# Deep Learning-based Skin Lesion Segmentation and Classification

Ayesha Rani[1], Haseeb Ullah[1], Hafiza Atika Shahab[1], Amaad Khalil[1], M. Abeer Irfan[1], Yaser Ali Shah[2],
[1]Computer System Engineering (University of Engineering and Technology Peshawar)
[2]Department of Compute Science (COMSATS University Islamabad, Attock Campus)
*Correspondence: ayesharani6788@gmail.com, haseebullahbj@gmail.com, atikashahab4455@gmail.com, amaadkhalil@uetpeshawar.edu.pk, abeer.irfan@uetpeshawar.edu.pk, Yaser@cuiatk.edu.pk

**Citation |** Rani. A, Ullah. H, Shahab. H. A, Khalil. A, Irfan. M. A, Shah. Y. A, "Deep Learning-based Skin Lesion Segmentation and Classification", IJIST, Special Issue pp 180-188, May 2024 DOI: https://doi.org/10.33411/IJIST/778
**Received |** May 04, 2024, **Revised |** May 16, 2024, **Accepted |** May 20, 2024, **Published |** May 26, 2024.

By using deep learning to automate skin lesion segmentation, this work aims to improve the classification of melanoma. By properly segmenting lesions and utilizing the U-Net algorithm's preprocessing capabilities, our research aims to improve the accuracy of skin cancer diagnosis. During preprocessing, raw dermoscopic pictures from the HAM10000 dataset are enhanced and normalized early. Next, the U-Net model is used to accurately segment lesions. Advanced deep learning approaches are applied after segmentation segmented images are subjected to classification, such as Convolutional Neural Networks (CNN) and Vision Transformer (VIT) models. The VIT model demonstrated a high training accuracy of 0.94, indicating its effectiveness in learning from the training data. However, its validation and testing accuracies were at 0.73. The CNN model showed a training accuracy of 0.95, implying its ability to learn the training data effectively. However, its validation and testing accuracies were at 0.73. This all-encompassing method not only improves dermatological image analysis's dependability and effectiveness, but it also shows promise for enhancing clinical outcomes in the diagnosis and management of different forms of skin cancer. Our work is a significant step toward the creation of more reliable techniques in this important area, opening the door for improvements in patient care and healthcare diagnostics.

**Keywords:** Melanoma, Skin cancer, lesion segmentation, deep learning, lesion classification, image processing, vision transformer.

**Introduction:**

Skin cancer originates in skin cells due to the uncontrollable growth of skin cells. Skin cancer is primarily caused by UV radiation damage to the cell's DNA, which is passed on by exposure to sunlight or artificial UV radiation sources. Melanoma, squamous cell carcinoma (SCC), and basal cell carcinoma (BCC) are the most common forms of skin cancer. According to [1], each year the number of cases of melanoma identified, increased by 53% between 2008 and 2018. If the skin cancer is detected later, the survival rate is less than 14%. Over the next decade, there will likely be an increase in the death rate. On the other hand, if skin cancer is diagnosed in the early stages, the survival rate will increase to 97%.

The increasing global prevalence of skin cancer highlights the urgent need for efficient techniques for both detection and categorization. With the advancement of technology, computer-aided diagnostic systems started to be used in the identification of skin cancer. Inspired by the critical need for detecting skin cancer in an early stage, scientists began to employ machine learning methods in unconventional ways [2]. Deep learning, a subset of AI (artificial intelligence), has been integrated into skin cancer detection in recent years, raising the bar for accuracy and efficiency in skin cancer diagnosis through automated, data-driven insights [3].

For the diagnosis of skin cancer segmentation is the most critical step for extracting features and isolating the skin lesion. The conventional method for classifying skin lesion images involves segmenting and pre-processing the image. After extracting characteristics from the region of interest, the lesion is classified using different classifiers [4]. The segmentation approach reduces the image processing steps. Seeja et al. [5] suggest that deep learning can also improve diagnostic efficiency by simplifying the process of interpreting images by using a model to extract representative features from lesions by combining the input image with a segmentation mask.

In this paper, we present a novel approach for the segmentation of skin cancer utilizing the deep convolutional neural network based on the U-Net algorithm in preprocessing. This work focuses on the segmentation and classification of different seven classes of skin cancer by employing the HAM10000 Dataset for training and evaluation. The experimental results showcase the effectiveness of our proposed method achieving the notable segmentation accuracy of seven classes of skin cancer. In the pre-processing step, we utilized the U-Net architecture for segmentation. For classification, we employed VIT and CNN architectures.

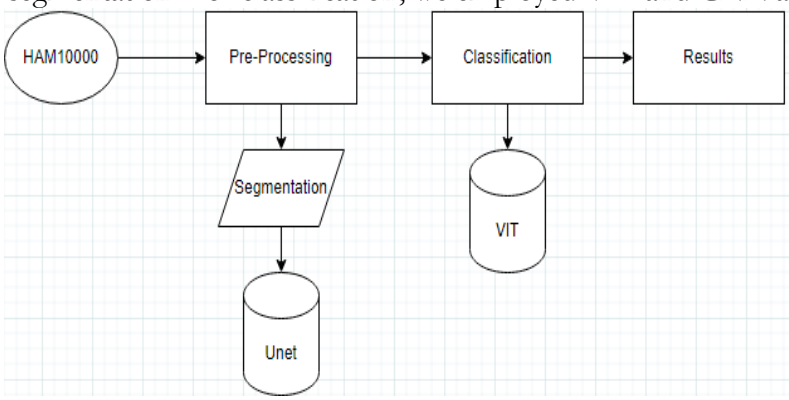



**Figure 1**: Flow chart of proposed methodology

**Related Work:**

Rafael Luz Ara´ujo et.al [6] proposed a segmentation method for melanoma skin lesions using modified U-net along with post-processing techniques. The research was conducted on two datasets, PH2 and DermIS involving acquisition and segmentation with a highly effective U-net network, and for further improvement, they performed post-processing techniques that disconnected extra regions, filled holes, and removed loose regions. In the PH2 dataset, they were able to acquire a dice coefficient of 0.933, and in the DermIS dataset, 0.872.

For the fast and accurate detection and segmentation of melanoma skin lesions, M. Taghizadeh et al. [7] proposed a method using SegNet and Yolov3 Based on Deep Transfer Learning. He suggests a two-phase procedure for melanoma detection. F-YoloV3 is applied for melanoma localization and F-SegNet for segmentation. Bisla et al. [8] suggested a deep learning model for both classification and segmentation for the diagnosis of skin lesions. The segmentation technique was to mask away portions of the image that weren't needed. He utilized the U-net architecture for segmentation. To improve the accuracy of the automated diagnosis of melanoma through deep learning image segmentation technique, Aleksandra Dzieniszeewska et al [4] combined the segmentation mask and skin lesion images using Gaussian blur. They employed the deeplabV3 and U-net segmentation networks for the segmentation process. On the combined ISC dataset, they obtained an accuracy of 84.85%. The classification performance of melanoma was enhanced by Seeja R D, and Suresh A [5] using deep learning-based automatic skin lesion segmentation. The process unfolds in three stages: Segmentation using U-Net, Feature Extraction (color, texture, shape) using HOG, LBP, Edge Histogram, and Gabor methods. , and Classification using SVM, Random Forest, K-NN and naïve Bayes. The SVM classifier yields the best result on the ISBI 2016 dataset based on F1-score and accuracy. For image segmentation, they achieved a Dice co-efficiency value of 77.59% and the SVM classifier produced 85.19 % accuracy. It is observed that classification with segmentation achieves much better accuracy, sensitivity, and specificity of the model compared to classification unsegmented images.

Mehwish Zafar et al. [9] proposed a segmentation model in which features are extracted through a pre-trained MobilebetV2 model. This model acts as a base of Deeplabv3+ for boundary extraction. Using the ISIC 2016, 2017, 2018, and PH2 datasets, the suggested segmentation strategy is assessed based on Mean Accuracy, Global Accuracy, BF Score, Weighted IoU, and Mean IoU. These metrics yield global accuracy values of 0.97481, 0.97297, 0.98642, and 0.95914, respectively. Nojus Dimša et al [10] suggested the automatic segmentation of skin lesions using deep learning, which explores the crucial field of melanoma diagnosis and highlights the important significance of early detection. MultiResUNet outperforms U-Net++ by a tiny amount (0.86%), all things considered, these U-Net variations show promise for improving the traditional U-Net model in skin lesion segmentation. Still, multi-class segmentation in skin lesions is a difficult field of study due to its complexity.

Researchers proposed a deep learning-based approach, comprising fuzzy k-means clustering (FKM) and region-based convolutional neural network (RCNN), to classify skin melanoma at an early phase [11]. The PH2, ISBI-2016, and ISIC-2017 datasets were utilized to evaluate the efficacy of the methodology that was offered. It beat existing state-of-the-art approaches, according to the data, with an average accuracy of 95.40%, 93.1%, and 95.6%. Mohammad Ali Kadampur et al [1] proposed a skin cancer detection method using Deep Learning Studio (DLS). Their approach involved data preparation, model construction, tuning, and deployment as a REST API. The DLS model achieved an exceptional AUC of 99.7% in skin cancer detection, demonstrating its effectiveness and ease of use. Ameri A [12] presented a robust deep-convolutional neural network utilizing AlexNet to classify skin lesions as benign or malignant. This model shows significant potential for assisting dermatologists in skin cancer detection with a classification accuracy of 84%, sensitivity of 81%, and specificity of 88%.

Using the Vision Transformer architecture, [13] research presents a unique method for classifying skin cancer that achieves an impressive 96.15% accuracy on the HAM10000 dataset. The work outperforms conventional deep learning techniques by utilizing pre-trained models like ViT patch-32, which has promising potential for dermatological diagnostics. The model's performance is further improved by using the Segment Anything Model for lesion segmentation, proving its usefulness in computer-aided skin cancer diagnosis. This study

demonstrates the noteworthy advancements in deep learning applications for the interpretation of medical images, especially in the field of dermatology, which enhances patient outcomes by enabling prompt and precise diagnosis.

**Material and Methods:**

In our approach, we utilize deep learning techniques for automatic skin lesion segmentation using the U-Net algorithm as a preprocessing step to improve the classification of melanoma. Convolutional neural network architecture U-Net performs exceptionally well in biomedical image segmentation applications [4][5][14]. We employ the HAM10000 Skin Cancer dataset, which consists of various classifications of skin lesions, including melanocytic nevi, melanoma, basal cell carcinoma, actinic keratoses, vascular lesions, and dermatofibroma. The methodology involves the following steps:

**Dataset:**

We use the HAM10000 dataset, which consists of 10,015 dermoscopic pictures with extensive clinical metadata, in our study. This collection includes a variety of skin lesion representations, including benign nevi, malignant melanoma, and other dermatological diseases [14]. The HAM10000 dataset makes it easier to train and assess machine learning models for automated skin lesion recognition and segmentation because of its diverse population sources and clinical contexts.

**Process:**

Some separate subtasks that are suited for various input skin image types comprise the classification process:

- **Unaltered Lesion Classification:** In this subtask, skin lesions are categorized without any segmentation or preprocessing. It acts as a reference point for comparing segmented lesion classifications.
- **U-Net Segmented Lesion Classification:** Lesions are automatically separated from input images by utilizing the U-Net segmentation model. The U-Net architecture is trained to precisely identify skin lesions and is well-known for its efficacy in biomedical image segmentation.

**Pre-Processing:**

Our preprocessing stage is used for segmentation tasks by preparing the raw dermoscopic pictures. By means of the given service, we analyze the data to remove biases, the adjustment of the intensity levels, and also improve the image quality. The delineation procedure is performed in the next step to draw a border around a skin lesion. We will before segregation and classification do certain preprocessing tasks so as to make an accurate and consistent process.

**Data Cleaning**

To ensure data integrity and dependability for our research, we used data cleaning techniques during the preprocessing step to correct mistakes, inconsistencies, and missing values in the dataset.

**Distribution of 7 Different Classes:**

Subsequently, we Analyze the dataset's distribution of the seven distinct groups of skin lesions to identify any class imbalances.

**Addressing Class Imbalance:**

To balance the distribution, we decide to employ both up-sampling and down-sampling methods. A more equal distribution of the classes resulted in down-sampling, in which randomly selected samples of samples from the majority class were used to trim the size of the minority class. Moreover, we ensured that each class of the minority samples was replicated using such methods as up-sampling techniques to the level of attaining the balance of the dataset and a proper representation of each class.

**Analysis of Spatial and Demographic Factors:**

We performed an analysis to understand the distribution of skin lesions in several localized fields (e.g., arms, legs, torso) as well as the demographic features (gender, age) of skin lesioned individuals. The goal of the comprehensive study was to find geographic patterns or variances in the distribution of lesions and demographic trends or correlations, offering significant data for the classification of skin lesions.

**Segmentation:**

A crucial preprocessing step in our work is segmentation, which is used to precisely identify skin lesions from dermoscopic pictures. Each of the numerous subtasks that make up the segmentation approach contributes to the process's improvement and optimization.

**Model Definition and Training:**

Our approach relies heavily on the U-Net segmentation model to precisely identify skin lesions from dermoscopic images. The U-Net architecture, which is well-known for its effectiveness in biomedical image segmentation, is described inside a function that enables the customization of parameters such as the number of epochs. By iteratively modifying its parameters throughout the training phase and utilizing the rich contextual data that its architecture captures, the U-Net model gains the ability to accurately recognize skin lesions.

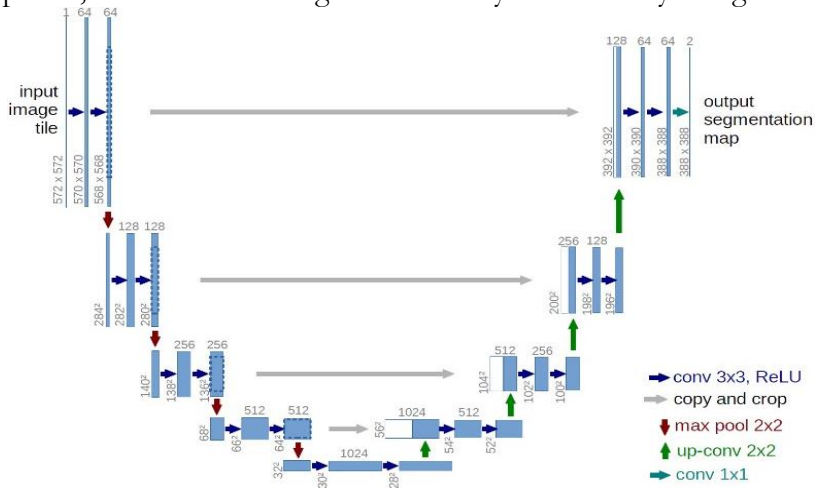



**Figure 2**: Illustration of U-Net Architecture [4]

**Data Loading and Preparation**

To provide consistency and structure in data handling, a function is defined to load the dataset in a sorted order. The dataset is divided into training and test sets to evaluate model performance impartially and prevent biases.

**Evaluation Metrics:**

As evaluation metrics, intersection over union (IoU), Jaccard Index, and dice coefficient are used to measure how accurate and similar segmented lesions are to ground truth annotations. The Jaccard Index, also known as the Jaccard similarity coefficient, is calculated as the ratio of the size of the intersection between the two sets to the size of their union. It is essentially the same as IoU but sometimes computed slightly differently. A statistic used to assess the similarity and diversity of sample sets is the Jaccard index, sometimes referred to as the Jaccard similarity coefficient and Intersection over Union. Like accuracy, the Dice score penalizes for false positives that the algorithm detects in addition to counting how many positives you find.

$$\text{Dice} = 2 * \frac{\text{tp}}{\text{tp} + \text{fp}} + (\text{tp} + \text{fn})$$

**Visualization and Post-Processing:**

To evaluate the implemented model for skin lesion segmentation, predictions are made on unknown data that have not been trained upon. We apply extra postprocessing approaches,

for example, to improve lesion boundary feature visibility and then to adjust segmentation mask projections afterward the segmentation process. The proposed method is also suitable for smoothing segmentation masks and increasing visual interpretation accuracy with imprecise lesion boundaries. Thus, these after-processing procedures make the skin cancer detection method based upon the accurate segmentation results both reliable and aesthetically appealing.

**Application of Masks:**

To precisely define skin lesions for subsequent classification tasks, we employed the segmentation masks produced by the U-Net model in the last segmentation stage. Precise lesion borders were provided by the segmentation masks, which made feature extraction and classification easier. the application of segmentation masks generated by the U-Net model for accurate delineation of skin lesions is illustrated in Figure 3.

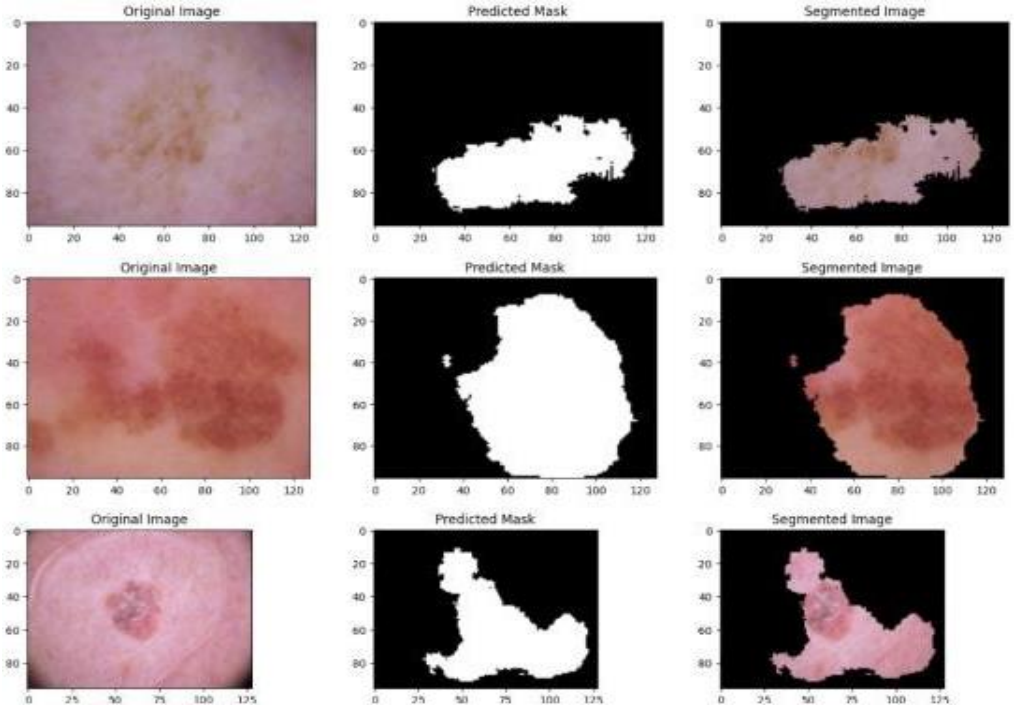



**Figure 3:** Predicted Mask and Segmented Images

**Classification Using VIT:**

Several separate segments were produced when we trained a model to segment VIT images. The VIT model was then trained to distinguish between melanoma and melanocytic nevi using these segments. Accuracy measurements displayed in tables and figures helped to clarify the findings. The segmentation-based classification strategy is shown to be reliable and successful by this thorough analysis. Significant progress toward improving our understanding and identification of skin lesions has been accomplished through this research, likely leading to improvements in the results of dermatological healthcare.

To train our classification model, segmented images were sent to a Vision Transformer (VIT) model. This method took advantage of VIT models' ability to efficiently handle image segmentation tasks. Our model was trained with the segmented representations to discriminate between two classes: melanoma and melanocytic nevi. This approach demonstrates a possible path toward increasing binary classification accuracy in medical image analysis by leveraging the advantages of both transformer-based models and segmentation. From this model, the training accuracy is 0.73, and the validation accuracy is 0.73.

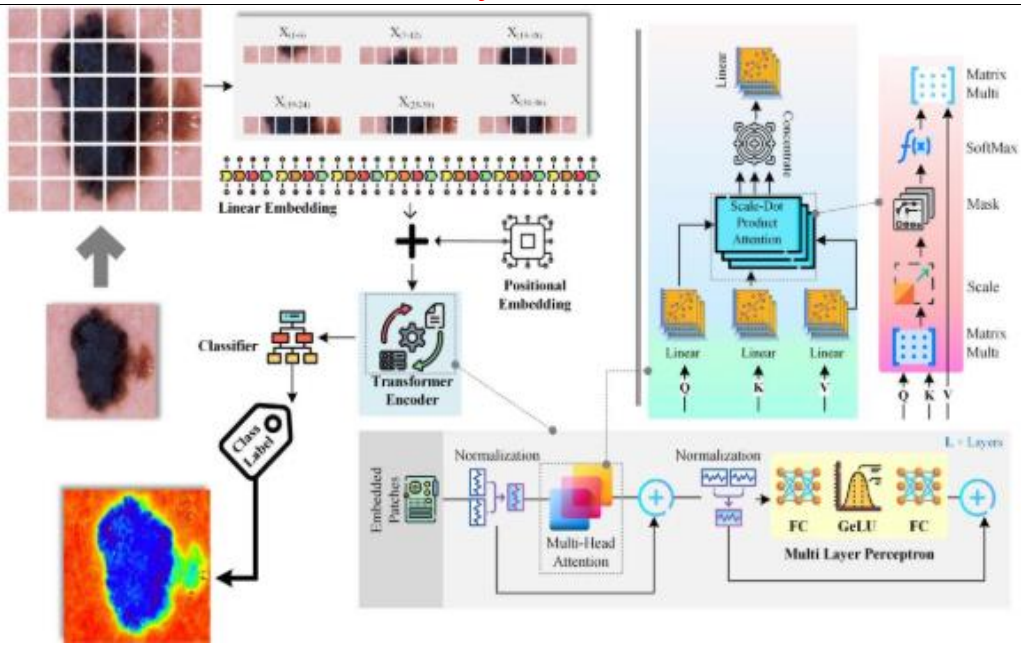



**Figure 4:** (Vision Transformer-based Skin Cancer Classification Model) [13]




**Figure 5:** (Training & validation loss)

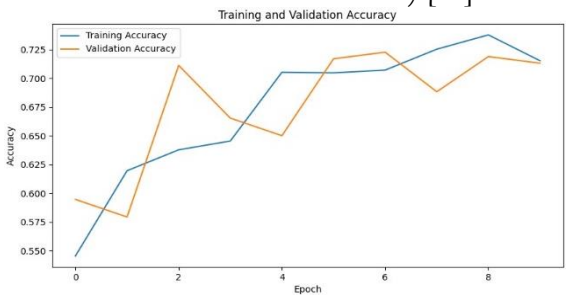



**Figure 6:** (Training & validation Accuracy)

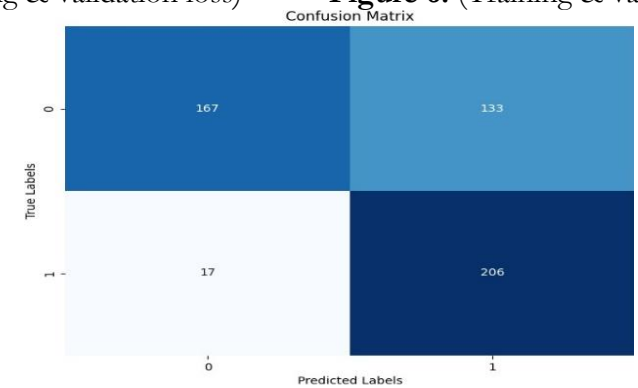



**Figure 7**: Confusion Matrix

The confusion matrix above Thirteen of the actual class one 167 images were wrongly classified as class 0 133 were accurately predicted as class 1 and 206 images that were truly class 0 were accurately classified, whereas 17 images were mistakenly classed as class 1.

**Classification Using CNN:**

Because of dataset imbalances, we trained a CNN model for a 7-class classification challenge using the segmented images. This all-inclusive strategy sought to address the data's unpredictability more successfully. A comprehensive table and figure below give the results of a thorough analysis of the resulting categories, together with pertinent metrics. By providing insights into how CNN models can handle unbalanced datasets and improve classification

accuracy across several classes, this approach makes a substantial contribution to the area of medical image analysis.

It uses pre-trained weights from ImageNet to initialize Mobile Net, leaving out the fully connected layers. It reduces geographic dimensions by adding a layer of global average pooling. We add a completely connected layer with ReLU activation and 256 units at the end. It includes SoftMax activation and an output layer with 7 units (for a classification problem with 7 classes). We build a new model with the custom classification layers and the input from the Mobile Net. Essentially, it adds unique layers for classification to modify the robust Mobile Net architecture for a particular picture classification application.

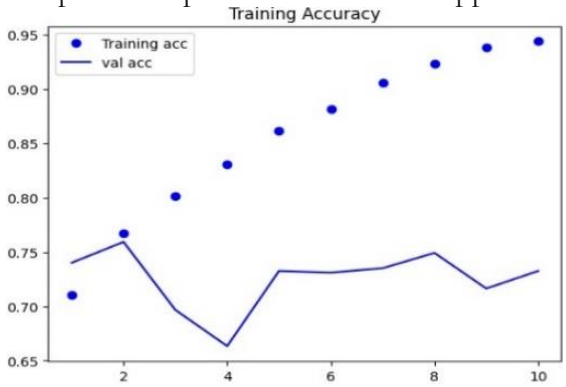



**Figure 8:** Training and Validation Accuracy

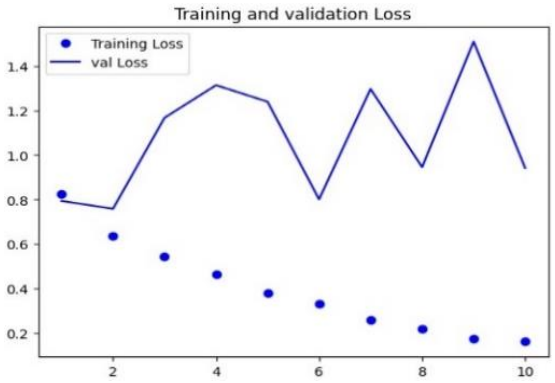



**Figure 9:** Training and Validation Loss

**Result and Discussion:**

The VIT model demonstrated a high training accuracy of 0.94, indicating its effectiveness in learning from the training data. However, its validation and testing accuracies were at 0.73. The VIT model still achieved a notable testing accuracy, indicating its capability to generalize well to unseen data.

**Table 1: Classification Using VIT Model**

| Training Accuracy | Validation Accuracy | Testing Accuracy |
|---|---|---|
| 0.73 | 0.71 | 0.71 |

On the other hand, the CNN model showed a comparable training accuracy of 0.95 to the VIT model, implying its ability to learn the training data effectively. However, its validation and testing accuracies were lower at 0.73 respectively.

**Table 2: Classification Using CNN Model**

| Training Accuracy | Validation Accuracy | Testing Accuracy |
|---|---|---|
| 0.94 | 0.73 | 0.73 |

**Conclusion:**

In conclusion, this work presents a method for skin lesion segmentation and classification using state-of-the-art deep learning techniques. Using Convolutional Neural Networks (CNN) and Vision Transformer (VIT) models for classification and the U-Net technique for segmentation, we have shown encouraging results in correctly classifying different kinds of skin lesions. The metrics acquired for segmentation accuracy verify the efficacy of our method in medical picture analysis. Our approach has a lot of potential to boost clinical outcomes by increasing the precision and dependability of skin cancer diagnosis. better work will concentrate on enlarging the classification task to support a wider variety of lesion classes, and continuous efforts will be made to better hone and optimize our methodology. This work represents a major advancement in the use of deep learning in dermatological analysis and emphasizes the need for ongoing research and development in this area.

**Acknowledgement:** Acknowledgments are considered necessary.

**Author's Contribution:** The corresponding author should explain the contribution of each co-author completely.

**Conflict of Interest:** Authors are advised to explain that there exists no conflict of interest for publishing this manuscript in IJIST.

**Project Details:** If this research was conducted as a result of a project, please give details like project number, project cost completion date, etc.

**References**:


[1] M. A. Kadampur and S. Al Riyaee, "Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images," Informatics Med. Unlocked, vol. 18, p. 100282, Jan. 2020, doi: 10.1016/J.IMU.2019.100282.

[2] M. A. Kassem, K. M. Hosny, R. Damaševičius, and M. M. Eltoukhy, "Machine Learning and Deep Learning Methods for Skin Lesion Classification and Diagnosis: A Systematic Review," Diagnostics 2021, Vol. 11, Page 1390, vol. 11, no. 8, p. 1390, Jul. 2021, doi: 10.3390/DIAGNOSTICS11081390.

[3] R. A. Mehr and A. Ameri, "Skin Cancer Detection Based on Deep Learning," J. Biomed. Phys. Eng., vol. 12, no. 6, p. 559, Dec. 2022, doi: 10.31661/JBPE.V0I0.2207-1517.

[4] A. Dzieniszewska, P. Garbat, and R. Piramidowicz, "Skin Lesion Classification Based on Segmented Image," 2023 12th Int. Conf. Image Process. Theory, Tools Appl. IPTA 2023, 2023, doi: 10.1109/IPTA59101.2023.10320004.

[5] R. D. Seeja and A. Suresh, "Deep Learning Based Skin Lesion Segmentation and Classification of Melanoma Using Support Vector Machine (SVM)," Asian Pac. J. Cancer Prev., vol. 20, no. 5, p. 1555, 2019, doi: 10.31557/APJCP.2019.20.5.1555.

[6] R. L. Araujo, R. D. A. L. Rabelo, J. J. P. C. Rodrigues, and R. R. V. E. Silva, "Automatic segmentation of melanoma skin cancer using deep learning," 2020 IEEE Int. Conf. E-Health Networking, Appl. Serv. Heal. 2020, Mar. 2021, doi: 10.1109/HEALTHCOM49281.2021.9398926.

[7] M. Taghizadeh and K. Mohammadi, "The Fast and Accurate Approach to Detection and Segmentation of Melanoma Skin Cancer using Fine-tuned Yolov3 and SegNet Based on Deep Transfer Learning," Oct. 2022, Accessed: May 06, 2024. [Online]. Available: https://arxiv.org/abs/2210.05167v2

[8] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, and D. Feng, "Dermoscopic Image Segmentation via Multistage Fully Convolutional Networks," IEEE Trans. Biomed. Eng., vol. 64, no. 9, pp. 2065–2074, Sep. 2017, doi: 10.1109/TBME.2017.2712771.

[9] M. Zafar, J. Amin, M. Sharif, M. A. Anjum, G. A. Mallah, and S. Kadry, "DeepLabv3+-Based Segmentation and Best Features Selection Using Slime Mould Algorithm for Multi-Class Skin Lesion Classification," Math. 2023, Vol. 11, Page 364, vol. 11, no. 2, p. 364, Jan. 2023, doi: 10.3390/MATH11020364.

[10] N. Dimša and A. Paulauskaitė-Tarasevičienė, "Melanoma multi class segmentation using different U-Net type architectures," 2021.

[11] M. Nawaz et al., "Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering," Microsc. Res. Tech., vol. 85, no. 1, pp. 339–351, Jan. 2022, doi: 10.1002/JEMT.23908.

[12] A. Ameri, "A Deep Learning Approach to Skin Cancer Detection in Dermoscopy Images," J. Biomed. Phys. Eng., vol. 10, no. 6, p. 801, 2020, doi: 10.31661/JBPE.V0I0.2004-1107.

[13] G. M. S. Himel, M. M. Islam, K. A. Al-Aff, S. I. Karim, and M. K. U. Sikder, "Skin Cancer Segmentation and Classification Using Vision Transformer for Automatic Analysis in Dermatoscopy-Based Noninvasive Digital System," Int. J. Biomed. Imaging, vol. 2024, 2024, doi: 10.1155/2024/3022192.

[14] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Sci. Data 2018 51, vol. 5, no. 1, pp. 1–9, Aug. 2018, doi: 10.1038/sdata.2018.161.




Copyright © by authors and 50Sea. This work is licensed under Creative Commons Attribution 4.0 International License.