reduce class ambiguity in the training of the model. The proposed framework integrates transfer learning, attention-guided feature refinement, and class balancing techniques to improve the classification performance of various skin lesion categories.

To validate the robustness of the model across all lesion classes, a comprehensive assessment is carried out using accuracy, precision, recall, F1-score, and confusion matrix analysis. Thus, the proposed framework aims to provide an accurate, computationally efficient and interpretable solution for multiclass skin cancer diagnosis, thereby addressing key research gaps identified in the literature review.

3. Proposed Methodology

3.1. Dataset Description

The dermoscopic images used in this study were obtained from a publicly available skin lesion dataset hosted on Kaggle and derived from the International Skin Imaging Collaboration (ISIC) archive [21]. The original database consisted of 2,357 high-resolution images of skin lesions in different diagnostic categories, including melanoma. During preprocessing, melanoma images were removed to maintain the dataset consistency and to align with the goals of multiclass non-melanoma skin lesion analysis. The final dataset included eight categories of skin lesions: actinic keratosis, basal cell carcinoma, dermatofibroma, nevus, pigmented benign keratosis, seborrheic keratosis, squamous cell carcinoma and vascular lesions.

The selected categories contain benign and malignant skin diseases with significant variations in shape, color, texture and boundary of the lesion. This diversity poses a big challenge for automated classification and is very close to real-life clinical situations. Some categories were observed to have a slight class imbalance, especially nevus and pigmented benign keratosis, which contained a relatively larger number of images compared to other classes.

The clinical diversity and the visual complexity of the dataset make it ideal for developing robust deep learning models for computer-aided diagnosis of skin lesions. The inclusion of multiple dermatological conditions allows for a comprehensive testing of the proposed framework in differentiating visually similar lesion types and aiding in early disease detection.

Figure 1 shows examples of skin lesion images in the skin lesion database showing different types of lesions: actinic keratosis, basal cell carcinoma, dermatofibroma, nevus, seborrheic keratosis, benign pigmented keratosis, squamous cell carcinoma, and vascular lesion. These images are characterized by various colors, textures, shapes, and contours of the lesions. The images feature grid-based organization for showing distinctive visual patterns, which serve as training material for automated systems that classify skin lesions. Standardized clinical photographs with pixel coordinates compose the axes of the analyzed images which serve the purpose of computational dermatological research.

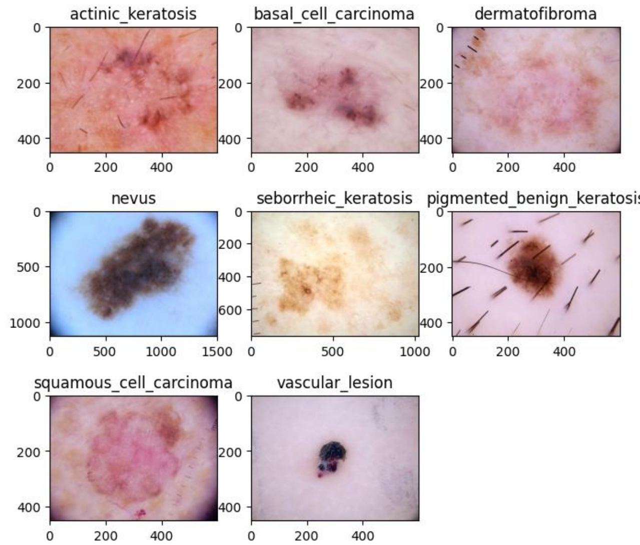

Fig. 1 Sample dermoscopic images of different skin lesion categories used in the study

3.2. Preprocessing of the Data

Data preprocessing is a critical step for building robust deep learning models for automated skin lesion classification. Before the model training, several preprocessing operations were applied to the dermoscopic images to ensure data consistency and improve the classification performance.

Each image was initially assigned to its respective skin lesion category. As the problem is a multiclass classification problem, categorical labels were converted into one-hot encoded vectors so as to make it compatible with deep learning algorithms using the Label Binarizer technique. Subsequently, partition the dataset into training and testing sets through the use of a stratified sampling method to maintain the class distribution for all lesion categories.

All images were resized to a fixed resolution of 128 × 128 pixels using the OpenCV library to maintain the consistency of the model input. Resizing the image reduces the computational complexity and makes it compatible with the deep learning framework used in this study.

Grayscale conversion was not performed, as, unlike the usual image-processing pipelines, color information is a major factor in differentiating different dermatological conditions. Thus, all dermoscopic images were stored in the RGB format to keep the important chromatic characteristics of skin lesions.

The pre-processed dataset was saved in an optimized format for fast batch loading during training. The proposed preprocessing pipeline enhances data consistency, preserves clinically relevant information, and provides a solid foundation for automated skin lesion classification.

3.3. Data Balancing Strategy

The original skin lesion dataset had a class imbalance problem, i.e., some lesion categories had significantly fewer samples than others. Such an imbalance can lead to bias in deep learning models towards majority classes, thus resulting in degraded classification performance on minority lesion classes. This problem is particularly relevant in medical image analysis, where rare classes can be of high clinical importance.

Before training the models, the Synthetic Minority Oversampling Technique (SMOTE) was used to address the issue of class imbalance [22]. First, the image data were transformed into vectorized forms, in order to be able to apply SMOTE. Then, synthetic samples were generated in the feature space using the nearest-neighbour approach to augment the minority classes. This technique improves the class representation while preserving the underlying data characteristics. Then, the balanced samples were reshaped into the image format to perform further deep learning experiments.

SMOTE was used to improve the performance of minority lesion categories and reduce model bias during training. A more balanced dataset allows the classifier to learn discriminative features for all lesion classes, which in turn improves the generalization and classification stability. The proposed deep learning framework was then trained and tested on the balanced data set.

3.4. Proposed Attention-Guided Xception Framework

In this work, an Attention-Guided Xception architecture is proposed for robust multiclass skin lesion classification. The framework is designed to automatically extract discriminative features from dermoscopic images by highlighting diagnostically relevant regions through an attention mechanism. The entire architecture incorporates transfer learning, feature refinement, and multiclass classification for enhanced diagnostic performance of different types of skin lesions.

The Xception network is the backbone model because of its efficient depthwise separable convolutions, which can significantly reduce the computational complexity while maintaining the high feature extraction capability [23]. Unlike traditional convolutional neural networks, Xception decouples the convolutional operation into depthwise and pointwise convolution, which helps to learn spatial and channel-wise representations efficiently. Moreover, transfer learning with ImageNet pre-trained weights was used to accelerate convergence and improve generalization performance.

The Xception acts as a backbone that extracts a high-dimensional feature representation (F) for a given input dermoscopic image (I), which can be expressed as:

$$[F=\text{Xception}(I)] \tag{1}$$

(F) represents the extracted deep feature maps.

To further improve the lesion representation, an attention mechanism was added to highlight the informative regions while suppressing irrelevant background features [17]. The proposed attention module learns adaptive weights of features to enable the network focus on lesion characters that are clinically meaningful. The attention weights are calculated as:

$$A=\sigma(Wa\cdot F+b) \tag{2}$$

where:

- $\boldsymbol{A}$= attention map
- $\boldsymbol{W_a}$= learnable attention weight matrix
- $\boldsymbol{F}$= extracted feature map from Xception
- $\boldsymbol{b}$= bias term
- $\boldsymbol{\sigma}$= activation function (typically Sigmoid or Softmax)

Here, $\boldsymbol{W_a} \cdot \boldsymbol{F}$means that the extracted features are multiplied by trainable weights to compute the importance of different regions in the feature map. The extracted features are then multiplied element-wise with the attention weights to obtain the refined feature representation as follows

$$\boldsymbol{F_{att}}=F\odot A \tag{3}$$

where:

- $\boldsymbol{F_{att}}$= attention-refined feature map
- $\boldsymbol{F}$= extracted feature map from the Xception network
- $\odot$= element-wise multiplication
- $\boldsymbol{A}$= attention map

After feature refinement, Global Average Pooling (GAP) is applied to reduce the dimensionality of the features while preserving the salient information. The resulting feature vector is then fed through fully connected layers and a Softmax classifier to produce the final multiclass predictions. The Softmax function is defined as

$$\boldsymbol{P(yi)} = \frac{e^{Zi}}{\sum_{j=1}^{C} e^{Zj}} \tag{4}$$

where P(yi) represents the probability of assigning an input image to class i, Zi denotes the output score for class i, C represents the total number of classes, and e is the exponential function. The Softmax function converts the classifier outputs into normalized probability values, ensuring that the sum of probabilities across all classes equals one.

The proposed Attention-Guided Xception framework effectively integrates deep feature extraction and adaptive attention learning to improve the robustness of lesion discrimination and classification. The framework targets clinically relevant regions of interest to enhance the prediction performance and to provide a reliable computer-aided diagnostic tool for automatic skin lesion analysis.

In this study, C = 8, corresponding to the eight skin lesion categories considered for multiclass classification.

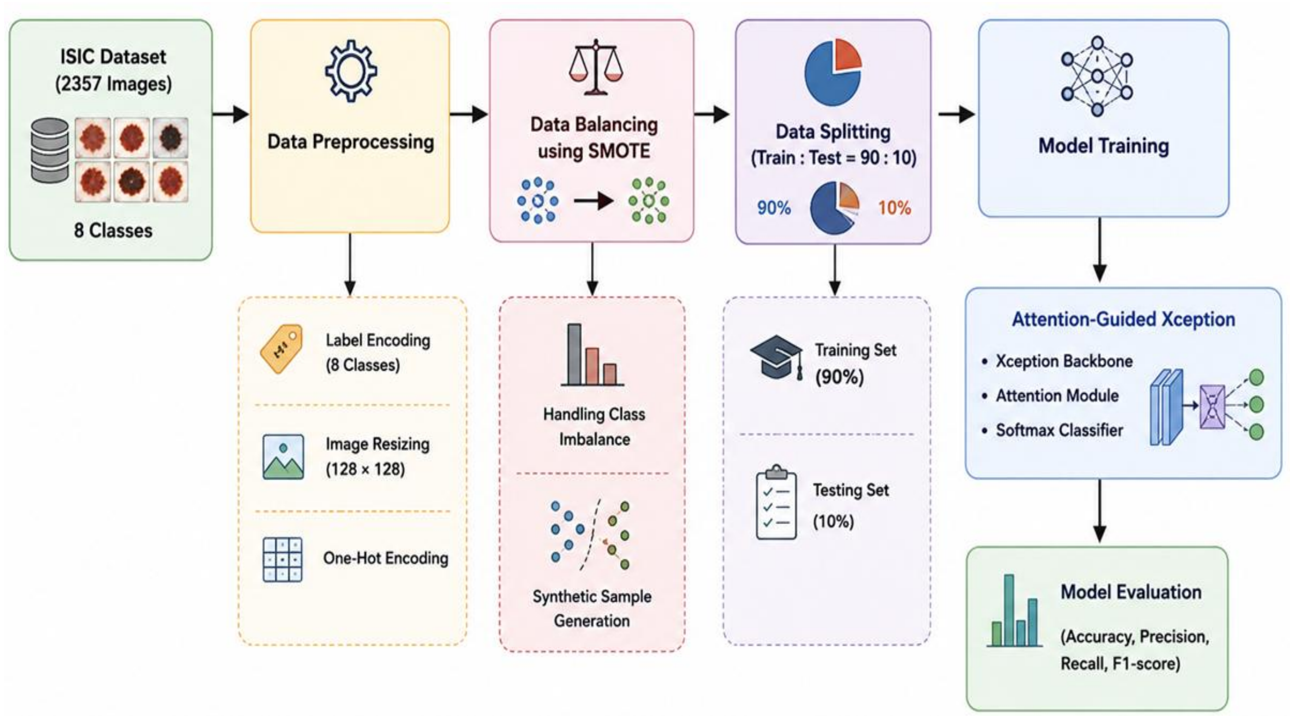

Fig. 2. Workflow of the proposed Attention-Guided Xception framework for multiclass skin lesion classification

Figure 2 shows the general workflow of the proposed Attention-Guided Xception framework for multiclass skin lesion classification. First, 2,357 dermoscopic images were collected from ISIC dataset and classified into eight lesion categories after removing the samples of melanoma. During pre-processing, label encoding, resizing of images to 128 × 128 pixels and one-hot encoding were performed to maintain the consistency of data and compatibility with deep learning algorithms.

To solve the problem of class imbalance between different categories of lesions, the Synthetic Minority Oversampling Technique (SMOTE) was used to generate synthetic samples for the minority classes to improve the class representation. Then, the balanced dataset was divided into training and test subsets with a 90:10 ratio with stratification.

In the training phase, the proposed Attention-Guided Xception model is used, which integrates the Xception backbone network with an attention mechanism to enhance feature extraction and emphasize diagnostically relevant lesion regions. Finally, the model performance was evaluated using standard classification metrics such as accuracy, precision, recall, and F1-score to comprehensively evaluate the effectiveness of the proposed framework in multiclass skin lesion classification.

3.5. Model Training and Hyperparameter Settings

The proposed Attention-Guided Xception framework is implemented in Python with TensorFlow/Keras libraries. The model was trained on a system with GPU acceleration to improve computational efficiency. Transfer learning by initializing the Xception backbone with weights pre-trained on ImageNet to speed up convergence and improve feature learning was used.

The balanced dataset resulting from SMOTE processing was partitioned into training and testing subsets through a stratified sampling approach with a 90:10 ratio. Adam optimizer was used during the training because of its adaptive learning property and faster convergence. Considering the multi-class classification nature of the problem, the categorical cross-entropy loss function was utilized. The model parameters were optimized iteratively during training to minimize the classification loss and maximize the predictive performance. The images were processed in mini-batches during the training to improve the memory utilization and the learning stability.

Hyperparameters for training the Attention-guided Xception model is summarized in Table 2 below. The image size of 128 × 128 × 3 pixels was chosen to allow for effective computation while retaining the essential characteristics of the lesions. The Xception architecture was used for feature extraction because of the deep separable convolutions in the model that made it more capable of feature extraction in this paper. Adam optimizer was used because of its adaptability and faster rate of convergence. Categorical cross-entropy was used to solve the multiclass classification problem. For a reliable model evaluation, the dataset was split into a training and testing set in a ratio of 90:10. Furthermore, the attention mechanism made the model focus on important regions within lesions, which improved the feature extraction performance.

Table 2. summarizes the hyperparameter settings

Parameter	Value
Input Image Size	128x128x3
Optimizer	Adam
Loss Function	Categorical Cross-Entropy
Output Classes	8
Train-Test Split	90:10
Feature Extractor	Xception
Attention Mechanism	Attention Layer

3.6 Performance Evaluation

The suggested Attention-Guided Xception architecture was validated using the traditional classification metrics, accuracy, precision, recall, and F1-score. This evaluation method is very important, since it provides a comprehensive analysis of the model's efficiency in classifying skin lesion types. it includes

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

with equations:

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP + TN}}{\boldsymbol{TP + TN + FP + FN}}$$

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP + FP}}$$

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP + FN}}$$

$$\boldsymbol{F1} = \frac{\boldsymbol{2 \times Precision \times Recall}}{\boldsymbol{Precision + Recall}}$$

where *TP* is True Positive, *TN* is True Negative, *FP* is False Positive, and *FN* is False Negative

The performance of the proposed model is measured using accuracy as a metric to measure the percentage of accurate classification against the total number of predictions made. Precision is the measure of the correctness of predictions on positive classes and indicates the reliability of the classification model. To measure the ability of the proposed model to recognize real positive instances, the sensitivity (recall). This measure is important when applying the proposed method in medical diagnostics, where undetected cases can be fatal. F1 score The F1 score is the harmonic mean of the above two measures: precision and sensitivity.

4. Results and Discussion

The following discussion is devoted to the discussion of the experimental results obtained from the experiments carried out with the proposed models on the ISIC skin lesion dataset. The performance metrics of the models, like convolutional neural networks, VGG16, EfficientNet-B2, Xception, and attention-guided Xception, will be analyzed and compared using different performance metrics like accuracy, precision, recall, and F1-Score. In addition to the above, a class-based classification analysis was carried out through the construction of the confusion matrix.

4.1. Performance Analysis of the Xception Model

Table 3 presents the mapping between the numerical class labels used in the classification results and their corresponding skin lesion categories. This mapping facilitates the interpretation of the confusion matrices and classification reports presented in the subsequent sections.

Table 3. Mapping of class labels used in classification results.

Class Label	**Skin Lesion Category**
0	Actinic Keratosis (AK)
1	Basal Cell Carcinoma (BCC)
2	Dermatofibroma (DF)
3	Nevus (NV)
4	Pigmented Benign Keratosis (PBK)
5	Seborrheic Keratosis (SK)
6	Squamous Cell Carcinoma (SCC)
7	Vascular Lesion (VL)

The classification ability of the Xception model is analyzed by employing the balanced ISIC dataset having eight classes of skin lesions. As shown in Fig. 3, the confusion matrix is generated for the Xception classifier. The dominant entries along the main diagonal confirm that the classifier accurately identified the majority of skin lesion images. Class 5 and 7 showed the best classification results, which had 46 and 45 correctly predicted samples, respectively. Likewise, classes 0 and 2 displayed the same level of prediction accuracy with 43 and 42 correct predictions, respectively.

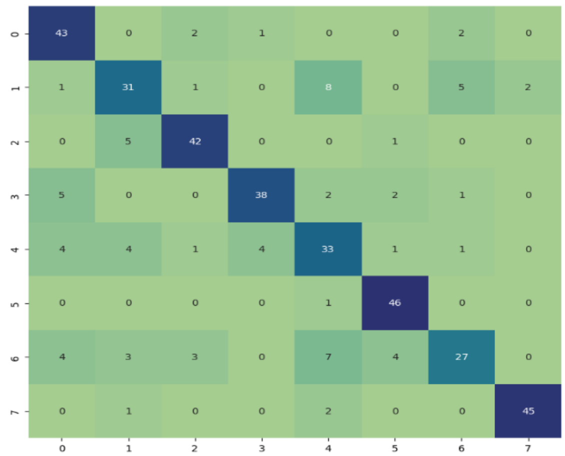

Fig.3 Confusion Matrix (Xception Model)

Even though the performance of the Xception model can be considered satisfactory, there were cases when skin lesions from similar classes resulted in misclassification. For instance, classes 4 and 6 were sometimes confused with each other. Moreover, class 6 was also mistaken as class 5 at times. This may be because some skin lesion classes have identical dermatological properties, resulting in difficulty in differentiation.

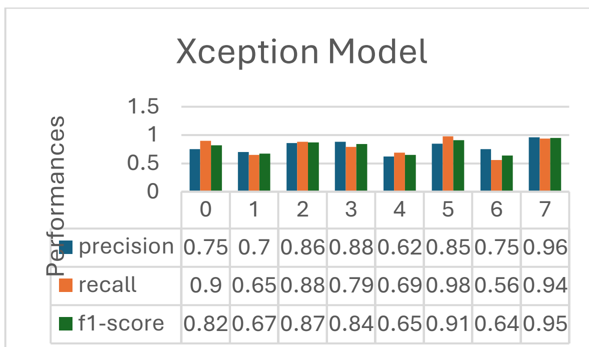

Fig.4 Performance Metrics Analysis Graph of Xception Model Architecture

Figure 4 shows the precision, recall, and F1-score for all lesion categories obtained by the Xception model. The model performed well on classes 2, 5 and 7, showing that it has learned features well for these types of lesions. However, class 6 had a lower recall relative to other classes, meaning some samples were misclassified into other classes. The class-wise performance variation shows the complexity of the multi-class skin lesion classification problem and the need for better feature discrimination mechanisms.

4.2 Performance Analysis of the Proposed Attention-Guided Xception Model

The proposed Attention-Guided Xception framework was evaluated on the balanced ISIC dataset, which includes eight skin lesion categories. Figure 5 shows the confusion matrix obtained from the developed model. In comparison to the default Xception architecture, the proposed framework shows a remarkable increment in the classification performance. This can be observed from the stronger diagonal values and the lower off-diagonal misclassifications.

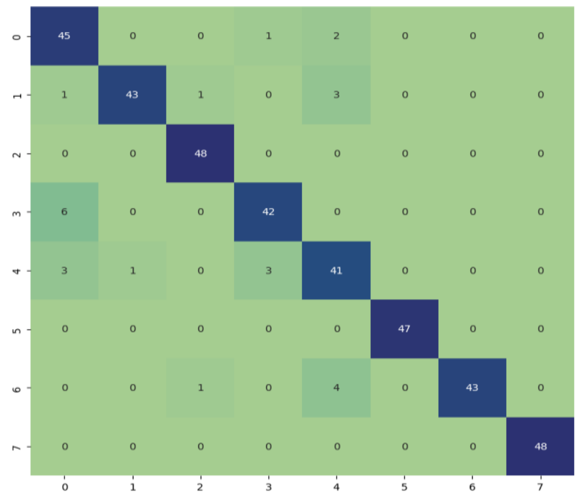

Fig. 5 Confusion Matrix (Xception with Attention Mechanism)

From the confusion matrix, the best performance of classification was obtained for classes 2 and 7, with 48 samples correctly classified each, and 47 correct predictions for class 5. Similarly, the classes 0, 1, 3 and 6 had 45, 43, 42 and 43 correct classifications, respectively. Some misclassifications were observed between visually similar lesion categories, indicating that the attention mechanism can improve feature discrimination and reduce inter-class confusion. The better classification performance could be attributed to the model being able to attend to the diagnostically important lesion regions and suppress the irrelevant background information.

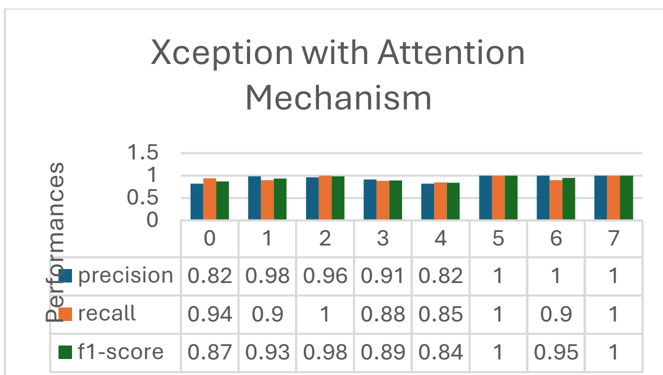

Fig. 6 Performance Metrics Analysis Graph of Xception with Attention Mechanism Model Architecture

Fig. 6. The precision, recall and F1-score of the proposed model for all eight lesion categories. The results show consistent performance across all classes, with some categories exhibiting near-perfect classification metrics. Classes 5, 6 and 7 had perfect classification precision and classes 2 and 7 had a recall of 1.0, showing very reliable classification performance. Notably, a significant improvement was observed for class 6 compared to the standard Xception model, where the precision increased from 0.75 to 1.0 and the recall improved from 0.56 to 0.90. These results demonstrate the effectiveness of the attention mechanism in improving feature representation and classification accuracy for difficult lesion classes.

In general, the proposed Attention-Guided Xception framework outperformed the baseline Xception model by reducing classification ambiguity and improving the identification of visually similar skin lesions. Results confirm the critical importance of attention-guided feature refinement for robust multiclass skin lesion classification and support the practical applicability of the proposed framework for computer-aided dermatological diagnosis.

4.3 Comparative Performance Analysis

To evaluate the effectiveness of the proposed framework, a comparative performance analysis was performed using five deep learning architectures namely CNN, VGG16, EfficientNet-B2, Xception and the proposed Attention-Guided Xception model. All the models were trained and tested on the balanced ISIC skin lesion dataset and under the same experimental conditions. The aim of this comparison was to evaluate the effect of different feature extraction techniques and to measure the effect of the attention mechanism in the performance of multiclass skin lesion classification.

Table 4. Accuracy comparison of different deep learning models for multiclass skin lesion classification

Model	Accuracy (%)
CNN	82
VGG16	74
EfficientNet-B2	84
Xception	80
Xception with Attention Mechanism	93

Table 4 summarizes the classification accuracy by the different deep learning algorithms used for the classification of lesions in the skin ISIC dataset. The conventional architecture algorithm, VGG16, achieved a relatively low classification accuracy of 74% due to its poor ability to capture complex characteristics of the lesion. CNN obtained the second-best classification performance with a high classification accuracy rate of 82% in the experiment. The third-best classification accuracy was produced by EfficientNet-B2, scoring a higher 84% due to its use of a compound scaling approach, enabling effective feature extraction. Xception had good feature learning capacity with the use of depthwise separable convolutions, but it obtained the fourth-highest accuracy (80%), which was a little less than the one achieved by EfficientNet-B2. From this observation, it can be seen that feature extraction is not sufficient to discriminate similar classes in multiclass classification problems.

Finally, the proposed Attention-Guided Xception framework achieved the best classification accuracy of 93%. Herein, the attention mechanism was helpful for better feature extraction by allowing the algorithm to focus only on lesion regions. Additionally, the usage of SMOTE to generate a balanced dataset gave the minority lesion classes equal representation, which improved classification performance and stability.

4.4 Comparison with Existing Studies

To further validate the effectiveness of the proposed framework, a comparison is made with prior published state-of-the-art studies on skin lesion classification. Selected studies used different deep learning architectures, including convolutional neural networks, transformer-based models, attention-guided frameworks, and hybrid models. The analysis compares the types of classification, the accuracy reported and the main limitations of the methods. The analysis presented here gives a comprehensive evaluation of the proposed Attention-Guided Xception framework and demonstrates its potential to achieve competitive performance in multiclass skin lesion classification.

Table 5. Comparative analysis of the proposed framework with the latest state-of-the-art skin lesion classification studies.

Ref.	Study	Method	Classification Type	Accuracy (%)	Key Limitation
[7]	Karthik et al. (2024)	Swin Transformer + DGSNLA	Binary	94.21	Limited multiclass evaluation
[8]	Marie et al. (2024)	YOLOv9 + Faster R-CNN	Binary	96.20	Focused mainly on melanoma detection
[9]	Mridha et al. (2024)	Explainable CNN with Grad-CAM	Multiclass	92.00	Limited feature extraction capability

[12]	Gururaj et al. (2024)	DenseNet169 + ResNet50	Multiclass	91.20	No attention mechanism
[17]	Hosny et al. (2025)	EDA-ResNet50 with Dual Attention	Binary	93.18	Focused primarily on binary classification
[19]	Bibi et al. (2026)	FoMoSkinNet (FocalNet + LFNet)	Binary	98.85	Non-dermoscopic dataset and binary classification
Proposed	**Attention-Guided Xception Framework**	**Xception + Attention + SMOTE**	**Multiclass (8 Classes)**	**93.00**	**Computational optimization can be further improved**

The proposed Attention-Guided Xception framework is compared with existing skin lesion classification approaches in Table 5. However, some studies claimed higher classification accuracies, but mostly concentrated on binary classification, the detection of melanoma or a limited number of disease categories. Instead, the proposed framework addresses the more challenging problem of multiclass classification with eight clinically relevant non-melanoma skin lesion classes that are visually very similar.

The accuracy of 93% achieved demonstrates the effectiveness of the combination of attention-guided feature refinement and the Xception backbone network. In contrast to the traditional CNN based methods, the proposed framework pays more attention to the diagnostically significant lesion regions, which favors the feature discrimination and inter-class confusion. Additionally, SMOTE-based balancing increased the number of samples in the minority class and improved the classification stability across all lesion categories.

The proposed framework has a simple architecture and computational efficiency, and obtains competitive performance. Findings show that attention-guided learning offers a viable and robust solution for automated classification of skin lesions and underlines its potential use in computer-aided diagnosis in dermatology.

4.5. Discussion and Clinical Implications

Experimental results show that the proposed Attention-Guided Xception framework can improve the performance of skin lesion multiclass classification. The framework could attend to meaningful lesion regions by introducing the attention-based approach which enhances feature extraction and reduces misclassification rates in terms of visually similar lesions. These advantages have been successfully applied to the task as seen in the confusion matrix analysis and the high classification scores produced by the proposed model.

In this work, SMOTE was used to balance class representation in the training process, improving lesion recognition in minority groups. Moreover, in the preprocessing, removed melanoma images from the dataset, so the model had to work with only eight clinically significant non-melanoma skin lesion classes. Thus, class confusion was minimized for better and more consistent features obtained over all lesion types.

The comparative analysis performed in this experiment demonstrated that the proposed framework outperformed the baseline models under test. Moreover, the model was competitive with the state-of-the-art solutions despite its relative simplicity. The main factors behind this achievement were the attention-guided mechanism, transfer learning with the Xception network, and balanced class distribution.

5. Conclusion and Future Work

In this paper, an Attention Guided Xception framework is proposed for multiclass skin lesion classification using dermoscopic images of the ISIC dataset. The dataset was filtered to contain eight clinically relevant categories of non-melanoma skin lesions, and the class imbalance was addressed using the Synthetic Minority Oversampling

Technique (SMOTE). A thorough comparative study using CNN, VGG16, EfficientNet-B2, Xception, and the proposed Attention-Guided Xception model was performed.

The experimental results demonstrated that the proposed framework obtained the highest classification accuracy of 93%, outperforming the evaluated baseline models. The attention mechanism enabled the model to concentrate on the lesion areas important for diagnosis, leading to a better representation of features and reduced confusion between classes. Besides, the use of balanced training data and transfer learning by means of the Xception backbone contributed to better classification stability and generalization performance.

The results confirm the effectiveness of the proposed framework as an efficient and reliable method for automated skin lesion classification and may enable computer-aided dermatological diagnosis and early disease detection.

Future work will be evaluating the framework on bigger and more diverse dermatology datasets, including HAM10000 and ISIC 2020. Future work will include developing more sophisticated transformer-based models, multimodal learning techniques that incorporate clinical metadata, and explainable artificial intelligence methods such as Grad-CAM and SHAP to enhance model interpretability and clinical applicability.

Conflicts of Interest

"The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper."

Funding Statement

The authors received no specific funding for this research, authorship, and/or publication of this article.

Acknowledgments

The author would like to express the heartiest thanks to the supervisor for guidance, suggestions and encouragement all along the period of this research. Special thanks to the institute for providing a suitable environment and necessary facilities to accomplish this research. The author wishes to thank, finally, the members of the family for the encouragement and support during this phase of research.

References

1. Y. G. Xu, J. L. Aylward, A. M. Swanson, V. S. Spiegelman, E. R. Vanness, J. M. C. Teng,S. N. Snow and G. S. Wood, "Nonmelanoma Skin Cancers: Basal Cell and Squamous Cell Carcinomas," in Dermatology, Elsevier, 2020, pp. 1052–1073.e8. doi: 10.1016/B978-0-323-47674-4.00067-0.
2. R. Rajan, "Sunscreens and Skin Cancer: An Update," in Sunscreens for Skin of Color. Singapore: Springer Nature Singapore, 2024, pp. 73–90. doi: 10.1007/978-981-97-3195-4_4.
3. K. C. Gogebakan, E. G. Berry, A. C. Geller, K. Sonmez, S. A. Leachman, and R. Etzioni, "Strategizing Screening for Melanoma in an Era of Novel Treatments: A Model-Based Approach," Cancer Epidemiology, Biomarkers & Prevention, vol. 29, no. 12, pp. 2599–2607, 2020. https://doi.org/10.1186/s12911-025-03074-9
4. C. C. Oh, A. Jin, and W. P. Koh, "Trends of Cutaneous Basal Cell Carcinoma, Squamous Cell Carcinoma, and Melanoma among the Chinese, Malays, and Indians in Singapore from 1968 to 2016," JAAD International, vol. 4, pp. 39–45, 2021. https://doi.org/10.1016/j.jdin.2021.05.006
5. P. R. Kshirsagar, H. Manoharan, S. Shitharth, A. M. Alshareef, N. Albishry, and P. K. Balachandran, "Deep Learning Approaches for the Prognosis of Automated Skin Disease," Life, vol. 12, no. 3, p. 426, 2022. doi: 10.3390/life12030426.
6. L. Riaz, H. M. Qadir, G. Ali, M. Ali, M. A. Raza, A. D. Jurcut, and J. Ali, "A Comprehensive Joint Learning System to Detect Skin Cancer," IEEE Access, vol. 11, pp. 79434–79444, 2023. doi: 10.1109/ACCESS.2023.3297644.
7. R. Karthik, R. Menaka, S. Atre, J. Cho, and S. Veerappampalayam Easwaramoorthy, "A Hybrid Deep Learning Approach for Skin Cancer Classification Using Swin Transformer and Dense Group Shuffle Non-Local Attention Network," IEEE Access, vol. 12, pp. 158040–158051, 2024. doi: 10.1109/ACCESS.2024.3485507.
8. M. I. Marie, M. S. Elredeny, and A. E. Yakoub, "A Hybrid Model for Early Melanoma Detection: Integrating YOLOv9 and Faster R-CNN for Enhanced Diagnostic Accuracy," IEEE Access, vol. 13, pp. 120327–120344, 2025, doi: 10.1109/ACCESS.2025.3587625.
9. K. Mridha, M. M. Uddin, J. Shin, S. Khadka, and M. F. Mridha, "An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart Healthcare System," IEEE Access, vol. 11, pp. 41003–41018, 2023, doi: 10.1109/ACCESS.2023.3269694.
10. K. M. Hosny, D. Elshoura, E. R. Mohamed, E. Vrochidou, and G. A. Papakostas, "Deep Learning and Optimization-Based Methods for Skin Lesions Segmentation: A Review," IEEE Access, vol. 11, pp. 85467–85488, 2023. doi: 10.1109/ACCESS.2023.3303961.

11. S. S. Noronha, M. A. Mehta, D. Garg, K. Kotecha and A. Abraham, "Deep Learning-Based Dermatological Condition Detection: A Systematic Review with Recent Methods, Datasets, Challenges, and Future Directions," in IEEE Access, vol. 11, pp. 140348-140381, 2023, doi: 10.1109/ACCESS.2023.3339635.
12. H. L. Gururaj, N. Manju, A. Nagarjun, V. N. Manjunath Aradhya, and F. Flammini, “Deep Skin: A Deep Learning Approach for Skin Cancer Classification,” IEEE Access, vol. 11, pp. 50205–50214, 2023, doi: 10.1109/ACCESS.2023.3274848.
13. R. Mittal, F. Jeribi, R. John Martin, V. Malik, S. J. Menachery, and J. Singh, “DermCDSM: Clinical Decision Support Model for Dermatosis Using Systematic Approaches of Machine Learning and Deep Learning,” IEEE Access, vol. 12, pp. 47319–47337, 2024, doi: 10.1109/ACCESS.2024.3373539.
14. A. Magdy, H. Hussein, R. F. Abdel-Kader, and K. Abd El Salam, “Performance Enhancement of Skin Cancer Classification Using Computer Vision,” IEEE Access, vol. 11, pp. 72120–72133, 2023, doi: 10.1109/ACCESS.2023.3294974.
15. M. M. Musthafa, M. T. R., V. Kumar, and S. Guluwadi, “Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification,” BMC Medical Imaging, vol. 24, no. 201, 2024, doi: 10.1186/s12880-024-01356-8.
16. A. D. Khalaf, A. A. Halin, H. Hamdan, and N. Manshor, “Segmentation and Classification of Skin Cancer Diseases Based on Deep Learning: Challenges and Future Directions,” IEEE Access, vol. 13, pp. 90163–90184, 2025, doi: 10.1109/ACCESS.2025.3569170.
17. M. Hosny, I. A. Elgendy, S. A. Chelloug, and M. A. Albashrawi, “Attention-Based Convolutional Neural Network Model for Skin Cancer Classification,” IEEE Access, vol. 13, pp. 172027–172050, 2025, doi: 10.1109/ACCESS.2025.3616022.
18. A. Saeed, K. Shehzad, S. Ahmed, and A. T. Azar, “EG-VAN: A Global and Local Attention-Based Dual-Branch Ensemble Network with Advanced Color Balancing for Multi-Class Skin Cancer Recognition,” IEEE Access, vol. 13, pp. 72852–72874, 2025, doi: 10.1109/ACCESS.2025.3561240
19. F. Bibi, Q. Abbas, and S. M. Naqi, “FoMoSkinNet: A Dual-Stream Deep Learning Model for Early Detection and Classification of Non-Dermoscopic Skin Lesions via Focal Modulation and Local Feature Integration,” IEEE Access, vol. 14, pp. 41998–42014, 2026, doi: 10.1109/ACCESS.2026.3674219.
20. M. K. Kar, V. Venugopal, and B. N. Anoop, “LeSegGAN: A Hybrid Attention-Based GAN for Accurate Lesion Segmentation in Dermatological Images,” IEEE Access, vol. 13, pp. 177019–177035, 2025, doi: 10.1109/ACCESS.2025.3621107.
21. International Skin Imaging Collaboration (ISIC) Archive, “ISIC Archive: International Skin Imaging Collaboration Skin Lesion Database,” Available: https://www.isic-archive.com, Accessed: January 2026.
22. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic Minority Over-sampling Technique,” Journal of Artificial Intelligence Research, vol. 16, pp. 321–357, 2002. https://doi.org/10.1613/jair.953
23. F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 1800-1807, doi: 10.1109/CVPR.2017.195.