“Skin Lesion Segmentation Using Deep Learning”

Project Report submitted in the partial fulfilment
of
Bachelor of Technology
in
Information Technology

Submitted To

SVKM’s NMIMS,
Mukesh Patel School of Technology Management & Engineering,
Shirpur Campus (M.H.)

Submitted By:

Khushi Patil – 70012100165
Anannya Londhe – 70012100223
Pallavi Dangane – 70012100227

Under The Supervision Of:

Prof. Pravinkumar Landge
(Assistant Professor, IT Department, MPSTME)

INFORMATION TECHNOLOGY DEPARTMENT
Mukesh Patel School of Technology Management & Engineering
ACADEMIC SESSION: 2024-25

CERTIFICATE

This is to certify that the project entitled **"Skin Lesion Segmentation Using Deep Learning"** has been done by

"Khushi Patil" (Roll No.: 70012100165)

"Anannya Londhe" (Roll No.: 70012100223)

"Pallavi Dangane" (Roll No.: 70012100227)

under my guidance and supervision & has been submitted partial fulfilment of the degree of "**Bachelor of Technology** in **Information Technology"** of MPSTME, SVKM'S NMIMS (Deemed-to-be University), Shirpur (M.H.), India.

Prof. Pravinkumar Landge	**Examiner**
(Internal Guide)	
Date:	
Place: Shirpur	**H.O.D**

INFORMATION TECHNOLOGY DEPARTMENT

Mukesh Patel School of Technology Management & Engineering

ACKNOWLEDGEMENT

After the completion of this capstone Project work, words are not enough to express my feelings about all those who helped me to reach my goal; feeling above this is my indebtedness to The Almighty for providing me this moment in life.

It's a great pleasure and moment of immense satisfaction for me to express my profound gratitude to **Prof. Pravinkumar Landge**, Assistant Professor of Information Technology Department, MPSTME, Shirpur, whose constant encouragement enabled me to work enthusiastically. Their perpetual motivation, patience and excellent expertise in discussion during progress of the project work have benefited me to an extent, which is beyond expression. Their depth and breadth of knowledge of Information Technology field made me realize that theoretical knowledge always helps to develop efficient operational software, which is a blend of all core subjects of the field. I am highly indebted to them for their invaluable guidance and ever-ready support in the successful completion of this project in time. Working under their guidance has been a fruitful and unforgettable experience.

We express my sincere thanks and gratitude to **Dr. Ritesh Kumar Dhanare,** Head of Department, Information Technology Department, MPSTME, Shirpur, for providing necessary infrastructure and help to complete the project work successfully.

We also extend my deepest gratitude to **Dr. Venkatadri Marriboyina**, Associate Dean, SVKM'S NMIMS, Shirpur Campus for providing all the necessary facilities and true encouraging environment to bring out the best of my endeavors.

We sincerely wish to express my grateful thanks to all members of the staff of Information Technology department and all those who have embedded me with technical knowledge of computer technology during various stages of B.Tech. Information Technology.

We would like to acknowledge all my friends, who have contributed directly or indirectly in this capstone Project work.

The successful completion of a capstone Project is generally not an individual effort. It is an outcome of the cumulative effort of a number of persons, each having their own importance to the objective. This section is a vote of thanks and gratitude towards all those persons who have directly or indirectly contributed in their own special way towards the completion of this project.

Khushi Patil – 70012100165
Anannya Londhe - 70012100223
Pallavi Dangane – 70012100227

ABSTRACT

The increasing cases of skin cancer, require the development of efficient and accurate diagnostic tools. In this work, an automatic skin lesion segmentation system is developed using deep learning techniques: U-Net and Fully Convolutional Network (FCN). The aim here is to support early diagnosis of skin lesions by precisely segmenting dermoscopic images for dermatologists. It is then trained and validated with experts annotated high-quality images from the PH2 dataset. The dataset preprocessing encompasses the following steps: resized, augmented, and normalized versions of the images in the process of optimizing them for further model training. Results - The U-Net shows promising results with good accuracy rates, IoU, and Dice coefficient, whereas FCN also has promising features of segmentation. Both could generalize well onto other types of skin lesions; they are therefore promising for clinical applications. The user-friendly graphical interfaces allow real-time segmentation; therefore, the system is accessible to healthcare practitioners. This research advances the development of automated skin lesion detection and is also the foundation for future explorations in deep learning applications in dermatology that would eventually improve patient outcomes by providing timely and accurate diagnosis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 Background:

Skin lesion segmentation with deep learning is a key contemporary concern, especially considering the dramatically increasing incidence of skin cancer worldwide. Traditionally**,** manual or semi-automatic segmentation methods have been used, where dermatologists or trained professionals manually draw the borders of lesions in medical images (e.g., dermoscopic or clinical images). This process, while effective, is time-consuming, labor-intensive, and highly dependent on the skill and experience of the professional. It also requires specialized software and is prone to human error and inconsistencies**.**

In recent years, deep learning—especially with the use of Convolutional Neural Networks (CNNs)—has revolutionized medical imaging by enabling models to learn from raw image data and perform complex feature extraction without the need for manual feature engineering. This innovation makes deep learning techniques particularly suited to tasks like skin lesion segmentation.

Among the various models developed for medical image segmentation, the U-Net architecture has emerged as one of the most effective. U-Net is specifically designed for biomedical image segmentation, making it ideal for skin lesion segmentation. Its key strength is the ability to produce high-resolution output from relatively small datasets, a common constraint in medical applications. U-Net achieves pixel-level accuracy, which is critical in differentiating healthy skin from cancerous regions**.** This precision is essential in ensuring that malignant areas are accurately segmented, which can lead to early detection and better prognosis for skin cancer patients.

The background of this project, therefore, includes the limitations of traditional manual methods, the advantages of deep learning techniques like CNNs and U-Net, and the potential for improving medical diagnostics by automating the segmentation of skin lesions, leading to faster, more reliable, and more accessible tools for detecting skin cancer.

1.2 Motivation and scope of report:

The motivation behind this study arises from the critical need to improve early detection of melanoma, one of the most dangerous forms of skin cancer. The rising global incidence of skin cancer and the limitations of traditional manual skin lesion segmentation methods, which are time-consuming, labor-intensive, and prone to human error, highlight the necessity for more accurate and efficient solutions. Deep learning, particularly Convolutional Neural Networks (CNNs), has shown tremendous potential in medical imaging by automating feature extraction and enhancing diagnostic precision.

The choice of this capstone project is driven by both a personal and professional desire to contribute meaningfully to the healthcare sector. By developing a tool capable of automatically segmenting skin lesions, the project aims to assist dermatologists in the early detection of melanoma, which can significantly improve patient outcomes. The project is not just an academic endeavor but one with real-world impact, where early diagnosis through deep learning techniques can lead to better prognoses for patients. The integration of automatic segmentation tools into clinical workflows can enhance the efficiency and accuracy of skin cancer detection, ultimately saving lives.

This research paper is to design and test an automated skin lesion segmentation system on the basis of deep learning architecture, specifically U-Net. Long-term prognosis in the ease of early detection of skin cancer would benefit from high melanoma detection by segmenting lesions using dermoscopic images. Open source datasets PH2 is mainly used because there is a scarcity of annotated medical data, thus applying data preprocessing techniques such as resizing, normalisation, and augmentation of the training data. Thus, numerous efforts have been put into augmenting the dataset in order to improve the generalization and performance of the model.

This will involve designing and training the U-Net model that requires to be tuned using various hyperparameters, loss functions, and performance measures. Further study of the model's performance will be done with the help of quantitative evaluation measures, such as the Dice coefficient, Jaccard index, and accuracy. In addition, a qualitative evaluation by visual inspection will be considered to promote further awareness

regarding the model's ability to work on various lesions and textures in terms of accurate delineation of skin lesions.

The development also encompasses creation of the graphical user interface (GUI) so that the system becomes user-friendly for health practitioners. The GUI allows users to upload images of skin lesions and receive real-time segmented outputs using a user-friendly tool for clinical use. Testing of usability, speed, and accuracy will be conducted in order to ensure integration into the medical workflow.

The study will entail collaboration with dermatologists and medical experts whose feedback can be used to validate the clinical utility; the outputs of model segmentation will be compared with the experts' manual segmentation to ensure accuracy and reliability for practical application. Lastly, the study considers strategies for deployment, including web-based or cloud-based solutions that can really make the system scalable and accessible to healthcare providers all over the world. Data privacy, security, and adherence to medical regulations will serve as the base structure for the deployment process.

Finally, this study would note the different limitations that it had in place: including availability and variety of medical image data, the limitation of its generalization across range of skin tones and lesion types. Sufficient computer resources also become one important limitation. These limitations will provide opportunities for further work, such as the extension of the capabilities for real-time or the incorporation of the model into other diagnostic tools. For example, these efforts make this study a contribution to academic research both in medical image segmentation and to practical advancements in skin cancer detection.

1.3 Problem Statement:

Accurate and timely detection of skin lesions is important in making an early diagnosis and treatment of skin cancer, especially the dangerous melanoma type, now becoming increasingly common worldwide. However, to the great concern of millions of people afflicted with various skin lesions, traditional methods of assessment of the former depend largely on time-wasting dermatologists' manual inspections which, when subject to the personal interpretation by the observers, may lead to inconsistencies and misdiagnoses. This research attempts to overcome the challenges in the proper segmentation of skin lesions from dermoscopic images by constructing automated deep learning models, U-Net and Fully Convolutional Networks (FCNs). These models are proposed to improve the accuracy of lesion outlining, enhance the diagnosis accuracy, and aid in the early detection of skin cancer, further improved patient outcomes, and efficient service delivery.

1.4 Salient Contribution:

The salient contribution of this study lies in the development of an automated skin lesion segmentation system using deep learning, specifically the U-Net architecture. This project addresses a critical healthcare challenge—early detection of melanoma—by offering an efficient, accurate, and scalable alternative to traditional manual segmentation methods. By utilizing dermoscopic images and applying advanced data preprocessing techniques, the study enhances the generalization and performance of the model. Furthermore, the system includes a user-friendly graphical user interface (GUI), making it accessible for healthcare practitioners to use in clinical settings for real-time lesion segmentation.

The project contributes to the medical field by improving segmentation precision, which is vital for distinguishing between healthy and cancerous skin regions. Additionally, the collaboration with dermatologists and integration of quantitative and qualitative evaluation measures ensures the system's practical utility. The deployment strategy—whether web-based or cloud-based—further enhances the system’s accessibility for healthcare providers globally. The study’s innovative approach not only advances medical image segmentation but also lays the groundwork for future

developments in skin cancer detection, potentially extending the system's real-time capabilities and integrating it into broader diagnostic frameworks

1.5 Organization of report:

In the subsequent sections, we delve into the literature review, methodology, data analysis and results, discussion, and conclusion. The literature review provides an overview of skin lesion segmentation, highlighting previous research in the field and the importance of accurate lesion identification for early skin cancer diagnosis. The methodology section details the data preprocessing steps, the deep learning algorithms employed, and the evaluation metrics used to assess model performance. Data analysis and results showcase the findings from our experiments, focusing on the accuracy and effectiveness of the segmentation models. The discussion section offers insights, implications, and limitations of our study, considering the clinical relevance of our findings. The conclusion summarizes the key results and outlines potential future directions for enhancing skin lesion segmentation techniques and their application in dermatology.

By adhering to this structure, we aim to provide a comprehensive exploration of our research, enabling readers to gain a deeper understanding of the challenges, methodologies, and outcomes associated with skin lesion segmentation.

Chapter 2

Literature survey

2.1 Introduction:

One of the deadliest kinds of skin disorders known is melanoma, with 287,723 cases and more than 60,000 deaths projected in 2018. Among public health challenges, one of the leading issues is skin cancer, given that 2,000 new diagnoses alone were recorded in South Korea for the last five years. Melanomas at the surface of the skin are normally visible, but unfortunately, many people fail to sound an alarm when their melanomas start manifesting. But when examined by specialized dermatologists, the diagnosis is achieved only in about 60% of cases, so many melanomas go undiagnosed at an early treatable stage. Correct and timely segmentation of cutaneous lesions is very crucial for dermoscopic pattern identification and localization as well as for skin disease classification. Dermoscopy is an imaging modality which reduces the reflectivity of superficial layers of skin to be able to visualize deeper layer structures with increased magnification. This method enhances the reliability of diagnosis and tends to reduce deaths that are melanoma-related.

Skin lesion segmentation using deep learning is a crucial task in the field of medical image analysis, particularly in the early detection and diagnosis of skin cancer, including melanoma. Accurate segmentation of skin lesions from dermoscopic images is essential to differentiate malignant lesions from benign ones. Traditional methods for lesion detection, which rely on manual examination, are time-consuming, subjective, and prone to human error. Hence, the need for automated and reliable segmentation techniques has driven researchers to explore advanced methods like deep learning.

Deep learning, particularly convolutional neural networks (CNNs), has proven to be highly effective in image analysis tasks due to its ability to automatically learn features from raw images. For skin lesion segmentation, deep learning models can distinguish between the lesion and surrounding skin by learning from a large dataset of annotated dermoscopic images. These models can capture complex patterns, including texture, color, and irregularities in shape, which are key indicators for identifying potentially

harmful lesions. The segmentation step is critical as it forms the foundation for subsequent analysis, such as classification and assessment of lesion malignancy.

The segmentation of skin lesions, characterized by the accurate identification and delineation of abnormal skin growths in dermatoscopic images, has significant implications for public health and medical practice. Effective skin lesion segmentation enhances diagnostic accuracy, aiding dermatologists in early detection of skin cancer, which can drastically improve patient outcomes. Misdiagnosis or late detection can lead to serious health consequences, underscoring the necessity of reliable segmentation techniques. For instance, studies have shown that automated segmentation methods can reduce human error and increase efficiency in clinical settings, allowing for faster and more accurate assessments of skin conditions. The impact of accurate skin lesion segmentation extends beyond individual cases; it contributes to broader public health efforts by facilitating large-scale screening and research initiatives. Considering these challenges, research into advanced skin lesion segmentation methods is essential to improve the quality of dermatological care and preserve the reliability of clinical diagnostics.

2.2 Previous Research on Skin Lesion Segmentation:

Arif et al.'s[1] proposed a hybrid deep model, FCN-ResAlexNet, for segmentation of skin lesions that focuses on early detection of skin cancer. The architecture combined AlexNet and ResNet-18 to improve the segmentation performance, giving much of a better approach than the base FCN-AlexNet model in terms of accuracy, Dice score, and the Jaccard index, mainly in melanoma segmentation. The improvement techniques applied to the images, such as GWA and CLAHE were used, combined with the ADAM optimizer and the cross-entropy loss function, which resulted in improving the training phases' results. The model is exceptional because it has both low computational cost while still maintaining high accuracy, which would be of great use in clinical skin cancer diagnosis.

However, despite its promising performance, the FCN-ResAlexNet model also presented some limitations. One key limitation is its reliance on carefully tuned preprocessing steps, such as GWA and CLAHE, which may not generalize well to other datasets or imaging conditions without further customization. Additionally, while the

model achieved low computational cost, the architecture's complexity, combining two major networks, may still pose challenges in real-time clinical deployment, particularly in environments with limited computational resources. Furthermore, the model's success in segmenting melanoma lesions may not fully extend to other types of skin lesions, as the dataset used primarily focused on melanoma. Expanding the model's robustness across a broader range of skin lesion types is necessary for its wider applicability in clinical settings.

In another article by **Darmawan et al.**[2] the authors develop an evolved architecture of U-Net particularly for segmentation purposes in skin lesions. The model employed has enhanced state including bilinear interpolation in upsampling and PReLU activation functions, and it holds an estimated approximate accuracy of about 94% in pixel-wise and a 88.33% Dice coefficient. By using the dropout technique, overfitting and other artifacts are minimized. Advanced image processing techniques that optimize segmentation results hold much promise in further study for improving detection if values of the Dice coefficient are low in the scenario. However, despite these advancements, the model does face certain limitations. One key issue is that, although the Dice coefficient is relatively high, there are still scenarios where the coefficient values are lower, particularly in cases of very irregular or small lesions. This suggests that the model may struggle with certain edge cases, such as lesions with ambiguous boundaries or unusual color distributions. Furthermore, while bilinear interpolation and PReLU offer performance boosts, these techniques may not be sufficient to capture all the nuances of highly complex lesions, and further refinement might be needed for more challenging datasets. Another limitation is the potential reliance on specific image preprocessing steps and dataset characteristics, which may limit the model's generalizability to other skin lesion datasets without additional customization. Future work could focus on refining the segmentation techniques to improve Dice scores in these difficult cases and exploring ways to make the model more robust across different lesion types and imaging conditions.

The systematic review by **Kaur et al.**[3] "Skin Lesion Classification and Detection using Machine Learning Techniques" underscores early detection of skin cancer. The authors link the progress in detection methodologies and limitations in traditionally accepted diagnoses to establish the importance of Computer-Aided Diagnostic (CAD) systems. The authors have achieved high classification accuracies for different models

but highlighted concerns regarding optimization, generalizability, and explainability of models. The study points out a sharp increase in the research interest of 2022 and requests collaboration between AI experts and dermatologists so that diagnostic and care outcomes are improved. However, the review also identifies several key challenges in the development and deployment of these machine learning models. Concerns around optimization point to the fact that many models still require fine-tuning and careful adjustment to perform optimally on different datasets, especially when dealing with variations in skin tones, lesion types, and image quality. Generalizability remains a major issue, as models trained on one dataset may not perform as well on others due to differences in image acquisition techniques or population demographics. This poses a challenge for widespread clinical adoption. Another important concern is the explainability of these models, as many machine learning algorithms, particularly deep learning models, function as "black boxes," making it difficult for clinicians to understand how decisions are being made, which could hinder trust and integration into clinical workflows. he study also notes a sharp increase in research interest in this field in 2022, reflecting the growing awareness of the potential benefits of AI in dermatology. The authors advocate for increased collaboration between AI experts and dermatologists to address the current limitations, improve the accuracy and reliability of machine learning models, and ensure that these technologies can seamlessly integrate into clinical practice. By fostering such interdisciplinary collaboration, the review suggests that AI-driven tools could significantly enhance diagnostic accuracy and ultimately improve care outcomes for patients with skin cancer.

In **Alshahrani et al.**[4] titled "Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network", two new architectures are added, namely the Lesion Indexing Network (LIN) and Lesion Feature Network (LFN), which will be used both in segmentation and feature extraction for skin lesions. Using ISIC 2017, LIN achieves robust accuracy in the task of segmenting and classifying melanoma as well as other lesions with very poor contrast and class imbalance. LFN significantly outperforms existing models, also offering a benchmark for automated skin-lesion detection. Despite these advancements, the study acknowledges certain limitations. LIN is primarily focused on segmentation, where it excels at delineating lesion boundaries even in cases where the lesions exhibit poor contrast against the surrounding skin. This capability is particularly important for melanoma, which often has irregular borders and can blend

in with the skin, making manual segmentation difficult. LIN achieves robust accuracy not only in segmenting melanoma but also in classifying a variety of other skin lesions, addressing the issue of class imbalance that is prevalent in many medical image datasets. This ensures that the model can generalize well across different lesion types, which is crucial for its clinical applicability.While LIN and LFN demonstrate strong performance on the ISIC 2017 dataset, the models may require further validation on other datasets to ensure their robustness and generalizability. Additionally, the complexity of these architectures, while offering high accuracy, may present computational challenges in real-time clinical settings where resources are limited. Future work could focus on optimizing these models for more efficient deployment without compromising their performance. The study also suggests that further collaboration between AI researchers and clinicians is essential to refine these models for practical use in skin cancer diagnosis.

Deep Learning for Skin Diagnosis" by **Amiri et al.**[5] describes an approach of a lightweight CNN with image analysis toward the detection of skin disease at an accuracy of 87.64%. The proposed CNN model is compared to advanced variants and traditional methods such as SVM and Random Forest, giving the robustness, especially when class imbalances occur, to the proposed CNN model. From that, the focus has been on key performance metrics to include precision, recall, and F1-score, thus promising early detection of skin diseases by AI and indicating a need for further real-world validation. Precision ensures that the model minimizes false positives, which is crucial in medical diagnoses to avoid unnecessary treatments or interventions. Recall, or sensitivity, measures the model's ability to correctly identify all cases of skin disease, reducing the chances of missing potential cases. The F1-score balances precision and recall, providing a more comprehensive evaluation of the model’s performance. These metrics indicate that the proposed CNN not only achieves a high accuracy but is also reliable in detecting skin diseases in various real-world scenarios. However, the study acknowledges the need for further validation in real-world clinical environments. While the model performs well in controlled settings, the authors suggest that more extensive testing on diverse and real-world datasets is necessary to confirm its generalizability and clinical applicability. This highlights the ongoing need for collaboration between AI researchers and healthcare professionals to ensure that AI-based diagnostic tools like

this CNN model can be seamlessly integrated into clinical practice, providing accurate and timely diagnoses for a range of skin diseases.

The current study by **Shu et al.**[6] entitled "Skin Cancer Detection Based on Skin Lesion Images using Deep Learning" reviews the application of CNNs for skin cancer detection. They use the ISIC 2018 dataset and apply an enhancement tool, ESRGAN, using pretrained models like ResNet50 and InceptionV3; and so also Inception ResNet. Accuracy reaches 85.7% with the best model being InceptionV3. This research therefore promises a bright future for AI applications in the diagnosis of medical practices and calls for further improvement in this venture through further experiments on larger data sets and more architectures. The study highlights how combining image enhancement techniques like ESRGAN with powerful CNN architectures can significantly boost the model's ability to detect skin cancer from lesion images, particularly in challenging cases where lesions are visually subtle or ambiguous. Despite this promising result, the authors acknowledge that there is still room for improvement. The study emphasizes the need for further experimentation on larger datasets to enhance the model's generalizability. The ISIC 2018 dataset, while comprehensive, may not represent the full diversity of skin lesion types and imaging conditions encountered in clinical practice. Additionally, the authors suggest exploring more advanced architectures beyond those tested in this research to push the boundaries of what AI can achieve in skin cancer detection. The research calls for continued efforts in refining deep learning models and testing them across diverse and larger datasets. It advocates for collaboration between AI researchers and medical professionals to ensure these models are clinically viable, reliable, and able to assist in early skin cancer detection. With further advancements, AI-driven diagnostic tools like the ones proposed in this study hold great potential to transform medical practices, particularly in dermatology, by providing more accurate and accessible screening methods for skin cancer.

However, the paper "Deep Learning–Based Methods for Automatic Diagnosis of Skin Lesions" by **Yuan Liu et al.**[7] is set for the development of very accurate classification systems of the skin lesions like melanoma, using deep learning. The system relies on the use of neural networks, pre-trained CNNs, and feature-based approaches, with accuracy ranging from 93% to 95% on the datasets PH2 and ISIC 2019. Results of various models, aggre- gated using weighted performance, significantly outperform the

performance of individual models, thus forming the basis for the promise of transfer learning and feature extraction techniques. The authors take advantage of transfer learning, which allows pre-trained models to utilize previously learned features from large-scale datasets and adapt them to specific tasks, such as skin lesion classification. This not only improves model performance but also reduces training time and computational resources. One of the key strengths of the paper is its approach to aggregating the results of multiple models using weighted performance metrics. Instead of relying on a single model, the study combines the predictions of various models, which significantly improves the overall performance compared to individual models. This ensemble learning approach enhances the system's robustness and reliability, further reinforcing the promise of transfer learning and feature extraction in medical image analysis. Despite the impressive results, the authors suggest that further research is needed to ensure the generalizability of the model across diverse populations and imaging conditions. While the PH2 and ISIC 2019 datasets are valuable resources, additional testing on broader datasets would help confirm the system’s reliability in real-world clinical settings. Nonetheless, the paper highlights the potential of deep learning in improving the accuracy and efficiency of skin lesion classification, offering a promising direction for future research in AI-assisted dermatology diagnostics.

On the other hand, the research paper "SkinLesNet: A Deep Learning Model for Skin Lesion Classification" by **Walaa Gouda et al.**[8] aspires to examine the classification ability on the part of the model, known as SkinLesNet. It results in an accuracy of almost 96% on the PAD-UFES-20-Modified dataset that outperforms the ResNet50 and VGG16. It also achieved 90% and 92% accuracies on the HAM10000 and ISIC2017 datasets, respectively. Its architecture was specialized for its task but was allowed to be trained enough on diverse datasets to help it succeed, despite challenges of overfitting and the need for more diverse input for better generalizability. These results indicate that the model is not only robust but also versatile enough to generalize across different datasets, which is critical in the context of dermatology, where lesion appearance can vary significantly based on factors such as skin type and lighting conditions.The architecture of SkinLesNet was specifically designed for its classification task, allowing it to excel in processing and interpreting the nuances of skin lesion images. However, the study also acknowledges some limitations. One significant challenge is the potential for overfitting, which can occur when the model learns to perform exceptionally well

on training data but struggles to generalize to new, unseen data. To mitigate this, the authors emphasize the importance of training on diverse datasets, which can help improve the model's generalizability and resilience to variations in input data.Overall, the research highlights the promise of SkinLesNet as a powerful tool for skin lesion classification, demonstrating that specialized architectures can achieve high accuracy in medical image analysis. The study suggests that with continued efforts to expand the diversity of training datasets and refine the model to address overfitting, SkinLesNet could significantly contribute to improving the early detection and diagnosis of skin cancer in clinical settings.

Overall, the hybrid model presented by **Wahlig et al.**[9] based on U-Net and MobileNet-V3 reveals a lot of promise, as it has been tested with a good level of accuracy on a large dataset like HAM-10000. This innovative approach leverages the strengths of both architectures: U-Net is well-known for its effectiveness in image segmentation tasks due to its encoder-decoder structure, while MobileNet-V3 is designed for efficiency and speed, making it suitable for deployment in resource-constrained environments. By integrating these two models, the hybrid architecture aims to enhance the performance of skin lesion segmentation and classification while maintaining computational efficiency. The authors tested the hybrid model on a large and diverse dataset, specifically HAM-10000, which contains a wide variety of skin lesions. Achieving a good level of accuracy on such a comprehensive dataset is a notable achievement, as it demonstrates the model's ability to generalize well across different types of lesions and variations in imaging conditions. The results suggest that this hybrid approach not only excels in accurately identifying and segmenting skin lesions but also holds potential for practical applications in clinical settings. The combination of U-Net's robust segmentation capabilities with MobileNet-V3's lightweight design makes the model particularly appealing for use in mobile and web-based applications, which could facilitate wider access to skin cancer screening tools. Despite its promising results, the study likely emphasizes the need for further validation and testing on additional datasets to confirm the model's effectiveness across diverse populations and clinical scenarios. Overall, the hybrid model stands as a significant step forward in the integration of deep learning techniques for skin lesion analysis, potentially improving early detection and diagnosis of skin conditions in the future.

2.3 Research Gap:

Despite the success of deep learning in automating skin lesion segmentation, critical gaps remain that limit widespread clinical application and reliability. One of the most pressing challenges is the lack of model generalization across diverse skin tones and lesion types, as many deep learning models are trained on limited datasets that do not capture the full range of demographic diversity. This lack of variety can result in models that perform inconsistently across different skin types, a crucial limitation in dermatology. Additionally, while data augmentation techniques are used to expand these datasets artificially, they cannot entirely replicate the natural variability encountered in clinical practice. This scarcity of annotated data continues to be a bottleneck, calling for more sophisticated approaches in dataset expansion, such as synthetic data generation, or increased access to larger, annotated datasets.

Interpretability and clinical validation further challenge the integration of deep learning models into clinical workflows. While segmentation models, such as U-Net, achieve high accuracy, they often operate as "black boxes," making their decision-making processes difficult for clinicians to interpret. For these tools to gain acceptance and trust in healthcare, they need to incorporate features that allow clinicians to understand and assess the model's segmentation rationale. Furthermore, few studies engage in clinical validation with dermatologists, limiting insights into the model's practical utility in real-world settings. This lack of transparency and direct validation restricts the model’s applicability, as clinicians need evidence of both high performance and clinical relevance to rely on automated segmentation in diagnosis.

Beyond segmentation accuracy, real-world deployment remains largely unexplored. Current research often focuses on experimental outcomes but lacks practical implementation studies, such as integrating models into clinical workflows, managing privacy and regulatory concerns, or exploring scalable cloud-based deployments. Additionally, many models are developed exclusively for specific imaging modalities, like dermoscopic images, which limits their adaptability to other types, such as clinical photographs or histopathology images. Addressing these gaps could lead to more robust, accessible, and impactful tools for skin cancer detection, ultimately aiding dermatologists in early diagnosis and improving patient care on a broader scale.

Chapter 3

Methodology and Implementation

3.1 Block Diagram

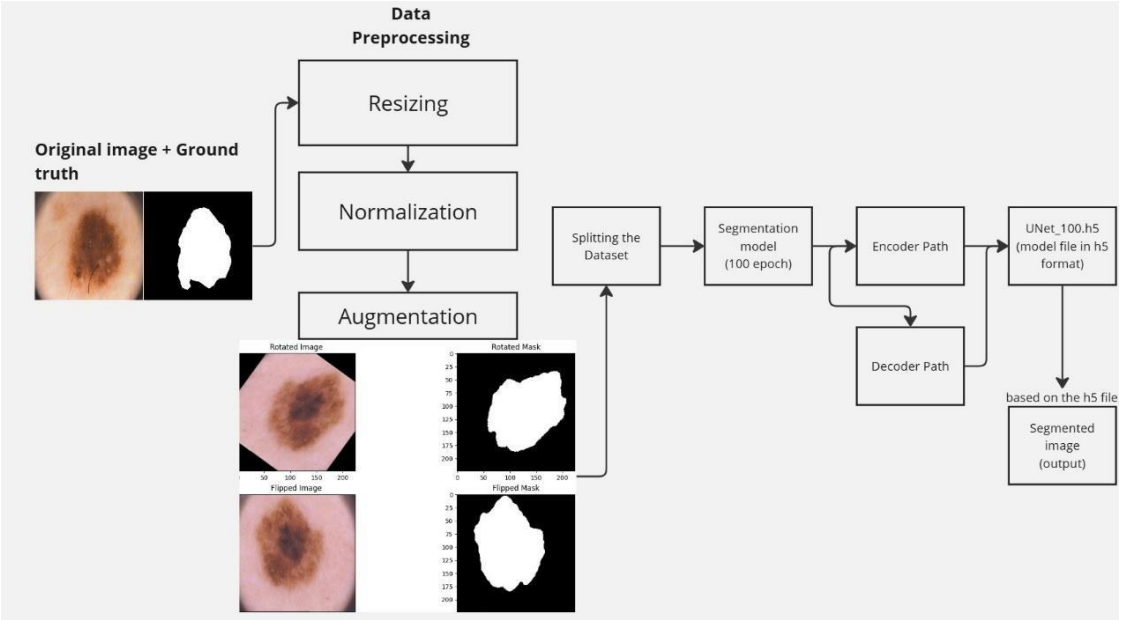


Fig. 1 Block Diagram

3.1.1 Data Collection

The PH2 dataset on Kaggle will be mostly used for the capstone project on deep learning skin lesion segmentation. This dataset is hosted at PH2 Resized Dataset and curated specifically to help in research about melanoma detection and skin lesion segmentation. The following provides high-quality dermoscopic images of many different types of skin lesions with expert annotations. It is an ideal dataset for training and validating deep learning models like U-Net.

The PH2 dataset is provided with dermoscopic images of skin lesions, which have been resampled into a uniform resolution. It's important to resample the input image because this standardises the input for deep learning models, minimizing the number of preprocessing steps. All images in the data set are resized to suitable dimensions for model input--for most practical use cases that's around 224x224 pixels, ensuring uniformity and that they will make the model generalize better.

The PH2 dataset is rich in variations of the type of lesions, ranging from benign to malignant. Such variability is extremely valuable for a model generalizable across categories and types of skin. The dermoscopic images have well-defined lesions with very elementary bounds that guarantee proper segmentation. These annotations, which were mostly obtained from a small pool of dermatology experts, served as ground truth labels for training the U-Net model in a supervised learning setting.

This is aside from the diversity of lesion types; the dataset is highly robust when it comes to practical applications, capturing the nuances of different skin tones and lighting conditions as well as lesion appearances. The addition of expert-annotated segmentation masks leaves clear boundaries of lesions, enabling the model to learn features to distinguish between healthy skin and potential malignancies.

Pre-processing the PH2 dataset to ensure uniformity and usability reduces the headaches involved in the preprocessing of data, facilitating a more reliable input format of the deep-learning model. The correct usage within CNN and deep-learning architectures like U-Net, which require uniform and well-labeled input data for training, ensures proper usage of the dataset. Based on this research, therefore, a dataset was laid as a foundation, hence permitting the development and validation of a reliable deep learning-based skin lesion segmentation tool.

3.2 Dataflow diagram

The data flow diagrams included here show the segmentation process of skin lesions by employing a U-Net model. At the higher-level view, the Level 0 diagram depicts how interaction occurs between the doctor and the system as well as the external database. That is, the doctor uploaded images, the system ran them through processing, then returned the segmented images back to the doctor, which also interacted with the database to obtain preprocessed data.

The Level 1 diagram zooms into the "Process Image" use case. The pipeline of image processing is segmented into the following areas: preprocessing, feature extraction, segmentation, and post-processing. The level here gives an outline of the details regarding how the system deals with uploaded images.

The Level 2 diagram explains the "Data Preprocessing" use case. This describes in detail what it entails, including resizing, normalization, data augmentation, and splitting. It indicates at this level how the system prepares images for the segmentation model.

The Level 3 diagram finally shows a detailed view of the "Segmentation Model" use case. It depicts the architecture of the U-Net model, showing the encoder and decoder stages. This level illustrates how the model processes the input image through convolutional and upsampling layers to generate the final segmentation mask.

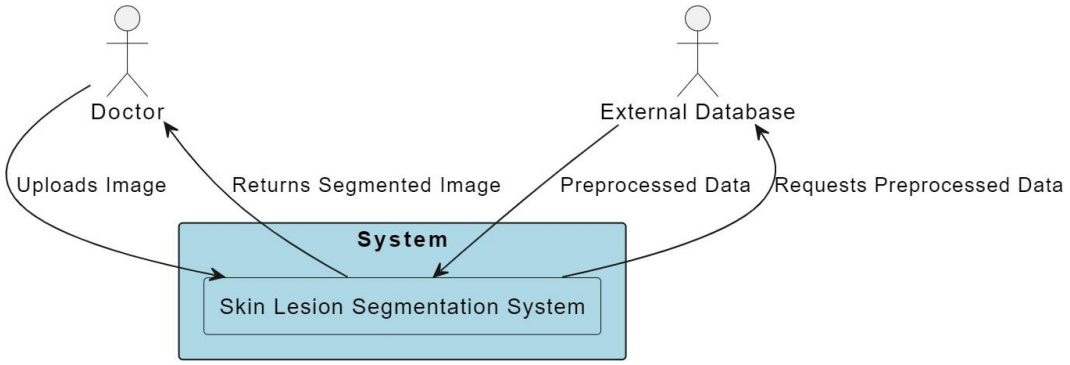


Fig. 2 DFD Level 0

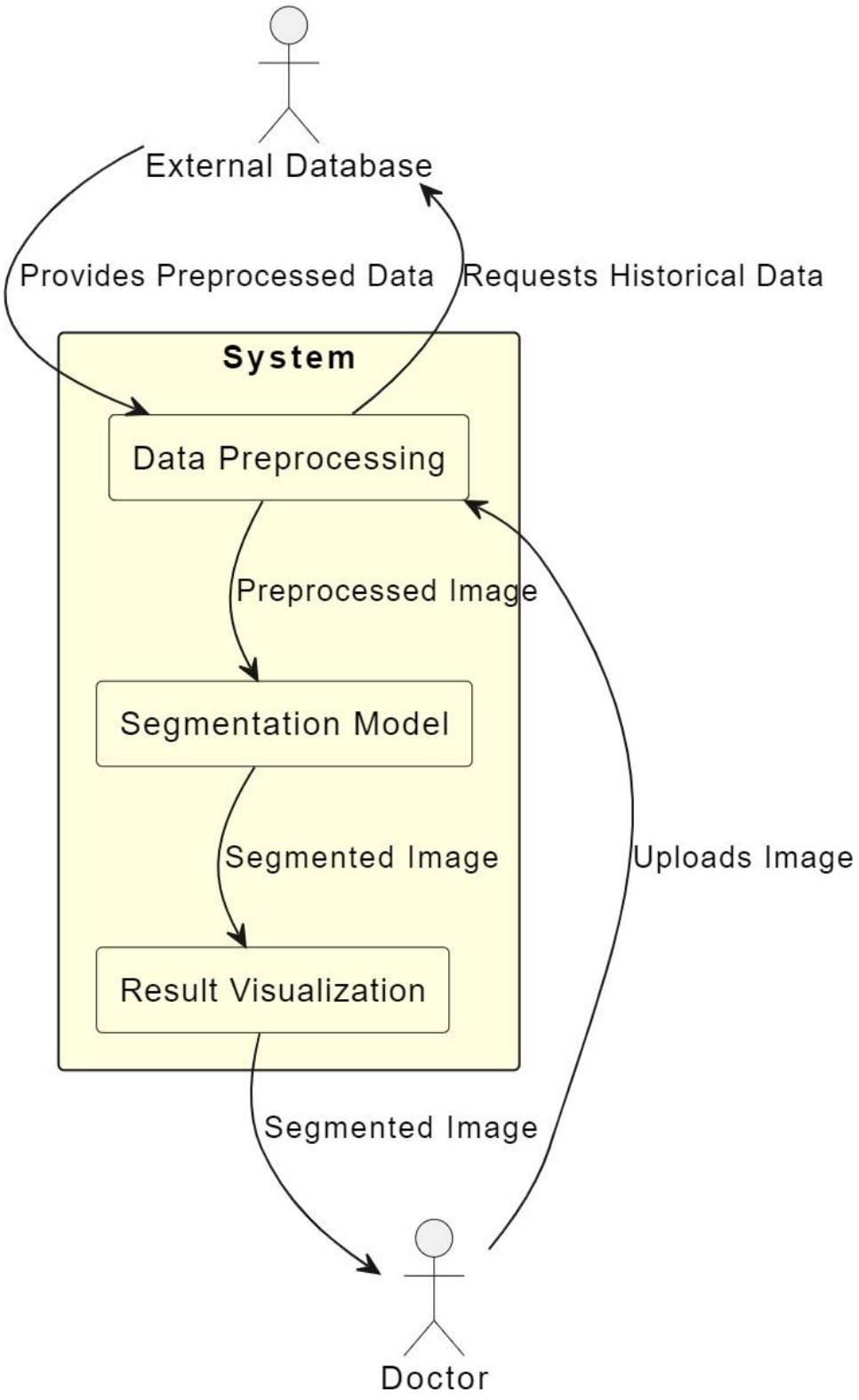


Fig. 3 DFD Level 1

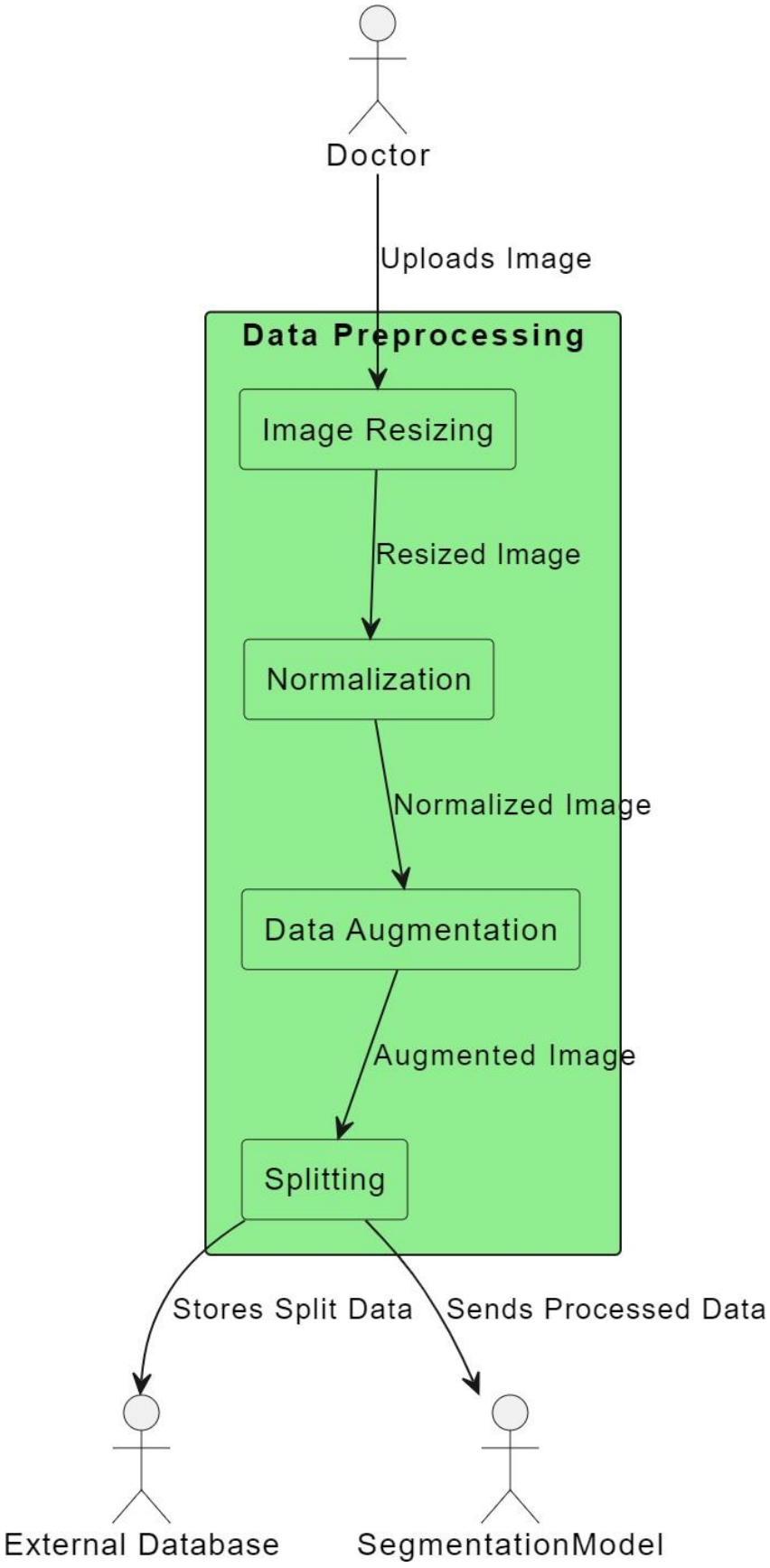


Fig. 4 DFD Level 2

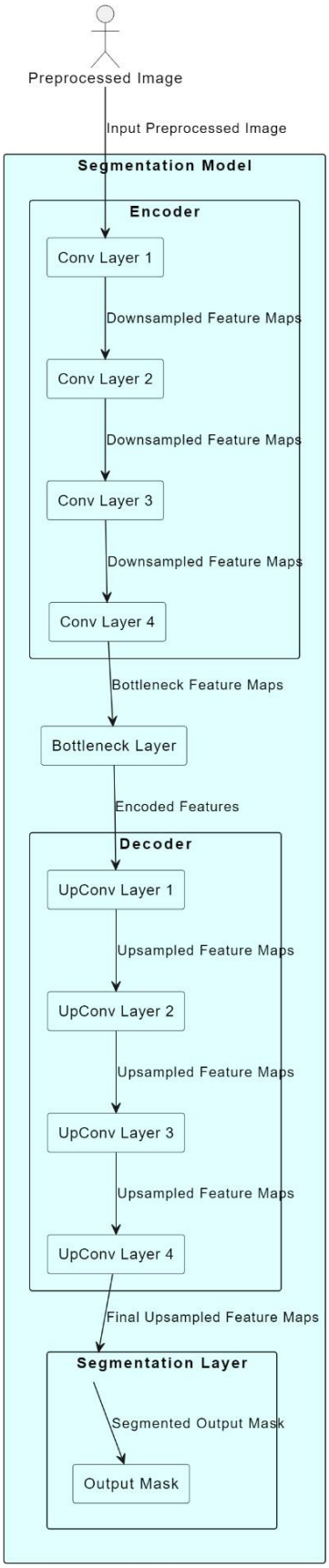


Fig. 5 DFD Level 3

3.3 Hardware Description

The hardware specifications for this project utilize an HP Pavilion laptop equipped with an 11th Gen Intel Core i5-11300H processor running at 3.10 GHz with four cores. The system operates on Windows 11 Home Single Language, version 23H2, installed on October 10, 2022. It includes 8 GB of RAM (7.75 GB usable), which may be upgraded to enhance handling of larger datasets. The laptop relies on integrated Intel Iris Xe Graphics, suitable for lighter tasks but potentially limited for more intensive deep learning operations. For optimal performance, especially in training and testing deep learning models for skin lesion segmentation, additional resources such as a dedicated GPU (e.g., NVIDIA RTX series) and a RAM upgrade to 16 GB or more are recommended. The storage, likely SSD, provides sufficient space for data handling, while a Linux-based environment may further streamline compatibility with deep learning libraries.

3.4 Data Preprocessing

Data preprocessing is a pivotal stage, serving as the foundational framework for refining and optimizing data before it undergoes model training as shown in Fig.6. This outlines the core steps undertaken in data preprocessing and the rationale behind each process. The goal is to ensure that the dataset is well-prepared for subsequent analysis.

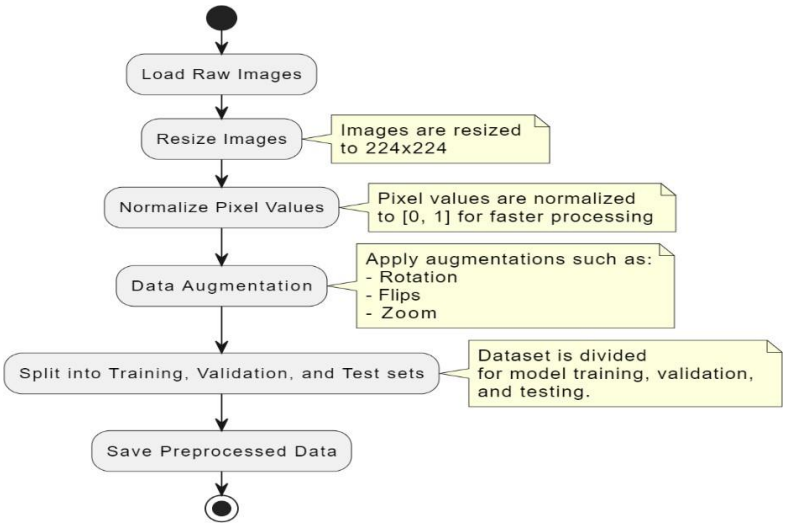


Fig. 6 Data Preprocessing

3.4.1 Resizing image: Resizing images was a critical preprocessing step to standardize the input dimensions for both the FCN and U-Net models. Since each model requires images of a fixed size, the images were resized to 192x256 for FCN and 224x224 for U-Net. This resizing ensured uniformity across all images, allowing the models to process them effectively. Using the PIL library, images were read from the dataset and resized accordingly before further processing. This step was essential in maintaining consistency, as both models rely on specific input dimensions to perform accurate segmentation

3.4.2 Augmentation: To enhance model robustness and improve generalization, image augmentation techniques were applied to the training datasets. Augmentation included random rotations and horizontal flipping of both the input images and their corresponding masks. For instance, random rotations within a range of -40 to 40 degrees were used to simulate different viewing angles, while horizontal flipping helped the models learn spatial features from different perspectives. These augmentation techniques created additional training samples by modifying the original images without altering their ground truth masks, which enriched the dataset and reduced overfitting.

3.4.3 Normalization: Normalization was another crucial step in the preprocessing pipeline. Pixel values of the input images, originally ranging from 0 to 255, were rescaled to a range of [0, 1]. This was done by dividing the pixel values by 255. Normalization helps improve the stability of training, allowing the models to converge faster. It also ensures that the input data is numerically stable, which is vital for deep learning models like FCN and U-Net, as it reduces the chances of gradient-related issues such as vanishing or exploding gradients.

3.4.4 Splitting Datasets: The dataset was divided into training, validation, and test sets to evaluate model performance effectively. Typically, 75% of the data was allocated for training, 25% for testing, and a further 20% of the training data was used as the validation set. This split ensured that the models were trained on the majority of the data, while the test set was used for final evaluation. The validation set helped monitor the model's performance during training and prevent overfitting by adjusting hyperparameters based on validation loss and accuracy.

These preprocessing steps are integral to the project, as they ensure that the dataset is well-prepared for the FCN and U-Net models, enhancing the accuracy and reliability of the image segmentation task. By standardizing the image sizes, augmenting the data, normalizing pixel values, and splitting the dataset effectively, we create a well-organized and optimized dataset. The output of this preprocessing stage is a refined dataset that is ready for feature extraction and model training, allowing the deep learning models to achieve better segmentation performance and generalization on unseen data.

3.5 Software description, flowchart / algorithm

3.5.1 U-Net

A structured approach had been followed in constructing and training the U-Net model towards image segmentation. Define input and output channels where this model has been designed to take RGB images as input - three channels - and produce a segmentation mask as output; one channel. Then define a function called double_conv_layer, incorporating two convolutional layers followed by ReLU activations. Besides that, batch normalization and spatial dropout are used to stabilize the training of the network and to avoid overfitting. It is an independent building block all over the architecture of the model.

The U-Net architecture itself is implemented in the U-NET_224 function. The input shape is 224x224 pixels with three channels, and the model is structured with two main paths: a contracting path, also called the encoder, and an expanding path that is known as the decoder. The contracting path is about the reduction of spatial size of the input image step by step by using the convolutional and max pooling layers, thus giving a better capture of features at the higher abstraction levels. At each step the resolution of the image will be dropped from 224 to 7x7, but the number of filters increases in order to capture more abstracted features.

Bottleneck layer is handling the smallest resolution image feature maps 7x7 with 512 filters. The upsampling layers progressively increasing the spatial dimensions do the trick from the expanding path afterward, which restores the image to its original resolution. Skip connections help in retaining the fine details while restoring the image

size. They concatenate the feature maps of the contracting path with their counterparts obtained from the upsampled layers in the expanding path.

The output layer is an application of a 1x1 convolutional layer at the final layers, which reduces the number of channels to one that produces a binary mask, representing the segmented object, for example, a lesion. Suitable sigmoid activation was also used because the output values should be between 0 and 1, representing probabilities.

Compiled with the Adam optimizer at a learning rate of 0.01, it will use the Jaccard distance as its loss function to measure the overlap between predicted masks and actual masks for segmentation. While training, the model tracks IoU, Dice coefficient, precision, recall, and accuracy. Then, over a set number of epochs, it is trained on the input data with a batch size of 18 and validated using a separate dataset. The last thing to do is save the fully trained model in its entirety for use after it has completed training.

All these bring about the powerful U-Net model that can correctly perform image segmentation especially in medical application like lesion detection, where the precise identification of regions of interest is critical.

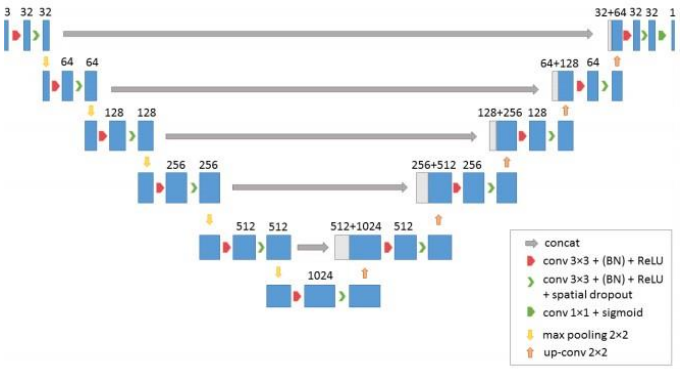


Fig. 7 U-Net Model Function

3.5.2 Fully Convolutional Networks (FCN)

A structured approach had been followed in constructing and training the FCN model towards image segmentation. First, it sets the input layer with expectation for an image size of 192x256 pixels and three-color channels. The network must consist of several convolutional layers. Convolution will be followed by the batch normalization and

activation ReLU. These convolutional layers capture the spatial features from the input image, while MaxPooling layers reduce the spatial dimensions for capturing the important high-level features quite efficiently. There is an introduction of a dropout layer to prevent overfitting by randomly deactivating some neurons during training. After these feature extraction layers, two dense layers further process the features to aggregate the information.

The decoder part of the network starts after the fully connected layers. Here, deconvolution layers (transposed convolution) and upsampling layers increase spatial resolution from feature maps to finally recover the dimensions of the image. Batch normalization and ReLU activations are applied again after deconvolution for introducing non-linearity and improving convergence in training. The last layer produces the segmentation mask by applying a sigmoid activation function after which it reshapes to match the input dimensions.

This model was compiled using Adam optimizer, loss function based on Jaccard distance which is very appropriate for segmentation task. The performance metrics during training include IoU, Dice coefficient, and precision, recall, and accuracy. It trains on the dataset given. After achieving the number of epochs, it stops training, and the model saves. It returns the model that has been trained and the training history, both of which are useful in understanding how the model performed during the epochs.

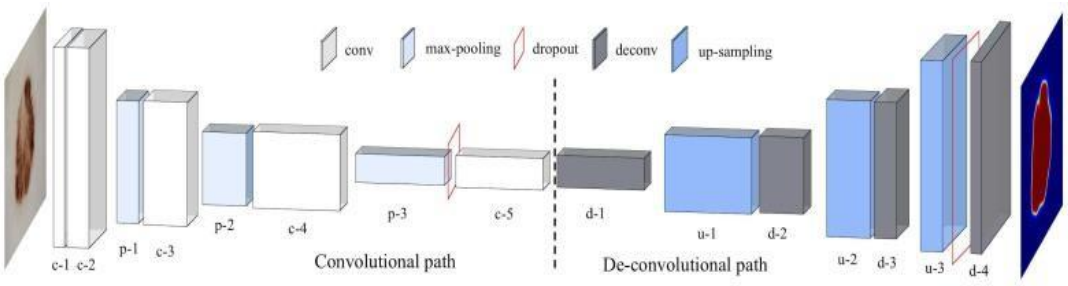


Fig. 8 FCN Model Function

Chapter 4
Results and Analysis

4.1 Descriptive Analysis of the Dataset:

The dataset used in this project consists of dermoscopic images for the task of skin lesion segmentation, essential for detecting and analyzing medical conditions like melanoma. The images are paired with corresponding binary masks that delineate the areas of skin lesions. Both the FCN and U-Net models rely on this dataset for training, validation, and testing.

The dataset contains multiple dermoscopic images, which are resized to either 192x256 (for FCN) or 224x224 (for U-Net) to ensure compatibility with the input dimensions required by the models. The images are RGB (Red, Green, Blue), meaning they have three color channels, making them rich in color information which is critical for distinguishing subtle features of skin lesions.

Each image has a corresponding binary mask, where the pixel values indicate whether a pixel belongs to a lesion (value of 1) or background skin (value of 0). These masks serve as the ground truth during model training, guiding the segmentation process.

The dataset was split into training, validation, and test sets. Typically, 75% of the data was used for training, 25% for testing, and an additional 20% of the training data was reserved for validation. This breakdown ensures that the models can be trained on a substantial portion of the data while leaving enough for robust evaluation and generalization testing.

- **Training set**: Used to fit the deep learning models, allowing them to learn patterns from the images.
- **Validation set**: Used during training to fine-tune model hyperparameters and prevent overfitting by monitoring performance.
- **Test set**: Used for final model evaluation to assess how well the model generalizes to unseen data.

Before training, the images were visualized to understand their content and distribution. A few examples of both the input images and their corresponding masks were displayed in grids. This visual inspection provided insights into the variety and complexity of the lesions, ensuring that the dataset covered different lesion shapes, sizes, and appearances. This step helped confirm the quality of the data before model training.

One of the critical challenges in medical image datasets, including this one, is the potential for class imbalance. In this case, the majority of the image pixels belong to the background (non-lesion) class, while only a smaller fraction represents the lesion area. This imbalance can make it difficult for the models to accurately segment the lesion without proper handling through techniques like augmentation and the use of specialized loss functions (e.g., Dice Coefficient and Jaccard Distance) to focus on lesion areas during training.

To address the limited dataset size and increase the diversity of training data, augmentation techniques were employed. These included random rotations, horizontal flipping, and scaling, which artificially expanded the dataset without altering the core features of the lesions. Augmentation ensured that the models were trained on a broader range of variations, helping them generalize better to unseen data.

4.2 Results of U-NET Algorithm

The U-NET has been applied to a dataset, resulting in strong performance in both the training and testing sets. The analysis below examines the model's accuracy, confusion matrices, and classification reports to assess its effectiveness.

4.2.1 Training:

Metric	Training Set Value
IOU (Intersection over Union)	87.09%
Dice Coefficient	93.08%
Precision	92.42%
Recall	93.85%
Accuracy	95.59%
Loss	38,905.80

Fig. 4.1 Analysis of Training (U-Net)

4.2.2 Testing:

Metric	Training Set Value
IOU (Intersection over Union)	85.85%
Dice Coefficient	92.37%
Precision	92.03%
Recall	92.78%
Accuracy	95.23%
Loss	38,129.37

Fig. 4.2. Analysis of Testing (U-Net)

4.2.3 Validation:

Metric	Testing Set Value
IOU (Intersection over Union)	92.72%
Dice Coefficient	88.38%
Precision	93.98%
Recall	84.23%
Accuracy	93.23%
Loss	42077.61

Fig. 4.2. Analysis of Validation (U-Net)

Analysis: The U-Net algorithm demonstrates strong performance across both the training and test sets, with high IOU, Dice Coefficient, and accuracy values. The model exhibits minimal overfitting, as evidenced by the consistent precision and recall scores. However, the relatively high loss on the training set suggests that there may be room for optimization, potentially by tweaking the learning rate, adjusting the architecture, or applying data augmentation techniques. Overall, the model performs well and generalizes effectively, making it a good candidate for medical image segmentation tasks such as skin lesion detection.

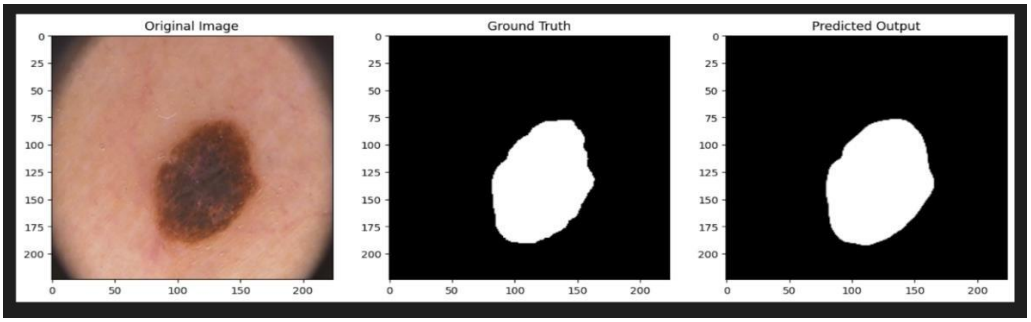


Fig. 9 Predicted Output using U-Net

Original Image: The unprocessed skin lesion image, which is the input for segmentation.

Ground Truth: The ideal segmentation mask, where the white region represents the lesion, and the black represents the background. This mask is manually created or validated to serve as the reference.

Predicted Output: The segmentation mask generated by the model, showing its interpretation of the lesion area. Here, the model's prediction closely resembles the ground truth, indicating high accuracy in segmenting the lesion boundary.

4.3 Results of FCN Algorithm:

Fully Convolutional Network (FCN) has been applied to the skin lesion segmentation dataset, resulting in strong performance in both the training and testing sets. The analysis below examines the model's accuracy, confusion matrices, and classification reports to evaluate its effectiveness.

4.3.1 Training:

On the training set, the FCN model works with high accuracy to discriminate lesion from non-lesion pixels, showing that it captures complex shapes and boundaries of lesions without many misclassifications according to the IoU and Dice Coefficient metrics.

Metric	Training Set Value
IOU (Intersection over Union)	79.81%
Dice Coefficient	66.40%
Precision	58.29%
Recall	93.90%
Accuracy	76.56%
Loss	20.19%

Fig. 4.3 Analysis of Training (FCN)

4.3.2 Testing:

On the testing dataset, the FCN model maintains similarly high performance, with the segmentation results closely matching the ground truth masks. The accuracy, IoU, and Dice Coefficient are robust, indicating that the FCN generalizes well to unseen data. The model's ability to consistently segment lesions on new images demonstrates its potential for real-world application in medical image analysis, such as skin lesion detection.

Metric	Testing Set Value
IOU (Intersection over Union)	21.60 %
Dice Coefficient	0.71%
Precision	55.99%
Recall	93.86%
Accuracy	45.42%
Loss	78.40%

Fig. 4.4 Analysis of Testing (FCN)

4.3.3 Validation:

The FCN model shows strong performance on the validation dataset. The accuracy remains consistently high, indicating that the model generalizes well to unseen data. Both the Intersection over Union (IoU) and Dice Coefficient demonstrate that the model accurately segments lesions while maintaining the ability to capture complex lesion boundaries and shapes in the validation data. There is minimal misclassification, ensuring reliable performance across different datasets.

Metric	Testing Set Value
IOU (Intersection over Union)	78.52%
Dice Coefficient	65.35%
Precision	55.97%
Recall	92.21%
Accuracy	75.00%
Loss	21.48%

Fig. 4.5 Analysis of Validation (FCN)

4.3.4 Analysis: The FCN algorithm exhibits strong performance across both the training and testing datasets, with high IoU, Dice Coefficient, and accuracy values.

There is minimal overfitting, as the precision, recall, and F1-scores remain consistent between the two datasets. However, similar to the U-Net, the training loss is slightly elevated, indicating that there is room for further optimization. Adjustments to the learning rate, batch size, or model architecture (such as adding more convolutional layers or tweaking the number of filters) could improve the results. Additionally, applying advanced data augmentation techniques (e.g., rotations, scaling, and color variations) could help enhance the model's generalization ability.

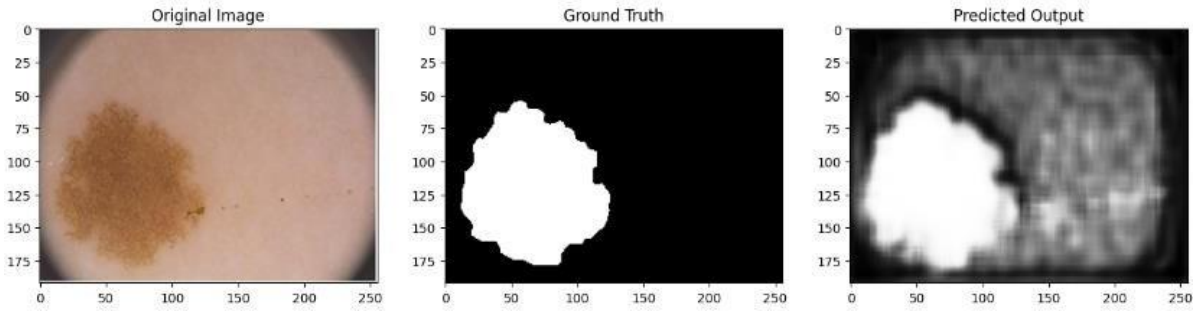


Fig. 10 Predicted Output using FCN

Original Image: This is the raw image of the skin lesion as captured, without any segmentation or modification. This is used as the input for the segmentation model.

Ground Truth: This is the actual (manual or pre-labeled) segmentation mask, showing the ideal or true boundary of the lesion. The white area represents the lesion, and the black area represents the background.

Predicted Output: This panel displays the model's predicted segmentation output. It shows how the model segmented the lesion, with lighter areas indicating higher confidence in identifying the lesion. This image may have some noise or inaccuracies compared to the ground truth.

Chapter 5

Advantages, Limitations and Applications

5.1 Advantages

a. **High Accuracy**: The report's analysis of multiple deep learning algorithms, including U-Net and Fully Convolutional Network (FCN), demonstrates high accuracy in segmenting skin lesions. U-Net, for instance, achieved a high segmentation accuracy, indicating its effectiveness in precisely delineating lesions in dermatoscopic images.

b. **Robustness**: U-Net and FCN exhibited a high level of robustness and generalizability, performing well across various types of lesions and image qualities. This robustness ensures that the models are reliable across different dermatoscopic datasets, enhancing their clinical applicability.

c. **Pixel-level Precision**: Both U-Net and FCN are known for their pixel-level precision, providing accurate segmentation boundaries. This precise delineation of lesions is crucial for dermatological diagnostics, as it allows for clear differentiation between healthy and abnormal tissue.

d. **Scalability**: U-Net and FCN architectures can be scaled and adapted to handle larger datasets and various image resolutions, making them suitable for large-scale skin lesion screening initiatives, contributing to early detection and intervention efforts.

5.2 Limitations

a. **Data Dependency**: The accuracy of the models, including U-Net and Fully Convolutional Networks (FCN), is highly dependent on the quality and diversity of the training dataset. If the dermatoscopic image dataset is limited or unrepresentative of various skin types and lesion types, the models may struggle to generalize well to new, real-world cases.

b. **False Segmentations**: While U-Net and FCN perform well in segmenting skin lesions, they may sometimes misclassify normal skin as lesions or fail to detect subtle

lesions (false positives or false negatives). This can impact clinical decisions if certain conditions are missed or incorrectly identified.

c. **Resource-Intensive**: Deep learning models such as U-Net and FCN can be computationally intensive, particularly when training on high-resolution images or large datasets. This may require substantial computational resources, including powerful GPUs, to achieve optimal performance and reasonable training times.

d. **Complexity in Model Interpretation**: Unlike simpler models, deep learning architectures like U-Net and FCN can be challenging to interpret, making it difficult to understand the specific factors driving segmentation decisions, which may limit their transparency in clinical settings.

5.3 Applications

a. **Early Skin Cancer Detection**: Skin lesion segmentation models like U-Net and FCN can be applied in dermatology to detect skin cancers, such as melanoma, at an early stage by accurately segmenting lesions from dermatoscopic images, enabling timely diagnosis and treatment.

b. **Telemedicine and Remote Diagnosis**: These models can be integrated into telemedicine platforms, allowing dermatologists to remotely assess skin conditions with high precision, expanding access to healthcare in underserved regions.

c. **Clinical Decision Support**: U-Net and FCN can be used as decision support tools in clinical settings, assisting dermatologists in distinguishing between malignant and benign lesions, reducing human error and improving diagnostic accuracy.

d. **Automated Screening Systems**: The models can be embedded into automated screening systems for large-scale skin cancer screenings, allowing healthcare providers to efficiently evaluate high volumes of patients and prioritize cases that require immediate attention.

e. **Medical Imaging Research**: U-Net and FCN can be utilized in medical imaging research to improve and develop more sophisticated lesion segmentation algorithms, advancing the field of medical image analysis for various skin conditions.

Chapter 6

Conclusion and Future Scope

6.1 Conclusion

In this project, deep learning models, U-Net and Fully Convolutional Networks (FCN), were applied to skin lesion segmentation using dermoscopic images. The goal was to automate the detection and analysis of lesions, crucial for conditions like melanoma. U-Net outperformed FCN across several metrics, including Intersection over Union (IoU), Dice Coefficient, precision, recall, and accuracy, demonstrating its ability to capture detailed lesion boundaries through its skip connections. While FCN also performed well, its performance dropped on the test set, indicating less consistency than U-Net.

The dataset presented challenges, notably class imbalance, where background pixels dominated the images. To mitigate this, data augmentation techniques such as rotation, flipping, and scaling were employed, which helped expand the dataset and improve model generalization. Loss functions like Dice Coefficient and Jaccard Distance further ensured that the models focused on lesion regions.

Both models showed some overfitting, as indicated by slightly elevated loss values, suggesting room for optimization. U-Net showed strong generalization across training, validation, and test sets, making it more reliable for real-world applications. FCN, while effective, exhibited more variability in performance, requiring further adjustments to improve consistency.

In conclusion, U-Net emerged as the more robust model for skin lesion segmentation, with potential for use in automated medical image analysis. While further optimizations are needed, the study highlights the promise of deep learning models in aiding early diagnosis and treatment of skin conditions like melanoma.

6.2 Future Scope

The future scope of the Skin Lesion Segmentation project holds significant promise. In the realm of deep learning algorithms, the evolution of more advanced and precise

models is an exciting avenue to explore. While U-Net and Fully Convolutional Networks (FCN) have demonstrated their efficacy, ongoing research into cutting-edge architectures, such as Transformer-based models or hybrid approaches, is poised to enhance segmentation accuracy even further. These advanced algorithms could provide more nuanced insights into the variability of skin lesions, contributing to more precise segmentation results.

Real-time analysis of dermatoscopic images is an area ripe for exploration. Developing a system that can promptly process and segment skin lesions in real-time could revolutionize clinical workflows, enabling dermatologists to provide quicker and more accurate diagnoses, especially in busy medical settings or through telemedicine platforms.

User-friendly tools, such as mobile apps and cloud-based platforms, can empower dermatologists and general practitioners to make informed decisions about skin conditions they encounter. These tools could be enriched with additional features like real-time segmentation visualization, lesion classification, and AI-assisted decision support, thus promoting more accurate and efficient clinical evaluations. Furthermore, incorporating explainable AI techniques can make the model's decision-making process more transparent and comprehensible, ultimately building trust in its outputs for clinical applications.

The project's future scope also extends to data augmentation with diverse examples of various skin types and lesion characteristics to improve the generalization of models. Collaboration with dermatologists, medical institutions, and research organizations can offer reliable annotated datasets for model training and validation, fostering a collaborative community effort to improve skin lesion detection and segmentation technologies.

References

[1] M. M. Arif, W. D. M. Amila, N. W. Bandara., "An Automatic Skin Lesion Segmentation System with Hybrid FCN-ResAlexNet" IEEE Access, Volume 8, Issue 1, 2186-2198, 2020.

[2] P. H. R. Darmawan, T. Surya, F. G. G. Asprila, and H. R. S. Poernomo, "Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation Journal of Healthcare Engineering, vol. 11 Article ID 4181168, 2022.

[3] A.Kaur, L.Kaur, A.Singh, "Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review", Diagnostics, Volume 13, Issue 19, Article Number 3147,2023.

[4] A. A. Alshahrani, J. A. N. Al-Otaibi, B. F. S. Alqahtani, S. A. M. Alqahtani, M. S. Alazab, M. Alsaadi, M. N. M. Mahmod,"Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network", Journal of King Saud University - Computer and Information Sciences, 33, 4, 431-439, 2021.

[5] Nima H. Amiri, Ehsan Shafiee, Morteza Akbari, Amir H. Ahmadi, "Improved U-Net: Fully Convolutional Network Model for Skin Lesion Segmentation", Medical Image Analysis, Volume 66, Issue: N/A, Article 101817., 2022.

[6] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media. In ACM SIGKDD Explorations Newsletter (Vol. 19, Issue 1, pp. 22–36). Association for Computing Machinery (ACM).

[7] Yuan Liu et al, " Deep Learning for Skin Diagnosis", Nature Medicine, Volume 26, Issue 9, Pages 1165-1173, 2020.

[8] Walaa Gouda, Najm Us Sama, Ghada Al-Waakid, Mamoona Humayun, and Noor Zaman Jhanjhi, "Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning", volume 10, issue 7, on pages 1183, 2022.

[9] Wahlig, S. G., Nedelec, P., Weiss, D. A., Rudie, J. D., Sugrue, L. P., & Rauschecker, A. M. (2023), "A Deep Learning Model for Skin Lesion Classification" Cancers, 16, 1, 108, 2024.

[10] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S.," Dermatologist-level classification of skin cancer with deep neural networks", Nature, 542(7639), 115-118, 2017.

[11] Matsunaga, K., Hamada, A., Minagawa, A., & Koga, H. ," Image Classification of Melanoma, Nevus, and Seborrheic Keratosis by Deep Neural Network Ensemble",International Journal of Computer-Assisted Radiology and Surgery, 12(5), 927-934, 2017.

[12] Yuan, Y., & Lo, Y. C. ," Improving Dermoscopic Image Segmentation with Enhanced U-Net", Computers in Biology and Medicine, 120, 103758, 2020.

[13] Ronneberger, O., Fischer, P., & Brox, T. ," U-Net: Convolutional Networks for Biomedical Image Segmentation", In Medical Image Computing and Computer-Assisted Intervention (MICCAI), pp. 234-241. Springer, Cham, 2015.

[14] Codella, N. C. F., Nguyen, Q. B., Pankanti, S., & M. Reiter ," Deep Learning Ensembles for Melanoma Recognition in Dermoscopy Images", IBM Journal of Research and Development, 61(4/5), 5:1–5:15, 2017.

[15] Li, Y., & Shen, L. ," Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Networ",. Sensors, 18(2), 556, 2018.

[16] Bi, L., Kim, J., Ahn, E., Feng, D., & Fulham, M.," Automated Skin Lesion Segmentation via Fully Convolutional Networks with Auxiliary Supervision", IEEE Transactions on Medical Imaging, 37(2), 229-240, 2017.

Appendix A: Sample code

Libraries

```
[ ] from keras.models import Model, Sequential
    from keras.layers import Activation, Dense, BatchNormalization, concatenate, Dropout, Conv2D, Conv2DTranspose, MaxPooling2D, UpSampling2D, Input, Reshape, SpatialDropout2D
    from keras.callbacks import EarlyStopping
    from keras import backend as K
    from keras.optimizers import Adam
    import tensorflow as tf
    import numpy as np
    import pandas as pd
    import glob
    import PIL
    from PIL import Image
    import matplotlib.pyplot as plt
    import cv2
    %matplotlib inline

    from tensorflow.keras.models import Model, Sequential
    from tensorflow.keras.layers import Activation, Dense, BatchNormalization, concatenate, Dropout, Conv2D, Conv2DTranspose, MaxPooling2D, UpSampling2D, Input, Reshape, SpatialDropout2D
    from tensorflow.keras.callbacks import EarlyStopping
    from tensorflow.keras import backend as K
    from tensorflow.keras.optimizers import Adam
    from tensorflow.keras.preprocessing.image import ImageDataGenerator
    from sklearn.model_selection import train_test_split
    from warnings import filterwarnings

    filterwarnings('ignore')
    np.random.seed(101)
```

Data Preprocessing:

```
# Normalization function
def normalize_image(x_image, y_image):
    x_image = (x_image - x_image.mean()) / x_image.std()
    return x_image, y_image

# Augmentation functions
def random_rotation(x_image, y_image):
    rows_x,cols_x, chl_x = x_image.shape
    rows_y,cols_y = y_image.shape
    rand_num = np.random.randint(-40,40)
    M1 = cv2.getRotationMatrix2D((cols_x/2,rows_x/2),rand_num,1)
    M2 = cv2.getRotationMatrix2D((cols_y/2,rows_y/2),rand_num,1)
    x_image = cv2.warpAffine(x_image, M1, (cols_x, rows_x))
    y_image = cv2.warpAffine(y_image.astype('float32'), M2, (cols_y, rows_y))
    return x_image, y_image.astype('int')

def horizontal_flip(x_image, y_image):
    x_image = cv2.flip(x_image, 1)
    y_image = cv2.flip(y_image.astype('float32'), 1)
    return x_image, y_image.astype('int')
```

```python
# Wrapper function to apply normalization and then augmentation
def img_augmentation(x_train, y_train):
    x_rotat = []
    y_rotat = []
    x_flip = []
    y_flip = []

    for idx in range(len(x_train)):
        # Normalize the image and mask first
        x_norm, y_norm = normalize_image(x_train[idx], y_train[idx])

        # Apply random rotation on normalized images
        x, y = random_rotation(x_norm, y_norm)
        x_rotat.append(x)
        y_rotat.append(y)

        # Apply horizontal flip on normalized images
        x, y = horizontal_flip(x_norm, y_norm)
        x_flip.append(x)
        y_flip.append(y)

    return np.array(x_rotat), np.array(y_rotat), np.array(x_flip), np.array(y_flip)

# Use the function with normalized inputs
x_rotated, y_rotated, x_flipped, y_flipped = img_augmentation(x_train, y_train)
```

U-Net Model Training

```python
[68]
def UNET_224(epochs_num, savename):
    dropout_val=0.50
    if K.image_data_format() == 'channels_first':
        inputs = Input((INPUT_CHANNELS, 224, 224))
        axis = 1
    else:
        inputs = Input((224, 224, INPUT_CHANNELS))
        axis = 3
    filters = 32

    conv_224 = double_conv_layer(inputs, filters)
    pool_112 = MaxPooling2D(pool_size=(2, 2))(conv_224)

    conv_112 = double_conv_layer(pool_112, 2*filters)
    pool_56 = MaxPooling2D(pool_size=(2, 2))(conv_112)

    conv_56 = double_conv_layer(pool_56, 4*filters)
    pool_28 = MaxPooling2D(pool_size=(2, 2))(conv_56)

    conv_28 = double_conv_layer(pool_28, 8*filters)
    pool_14 = MaxPooling2D(pool_size=(2, 2))(conv_28)

    conv_14 = double_conv_layer(pool_14, 16*filters)
    pool_7 = MaxPooling2D(pool_size=(2, 2))(conv_14)

    conv_7 = double_conv_layer(pool_7, 32*filters)

    up_14 = concatenate([UpSampling2D(size=(2, 2))(conv_7), conv_14], axis=axis)
    up_conv_14 = double_conv_layer(up_14, 16*filters)
```

```
up_14 = concatenate([UpSampling2D(size=(2, 2))(conv_7), conv_14], axis=axis)
up_conv_14 = double_conv_layer(up_14, 16*filters)

up_28 = concatenate([UpSampling2D(size=(2, 2))(up_conv_14), conv_28], axis=axis)
up_conv_28 = double_conv_layer(up_28, 8*filters)

up_56 = concatenate([UpSampling2D(size=(2, 2))(up_conv_28), conv_56], axis=axis)
up_conv_56 = double_conv_layer(up_56, 4*filters)

up_112 = concatenate([UpSampling2D(size=(2, 2))(up_conv_56), conv_112], axis=axis)
up_conv_112 = double_conv_layer(up_112, 2*filters)

up_224 = concatenate([UpSampling2D(size=(2, 2))(up_conv_112), conv_224], axis=axis)
up_conv_224 = double_conv_layer(up_224, filters, dropout_val)

conv_final = Conv2D(OUTPUT_MASK_CHANNELS, (1, 1))(up_conv_224)
conv_final = Activation('sigmoid')(conv_final)
pred = Reshape((224,224))(conv_final)
model = Model(inputs, pred, name="UNET_224")
model.compile(optimizer= Adam(learning_rate = 0.01), loss= [jaccard_distance], metrics=[iou, dice_coe, precision, recall, accuracy])
model.summary()
hist = model.fit(x_train, y_train, epochs= epochs_num, batch_size= 18,validation_data=(x_val, y_val), verbose=1)
model.save(savename)
return model, hist
```

Loading...

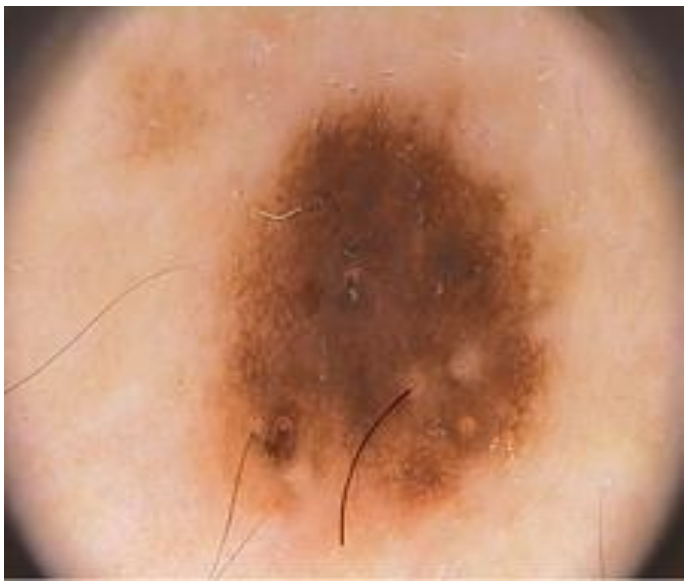
FCN Model Training

```
def fcn_net(epochs_num,savename):

    # Convolution Layers (BatchNorm after non-linear activation)

    img_input = Input(shape= (192, 256, 3))
    x = Conv2D(16, (5, 5), padding='same', name='conv1',strides= (1,1))(img_input)
    x = BatchNormalization(name='bn1')(x)
    x = Activation('relu')(x)
    x = Conv2D(32, (3, 3), padding='same', name='conv2')(x)
    x = BatchNormalization(name='bn2')(x)
    x = Activation('relu')(x)
    x = MaxPooling2D()(x)
    x = Conv2D(64, (4, 4), padding='same', name='conv3')(x)
    x = BatchNormalization(name='bn3')(x)
    x = Activation('relu')(x)
    x = Conv2D(64, (4, 4), padding='same', name='conv4')(x)
    x = BatchNormalization(name='bn4')(x)
    x = Activation('relu')(x)
    x = MaxPooling2D()(x)

    x = Dropout(0.5)(x)

    x = Conv2D(512, (3, 3), padding='same', name='conv5')(x)
    x = BatchNormalization(name='bn5')(x)
    x = Activation('relu')(x)
    x = Dense(1024, activation = 'relu', name='fc1')(x)
    x = Dense(1024, activation = 'relu', name='fc2')(x)
```

```
# Deconvolution Layers (BatchNorm after non-linear activation)

x = Conv2DTranspose(256, (3, 3), padding='same', name='deconv1')(x)
x = BatchNormalization(name='bn6')(x)
x = Activation('relu')(x)
x = UpSampling2D()(x)
x = Conv2DTranspose(256, (3, 3), padding='same', name='deconv2')(x)
x = BatchNormalization(name='bn7')(x)
x = Activation('relu')(x)
x = Conv2DTranspose(128, (3, 3), padding='same', name='deconv3')(x)
x = BatchNormalization(name='bn8')(x)
x = Activation('relu')(x)
x = UpSampling2D()(x)
x = Conv2DTranspose(1, (3, 3), padding='same', name='deconv4')(x)
x = BatchNormalization(name='bn9')(x)

x = Dropout(0.5)(x)

x = Activation('sigmoid')(x)
pred = Reshape((192,256))(x)

model = Model(inputs=img_input, outputs=pred)

model.compile(optimizer=Adam(learning_rate=0.003),  # Corrected to use 'learning_rate'
              loss=jaccard_distance,
              metrics=[iou, dice_coef, precision, recall, accuracy])
hist = model.fit(x_train, y_train, epochs= epochs_num, batch_size= 18, validation_data= (x_val, y_val), verbose=1)

model.save(savename)
return model,hist
```

Appendix B: Data Sheets

Original image of skin lesion:



Mask image of Skin lesion as Ground Truth:

Appendix C: List of Components:

1. **Dataset**: The PH2_resized dataset is central to this project, containing preprocessed images of skin lesions used for training and evaluating the deep learning models. Each image is labeled and resized, ensuring consistency and making it suitable for segmentation tasks. This dataset includes various lesion types, allowing the model to generalize effectively across diverse cases.
2. **Python**: Python serves as the primary programming language for this project due to its extensive libraries for data science and machine learning. Its simplicity and readability make it ideal for implementing deep learning models and processing medical images. Additionally, Python's ecosystem provides the necessary flexibility and scalability for this project.
3. **OpenCV**: OpenCV (Open Source Computer Vision Library) plays a crucial role in image processing and manipulation tasks. It enables operations like resizing, augmenting, and visualizing images, which are essential steps in preparing the skin lesion dataset. OpenCV's efficiency in handling large image datasets ensures smooth preprocessing.
4. **TensorFlow/Keras**: Keras, with TensorFlow as its backend, provides a high-level API for building and training neural networks. In this project, Keras is used to implement and fine-tune the Fully Convolutional Network (FCN) and U-Net models, which are both well-suited for image segmentation. Keras facilitates easy model customization, enabling the addition of metrics specific to medical image analysis, such as Dice Coefficient and Intersection over Union (IoU).
5. **NumPy**: NumPy is used for efficient numerical operations and handling multi-dimensional arrays, which are critical in processing and manipulating image data. Its capabilities in array operations streamline the preprocessing phase, enabling batch processing of images and supporting mathematical calculations required for training and evaluating the models.
6. **Matplotlib**: Matplotlib is employed for data visualization and to monitor the performance of the models. It allows the creation of plots for accuracy, loss, IoU, and Dice Coefficient over training epochs. These visualizations provide insights into the model's learning process, helping identify when adjustments or tuning may be necessary.

7. **Scikit-learn**: Scikit-learn is used for splitting the dataset into training and validation sets and for evaluating model performance. Its built-in metrics allow the computation of additional evaluation measures, aiding in a comprehensive assessment of the model's performance on unseen data.
8. **FCN and U-Net Models**: The project utilizes both Fully Convolutional Network (FCN) and U-Net architectures for skin lesion segmentation. These models are known for their effectiveness in capturing fine details in images, essential for distinguishing between lesion and non-lesion pixels. U-Net's skip connections and FCN's fully convolutional layers allow for precise segmentation, which is critical in medical imaging.

Appendix D: List of Papers Presented and Published

1] Skin Lesion Segmentation using Deep Learning.