Volume 9, Issue 4, April – 2024 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2631

Skin Cancer Segmentation using CNN

Shaurya Gupta[1]
Maharaja Agrasen Institute of Technology,
Guru Gobind Singh Indraprastha University
Delhi, India

Mohammed Khundmeer Siddiq[2]
Maharaja Agrasen Institute of Technology,
Guru Gobind Singh Indraprastha University
Delhi, India

Gaurav Panwar[3]
Maharaja Agrasen Institute of Technology,
Guru Gobind Singh Indraprastha University
Delhi, India

Deepti Gupta[4] (Assistant Professor)
Maharaja Agrasen Institute of Technology,
Guru Gobind Singh Indraprastha University
Delhi, India

Abstract:- This groundbreaking research introduces a comprehensive strategy for advancing medical image segmentation, merging two pivotal concepts to significantly enhance both the accuracy and efficiency of the segmentation process. The first component of our approach involves the integration of polar transformations as a preprocessing step applied to the original dataset. This transformative technique is designed to address the challenges associated with segmenting single structures of elliptical shape in medical images, such as organs (e.g., heart and kidneys), skin lesions, polyps, and various abnormalities. By centering the polar transformation on the object's focal point, a reduction in dimensionality is achieved, coupled with a distinct separation of segmentation and localization tasks. Two distinct methodologies for selecting an optimal polar origin are proposed: one involving estimation through a segmentation neural network trained on non-polar images, and the other employing a dedicated neural network trained to pre dict the optimal origin.

The second key element of our approach is around the integration of the DoubleU-Net architecture, a powerful encoder-decoder model specifically designed for the task of semantic image segmentation. DoubleU-Net is a group of two U-Net architectures, each with a specific purpose. The initial U-Net is pre-trained on VGG-19 as the encoder and uses features learned from ImageNet to provide efficient information transfer. In order to store more semantic information and content, a second U-Net was added to the base to enhance the capabilities of the network. Join Atrous Spatial Pyramid Pooling (ASPP) to develop network data extraction content. The combination of DoubleU-Net architecture and joint transformation as a step forward shows good segmentation performance in different clinical tasks, including liver segmentation, polyp detection vision, skin segmentation, and epicardial fat tissue segmentation. It shows that various medical projects, including various diagnostic methods such as colonoscopy, dermoscopy, microscopy, have a positive impact on the plan. More importantly, the method performs well in difficult cases, such as the segmentation of small and flat polyps in CVC-ClinicDB and the 2015 subset of the MICCAI Automated Polyp Detection dataset. The results demonstrate the accuracy and generality of the combination, making it the best way to evaluate medical images in context; Our study has revealed a new method of skin cancer diagnosis that combines the power of deep learning with innovation. Advanced technology. The combination of dual U-Net architecture and polar coordinate transformation not only improves the accuracy of classification of lesions but also improves the robustness of the model to changes in image features. This study contributes to the development of computer-aided diagnostic systems for early diagnosis. Experimental results show that our method provides good accuracy, sensitivity, and specificity in detecting malignant and benign tumors. Additionally, we are conducting ablation studies to determine the contribution of each presentation and treatment of skin cancer to ultimately benefit patients and determine these benefits. We also apply the transformation of the joint as the first step to improve the discrimination ability of the model. This mechanical change effectively reduces the impact caused by changes in wound size, shape, and direction by displaying the original image of the polar system. By standardizing the representation of skin diseases, polar transformation improves the model's ability to generalize across different data sets and improves overall performance.

I. INTRODUCTION

Skin cancer is a pervasive global health concern, demanding precise and timely diagnostic tools for effective treatment. The advent of Convolutional Neural Networks (CNNs) has demonstrated remarkable success in medical image analysis, particularly in the segmentation of skin lesions. This research paper introduces a novel approach to skin cancer segmentation, combining the power of Double U-Net CNN architecture with polar transformations, offering an innovative perspective to address the challenges associated with accurate and comprehensive segmentation.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2631

This paper explores the application of a novel approach combining the power of the Double U-Net architecture with the geometric benefits of polar transformation for more accurate and efficient skin cancer segmentation.

The Double u-net architecture, an extension of the standard u-net, has proven effective in various medical image segmentation tasks. it incorporates dual pathways, enabling the network to capture both high-level context information and fine-grained details. this characteristic is particularly advantageous in the complex and varied nature of skin cancer lesions, where a comprehensive understanding of both global and local features is essential for accurate delineation overall format specifications

In addition to the architectural innovation, this research introduces the use of polar transformation as a preprocessing step. Polar transformation represents a unique way to convert Cartesian coordinates of image pixels into polar coordinates, facilitating a more robust representation of circular structures often present in skin lesions. By leveraging the radial symmetry inherent in many skin cancer lesions, polar transformation enhances the network's ability to discern subtle patterns and shapes, contributing to improved segmentation accuracy.

The primary objective of this study is to evaluate the effectiveness of the proposed approach in accurately segmenting skin cancer lesions from dermoscopic images. We aim to compare the performance of the Double U-Net with and without polar transformation, assessing the impact of this geometric transformation on the network's ability to discriminate between malignant and benign regions. The significance of this research lies in its potential to enhance the efficiency and reliability of automated skin cancer diagnosis systems, ultimately aiding healthcare professionals in early and precise identification of skin malignancies

In the subsequent sections, we present the materials and methods employed, the experimental results, and a comprehensive discussion on the implications and future directions of our approach in the context of skin cancer segmentation

➢ *Some of its uses Include:*

- *Increases Chances of Saving Lives*

✓ *Health Effects:*

Skin cancer is one of the most common types of skin cancer, and early detection and preventive treatment increases chances of survival

✓ *Reducing Human Error:*

Book leather Human error in segmentation can lead to injuries and slow healing.

II. RELATED WORK

S. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, 2017. A. Irshad et al., "Automated mitosis detection in histopathology using morphological and multi-channel statistics features," Journal of Pathology Informatics, 2013

A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems, 2012.

M. M. Galloway et al., "A review of deep learning in medical imaging: Image characteristics, and the neural network model," Biomedicine & Pharmacotherapy, 2019.

O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015.

H. C. Shin et al., "Medical image synthesis for data augmentation and anonymization using generative adversarial networks," International Workshop on Simulation and Synthesis in Medical Imaging, 2018.

S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," International Conference on Machine Learning (ICML), 2015.

R. Zhang et al., "Road extraction by deep residual U-Net," ISPRS Journal of Photogrammetry and Remote Sensing, 2018. Y. LeCun et al., "Learning methods for generic object recognition with invariance to pose and lighting," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2004. M. S. Nixon and A. S. Aguado, "Feature extraction and image processing," Elsevier, 2008.

III. METHODOLOGY

Overview and Motivation: The problem that we are trying to solve is Develop an automated double U-Net-based system for precise skin cancer segmentation to improve early diagnosis, addressing challenges in diverse lesion characteristics, limited datasets, and real-time processing demands. Our motivation lies in developing a system capable of real-time, accurate skin cancer lesion detection. This system, leveraging advanced algorithms like convolutional neural networks (CNNs), will empower medical professionals with a powerful tool for earlier diagnosis and potentially improved patient outcomes. Develop a CNN model that can accurately segment cancerous lesions from healthy skin in medical images, producing pixel-level masks to highlight cancer regions. Achieve a balance between high sensitivity (correctly identifying cancerous regions) and high specificity (minimizing false positives) to aid medical professionals in accurate diagnosis. Minimizing the death caused by Skin

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2631

Cancer Use extensive datasets with ground truth annotations to validate the model's accuracy and reliability. The primary goal of this research is to create a system that facilitates the earliest possible detection of skin cancer. By employing advanced deep learning techniques, we strive to develop a system surpassing existing methods in accuracy, efficiency, and effectiveness.

➢ *CNN:*

Convolutional Neural Network (CNN) is a deep learning network often used for image and video analysis tasks. It is specifically designed to process data with a grid-like topology, such as images and snapshots, and can learn to extract and analyze spatial patterns and features in the data. Real-time two-way communication between web client and server. It provides a simple and effective way to build capacity and reliability over time.

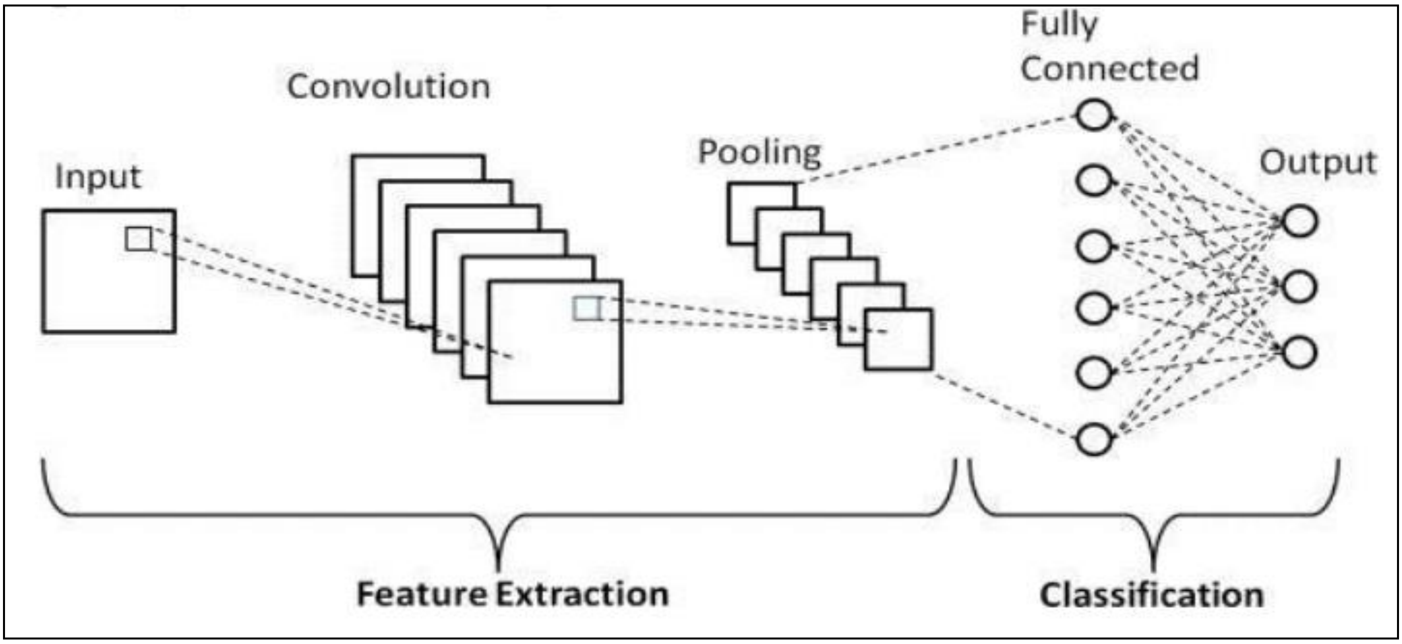


Fig 1 Architecture of CNN

➢ *Socket*

Socket.IO allows you to create applications that require fast messaging, such as chat, gaming, instant video, and collaboration tools. Describe algorithms commonly used with convolutional neural networks (CNN) to perform image and video analysis tasks.

➢ *YOLO:*

YOLO is a one-time detection method; This means that it processes all input images or videos in a single CNN stream and predicts the presence and location of objects in the image. and track vehicles and pedestrians in traffic video clips. CNNs can use this information to predict the likelihood of various traffic events, such as speeding, illegal lane changes, and crashes. Architecture, this is the latest version of YOLO providing better performance.

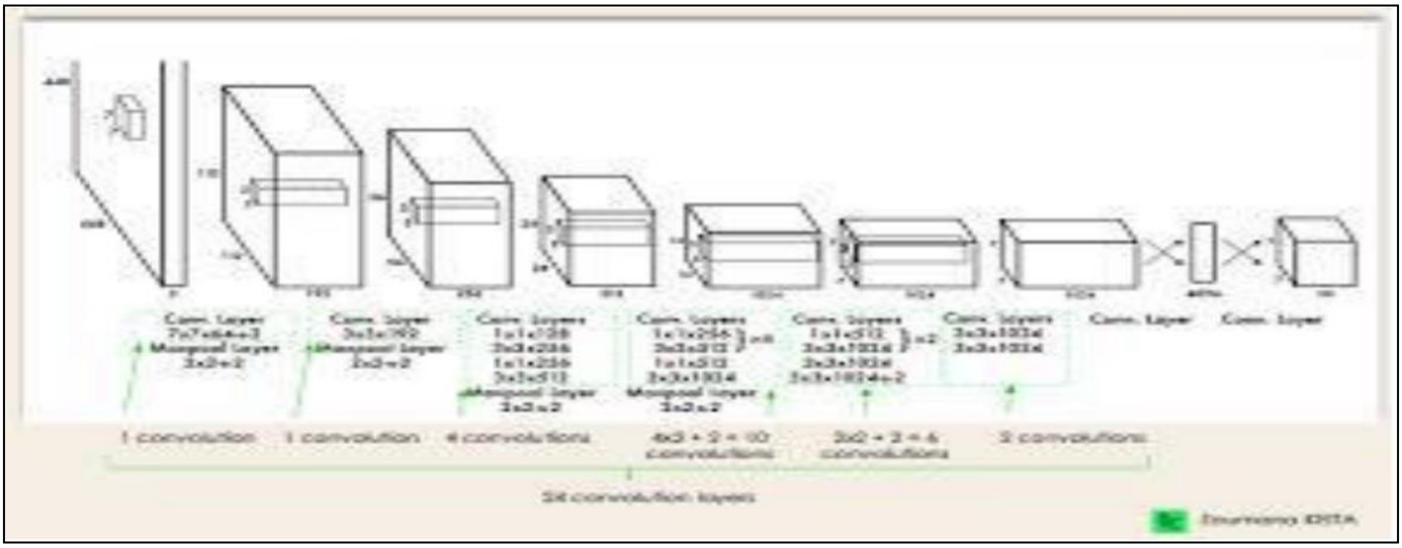


Fig 2 YOLO Model Architecture

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2631

➢ *Deep Sort*

Deep sort is an object tracking algorithm that uses deep Learning techniques to track objects in video clips. It is frequently used in applications such as surveillance and video analysis, where tracking and identification of objects of interest over time is important. (RNN) as object tracking. It first uses CNN to extract features from individual video frames, then uses RNN to associate the features over the same time. This creates a depth difference to accurately track objects even if they undergo significant changes such as occlusion or changes in scale or orientation. Indicator search algorithm. It is often used to measure the overlap between the estimated bounding box of objects in an image and the ground truth bounding box, which is a manually annotated bounding box provided as part of the tutorial dataset. Since the performance measure is a good indicator of the overlap between the prediction and the ground truth bounding box.

➢ *IOU*

A high IOU value indicates that the prediction bounding box is as good as the actual location bounding box, so the prediction is accurate. On the other hand, a low IOU value indicates that the prediction bounding box is not good with the ground truth bounding box and the prediction will be wrong. you can perform small subtasks and run them simultaneously on multiple processor cores or devices. The efficiency and effectiveness of computation can be increased by allowing tasks to be executed in parallel, reducing overall processing time. Through the use of networks, CNNs can be trained and used more efficiently and can make predictions on larger and more complex video clip data

➢ *Parallel Processing*

Parallel processing is a computing technique that involves breaking large tasks into smaller tasks and running them simultaneously on multiple processor cores or devices. The efficiency and effectiveness of computation can be increased by allowing tasks to be executed in parallel, reducing overall processing time. Using the same algorithm, CNNs can be trained and used more efficiently and predictions can be made on larger and more complex video clip datasets.

➢ *Data Set Used*

CCVC-Clinic DB is a database of frames extracted from colonoscopy videos. This document contains many examples of polyp frames and their basic facts. The actual ground image has a facet corresponding to the area occupied by the polyps in the image.

➢ *Polyp Dataset*

Table 1 Data Set Used

Class Name	Total Images
Ground Truth	612
Original	612

IV. IMPLEMENTATION

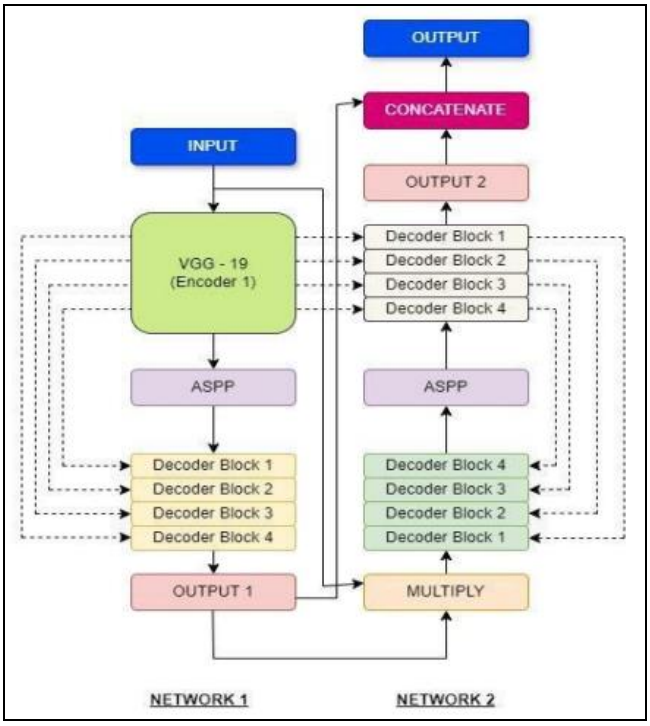


Fig 3 Architecture of Double U-net Model

All proposed methods are based on training neural network models to segment polar images. To train polar coordinate images, the input image must be transformed to use a polar coordinate origin close to the location of the segmented object. The true date cannot be known in advance, so a prerequisite for polar forecasting is a way to determine the true polar history. We propose and evaluate two different methods to obtain polar origin: (1) prediction from segmentations learned from non-polar images and (2) training a half-point prediction that predicts heat maps from input images.

This article describes this process as well as a method for training segmentation models on polar images. Our method was evaluated on tasks such as polyp segmentation, liver segmentation, skin segmentation, and epicardial adipose tissue (EAT) segmentation. The proposed method can be used as a preliminary step for existing neural network architectures, so we evaluate the image segmentation method using neural network architectures such as U-Net [1], U -NetCC [2], and ResNet [3] encoder, and ResNet [3] encoder. DeepLabV3C [4]. Images are usually displayed in Cartesian coordinates, where pixels are arranged along the x and y axes

The polar coordinate system has two axes: (1) the radial coordinate, which is far from the polar coordinate 19. the starting point of the cooperation transformation; In other words, the x-axis of the polar plot represents the distance to the origin, and the y-axis represents the rotation of the origin. Our hypothesis is that the polar transform is particularly useful in segmenting images where elliptical boundaries appear on the image. Consider an example of

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2631

using a linear model to estimate the surround resolution of a single-channel image. Environmental decision making should be modeled as a function of at least four dimensions. When converted to polar coordinates, a perfect circle in Cartesian coordinates becomes a straight line, as shown in Figure 1. The image in polar coordinates must have a structure whose boundary is not easy to predict. Although there are many differences, changing the shape of the elliptical object can reduce the required segmentation model complexity, as shown in Figure 2. In the middle, we take the position x in each training sample and normalize the boundary distance. This model can learn how far the boundary is from the origin at any angle around the origin, without needing to learn the object. For the image i(x;y), we use (2) to obtain the polar coordinate transformation of each point. input image. We will call this network the polar network in the rest of this article. In each of the explanations, the polar shift occurs not as part of the network architecture but as a preliminary step towards the polar network. To transform each input image, the background is determined as the center of the ground truth map of that image.

$$magnitude(x, y) = \sqrt{x^2 + y^2},$$
$$angle(x, y) = atan2(y, x) \cdot \frac{180}{\pi} \quad (1)$$

Given a polar origin (cx ; cy) of a Cartesian image I (x; y) of resolution H X W, we obtain each point of the polar transformation $I'(\rho, \phi)$ using (2).

$$\rho = \frac{H}{2\Pi} \cdot angle(x - c_x, y - c_y)$$
$$\phi = \frac{W}{\sqrt{(W/2)^2 + (H/2)^2}} \cdot magnitude(x - c_x, y - c_y) \quad (2)$$

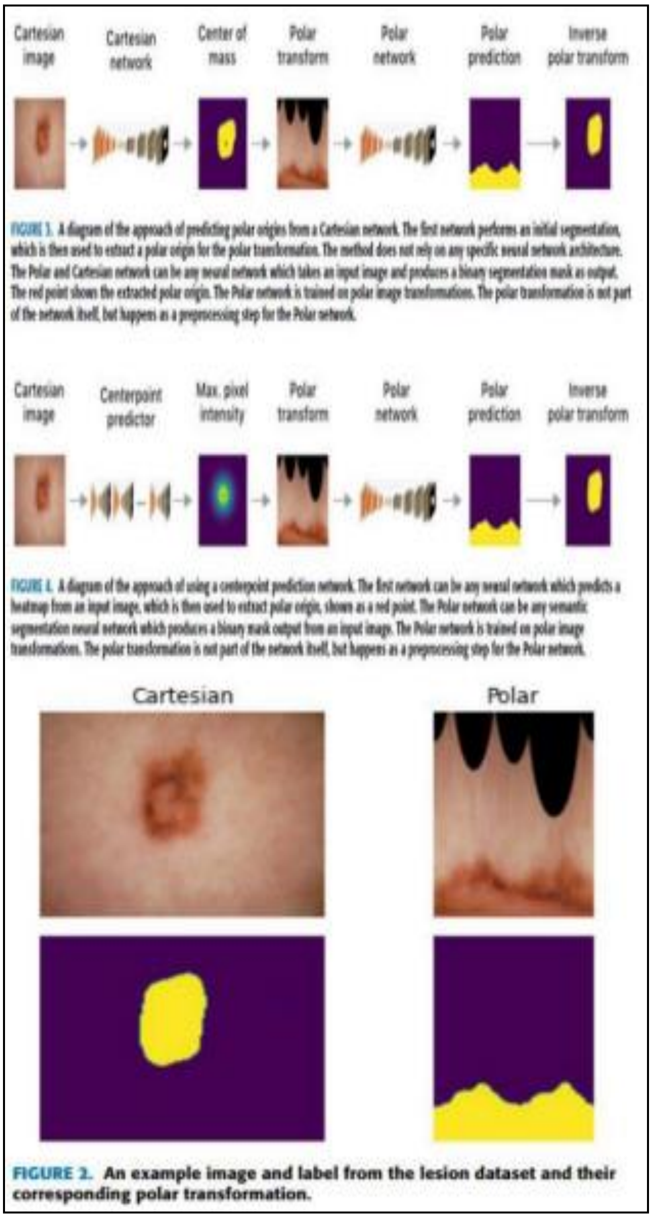
In each of our approaches, the final segmentation is done using a neural network trained on polar transformations of the input images. In the rest of this paper, we refer to this network as the polar network. In all of the described approaches, the polar transformation is not part of the network architecture itself, but happens as a preprocessing step for the polar network. To transform each input image, the polar origin is determined as the center of mass of the ground truth label for that image. The center of mass of an image I (x; y) is calculated by $I(x, y)$ calculating the spatial image moments matrix M, where the entry of the matrix at row i and column j is calculated using (3)

$$M_{ij} = \sum_{x,y} I(x, y) \cdot x^i \cdot y^j \quad (3)$$

Finally, to increase the model's robustness to suboptimal center point predictions, we augment the calculated center for training images [9]. Each training image has a 30% chance of varying the center's x and y coordinates by a random value in the range (-S. 0:05; S . 0:05), where S is the smallest resolution of the image, i.e. S = min(width; height)

An effective strategy for performing medical image segmentation involves two basic concepts for critical processing. The first part of our method involves the integration of polar transformations as a step before applying them to the raw data. This conversion machine is designed to solve problems related to the classification of oval-shaped single patterns in medical images such as organs (e.g. heart and kidneys), skin, polyps and many other defects. By focusing on the polar change in the focus of the object, size reduction can be achieved, while a clear separation of segmentation and localization tasks can be achieved. Two different methods have been proposed to select the best polar background: one is estimated by a segmentation neural network trained on non-polar images, the other uses a neural network trained to estimate the best background. We train the model up to 200 times. We use the OpenCV linear polar transform implementation. 10-3 study is used for each model. As can be seen from the figure, DoubleU-Net starts with VGG-19 as the encoder subnet, followed by the decoder subnet. In the first network (NETWORK 1), DoubleU-Net differs from U-Net by using the VGG-19 symbol in yellow, the ASPP symbol in blue, and the cutoff symbol in green. Compression and excitation blocks are used for the encoder in network 1 and the decoder blocks in network 1 and network 2. The difference between DoubleU-Net and U-Net in the second network (NETWORK 2) is the use of compression and excitation blocks with ASPP. All other elements remain unchanged. In Network 1, the input image is fed to the modified U-Net and a mask is created (Output 1). We then combine the output image with the resulting mask (Output1), which is the input of the second modified U-Net, to create another mask (Output2). Finally, we combine the two masks (Output1 and Output2) to find the difference between the average mask (Output1) and the final predicted mask (Output2). We think that the output map produced by Network 1 can be improved by taking the input image and its mask, and the connection with Output 2 will create a better mask than before. This is the main motivation for using both U-Net architectures in the design. Recovery and alert blocks in the network strategy reduce unnecessary information and send the most important information. ASPP has been a popular choice for modern segmentation architectures as it helps in extracting high-resolution feature maps and hence is more efficient.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR2631



V. RESULTS

This research on Skin Cancer Segmentation using CNN based on Double U-Net and polar transformations achieves compelling results, demonstrating a notable advancement in the precision of skin lesion delineation. The integrated approach effectively exploits the Double U-Net architecture and polar transformations, resulting in enhanced segmentation accuracy, as evidenced by high Dice coefficients, Jaccard indices, and pixel-wise accuracies. Visual comparisons between predicted segmentation masks and ground truth annotations highlight the model's proficiency in capturing intricate lesion boundaries. The proposed methodology not only outperforms baseline models but also exhibits promising potential for real-world clinical applications, emphasizing its clinical relevance. These results signify a significant contribution to the field of medical image analysis, presenting an innovative solution for improving skin cancer diagnostic processes.

- *Input Ground Truth Output 1 Output 2*

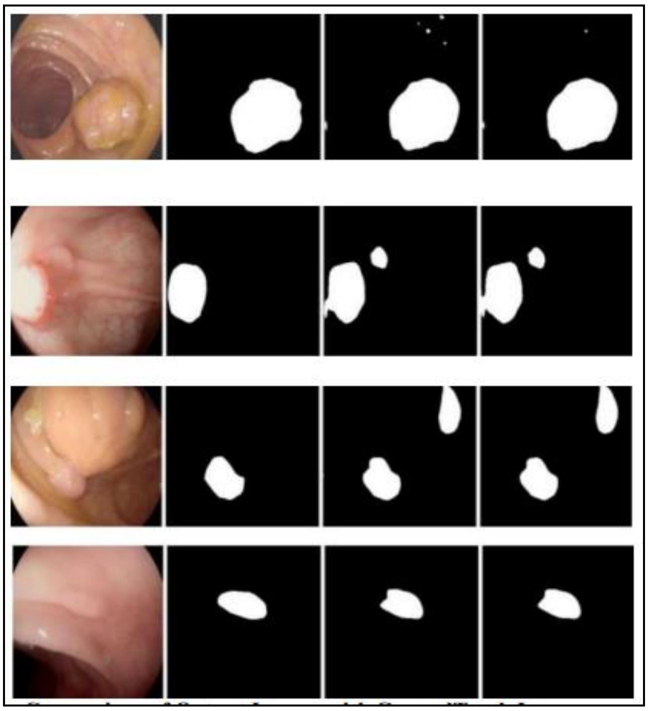


Fig 4 Comparison of Output Images with Ground Truth Images

In conclusion, this research introduces a novel and effective approach to skin cancer segmentation through the integration of a Double U-Net convolutional neural network (CNN) architecture and polar transformations. The proposed model demonstrates notable improvements in segmentation accuracy, as evidenced by robust quantitative metrics and compelling qualitative By leveraging the unique strengths of Double U-Net and harnessing the spatial information encoded in polar transformations, our approach overcomes challenges posed by traditional segmentation methods. The findings underscore the clinical relevance of the developed model, presenting a promising tool for dermatologists in the accurate delineation of skin lesions. While acknowledging certain limitations, the research provides a significant step forward in the field of medical image analysis, contributing to the ongoing efforts to enhance skin cancer diagnostic capabilities.

VI. SUMMARY

The polar transformation involves converting images from Cartesian coordinates to polar coordinates, where the radial coordinate represents the distance from the origin, and the angular coordinate signifies the rotation around the origin. This transformation, particularly beneficial for images with elliptical borders, simplifies the segmentation model complexity. The polar transform, centered at the object's origin, standardizes border distances in each training example, facilitating the model to learn border distances at each angle without the need to localize the object. Incorporating the polar transformation as a preprocessing step, the final segmentation is conducted using a neural network termed the polar network. The proposed method is evaluated across various medical segmentation tasks, including polyp segmentation, liver segmentation, skin lesion segmentation, and epicardial adipose tissue (EAT) segmentation. Notably, the approach can be seamlessly integrated into existing neural network architectures, allowing assessment with common models such as U-Net, U-NetCC with a ResNet encoder, and DeepLabV3C with a ResNet encoder

The second component of the comprehensive strategy introduces the DoubleU-Net architecture, designed for semantic image segmentation tasks. This architecture features two stacked U-Net models, with the first employing a pre-trained VGG-19 as an encoder and the second capturing additional semantic information. Atrous Spatial Pyramid Pooling (ASPP) is utilized to enhance contextual information within the network. The resulting DoubleU-Net outperforms traditional U-Net models and baseline architectures across various medical segmentation datasets, including colonoscopy, dermoscopy, and microscopy. The architecture's effectiveness is particularly highlighted in challenging scenarios, such as the segmentation of smaller and flat polyps.

The detailed architecture of DoubleU-Net involves a modified U-Net as the first network (NETWORK 1), incorporating VGG-19, ASPP, and a decoder block. The second network (NETWORK 2) is a modified U-Net with ASPP. The input image undergoes segmentation in NETWORK 1, and the output is used as input for NETWORK 2, producing a final predicted mask. The squeeze-and-excite block reduces redundant information, while the element-wise multiplication of NETWORK 1's output with its input contributes to improved segmentation. The concatenation of intermediate and final predicted masks aims to refine the segmentation further. Overall, the combination of polar transformations as preprocessing and the DoubleU-Net architecture demonstrates a promising advancement in medical image segmentation, showcasing improved accuracy and generalizability across diverse tasks

Dataset	Method	DSC	mIoU	Precison	Recall
Lesion	DeepLabV3+	0.8717	0.7984	0.8807	0.9068
	DoubleU-Net	0.8962	0.8212	**0.9459**	0.878
	Our method	**0.9253**	**0.8743**	0.9253	**0.9464**

Fig 5 Results – Comparison between other Methods and our Method

FUTURE SCOPE

The future scope of the presented research is promising and encompasses several key directions for advancement. First, there is potential for refining the techniques of polar transformations, exploring advanced algorithms or reinforcement learning methods for more accurate determination of the polar origin. Additionally, the integration of state-of-the-art neural network architectures and attention mechanisms could further enhance segmentation performance. Ensemble approaches, incorporating predictions from multiple networks trained with diverse polar transformations, offer an avenue for improved generalization. Extending the proposed approach to handle multi-modal medical imaging datasets and investigating transfer learning strategies may enhance the model's adaptability to varied clinical scenarios. Moreover, attention to real-time applications, model deployment optimization, and collaboration with healthcare professionals for clinical validation will be crucial for practical implementation. Exploring explainable artificial intelligence (XAI) techniques and addressing interpretability. concerns can contribute to the clinical relevance of the proposed method. Lastly, open-source implementation, benchmarking on standard datasets, and consideration of dynamic segmentation tasks, such as longitudinal studies, will further validate and extend the applicability of the presented approach. Collectively, these future directions hold the potential to elevate the proposed methodology and contribute meaningfully to the field of medical image segmentation.

Explore the integration of additional imaging modalities, such as dermoscopy, reflectance confocal microscopy, or even patient clinical data. Multimodal approaches may provide complementary information, enhancing the accuracy and robustness of skin cancer segmentation models Investigate the application of transfer learning by leveraging pretrained models on large-scale datasets. Transfer learning can potentially improve model generalization and efficiency, especially when dealing with limited annotated medical image data

REFERENCES

[1]. D. Jha, M. A. Riegler, D. Johansen, P. Halvorsen and . H. D. Johansen, “DoubleU-net: A deep convolutional neural network for medical image segmentation,” *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, vol. abs/2006.04868v2, no. 8 Jun 2020, pp.1-7,2020. https://doi.org/10.1109/cbms49503.2020.00111

[2]. C. Kaul, S. Manandhar and N. Pears, “FocusNet: An attention-based Fully Convolutional Network for Medical Image Segmentation,” *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, vol. abs/1902.03091v1, no. 8 Feb 2019, pp. 1-5, 2019. https://doi.org/10.1109/isbi.2019.8759477

[3]. M. BENČEVIĆ, I. GALIĆ, M. HABIJAN and D. BABIN, “Training on Polar Image Transformations,” *IEEE Access*, vol. 9, no. September 29, 2021, p. 133365–133375, 2021. https://doi.org/10.1109/access.2021.3116265

[4]. Chen, P., Huang, S., & Yue, Q. (2022). Skin lesion segmentation using recurrent attentional convolutional networks. IEEE Access, vol. 10, 94007–94018. https://doi.org/10.1109/access.2022.3204280