	# Hacettepe University Computer Engineering Department **BBM479/480 End of Project Report**

Project Details

Title	Pressure Ulcers
Supervisor	Cemil Zalluhoğlu

Group Members

	Full Name	Student ID
1	Mert Tazeoğlu	21946606
2	Yusuf Efe Kalaycı	21827517
3	Asım Ateş	21827076
4		

Abstract of the Project (/ 10 Points)

Explain the whole project shortly including the introduction of the field, the problem statement, your proposed solution and the methods you applied, your results and their discussion, expected impact and possible future directions. The abstract should be between 250-500 words.

Pressure ulcers, also known as bedsores or pressure sores, pose a significant challenge in medical care, particularly for patients with mobility issues. These injuries result from prolonged pressure on the skin, leading to varying degrees of tissue damage, classified into several grades. The early detection and accurate classification of pressure ulcers are crucial for effective treatment and prevention of progression. However, manual assessment of these ulcers is subjective and depends heavily on the clinician's expertise, which varies widely and may lead to inconsistent evaluations.

This project addresses the need for an objective and standardized tool by developing a deep learning-based system for the segmentation and classification of pressure ulcers from medical images. The proposed solution utilizes state-of-the-art convolutional neural networks (CNNs) to segment pressure ulcers from surrounding healthy tissue and subsequently classify them into their respective grades based on severity.

The methodology involves training a segmentation model using U-Net architecture, renowned for its efficiency in medical image analysis. The model processes input images to delineate the exact boundaries of ulcers. Following segmentation, a classification model built on the VGG-16 architecture categorizes the identified ulcers into one of four grades. This dual-stage approach ensures high precision in both localization and classification tasks. In addition, during this time many operations are carried out in the background, especially in the image processing area.

Our dataset comprises thousands of annotated images from hospital archives, which have been rigorously preprocessed to enhance model training. Actually we used two different datasets named 'Ventura Dataset' and 'Elazığ Dataset'. Initial results demonstrate an accuracy of over %83 accuracy rate in classification tasks, just similar to traditional manual methods. The discussion of results highlights the models' ability to handle various imaging conditions and ulcer stages, showcasing robustness and scalability. To further boost the classification performance, the models incorporate a region of interest (ROI) technique, focusing on specific areas within the images that are most indicative of ulcer severity. This targeted approach enhances the precision and accuracy of the classification phase, ensuring that the most critical aspects of the ulcers are considered during model training.

The impact of this project is twofold: enhancing patient care through timely and accurate ulcer management and providing a tool for continuous monitoring and assessment in clinical settings. Additionally, the automation reduces the workload on healthcare professionals and standardizes ulcer care across different facilities.

Future directions include expanding the dataset to include more diverse patient demographics and ulcer types to further improve the model's generalizability. Integrating this technology into mobile devices for use in telemedicine and home care settings could also broaden its applicability, making expert-level ulcer assessment accessible in remote areas. This project not only advances the field of dermatological imaging but also sets a precedent for the application of artificial intelligence in preventive medicine.

Introduction, Problem Definition & Literature Review (/ 20 Points)

Introduce the field of your project, define your problem (as clearly as possible), review the literature (cite the papers) by explaining the proposed solutions to this problem together with limitations of these problems, lastly write your hypothesis (or research question) and summarize your proposed solution in a paragraph. Please use a scientific language (you may assume the style from the studies you cited in your literature review). You may borrow parts from your previous reports but update them with the information you obtained during the course of the project. This section should be between 750-1500 words.

Introduction

Pressure ulcers, also known as bedsores or pressure sores, are localized injuries to the skin and/or underlying tissue resulting from prolonged pressure or friction. They often occur in patients with limited mobility, such as those who are bedridden or confined to a wheelchair. The occurrence of pressure ulcers is a significant concern in healthcare settings due to the associated morbidity, prolonged hospital stays, increased healthcare costs, and diminished quality of life for patients. Medical image segmentation and classification have emerged as pivotal tools in the early detection and management of pressure ulcers, enabling healthcare professionals to provide timely and appropriate interventions.

Problem Definition

The primary challenge in the management of pressure ulcers lies in their early detection and accurate classification. Traditional methods rely heavily on manual inspection, which can be subjective and prone to variability among healthcare providers. This project aims to develop an automated system for the segmentation and classification of pressure ulcers from medical images. By leveraging advanced machine learning and deep learning techniques, the project seeks to enhance the accuracy and efficiency of pressure ulcer detection and classification, thereby improving patient outcomes and reducing the burden on healthcare systems.

Literature Review

Several studies have addressed the problem of medical image segmentation and classification using various imaging modalities and machine learning algorithms. The literature reveals a progression from traditional image processing techniques to more sophisticated approaches.

1 - Traditional Image Processing Approaches: In the initial phases of pressure ulcer detection, traditional image processing techniques were predominantly used. These methods included edge detection, thresholding, and morphological operations. These techniques focused on leveraging color and texture features to distinguish between the different stages of pressure ulcers. Despite their early utility, these approaches often encountered challenges such as variations in lighting, skin tones, and ulcer appearances, which could lead to inconsistent and unreliable results.

2 - Machine Learning Techniques: As the field evolved with advancements in technology, machine learning offered new avenues for improving the accuracy of pressure ulcer classification. Researchers incorporated advanced classifiers like support vector machines (SVM) and random forests into their work. These newer approaches combined color, texture, and shape features to more effectively classify pressure ulcers into their respective stages. Although these machine learning methods represented a significant improvement over traditional image processing techniques, they still involved manual feature extraction, a process that was both time-consuming and susceptible to human error.

3 - Deep Learning Approaches: Recent developments in the field have been significantly influenced by advancements in deep learning, especially through the use of convolutional neural networks (CNNs). These networks have excelled in medical image analysis, primarily because they automatically learn relevant features directly from raw image data, greatly enhancing analysis efficiency and accuracy. For example, deep learning techniques have been employed for both the segmentation and classification of pressure ulcers, using architectures such as U-Net, which have shown high accuracy and robustness across various datasets. This approach has yielded superior performance in classifying the stages of pressure ulcers compared to traditional and other machine learning models, marking a shift towards more automated, reliable, and efficient detection and classification methods.

Hypothesis and Proposed Solution

The hypothesis of this project is that integrating a deep learning framework that combines both segmentation and classification tasks can significantly enhance the accuracy and efficiency of pressure ulcer detection and management. Leveraging a vast and varied dataset of medical images, the proposed solution aims to develop a robust model capable of generalizing well across various clinical environments.

The solution involves a methodical exploration of different machine learning models (AlexNet, DenseNet201, EfficientNet_B0, EfficientNet_B7, VGG16, MobileNet_V2, ResNet152 and GoogleNet) and parameters to identify the architecture that yields the highest classification success. This will be applied similarly for segmentation tasks, testing various U-Net models (especially Double U-Net and U-Net++) due to its widespread recommendation and use in medical imaging. The chosen U-Net model for segmentation will aim to precisely delineate ulcer boundaries, providing crucial data for the subsequent classification phase.

To further augment the classification accuracy, the project will incorporate the Attention Branch Master methodology. This sophisticated approach significantly enhances the classification process by employing segregated processing streams, which separately handle standard input and Region of Interest (ROI) input. This dual-stream setup allows for more focused analysis, where the standard input provides a broad contextual understanding, and the ROI input concentrates on specific areas that are critical for accurate diagnosis.

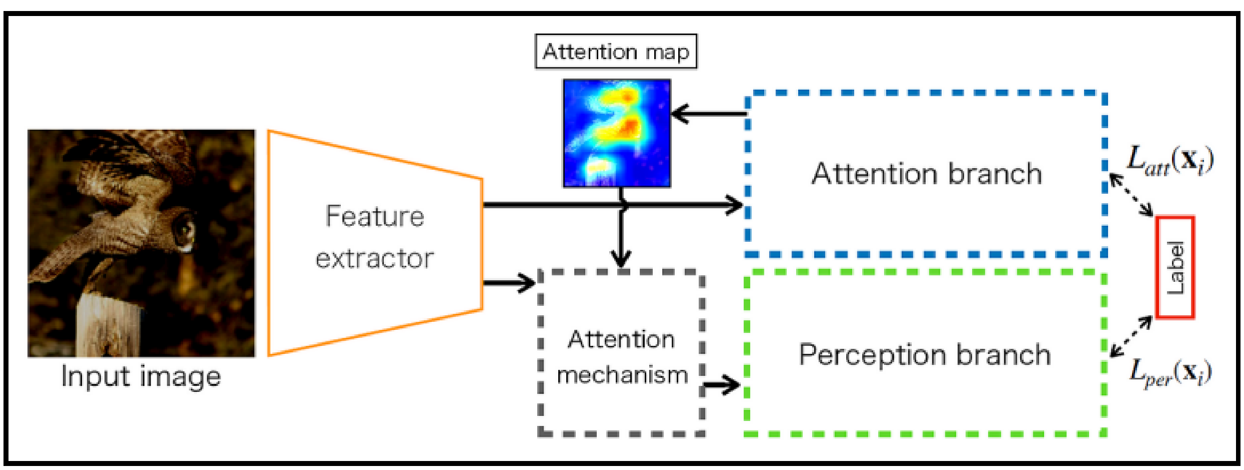

Figure-1: Working Principle of Attention Branch Master
(© Sik-Ho Tsang)

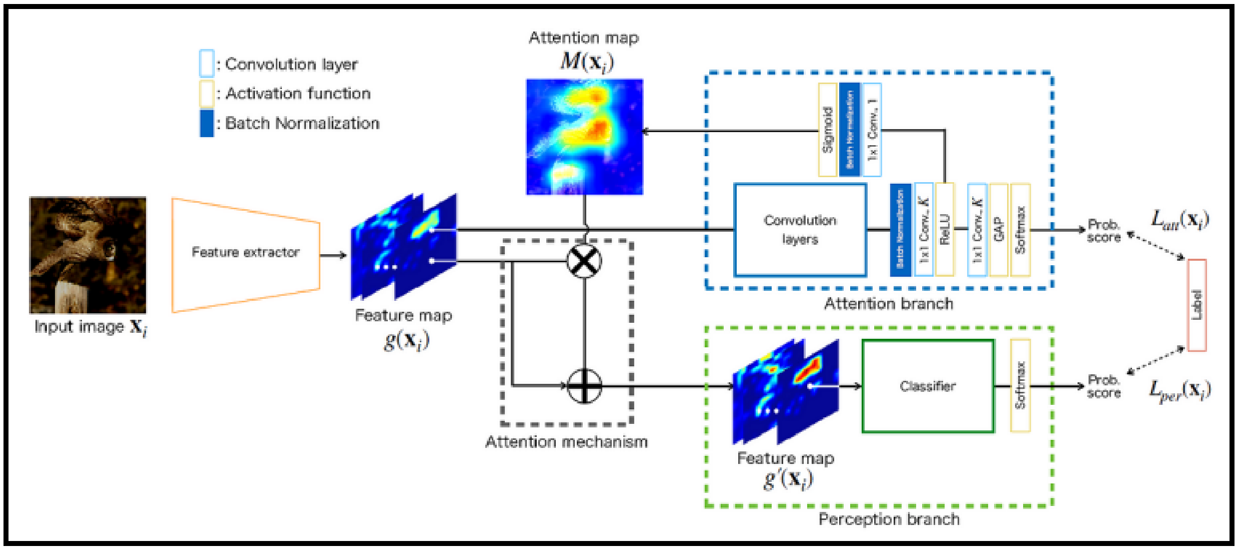

Figure-2: Detailed Working Principle of Attention Branch Master
(© Sik-Ho Tsang)

The Region of Interest (ROI) is a pivotal component in medical image analysis, particularly in tasks requiring detailed examination of specific anatomical or pathological features. In the context of pressure ulcer detection, ROI refers to the targeted sections of the image that likely contain ulcers. By focusing on these regions, the model can apply deeper, more concentrated computational resources to analyze subtle features and variations that might be indicative of different ulcer stages. This targeted approach not only improves the accuracy of the classification but also enhances the interpretability of the model by providing clear insights into which features influenced the decision-making process.

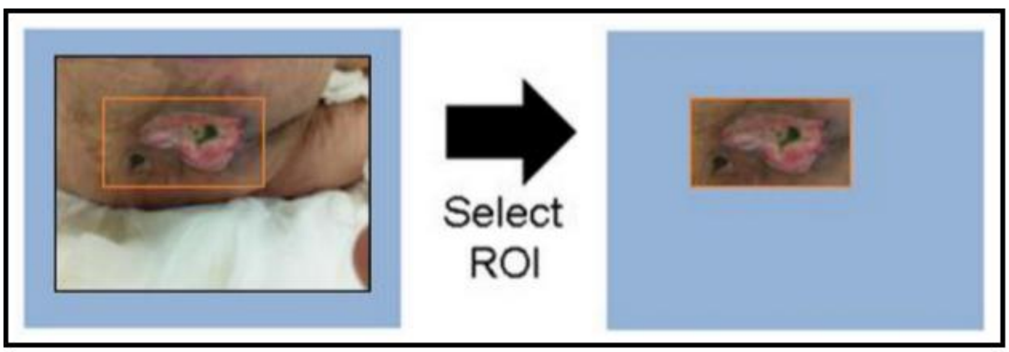

Figure-3: Working Principle of Region of Interest (RoI) in Image Processing

Moreover, the Attention Branch Master method integrates attention mechanisms that automatically identify and prioritize these ROIs during the training phase. This dynamic identification and prioritization process ensures that the model learns to recognize and focus on the most informative parts of an image, adapting to new and varying cases with greater efficacy. This methodology not only refines the model's performance but also provides a framework for continuous improvement as more data becomes available, ensuring that the system remains robust in different settings.

Methodology (/ 25 Points)

Explain the methodology you followed throughout the project in technical terms including datasets, data pre-processing and featurization (if relevant), computational models/algorithms you used or developed, system training/testing (if relevant), principles of model evaluation (not the results). Using equations, flow charts, etc. are encouraged. Use sub-headings for each topic. Please use a scientific language. You may borrow parts from your previous reports but update them with the information you obtained during the course of the project. This section should be between 1000-1500 words (add pages if necessary).

1 - Datasets

The project utilized two distinct medical image datasets containing representations of pressure ulcers, each curated to support both classification and segmentation tasks. The first dataset, named Ventura V2, includes a total of 1,210 images split into 911 training images and 299 test images, adhering to a 75% training and 25% testing distribution. The second dataset, referred to as Elazığ, comprises 1,099 images with 878 allocated for training and 221 designated for testing, maintaining an 80% training and 20% testing ratio. These splits were pre-established; therefore, we did not alter the ratios or the specific images within each split. Had the splits not been predetermined, we would have likely maintained the same distribution ratios to ensure a consistent approach to training and testing our models. The Ventura dataset was predominantly used throughout the project due to its comprehensive and diverse set of images suitable for our analytical needs.

Figure-4: Pressure Ulcer Images From 2 Different Datasets

2 - Data Pre-Processing and Featurization

2.1 - Creating Cropped Image Datasets

To optimize the training of our model, we produced a cropped image dataset by employing image cropping techniques. This was achieved by identifying and extracting the largest contour in each image, which typically corresponds to the area of the ulcer. Surrounding regions, considered extraneous and potentially distracting, were cropped out. This approach focuses the model on the most significant part of the images -the ulcers themselves- thereby enhancing its ability to learn relevant features efficiently. Ultimately, the effectiveness of this method was evaluated by testing the model on both the cropped and full images, aiming to ascertain improvements in model performance due to focused training on essential image areas.

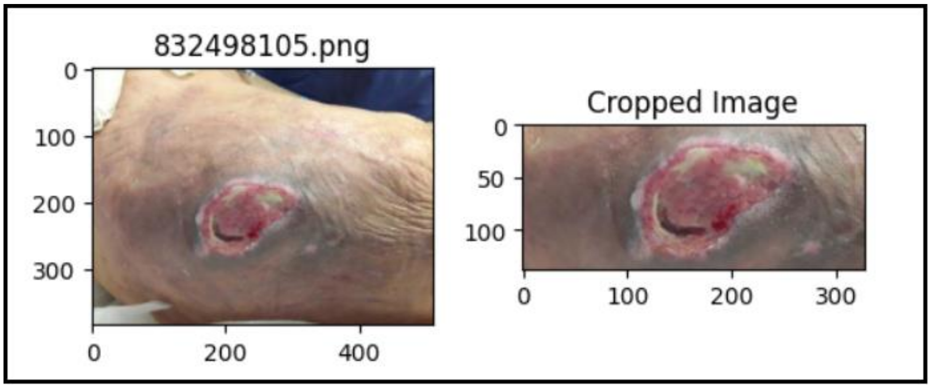

Figure-5: Original and Cropped Pressure Ulcer Images

2.2 - Creating Masked Image Datasets For Segmentation

For our segmentation model, we utilized two key techniques to create masked image datasets: Fill Mode and Bounding Box Mode. Fill Mode involves completely filling the identified ulcer regions to create distinct segmentation targets, training the model to recognize the entire affected area. Bounding Box Mode outlines the ulcer regions with surrounding it with lines, useful for initially locating ulcers and simplifying early training stages. These methods are benefical for preprocessing, with Fill Mode enhancing detailed recognition of ulcer shapes, and Bounding Box Mode helping focus the model on relevant areas during preliminary training phases. These image processes are done with extracting required data from JSON files.

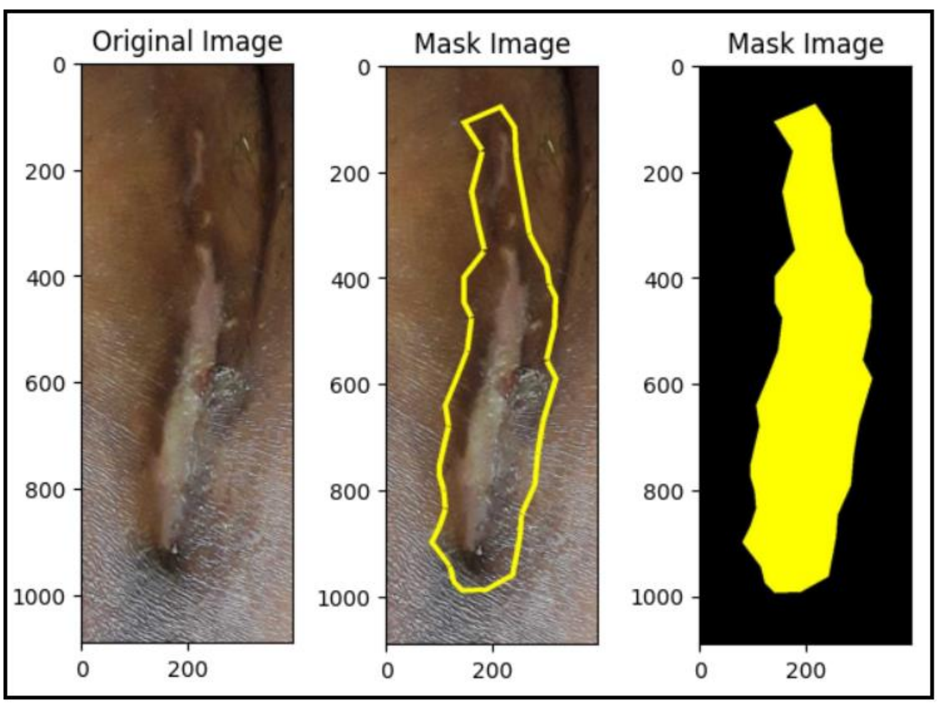

Figure-6: Original, Bound-Boxed Mask and Filled Mask Images

2.3 - Creating Augmented Image Datasets For Segmentation

To bolster the robustness and effectiveness of our segmentation model, we implemented a comprehensive image augmentation strategy using the Python library Albumentations. This strategy involved 26 key transformations that simulated various imaging conditions, significantly enhancing the diversity of our training data and thus improving the model's generalization capabilities. These augmentations, strategically chosen and applied, ensured the model was exposed to a broad spectrum of imaging scenarios during training. This not only facilitated more thorough training but also enabled the model to maintain accuracy across a range of challenging and unconventional image types, effectively preparing it for real-world applications.

Figure-7: Several Image Augmentation Techniques With Examples
(© https://github.com/albumentations-team/albumentations)

• **Geometric Transformations and Distortions:** 'PadIfNeeded', 'CenterCrop', and 'Crop' are used to adjust image dimensions and focus. To address orientation variability, 'HorizontalFlip', 'VerticalFlip', and 'RandomRotate90' were used for mirroring and rotating images. Additionally, 'Transpose' was applied to rearrange image axes, ensuring our model can interpret images from different orientations effectively. 'ElasticTransform', 'GridDistortion', and 'OpticalDistortion' are used to simulate realistic deformations typical in medical imaging scenarios. These transformations help the model adapt to anomalies and irregularities in image shapes and textures.

• **Image Composition and Color Adjustments:** The use of 'RandomSizedCrop' allowed for cropping variable regions, while 'Compose' and 'OneOf' combine multiple transformations to enhance the dataset's variability and robustness techniques such as 'CLAHE', 'RandomBrightnessContrast', 'RandomGamma', 'HueSaturationValue', and 'RGBShift' improve contrast and color variability. Additional tweaks with 'RandomBrightness' and 'RandomContrast' help fine-tune image visibility.

• **Augmentation Techniques for Realism and Robustness:** We introduced blurs ('MotionBlur', 'MedianBlur', 'GaussianBlur') and noise ('GaussNoise') to simulate common imaging artifacts. 'ChannelShuffle' was used to challenge the model with varied color channel information and 'CoarseDropout' was employed to randomly remove image regions, mimicking occlusions or missing data and preparing the model to handle incomplete information efficiently.

3- Training & Testing of Classification Model

After dataset pre-processing, we focused on the training and testing of classification models using the Ventura and Elazığ datasets. We experimented with eight different architectures, including VGG16, by modifying key parameters such as dropout rates, learning rates, and learning schedules to identify optimal settings. Our experiments resulted in a baseline table that compares the performance across the two datasets, providing a clear view of how each model configuration performed with the adjusted parameters.

Mimari	*Accuracy*	*Precision*	*Recall*	*F1-Score*	*Dropout value*
AlexNet	0.75	0.73	0.72	0.72	0.5
DenseNet201	0.76	**0.78**	0.69	0.71	0.7
EfficientNet_b0	0.77	0.75	0.76	0.75	0.3
EfficientNet_b7	0.72	0.69	0.72	0.69	0.5
Vgg16	**0.78**	0.77	**0.77**	**0.77**	0.5
MobileNetv2	**0.78**	0.76	**0.77**	0.76	0.7
ResNet152	**0.78**	0.77	0.73	0.74	0.5
GoogleNet	0.76	0.73	0.74	0.73	0.3

Figure-8: Best Classification Results For Ventura Dataset

Mimari	*Accuracy*	*Precision*	*Recall*	*F1-Score*	*Dropout value*
AlexNet	0.82	**0.83**	0.83	0.83	0.3
DenseNet201	0.82	**0.83**	0.83	**0.83**	0.7
EfficientNet_b0	**0.83**	**0.83**	**0.84**	**0.83**	0.7
EfficientNet_b7	0.80	0.80	0.81	0.80	0.5
Vgg16	**0.83**	**0.83**	0.83	**0.83**	0.3
MobileNetv2	0.82	**0.83**	0.83	**0.83**	0.7
ResNet152	0.81	0.82	0.81	0.81	0.7
GoogleNet	0.82	0.82	0.82	0.82	0.7

Figure-9: Best Classification Results For Elazığ Dataset

This analysis helped in understanding the most effective strategies for model optimization in medical imaging tasks. In our comprehensive study, the VGG16 model equipped with specific parameters of 0.5 dropout, a schedule interval of 100, gamma set at 0.1, and a learning rate of 0.001, emerged as the most effective across all tested architectures. This configuration yielded the best results during both training and testing phases. The other parameters remained constant across all models. Best model's results can be summarized as below:

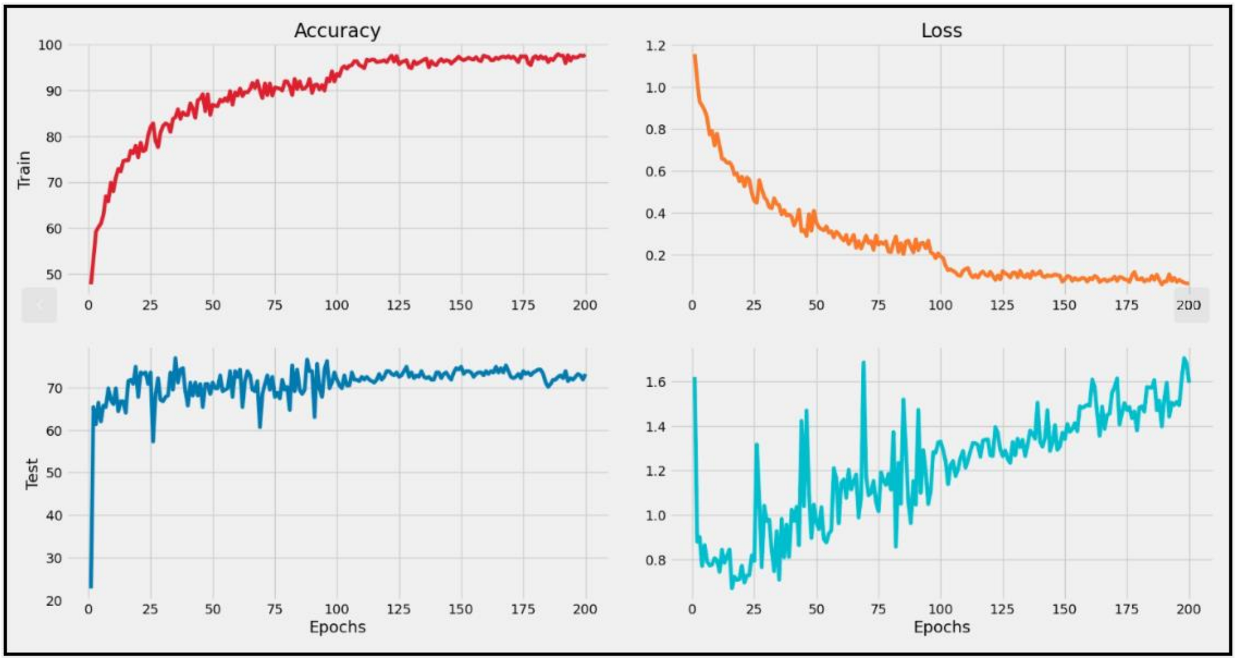

Figure-10: Train/Test Accuracies and Losses For Best VGG-16 Model

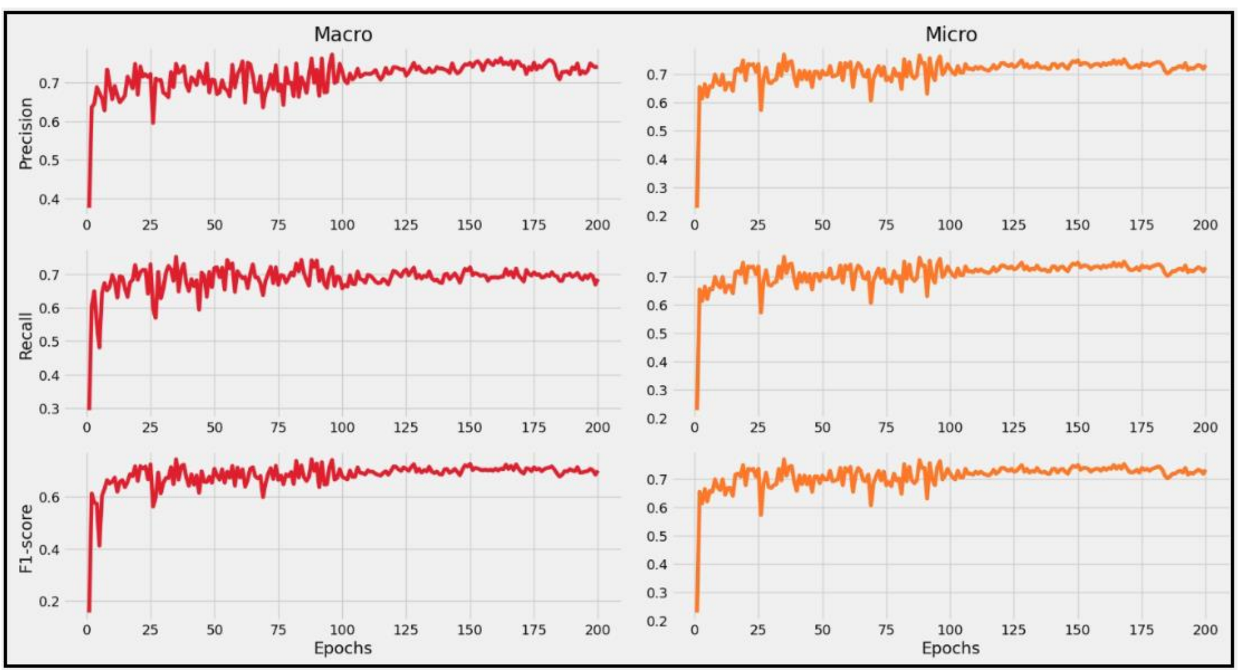

Figure-11: Test Micro/Macro Average - Precision/Recall/F1-Score For Best VGG-16 Model

4- Training & Testing of Segmentation Model

4.1 - Training & Testing Our Own Double U-Net Model From Scratch

A series of experimental combinations were conducted to ascertain the optimal parameters for enhancing the performance of the Double U-Net model. The parameters varied across different dimensions as summarized below:

• **Dataset Utilized:** Ventura dataset was used to ensure consistency and reliability of results.
• **Dataset Types:** Experiments were conducted on two distinct types of images within the dataset: bounding box annotated images and fill images, each presenting unique challenges.
• **Data Partitioning:** The data was split into training, validation, and testing sets in two different configurations to evaluate model robustness: an 80/10/10 split and a 70/30 split respectively.
• **Learning Rates:** Three different learning rates were tested (1E-02, 1E-03, and 0.0005) to identify the most effective rate for convergence without sacrificing the speed of training.
• **Batch Sizes:** To further refine our model, three different batch sizes (4, 8, 16) were examined, balancing between computational demand and memory capabilities.

No	Dataset Type	Data Split	Learn Rate	Batch Size	Train D.Coef	Train IOU	Validation IOU	Test D.Coef	Test IOU	Test Recall
1	Boundboxing	80-10-10	0,01	4	0.6521	0.4846	0.4594	0.5313	0.3619	0.4969
2	Boundboxing	80-10-10	0,0005	4	0.6523	0.4848	0.4591	0.6301	0.4603	0.9156
3	Boundboxing	80-10-10	0,001	4	0.6524	0.4849	0.4601	0.6341	0.4645	0.9162
4	Boundboxing	80-10-10	0,001	8	0.6520	0.4850	0.4567	0.6151	0.4444	0.7924
5	Boundboxing	80-10-10	0,001	16	0.6525	0.4845	0.4569	0.6125	0.4424	0.7073
6	Boundboxing	70-0-30	0,001	4	0.6531	0.4853	0.4605	0.6348	0.4648	0.9168
7	Filled	70-0-30	0,001	4	0.7929	0.6601	0.4803	0.6478	0.4820	0.9592

Figure-12: Experiment Results For Segmentation With Double U-Net

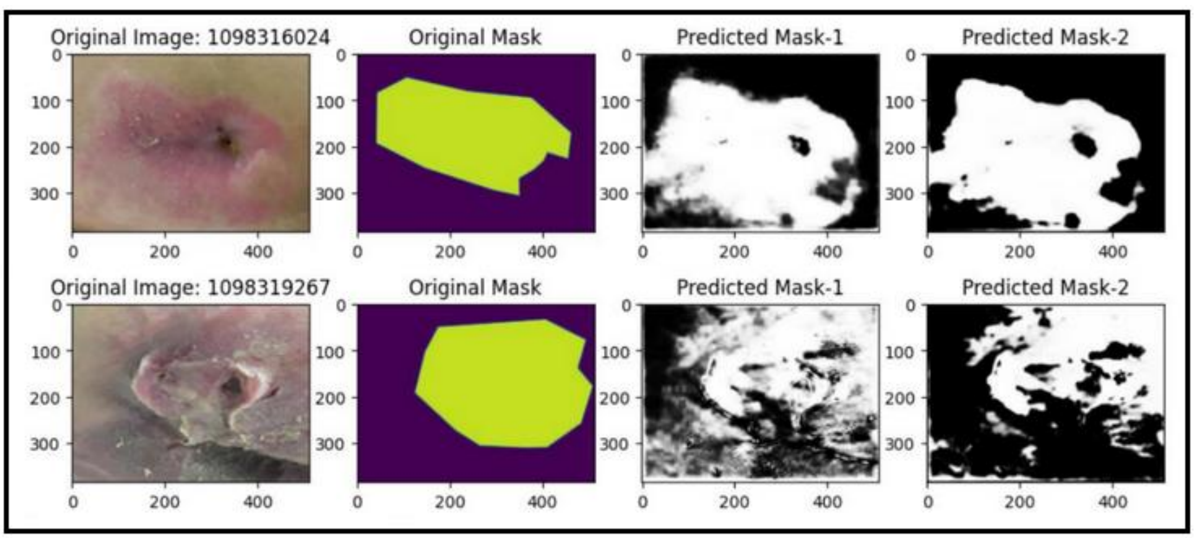

Figure-13: Some Test Results For Best Segmentation Model From Scratch

Based on the Dice_Coef and IOU metrics applied to the test dataset, the optimal configuration includes fill-type images, a learning rate of 1e-3, a batch size of 4, and a train/test split ratio of 70%-30%. It's important to highlight that experiments were repeated to ensure consistency.

4.2 - Training & Testing Our Own Double U-Net Model With Using Pre-trained Model

Despite the visually promising results, initial outputs were marred by noise (in part 4.1), suggesting that the small size of our dataset could be limiting performance. To address this, we adopted to use a pre-trained Double U-Net model, specifically trained with medical images. Since we could not find a satisfactory model in our research on the Internet, we sourced a Double U-Net segmentation model from the 'Skin Cancer Project Group', which had been trained on thousands of images. This strategic choice allowed us to leverage the refined parameters of an established model rather than starting from scratch. By initiating our training with these pre-trained parameters, we significantly improved the model's performance, achieving the desired outcomes effectively. This approach not only optimized our process but also underscored the importance of robust training sets in medical image segmentation.

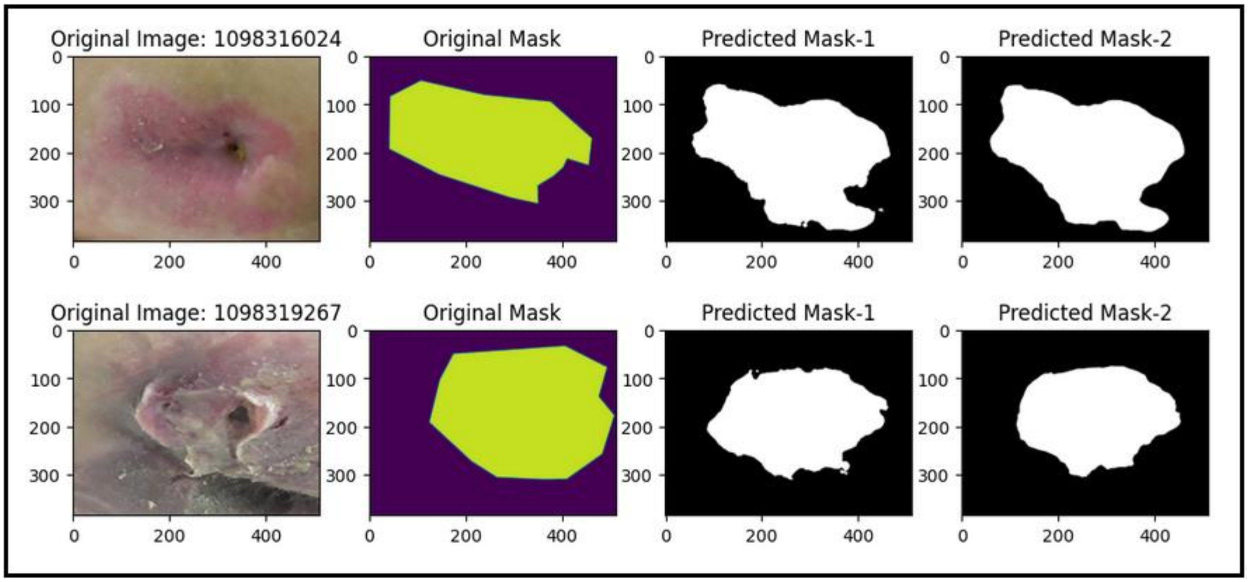

Figure-14: Some Test Results For Best Segmentation Model Based On Pretrained Model
(Note that, same images are used with in the figure-8 in order to show improvement.)

5- Boosting Classification Performance With ROI Model (Attention Branch Master)

5.1 - Extracting Contours and Storing Critical Coordinates

Identifying the largest contour related to the wound and noting the corner coordinates of the contour is essential. This process is typically necessary for optimizing the Region of Interest (ROI) for the Attention Branch Network (ABM) model. The initial step involves applying image segmentation techniques to highlight the area of the wound. Subsequently, contour detection algorithms are used to identify and extract the boundaries of the wound. The largest contour is then selected based on area size, which effectively captures the most significant wound-related shape in the image. The final step involves recording the corner coordinates of this contour. These coordinates are crucial as they define the precise boundaries of the ROI, which can then be fed into the ABM model to enhance its focus and accuracy in analyzing the wound within the pressure ulcer images. This methodology helps in refining the model's performance by providing it with targeted, high-relevance data. (This was used in approach-2 in part 5.2)

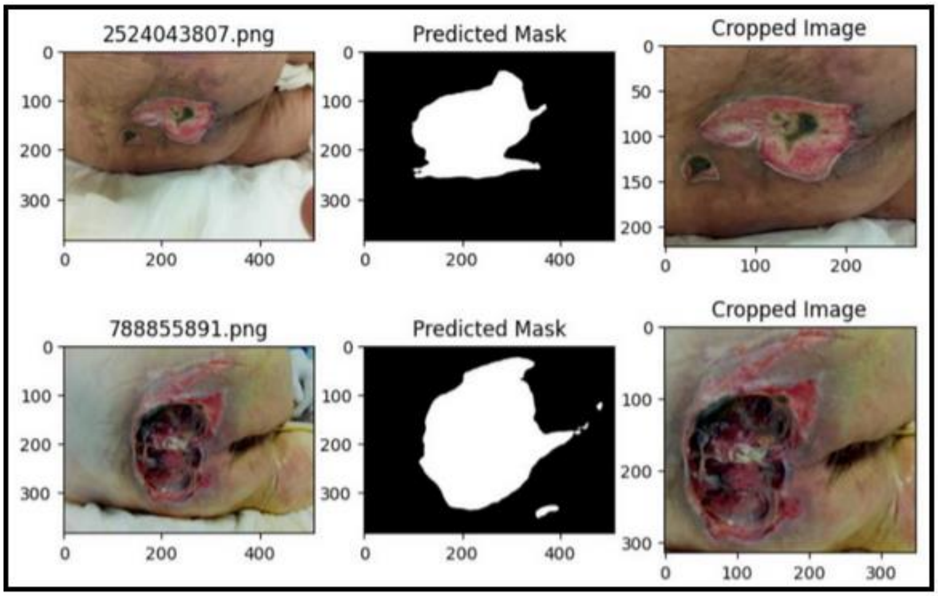

Figure-15: Producing Smaller Images With Cropping the Diseased Parts of Large Images

5.2 - Feeding Classification Algorithm With ROI Model

• **Approach-1:** Both ROI (cropped) and full-sized images are fed into the same model and used together, with specific weights assigned to each type of data to balance their contributions to the outcome.

• **Approach-2:** ROI's undergo a specialized processing procedure, and their contributions to the results are evaluated through an entirely separate classification pathway.

```
class MyModel(nn.Module):
    def __init__(self):

        (...)

        # RoI pooling layer
        self.roipool = ops.RoIPool(output_size=(7, 7), spatial_scale=0.01)

        (...)
        num_classes=4

        self.classifier_roipool = nn.Sequential(
            nn.Dropout(inplace=False),
            nn.Linear(147, num_classes)
        )

    def forward(self, x, y, rois):
        x1 = self.features(x)
        x1 = self.avgpool(x1)
        x1 = torch.flatten(x1, 1)
        x1 = self.classifier(x1)

        y = self.roipool(y, rois)
        y = torch.flatten(y, 1)
        print(y)
        y = self.classifier_roipool(y)

        return x1,y
```

Figure-16: Some Part Of Our Test Code Related To Attention Branch Master ROI Approach

Results & Discussion (/ 30 Points)

Explain your results in detail including system/model train/validation/optimization analysis, performance evaluation and comparison with the state-of-the-art (if relevant), ablation study (if relevant), a use-case analysis or the demo of the product (if relevant), and additional points related to your project. Also include the discussion of each piece of result (i.e., what would be the reason behind obtaining this outcome, what is the meaning of this result, etc.). Include figures and tables to summarize quantitative results. Use sub-headings for each topic. This section should be between 1000-2000 words (add pages if necessary).

Uncompleted Aspects of Project

1) Image Segmentation Experiments With U-Net++ Architecture
• **Whose Responsibility:** Asım
• **Detailed Description:** Despite Asım have found the code for image segmentation for U-Net++ architecture and have tested it with various medical images, he has failed to implement it for our 'Ventura' and 'Elazığ' datasets. Therefore, the experiments summarized for the Double U-Net architecture in Figure-12 have not been conducted for U-Net++. Consequently, even though we have used the Double U-Net architecture for segmentation tasks, we do not have sufficient data to determine which architecture provides better results.

2) Training Attention Branch Master Model With Roi-Pooling Layer
• **Whose Responsibility:** Asım
• **Detailed Description:** We encountered two significant issues with the attention branch master mechanism that utilizes ROI (which was discussed in section 5.2 on page 13 in this report). Firstly, the test accuracy rates for the model trained with attention were below the standard model's accuracy rate (78%), which was unexpected as the rate should have significantly improved under normal conditions. The second issue involved the non-saving of the weights for the ROI pooling layer in the trained attention model. Despite all layers being visible in our test code, the weights for the ROI pooling layer were inexplicably missing. As a result, we had to opt for the first approach mentioned in Section 5.2 of page 13 in our mobile application, instead of the second approach, despite having all training codes ready and test codes written. Therefore, we must acknowledge that the results in the 'attention' section of our mobile app are not sufficiently consistent.

Note: Except from these, everything related to the project has been successfully completed.

Model Optimization Analysis and Evaluation Principles

1) Classification Problem
In the context of medical imaging for pressure ulcers, applying machine learning models requires meticulous parameter tuning to optimize performance. In our project, using eight different models was crucial for addressing the nuances of this application. Parameter tuning, which involves adjusting the algorithm parameters to achieve the best performance from each model, was instrumental. Through rigorous testing and adjustments, these models collectively reached an accuracy of 78%. However, there was potential to elevate this further; projections and further researches on 3rd party resources indicated that with an optimal functioning of the attention branch network, accuracy could surpass 90%. Understanding the effectiveness of models in classifying medical images of pressure ulcers hinges on key metrics:

• **Accuracy:** This metric reflects the overall correctness of the model across all classes, measuring the proportion of true results (both true positives and true negatives) among the total number of cases examined. (Our best model has %78 test accuracy.)

• **Precision:** Precision is particularly crucial in medical imaging as it quantifies the rate of true positive predictions relative to all positive predictions made. High precision indicates a low rate of false positives, essential in clinical settings to avoid misdiagnoses. (Our best model has %77 test precision.)

• **Recall (Sensitivity):** This metric measures the model's ability to correctly identify all relevant instances of ulcers. A high recall rate is vital to ensure that all potential ulcers are detected, minimizing the risk of overlooking a critical condition. (Our best model has %77 test recall.)

• **F1 Score:** Combining precision and recall, the F1 score is the harmonic mean of these metrics. It is particularly useful when the class distribution is uneven, as it maintains a balance between precision and recall, providing a more holistic view of model performance. (Our best model has %77 test F1 score.)

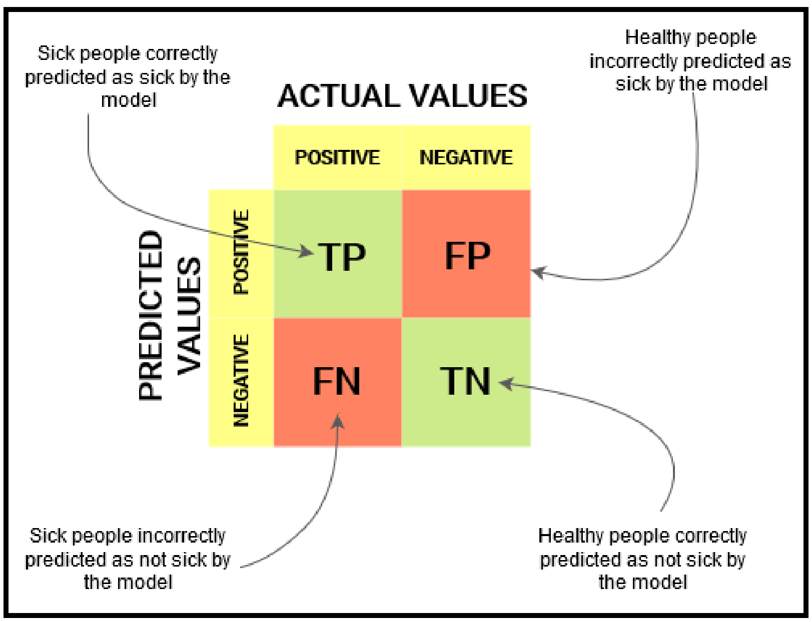

Figure-17: Precision Matrix For Machine Learning In Healthcare Area
(© Analytics Vidhya)

2) Segmentation Problem

Segmentation tasks in medical imaging, particularly for large images of pressure ulcers, require precise methods to accurately delineate affected areas. To ensure that our models effectively handle large images, which are common in clinical settings, we employed a strategy of training on cropped images focusing on the most crucial points. This approach helps the model learn from the most significant features of ulcers, improving its ability to generalize when applied to unseen, full-size images. To evaluate the effectiveness of our segmentation models, we incorporated several key metrics:

• **Dice Similarity Coefficient (DSC):** This metric gauges the similarity between the predicted segmentation and the ground truth, providing a measure of the model's precision in identifying ulcer regions. It is particularly valuable in medical imaging for ensuring that the segmented area closely matches the actual ulcerated tissue. (Our best model has 0.6478 test dice coefficient.)

• **Intersection over Union (IoU):** Also known as the Jaccard index, this metric assesses the overlap between the predicted segmentation area and the true area. A higher IoU indicates a more accurate model in terms of spatial overlap, which is critical for clinical accuracy. (Our best model has 0.4820 IOU value.)

• **Recall:** In the context of segmentation, recall reflects the model's ability to identify all relevant areas of interest within the image. High recall is essential for medical applications to ensure no affected area is missed during the analysis. (Our best model has 0.9592 test recall value.)

While metrics like Dice Coefficient, IoU, and recall provide quantitative measures of model performance, they can sometimes be misleading if considered in isolation. For instance, a high Dice score might still correspond with some clinically irrelevant areas being marked, or essential small lesions being overlooked. To counteract potential discrepancies between these metrics and actual model efficacy, we also conducted a thorough visual inspection of the segmented images. This dual approach of relying on both quantitative metrics and qualitative assessments through visual checks ensures that our model is robust and performs well even on large, complex images. A balance is essential to ensure both accuracy and robustness, and our optimized model has this kind of balance as you can see from image below:

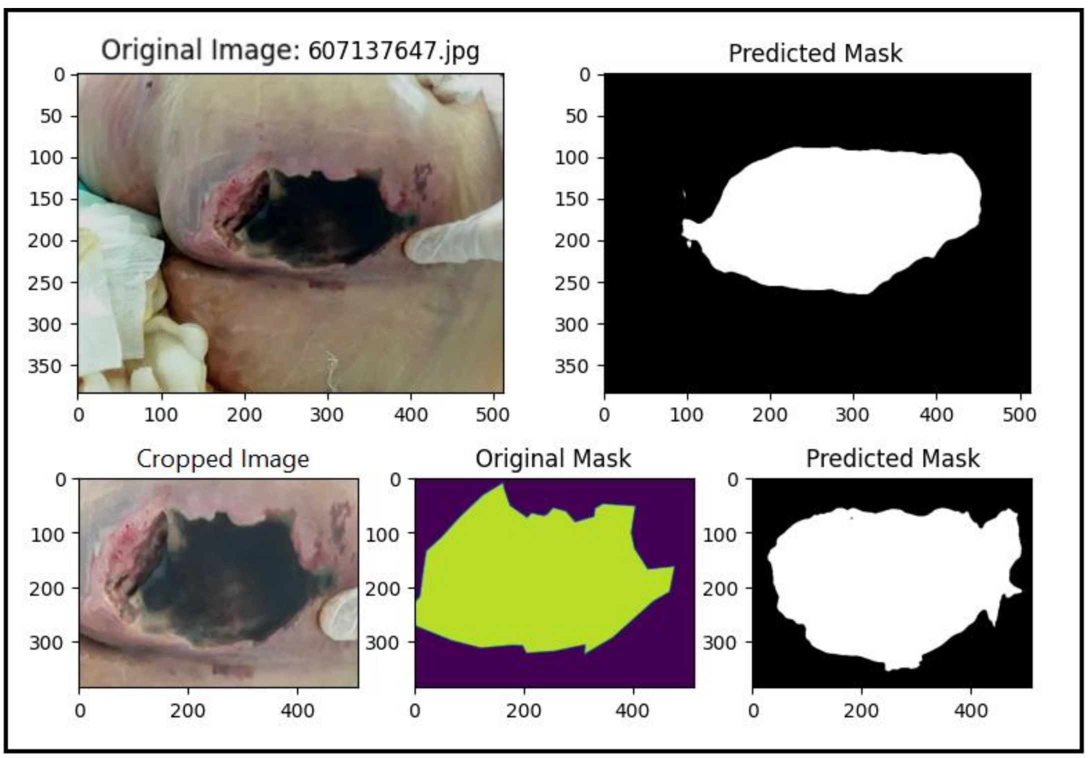

Figure-18: Segmentation Testing For Original and Cropped Images (Example-1)

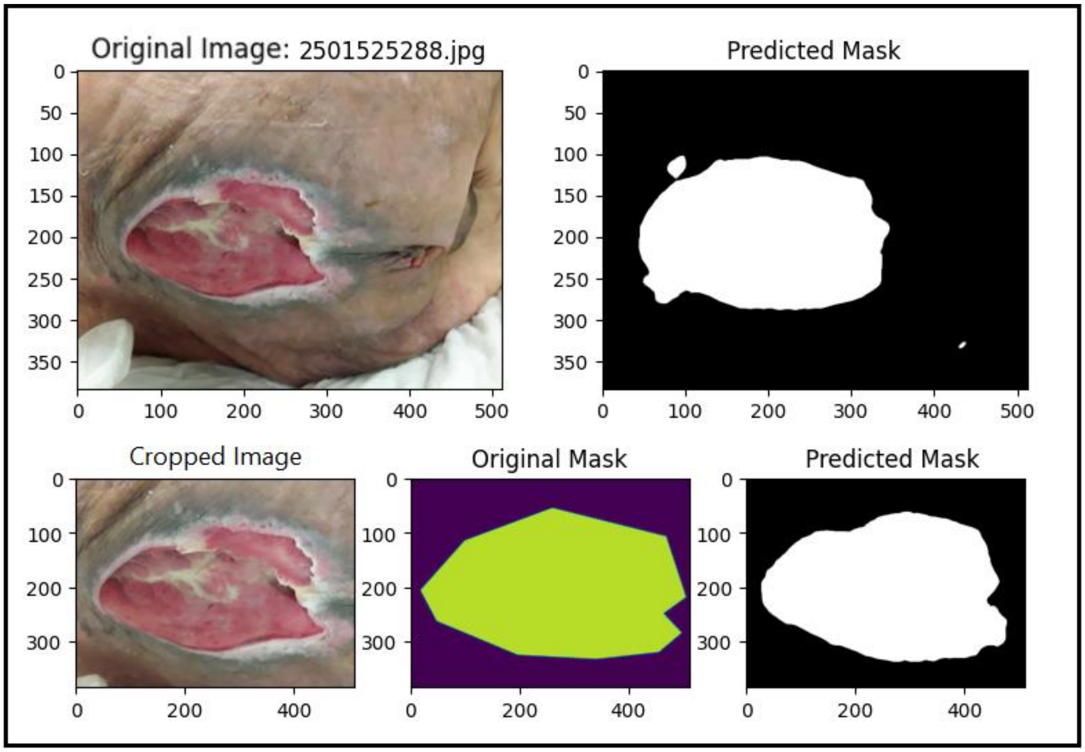

Figure-19: Segmentation Testing For Original and Cropped Images (Example-2)

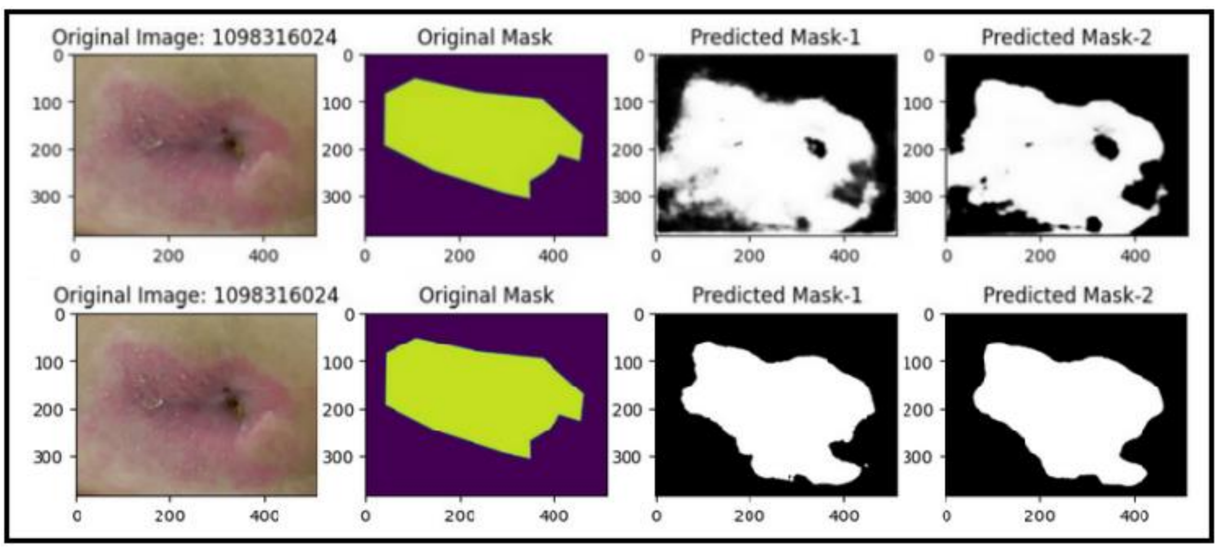

Figure-20: Segmentation Results Before and After Model Optimization For Same Image

Note: In addition to all this, several algorithms have been designed and coded for the classification section. These include an algorithm that uses seven different models to return the most sensible result for a single input image, another that processes a standard large-sized image by cropping the relevant area in the background and then feeds both the original and cropped images to the same model to determine the most sensible result. Additionally, a test code has been developed for the Attention Branch Master. Although only one of them was used in our mobile application.

Mobile Application Demo

Our newly developed mobile application offers a cutting-edge solution for image analysis, utilizing advanced machine learning algorithms to deliver precise and reliable results. Designed to be user-friendly, it supports both taking new photo and selecting photo from gallery, ensuring versatility across various use cases. This app integrates seamlessly with multiple models to analyze a single image, making it a valuable tool for both professional and personal use.

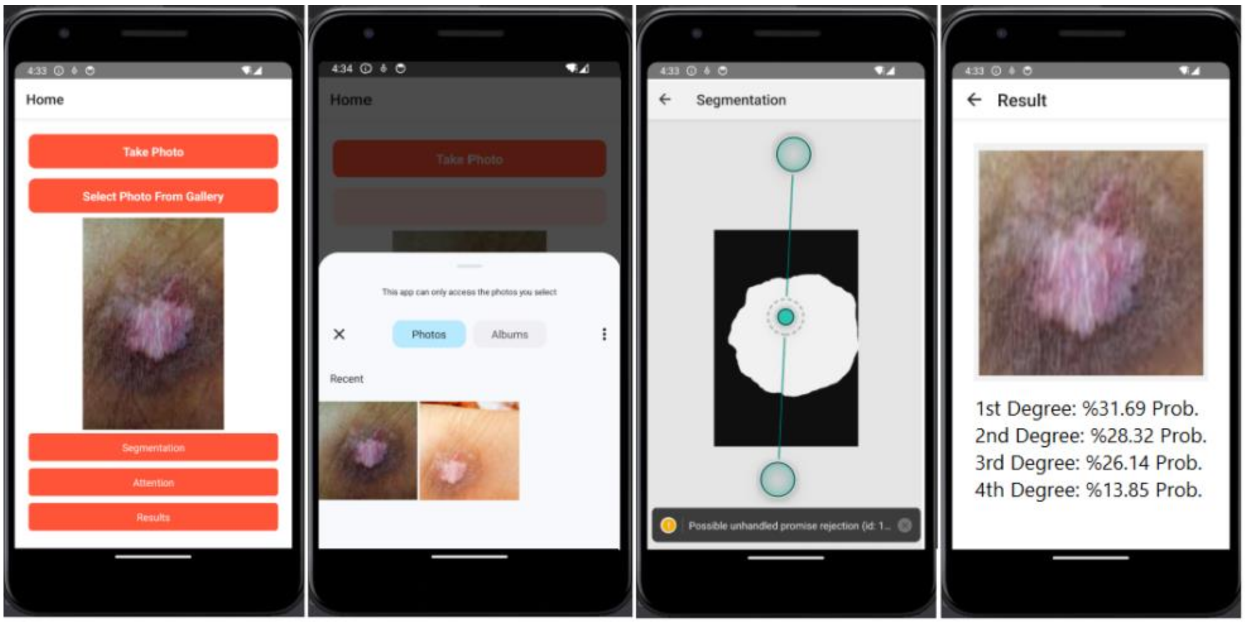

Figure-21: Our Application's Main, Upload Image, Segmentation and Classification Screens

You can find a detailed information about our mobile applications usage, in demo presentation video with that link: ....

Discussion of Results

In conclusion, the development of an automated system for medical image segmentation and classification with pioneering approach holds significant promise for advancing patient care and reducing healthcare costs. By building upon the advancements in deep learning and addressing the limitations of current approaches, this project aims to deliver a robust and clinically applicable solution. The successful implementation of this system will not only facilitate the early detection and precise classification of pressure ulcers but also support healthcare providers in making informed decisions, ultimately enhancing patient outcomes and quality of life.

Traceability Matrix

Team Member	Tasks That Member Handled During The 2 Semesters
Asım Ateş (21827076)	**Coding Tasks:** • Finding an effective u-net++ segmentation code **Other Tasks:** • Testing the ROI model to find the parameters that give the best accuracy • Preparation of a poster for final presentation

Yusuf Efe Kalaycı (21827517)	**Coding Tasks:** • Making changes and adjustments to the classification code • Combining all prepared codes and models into a mobile application (He is the only member of the team who knows mobile programming.) • Adding ‘take photo’ and ‘select photo from gallery’ feature to the app **Other Tasks:** • Finding the ideal parameters and thus creating a baseline table by making numerous experiments with all classification models (8 different models and for 2 different datasets) • Meanwhile, saving the results, graphs of metrics and the best models for classification • Helping Asım during the ROI experiments • Preparation of 2 minute-length video for final presentation
Mert Tazeoğlu (21946606)	**Coding Tasks:** **a- Image Processing Tasks** • Find an effective image data pre-processing module for segmentation (A module that performs 26 different data augmentation operations was found) • Creating 2 different types of masked image datasets for segmentation by extracting JSON files and with using original image datasets (fill mode and bound-boxing mode) • Writing codes that obtaining cropped images from full-size images in order to support the ROI function (includes processes such as detecting contours in the images, selecting the correct contour, cropping contour and storing the cropped coordinates in a .txt file for ROI operations) **b- Image Segmentation Tasks** • Finding an effective double u-net segmentation code • Finding the ideal parameters and thus creating a baseline table by making numerous experiments with double u-net model • Finding an pre-trained double u-net medical image segmentation model • Obtaining the best segmentation model by training the pre-trained model with the tuned ideal parameters • Repeating these operations that done for cropped images, for non-cropped images **c- Other Programming Tasks** • Writing several test codes for 7 different classification models 1- Development and coding of the algorithm that makes the most accurate prediction using 7 different models for an image 2- Development and coding of the algorithm that works with standard and ROI images and makes the accurate prediction with using only best model 3- Writing a test code for Attention Branch Master, which works with ROI **Other Tasks:** • Preparation of ‘End of Term Development Report’ • Preparation of ‘End of Project Report’ • Ensuring communication within the team • Preparation of a website for final presentation

The Impact and Future Directions (/ 15 Points)

Explain the potential (or current if exist) impacts of your outcome in terms of how the methods and results will be used in real life, how it will change an existing process, or where it will be published, etc. Also, explain what would be the next step if the project is continued in the future, what kind of qualitative and/or quantitative updates can be made, shorty, where this project can go from here? This section should be between 250-500 words.

The development and implementation of medical image segmentation and classification for pressure ulcers represent a significant advancement in clinical practices and healthcare technologies. This project, focusing on the automation of pressure ulcer detection and classification, has the potential to enhance diagnostic accuracy, expedite treatment decisions, and improve patient outcomes.

The methodology developed through this project can be integrated into hospital information systems to assist healthcare providers in early and accurate detection of pressure ulcers. By using advanced image segmentation techniques, our system can distinguish between different stages of ulcer development, allowing for timely and tailored treatment interventions. This not only improves the quality of patient care but also reduces the burden on healthcare professionals by automating routine tasks.Furthermore, the classification component of this system aids in the continuous monitoring of wound healing processes, enabling adjustments in treatment plans based on objective assessments. This real-time tracking could be particularly beneficial in long-term care facilities and for patients receiving home care, ensuring that they receive the best possible care based on up-to-date information.

The findings from this project are intended for publication in the end of year presentation at school and also Github. Additionally, presenting the results at medical conferences will facilitate the exchange of ideas with other researchers and clinicians, potentially leading to collaborative improvements and validation of the technology. Looking to the future, the project can expand in several ways. If this project is continued, several advancements are conceivable:

• **Enhanced Algorithms:** Future work could focus on refining the algorithms used for segmentation and classification. Incorporating deep learning models trained on a larger and more diverse dataset could improve the system's accuracy and robustness, particularly in handling cases with rare ulcer types or those in early stages.

• **Integration with Wearable Technologies:** Combining this technology with wearable sensors that monitor skin integrity could provide continuous, non-invasive assessments of at-risk patients. This integration would facilitate early intervention and potentially prevent the progression of ulcers. Such technology won't only supports clinicians in making more informed decisions but also enhance the quality of care provided to patients.

• **Expanded Scope:** While currently focused on pressure ulcers, the methodologies could be adapted for other types of skin lesions or conditions. Research into similar applications could extend the utility of the technology, making it a versatile tool in dermatological diagnostics.

• **Patient Engagement:** Developing an interface for patients to interact with the system could empower individuals by allowing them to monitor their own conditions. This would not only enhance patient engagement and education but also promote proactive health management.

• **Quantitative and Qualitative Research:** Ongoing research to gather both quantitative data (e.g., system performance metrics) and qualitative feedback (e.g., user satisfaction from patients and healthcare providers) will be crucial. This will help in continuously refining the system to better meet the needs of its users.