International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Convolutional Neural Network Model for Skin Cancer Diagnosis in a Dermatological Center

Victor Manuel Cabrejos-Yalán[*] , Ciro Rodriguez

Escuela Universitaria de Posgrado, Universidad Nacional Federico Villarreal (UNFV), Lima 15001, Peru

Corresponding Author Email: 2021007976@unfv.edu.pe

ABSTRACT

In confronting the global health challenge posed by skin cancer, early and accurate diagnosis is paramount. This research introduces an advanced Convolutional Neural Network (CNN) model optimized for skin cancer diagnosis using dermatological images. The innovation lies in applying state-of-the-art pruning techniques, specifically magnitude-based weight pruning and quantization, to refine the model's efficacy and computational efficiency. The model exhibited exceptional performance on a rigorous dataset, achieving an AUC (Area Under the Curve) value of 0.99. The acquired score indicates an exceptionally high degree of competence in distinguishing benign skin conditions from malignant ones. Critical performance indicators—with values of 0.9820 for precision, 0.9815 for recall, and 0.9812 for F1-score—offer supplementary substantiation concerning the dependability and accuracy of the model. Notably, the refined model maintained an impressive accuracy rate of 0.9815 post-pruning, validating the effectiveness of the pruning process. The employment of these pruning methods has substantially streamlined the model without compromising diagnostic accuracy, demonstrating the integration of machine learning can significantly enhance medical imaging. The findings of this study not only mark a leap forward in skin cancer diagnostics but also enrich the discourse on intelligent systems in healthcare, advocating for broader adoption and continued development.

1. INTRODUCTION

Cancer poses a substantial worldwide health concern, with skin cancer being particularly widespread as a result of extensive ultraviolet (UV) radiation exposure. Skin cancer ranks as the fourth most prevalent malignancy in Peru, according to the General Directorate of Epidemiology, accounting for approximately 10% of all cancer cases reported annually. Among the different types of skin cancer, basal cell carcinoma is the most common, followed by squamous cell carcinoma and melanoma, which although less frequent, accounts for the majority of skin cancer deaths due to its aggressive nature. The mortality rate for melanoma in Peru is estimated at 2.5 deaths per 100,000 people annually, with a higher incidence in regions with greater exposure to ultraviolet (UV) radiation, such as the coastal and highland areas. The challenges faced by healthcare professionals in Peru include limited access to specialized dermatological services, especially in rural regions, where early detection and treatment are often delayed due to geographical and infrastructural limitations. Additionally, there is a lack of widespread public awareness regarding the risks of UV exposure, further complicating prevention efforts. These factors emphasize the need for innovative diagnostic tools, such as the AI-based model proposed in this study, to improve early detection and facilitate more equitable healthcare access across the country [1]. The increasing incidence and prevalence of various types of skin cancer—such as melanoma, squamous cell carcinoma, and basal cell carcinoma—Highlight the critical need for effective diagnostic methods [2].

This article presents a novel artificial intelligence (AI) model that analyzes and classifies skin cancer types from medical images trained on Convolutional Neural Networks (CNNs). This study distinguishes itself by incorporating state-of-the-art pruning methods, such as quantization and magnitude-based weight pruning, in order to improve the efficacy and computational efficiency of the model [3].

Recent research has underscored the criticality of early detection in enhancing patient prognoses, specifically with regard to melanoma, which is classified as one of the most lethal varieties of skin cancer. Skin cancer comprised 5% of the total documented cancer cases in the United States in 2023, with melanoma incidence rates being approximately 21 cases per 100,000 people [4]. While the five-year survival rate for cutaneous melanoma is high when detected early, only 77.6% of cases are diagnosed at a localized stage, highlighting the critical importance of early detection [4].

The traditional diagnostic methods are valuable but subjective, varying based on the dermatologist's expertise [5]. This research aims to address these challenges by developing a more objective and consistent AI-based diagnostic tool, potentially alleviating the burden on specialized dermatological services and contributing to healthcare cost savings [5].

2. BACKGROUND AND RELATED WORK

This research contributes to the existing corpus of literature concerning the application of Convolutional Neural Networks (CNNs) in cutaneous cancer detection. An exhaustive bibliography was compiled through systematic searches in accordance with the protocol for a systematic review developed by Kitchenham et al. [6] as detailed in Tables 1-3.

Table 1. Systematic literature review (SLR) protocol for the study

Research Questions
Research Question 1: What are the current deep learning algorithms utilized in the domain of skin cancer detection for image classification?
Research Question 2: What are the fundamental principles, interconnections, characteristics, and constraints that are necessary for the implementation of Deep Learning in image classification as it pertains to the detection of skin cancer?
Research Question 3: What are the current methodologies employed to expedite the training process in Deep Learning with respect to image classification?

Search Protocol	
Search String	("Deep Learning" OR "Deep Neural Networks" OR "Artificial Intelligence" OR "Machine Learning") AND (("pruning" OR "quantization") OR ("Skin Cancer" OR "Melanoma" OR "Skin Lesion"))
Metadata for Search	Title; Abstract; Keywords
Selected Digital Libraries	Scopus, IEEE, ScienceDirect, SciVal, and Web of Science

Table 2. Systematic literature review (SLR) selection and quality criteria

Selection and Quality Criteria	
Inclusion Criteria	Three years have passed since the publication of the work. The research must provide documentation of the application of Deep Learning to classify images for the purpose of detecting skin cancer, or demonstrate the utilization of Pruning or Quantization methods to optimize a neural network model. Not merely an expert opinion or a lesson learned, but research must form the basis of the document. Publication in an indexed journal from the following databases is required: Scopus, IEEE, ScienceDirect, SciVal, or Web of Science.
Exclusion Criteria	A prologue, article summary, interview, news article, discussion letter, or poster are all unacceptable types of content. The work cannot currently be classified as a Finished journal article. The article's publication date must fall within the time frame of 2020 to 2023. Non-English must be used in the production. The submission should exclusively consist of a journal article. In Congress, not. The submission should not pertain to the domains of computer science or engineering.
Quality Criteria	The results must be directly applicable to the detection of skin cancer. The work must include the review of statistical tests and the availability of data or source code.

Table 3. State of the art summary

Title	Author	Country	Year
'Automatic Skin Cancer Detection in Dermoscopy Images Based on Ensemble Lightweight Deep Learning Network'	Wei, L., Ding, K., & Hu, H.	China	2020
'Early Skin Cancer Detection Using Deep Convolutional Neural Networks on Mobile Smartphone'	Emuoyibofarhe, J., & Ajisafe, D.	Nigeria and Germany	2020
'Towards Trustable Skin Cancer Diagnosis via Rewriting Model's Decision'	Yan, S., Zhang, Y., Zhang, X., Mahapatra, D., Chandra, S. S., Janda, M., Soyer, H. P., & Ge, Z.	Abu Dabi, EAU	2023
'Optimizing Deep Learning Networks for Edge Devices with an Instance of Skin Cancer and Corn Leaf Disease Dataset'	Sharmila, B., Santhosh, H., Parameshwara, S., Swamy, Baig, W. U., & Nanditha, S. V.	India	2023
'Iterative Magnitude Pruning-Based Light-Version of AlexNet for Skin Cancer Classification'	Medhat, S., Abdel-Galil, H., Aboutabl, A. E., & Saleh, H.	Egypt	2023

In a study introducing an innovative approach to categorize skin lesions, which are of utmost importance in the prompt identification of skin cancer, the following advanced deep CNN architectures were incorporated: Inception V3, Inception ResNet V2, and DenseNet 201 [7]. These architectures represent significant advancements in skin cancer diagnosis by leveraging the power of deep learning to improve classification accuracy. Similarly, an independent inquiry presented a non-invasive and interpretable technique for diagnosing melanoma using a combination of machine learning and deep learning models [8].

Previous studies have explored various machine learning techniques, such as SVM classifiers with feature extraction methods like the ABCD rule and Grey Level Co-occurrence Matrix (GLCM), to classify melanoma with notable accuracy. For instance, Pitchiah and Rajamanickam [9] applied these approaches to a dermoscopic image classification pipeline, highlighting the importance of optimized feature extraction in enhancing early melanoma detection.

Beyond traditional medical applications, CNNs have also been successfully implemented in agricultural diagnostics. Rachmad et al. [10] created a CNN-based model to detect corn leaf diseases, comparing architectures such as SqueezeNet, AlexNet, and ResNet-50. Their work reinforces CNNs' potential for high-accuracy classification across diverse tasks, an attribute that is equally beneficial in medical imaging. Additionally, Olayiwola et al. [11] applied CNNs to classify multi-class lung diseases using pre-trained models, demonstrating CNNs' ability to discern overlapping patterns in complex datasets. This highlights CNNs' suitability for diverse diagnostic needs, including skin cancer detection in our study. Finally, Lahouaoui et al. [12] demonstrated CNN performance in image classification by employing a fully convolutional architecture to diagnose pneumonia from X-ray images, achieving significant accuracy. This further illustrates CNNs' capacity for accurate, image-based diagnostics, aligning with

our goal of enhancing skin cancer detection through optimized CNN models. Together, these studies underscore the growing role of CNNs in improving diagnostic speed, accuracy, and interpretability across medical and non-medical domains.

Significant advancements in skin cancer classification have been achieved through the use of a lightweight version of AlexNet, optimized via iterative magnitude pruning (IMP) to address the computational demands of CNNs in resource-constrained environments [13]. By reducing the model's size while maintaining accuracy, this approach demonstrated the potential for CNN deployment on devices with limited processing power. Building on these improvements, our study further refines the application of IMP, enhancing both diagnostic accuracy and efficiency in real-world medical settings, ensuring the model remains practical for broader clinical use.

An extensive investigation was conducted to optimize deep neural networks for periphery devices by implementing pruning, weight clustering, and quantization, among other optimization techniques [14]. The aim of this research is to rectify any shortcomings in existing methodologies through the implementation of pruning. Pruning functions as a method to enhance the effectiveness of models and a procedure to improve the precision of skin cancer diagnosis.

Yan et al. [15] conducted a notable investigation into the dependability of deep neural network-based skin cancer diagnosis, introducing a novel framework that integrates human intervention during the model training phase. This approach allows users to interpret and adjust the model's decision-making logic, with the goal of improving both inference performance and reliability in clinical applications. By incorporating human insights, the framework enhances trust in AI-driven diagnostic systems while maintaining a high level of accuracy.

Emuoyibofarhe and Ajisafe [16] conducted a study wherein they compared three distinct CNNs designed for smartphone-based early detection of skin cancer. To enhance the generalizability of the model, data augmentation techniques and image normalization were incorporated into the methodology of this study. Emuoyibofarhe and Ajisafe's methodology signifies a substantial progression in the domain of mobile CNN applications designed to detect skin cancer. The present study's approach possesses the capacity to further develop and enhance this approach, thereby facilitating more precise and efficient clinical diagnosis.

Wei et al. [17] introduced an important approach for the automated detection of skin cancer using a lightweight deep learning network. Their strategy focuses on extracting discriminative features to reduce the number of parameters in the model, making it computationally efficient and suitable for deployment in resource-constrained environments. This emphasis on reducing model complexity while maintaining accuracy represents a valuable contribution to the field, particularly in settings where computational resources are limited, such as mobile health applications or remote clinics.

While all of this prior research has made significant strides in the application of CNNs for skin cancer detection, several limitations persist. For instance, the findings of Pratiwi et al. [7] and Alfi et al. [8] suggested that the focus is predominantly on binary classification, specifically differentiating between melanoma and benign lesions. This narrow focus limits their applicability in real-world clinical settings, where a broader range of skin cancer types, such as basal cell carcinoma and squamous cell carcinoma, must be detected. Our study

addresses this gap by using a more diverse dataset that covers multiple types of skin cancers, enhancing the model's versatility and clinical utility. Additionally, by emphasizing the reduction of mean squared error in processed images, we further enhance diagnostic accuracy through innovative pruning techniques.

According to the research by Medhat et al. [13], although iterative magnitude pruning (IMP) is employed to reduce model size, the method introduces a slight but critical drop in accuracy, which could be detrimental in medical diagnostics. This research improves upon this by integrating both magnitude-based weight pruning (MBWP) and quantization, a combination that allows us to maintain high accuracy while further optimizing the model's computational efficiency. These techniques not only refine the model's performance but also ensure scalability for deployment in resource-limited environments like rural clinics, without sacrificing diagnostic precision.

Additionally, Sharmila et al. [14] and Emuoyibofarhe and Ajisafe [16] emphasized computational efficiency for deployment on edge devices, but they overlook key aspects like real-time clinical integration and model interpretability. Our approach not only optimizes for edge devices through advanced pruning and quantization techniques, but also ensures interpretability, allowing healthcare professionals to trust and understand the model's predictions. This focus on practical deployment and transparency makes this research more viable for real-world clinical applications.

As reported by Wei et al. [17], despite the use of lightweight architectures for skin cancer detection, the study is limited by a small and imbalanced dataset, as well as a lack of cross-validation on diverse datasets. This research addressed this issue by using a larger and more diverse dataset, improving generalization across different skin types and conditions. To directly address the imbalance in class distribution, we implemented an oversampling technique using the imbalanced-learn library's RandomOverSampler to generate synthetic samples for the minority classes, ensuring balance across all skin cancer types. This step was crucial for improving model accuracy and generalization, particularly in medical datasets. Moreover, by combining MBWP with quantization, we reduce the computational load, making this model highly efficient and deployable in resource-constrained environments. While Wei et al. [17] focused on feature discrimination, this study enhances both interpretability and performance, ensuring the model is suitable for clinical adoption.

Lastly, in the study conducted by Yan et al. [15], while there is a strong focus on mitigating confounding factors with human-in-the-loop methodologies, the study does not address the need for computational efficiency, particularly in resource-constrained settings. This model achieves both efficiency and interpretability, streamlining the diagnostic process and minimizing reliance on continuous human intervention, thus offering a scalable solution for widespread clinical use. The integration of quantization further reduces inference and retraining time, making the model more suitable for real-time diagnostics.
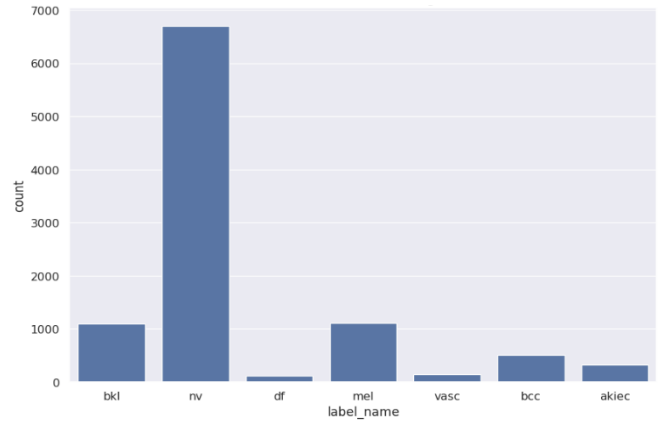
3. METHODOLOGY

3.1 Data collection and initial analysis

At the foundation of this investigation, a dataset comprising

10,015 dermatoscopic images was gathered from a renowned dermatological center. This dataset includes images from various skin cancer types, specifically: 6,705 images of Nevus, 1,113 images of Melanoma, 1,099 images of Keratosis, 514 images of Basal cell carcinoma, 327 images of Squamous cell carcinoma, 142 images of Vascular lesions, and 115 images of Dermatofibroma. This diverse distribution of images, as shown in Table 4, ensures that the model is trained on a comprehensive representation of skin cancer types, increasing its robustness and generalization capabilities. The significant number of Nevus images reflects the common nature of this benign condition, while the lower number of more severe conditions such as squamous cell carcinoma highlights the rarity but critical importance of detecting these malignant cases. This collection offers a significant diversity of skin conditions, crucial for the training of a Convolutional Neural Network (CNN) aimed at diagnosing skin cancer. The chosen dermatoscopic images utilize detailed visualization to allow for a more accurate diagnosis and enhanced model generalization.

Figure 1. The dataset's distribution of the various types of skin cancer

An in-depth exploration of the associated metadata for the 10,015 dataset images was conducted. This metadata, including critical information such as diagnosis ('dx'), diagnosis type ('dx_type'), age ('age'), sex ('sex'), and lesion localization ('localization'), provides a comprehensive understanding of the dataset's composition. This initial assessment ensures the training of a model that accurately reflects clinical realities, you can see the distribution on Table 4 and Figure 1.

Count plots visualization of the metadata will offer a comprehensible distribution of each categorical variable. Such visual aids facilitate the identification of any imbalances or patterns in the dataset, influencing preprocessing decisions and the interpretation of the model's outcomes, akin to the methodologies referenced in the related works [18].

Table 4. The dataset's distribution of the various types of skin cancer

Skin Cancer Class	Image Quantity
Nevus	6705
Melanoma	1113
Keratosis	1099
Basal cell carcinoma	514
Squamous cell carcinoma	327
Vascular lesions	142
Dermatofibroma	115

Table 5. Design parameters of base CNN

Step	Description	Dimensions Before	Dimensions After	Percentage of Data
Normalization	Pixel scale to [0,1]	10015×2352	10015×2352	100%
Reshaping for CNN	Adjust for CNN (28×28×3)	10015×2352	10015×28×28×3	100%
Data Split	Training and testing sets	N/A	N/A	Training 80%, Testing 20%

3.2 Preprocessing and oversampling of images

Preprocessing the dermatoscopic images is a crucial step to ready the data for the deep learning model. Employing libraries like Pandas and Numpy, image normalization was executed, standardizing inputs to aid the network's learning process. Given the uniform size of 28×28 pixels, no additional resizing was needed, streamlining the process.

The 2352 RGB pixel values of each image were normalized through division by 255, resulting in a scaling to a range of [0,1] that conforms to industry standards [19]. The data was subsequently restructured in accordance with the four-dimensional tensor of the CNN models. The data were partitioned into training and validation sets by applying an 80/20 division, as illustrated in Table 5. Training comprised 80% of the set, while validation comprised the remaining 20%. The utilization of a sequential methodology during the preprocessing stage effectively optimizes the data to enhance the functionalities of deep learning models.

To address the imbalance in skin cancer diagnosis class distribution, an oversampling technique was implemented. This step is vital in data preparation, particularly in medical datasets where balance across classes is crucial for model accuracy and generalization. Using the imbalanced-learn library's RandomOverSampler, synthetic samples of minority classes were generated to match the majority class's sample count [20], this can be seen on Table 6.

Table 6. Results after image oversampling

Type of Skin Cancer	Quantity of Images Prior	Quantity of Images After
Nevus	6705	5367
Melanoma	1113	5367
Keratosis	1099	5367
Basal cell carcinoma	514	5367
Squamous cell carcinoma	327	5367
Vascular lesions	142	5367
Dermatofibroma	115	5367

3.3 Design of the base Convolutional Neural Network model

After thorough preprocessing and optimization of the dataset, ensuring the highest quality of dermatological images, we moved to a pivotal phase in our research: constructing a robust base CNN model. Designing and initially training a base CNN model before implementing pruning techniques was essential to establish a benchmark for future evaluations. An effective base model is crucial to understand the inherent capabilities of the network before introducing the complexities of pruning. This allows us to capture the network's predictive essence regarding dermatological images and identify areas prone to efficiency improvements through pruning.

The base CNN architecture was selected not only to balance diagnostic accuracy, computational efficiency, and scalability for deployment in resource-constrained environments but also to provide greater control and explainability, essential in a medical setting. While pre-trained deep models like ResNet and DenseNet offer strong accuracy, their complexity, combined with high memory and computational requirements, limits their adaptability for real-time use on mobile devices or in clinical applications where transparency and control over model behavior are crucial.

By designing a custom CNN architecture, we gained the flexibility to continuously experiment and fine-tune the network to meet specific requirements. Custom models allow for better interpretability, as the inner workings of each layer and feature map can be fully understood and adjusted based on the medical context. This is particularly important in skin cancer detection, where clinicians need confidence in how the model arrives at its predictions.

Furthermore, constructing a base CNN model after optimizing image quality through mean squared error reduction ensures that any efficiency improvements through pruning do not compromise diagnostic accuracy. This sequential approach ensures that diagnostic integrity and quality remain the utmost priority throughout the model optimization process. The model structure, starts with an input layer accepting 28×28 pixel images with three channels (RGB), reflecting the standardized nature of previously processed dermatological images.

Multiple convolutional layers followed by MaxPooling layers are implemented. The convolutional layers, with filter sizes ranging from 32 to 256, are designed to capture a hierarchy of visual features from the simplest to the most complex. The 'same' padding ensures that the spatial size of the outputs is preserved, allowing the network to learn rich representations without losing image edge information. In the convolutional layers, the network implements the he_normal initializer and the ReLU activation function to reduce computation time and mitigate the vanishing gradient problem. The inclusion of BatchNormalization layers after each convolutional block and before each dense layer is a critical decision, speeding up training, improving network stability, and reducing weight initialization sensitivity. Following feature extraction and dimensionality reduction by convolutional and MaxPooling layers, the network flattens the feature maps to transition to a sequence of dense layers.

Table 7. Information on the hyperparameters of the CNN model

Parameter	Value
Optimizer	'Adam'
Loss Function	'Categorical Crossentropy'
Batch Size	128
Epochs	25
ReduceLROnPlateau - Monitor	5.50
ReduceLROnPlateau - Patience	2
ReduceLROnPlateau - Factor	0.5
ReduceLROnPlateau - Min LR	0.00001
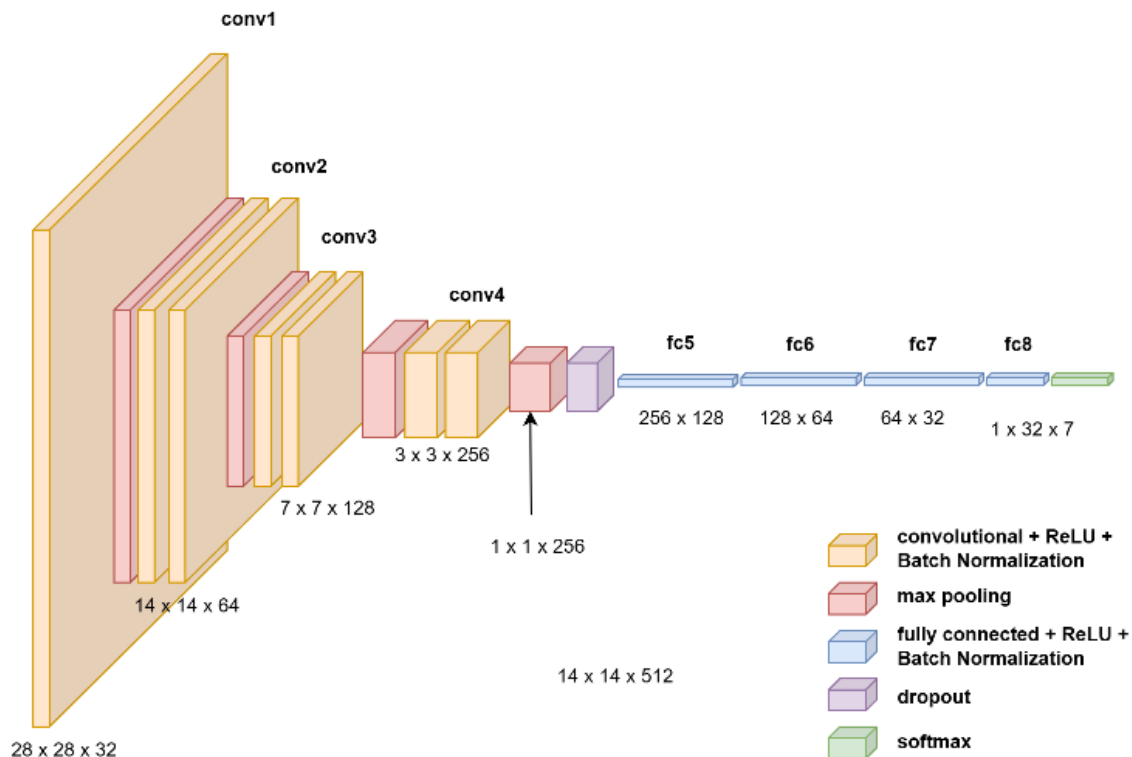
Figure 2. CNN architecture in detail

These layers, decreasing units from 256 to 32, act as classifiers learning complex patterns from extracted features. Dropout and L1 L2 regularization before dense layers prevent overfitting, ensuring the model generalizes well to new images unseen during training. Finally, the Softmax activation in the output layer distributes probability across the seven potential skin cancer diagnostic classes. Table 7 presents the hyperparameters selected for training the CNN model, with the 'adam' optimizer chosen for its efficiency in rapid convergence and automatic learning rate step size handling. The 'CategoricalCrossentropy' loss function suits multi-class classification problems like skin cancer detection by measuring the model's performance in assigning correct probability to the true label, the complete architecture can be seen on Figure 2.

3.4 Implementation and evaluation of pruning techniques - MBWP + Quantization

The execution of Magnitude Based Weight Pruning (MBWP) to optimize the CNN model adhered to a methodical strategy in order to decrease the model's dimensions and enhance performance while maintaining a high level of accuracy. This process was carried out using the TensorFlow Model Optimization Toolkit's sparsity module, which provides necessary tools for applying pruning techniques to TensorFlow and Keras models. The pruning parameters were defined using sparsity.PolynomialDecay, establishing a pruning schedule starting with an initial sparsity of 85%, increasing to a final sparsity of 95% from step 2000 to 5000, applied at 100-step intervals. This strategy aimed to significantly reduce the number of active model parameters, promoting a lighter model faster in inference.

For Quantization on the CNN, a post-training quantization approach was adopted, effectively reducing the model size and potentially speeding up inference. An optimization technique for deep learning models alters the precision of the number representations in the weights and, at times, activations. Specifically, it converts high-precision forms (such as 32-bit floating points) to 16-bit integers. The procedure was executed utilizing TensorFlow Lite's TFLiteConverter, which is specifically designed for deploying machine learning models on mobile and periphery devices. The pre-trained Keras model was converted to a TensorFlow Lite-compatible format, ready for quantization. The command converter.optimizations = [tf.lite.Optimize.DEFAULT] applied TensorFlow Lite's default quantization, balancing performance and precision. The quantized model was saved as a.tflite file, representing the model in an optimized format for resource-limited devices. TensorFlow Lite's interpreter was used to evaluate the quantified model's accuracy within the same Python environment.

In this project phase, the combination of two advanced optimization techniques, MBWP and Quantization, sought to improve the efficiency of the CNN model for image classification tasks. The goal was to find an optimal balance between model size reduction, inference acceleration, and precision preservation. Beginning with the sparsity-enhanced model, TensorFlow Lite was then leveraged to fine-tune the model's efficiency with 8-bit Quantization. This step marked a reduction from the previous 16-bit format, ensuring an even more compact model size suitable for deployment on resource-constrained devices. Throughout this process, the model was meticulously evaluated to ensure that precision remained intact. The TensorFlow Lite interpreter facilitated this evaluation within the Python environment, allowing for an assessment that mirrored real-world application conditions. In essence, the application of MBWP and Quantization underscored our commitment to a balanced optimization paradigm — where model size reduction and inferential speed were harmonized with the uncompromised precision of the CNN model. This innovative approach not only signified a stride towards resource-efficient AI deployment but also underscored the potential of machine learning in delivering accurate, clinical-grade diagnostic tools.

While implementing Magnitude-Based Weight Pruning (MBWP) and Quantization, several challenges were encountered, primarily related to maintaining model accuracy during the pruning process. As the sparsity level increased, a slight drop in accuracy was observed, particularly during the transition to higher sparsity levels (from 85% to 95%). To mitigate this, we applied a PolynomialDecay schedule, allowing gradual pruning and avoiding abrupt drops in performance. Another challenge was ensuring that Quantization, especially when moving from 16-bit to 8-bit, did not negatively impact the model's diagnostic precision. By leveraging TensorFlow Lite's default quantization optimizations, we were able to balance model size reduction and precision effectively. Careful post-quantization evaluations were carried out using the TensorFlow Lite interpreter to ensure real-world applicability (Table 8).

Table 8. Comparative table of the results of each pruning technique used

Technique	File Size (MB)	Accuracy	Efficiency Factor	% Decrease in Accuracy	% Size Reduction of File
Base Model	14.74	0.9864	1.00	-	-
MBWP	4.96	0.9772	2.97	< 1%	66%
Quantization (16-bits)	2.44	0.9866	6.04	0%	83%
MBWP + Quantization (8-bits)	1.24	0.9735	11.89	1.33%	92%

4. RESULTS

Upon reviewing the performance of the Convolutional Neural Network (CNN) prototype optimized with Magnitude Based Weight Pruning (MBWP) and Quantization, the comparative analysis revealed achievements in model efficiency without significant compromise to performance. While a minimal decrease in accuracy and F1-Score was observed, the reductions in file size and the boost in efficiency factor were remarkable. The model base of 14.74 MB was streamlined to an optimized prototype of just 0.39 MB, signifying a reduction in size by 97.36%. The optimized prototype achieved an efficiency factor of 37.79, suggesting potential advancements in storage and inference speed. This establishes the prototype as a feasible and exceptionally efficient solution for implementation in environments with limited resources. F1-Score, a harmonic measure of recall and precision, indicates a balance between the model's accuracy in

classifying positive instances and its avoidance of misclassifying negative instances. An F1-Score of 0.9812 in the Optimized Prototype denotes a highly effective model that maintains this equilibrium despite significant optimizations to reduce its size.

The performance of the ultimate CNN model was succinctly summed up in a single figure by employing the AUC (Area Under the Curve) metric. This metric evaluated the capability of the model to differentiate between positive and negative classes. An AUC score close to 1 indicates superior model performance with a high ability to differentiate between classes. Achieving an AUC of 0.99 underscored the exceptional capability of the model to make accurate classifications, suggesting high reliability in clinical contexts. This, coupled with a confusion matrix displaying an even distribution of correct classifications and minimal errors, validates the efficacy of the optimized CNN prototype. Integrating these metrics into the final discussion highlights the diagnostic accuracy of the model, its applicability in a real-world skin cancer diagnostic setting, and its potential to enhance patient satisfaction by reducing diagnostic waiting times and associated costs. Furthermore, the confusion matrix provides a valuable tool for dermatologists by offering a visual perspective on where the model excels and where it might require further improvements.

Assessment of overfitting and underfitting was conducted using learning curves, which demonstrated a healthy balance between bias and variance—crucial aspects to prevent underfitting and overfitting, respectively. The error curve illustrated a decreasing trend in the model's error rate during the training and validation phases over 25 epochs, indicating effective learning and improving precision as training progressed, the learning curves can be seen in Figure 3.

A rapid ascent to the upper left corner of the ROC curve for the final model indicated that it possessed a high degree of sensitivity and specificity. A value of 1.00 for the Area Under the Curve (AUC) indicates exceptional discriminatory ability. The substantial disparity between the ROC curve and the performance of a random classifier (illustrated by the dashed blue line) serves to emphasize the model's superiority. The proximity of the ROC curve to the upper-left vertex, as illustrated in Figure 4, indicates a significant proportion of true positives relative to false positives. This characteristic is especially advantageous in medical contexts such as skin cancer diagnosis, where precision is of the utmost importance.

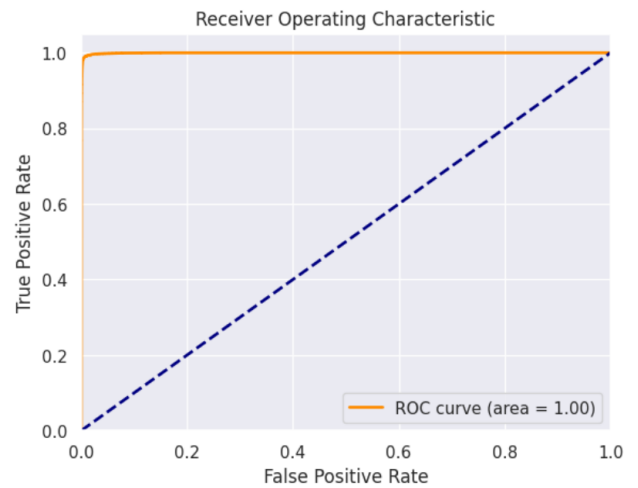**Figure 4.** Loss curves for CNN model

Figure 5. Cavity geometry

A rapid ascent to the upper left corner of the ROC curve for the final model indicated that it possessed a high degree of sensitivity and specificity. A value of 1.00 for the Area Under the Curve (AUC) indicates exceptional discriminatory ability. The substantial disparity between the ROC curve and the performance of a random classifier (illustrated by the dashed blue line) serves to emphasize the model's superiority. The proximity of the ROC curve to the upper-left vertex, as illustrated in Figure 5, indicates a significant proportion of true positives relative to false positives. This characteristic is especially advantageous in medical contexts such as skin cancer diagnosis, where precision is of the utmost importance.

5. ETHICAL CONSIDERATIONS

The integration of AI models into clinical settings, particularly for skin cancer diagnosis, raises important ethical considerations that must be addressed to ensure transparency, accountability, and patient trust. Transparency in the model's decision-making process is essential in a medical context. To ensure this, the model is designed with interpretable outputs, such as saliency maps and heatmaps, which visually highlight areas of the image that the model deems high-risk. These transparent outputs allow dermatologists to understand the reasoning behind the model's predictions, making it easier for

Figure 3. Accuracy curves for CNN model

them to trust the system and integrate it into their diagnostic workflows.

Accountability is also a key concern, as AI models assist in making potentially life-impacting decisions. While this model demonstrates high accuracy in predicting skin cancer, it is crucial that the final diagnostic responsibility remains with qualified dermatologists. The model is intended to function as a decision-support tool, providing supplementary information to assist clinicians rather than replace their judgment. This ensures that the ultimate responsibility for medical decisions stays with the human experts, preserving accountability in clinical practice.

Patient trust is crucial when introducing AI into healthcare. For AI systems to be fully accepted, patients need confidence in the diagnostic process. By providing interpretable outputs and clearly defining the model's role as an aid to clinicians, we aim to build this trust. Additionally, deploying the model in resource-limited settings can improve access to accurate diagnoses, reducing wait times and enhancing outcomes. The use of patient data, especially images, also raises concerns about privacy and security. All data used for training adheres to ethical guidelines and data protection laws, ensuring patient confidentiality throughout. The model operates within a secure infrastructure, safeguarding patient data during diagnosis and ensuring compliance with regulations.

6. CONCLUSIONS

The findings of this research illustrate the considerable capabilities of sophisticated Convolutional Neural Networks (CNN) when applied to dermatological imaging to precisely identify and categorize skin cancer. By incorporating cutting-edge pruning methods—Magnitude-Based Weight Pruning (MBWP) and Quantization—into this optimized CNN prototype, we have successfully struck a remarkable equilibrium between precision and efficiency. Notably, the prototype maintained a high F1-Score of 0.9812, reflecting the model's balanced precision and recall despite substantial reductions in model size and complexity.

Employing a comprehensive dataset of 10,015 dermatoscopic images, the model underwent rigorous preprocessing, including normalization and oversampling, to address class imbalances and enhance representational learning. The systematic application of data augmentation techniques played a pivotal role in ensuring a broad and diversified training scope, critical for a model's ability to generalize across various skin conditions.

The optimized CNN prototype marks a leap forward in medical diagnostic tools, evidencing a 97.36% reduction in model size which translated to enhanced operational efficiency without a significant compromise on diagnostic accuracy. The model exhibits exceptional discriminatory capability, as evidenced by its Area Under the Curve value of 0.99. This substantiates the model's high reliability when applied in clinical environments.

Furthermore, the model's ability to detect multiple types of skin cancer, including basal cell carcinoma, squamous cell carcinoma, and melanoma, makes it a versatile tool for dermatologists. Its high level of accuracy, combined with clear output explanations, allows for seamless integration into diagnostic workflows. This model could be integrated into digital dermatoscope devices, enabling real-time analysis of skin lesions during patient consultations. A dermatologist could capture a dermoscopic image, which would be immediately processed by the model. The output, in the form of a probability distribution across different cancer types, along with visual explanations (e.g., highlighting areas of concern on the image), could be directly displayed on a monitor or mobile device. This workflow allows the dermatologist to cross-reference the model's prediction probabilities with their own clinical observations, facilitating a faster decision-making process. The explainable nature of the model, especially when paired with saliency maps or heatmaps, ensures that the clinician understands why certain areas of the lesion are flagged as high-risk. This adds an extra layer of trust and transparency, making it easier for dermatologists to confidently use the model's output as a decision support tool rather than solely relying on automated predictions. Once deployed, it could be integrated into existing Electronic Health Record (EHR) systems, where it would flag lesions requiring follow-up or further testing. This would help streamline patient management.

For patients, the model's ability to provide early detection of various skin cancer types could significantly improve treatment outcomes, particularly in rural and underserved areas where access to specialized diagnostic tools is limited. By offering a scalable and efficient solution, the model has the potential to reduce mortality rates associated with late-stage skin cancer diagnoses, ultimately improving public health outcomes in both urban and remote regions.

REFERENCES

[1] Glenister, K., Witherspoon, S., Crouch, A. (2022). A qualitative descriptive study of a novel nurse-led skin cancer screening model in rural Australia. BMC Health Services Research, 22(1): 1019. https://doi.org/10.1186/s12913-022-08411-6

[2] Calniquer, G., Khanin, M., Ovadia, H., Linnewiel-Hermoni, K., Stepensky, D., Trachtenberg, A., Sedlov, T., Braverman, O., Levy, J., Sharoni, Y. (2021). Combined effects of carotenoids and polyphenols in balancing the response of skin cells to UV irradiation. Molecules, 26(7): 1931. https://doi.org/10.3390/molecules26071931

[3] Abdelaziz, A., Mahmoud, A.N. (2022). Skin cancer detection using deep learning and artificial intelligence: Incorporated model of deep features fusion. Fusion: Practice and Applications, 8(1): 8-15. https://doi.org/10.54216/fpa.080201

[4] Gouda, W., Sama, N.U., Al-Waakid, G., Humayun, M., Jhanjhi, N.Z. (2022). Detection of skin cancer based on skin lesion images using deep learning. Healthcare, 10(7): 1183. https://doi.org/10.3390/healthcare10071183

[5] Mazhar, T., Haq, I., Ditta, A., Mohsan, S.A.H., Rehman, F., Zafar, I., Gansau, J.A., Goh, L.P.W. (2023). The role of machine learning and deep learning approaches for the detection of skin cancer. Healthcare, 11(3): 415. https://doi.org/10.3390/healthcare11030415

[6] Kitchenham, B., Brereton, O.P., Budgen, D., Turner, M., Bailey, J., Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. Information & Software Technology, 51(1): 7-15. https://doi.org/10.1016/j.infsof.2008.09.009

[7] Pratiwi, R.A., Nurmaini, S., Rini, D.P., Rachmatullah, M.N., Darmawahyuni, A. (2021). Deep ensemble

learning for skin lesions classification with convolutional neural network. IAES International Journal of Artificial Intelligence, 10(3): 563. https://doi.org/10.11591/ijai.v10.i3.pp563-570

[8] Alfi, I.A., Rahman, M.M., Shorfuzzaman, M., Nazir, A. (2022). A Non-Invasive Interpretable diagnosis of melanoma skin cancer using deep learning and ensemble stacking of machine learning models. Diagnostics, 12(3): 726. https://doi.org/10.3390/diagnostics12030726

[9] Pitchiah, M.S., Rajamanickam, T. (2022). Efficient feature based melanoma skin image classification using machine learning approaches. Traitement du Signal, 39(5): 1663-1671. https://doi.org/10.18280/ts.390524

[10] Rachmad, A., Fuad, M., Rochman, E.M.S. (2023). Convolutional neural network-based classification model of corn leaf disease. Mathematical Modelling of Engineering Problems, 10(2): 530-536. https://doi.org/10.18280/mmep.100220

[11] Olayiwola, J.O., Badejo, J.A., Okokpujie, K., Awomoyi, M.E. (2023). Lung-related diseases classification using deep convolutional neural network. Mathematical Modelling of Engineering Problems, 10(4): 1097-1104. https://doi.org/10.18280/mmep.100401

[12] Lahouaoui, L., Abdelhak, D., Abderrahmane, B., Toufik, M. (2022). Image classification using a fully convolutional neural network CNN. Mathematical Modelling of Engineering Problems, 9(3): 771-778. https://doi.org/10.18280/mmep.090325

[13] Medhat, S., Abdel-Galil, H., Aboutabl, A.E., Saleh, H. (2024). Iterative magnitude pruning-based light-version of AlexNet for skin cancer classification. Neural Computing and Applications, 36(3): 1413-1428. https://doi.org/10.1007/s00521-023-09111-w

[14] Sharmila, B., Santhosh, H., Parameshwara, S., Swamy, Baig, W.U., Nanditha, S.V. (2023). Optimizing deep learning networks for edge devices with an instance of skin cancer and corn leaf disease dataset. SN Computer Science, 4(6): 793. https://doi.org/10.1007/s42979-023-02239-5

[15] Yan, S., Yu, Z., Zhang, X., Mahapatra, D., Chandra, S. S., Janda, M., Soyer, P., Ge, Z. (2023). Towards trustable skin cancer diagnosis via rewriting model's decision. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, Canada, pp. 11568-11577. https://doi.org/10.1109/cvpr52729.2023.01113

[16] Emuoyibofarhe, J., Ajisafe, D. (2020). Early skin cancer detection using deep convolutional neural networks on mobile smartphone. International Journal of Information Engineering and Electronic Business, 12(2): 21-27. https://doi.org/10.5815/ijieeb.2020.02

[17] Wei, L., Ding, K., Hu, H. (2020). Automatic skin cancer detection in dermoscopy images based on ensemble lightweight deep learning network. IEEE Access, 8: 99633-99647. https://doi.org/10.1109/access.2020.2997710

[18] Imran, A., Nasir, A., Bilal, M., Sun, G., Alzahrani, A., Almuhaimeed, A. (2022). Skin cancer detection using combined decision of deep learners. IEEE Access, 10: 118198-118212. https://doi.org/10.1109/access.2022.3220329

[19] Innani, S., Dutande, P., Baheti, B., Baid, U., Talbar, S.N. (2023). Deep learning based novel cascaded approach for skin lesion analysis. In Communications in Computer and Information Science, pp. 615-626. https://doi.org/10.1007/978-3-031-31407-0_46

[20] El-Ghany, S.A., Ibraheem, M.R., Alruwaili, M., Elmogy, M. (2021). Diagnosis of various skin cancer lesions based on fine-tuned RESNet50 deep network. Computers, Materials & Continua, 68(1): 117-135. https://doi.org/10.32604/cmc.2021.016102