2022

Artificial Intelligence-Powered Chronic Wound Management System: Towards Human Digital Twins

salih sarp
Virginia Commonwealth University

Artificial Intelligence-Powered Chronic Wound Management System: Towards Human Digital Twins

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

by

Salih Sarp

M.Sc. in Electrical Engineering, George Washington University, U.S.A.

Dissertation directed by

Yanxiao Zhao, Ph.D (Advisor)
Murat Kuzlu, Ph.D. (Co-Advisor)

Virginia Commonwealth University
Richmond, Virginia
August, 2022

Dedication

To my family.

Acknowledgments

I would like to express my deepest appreciation to my advisors, Dr. Yanxiao Zhao and Dr. Murat Kuzlu, who continuously provide valuable insights and motivate me to do my best. Dr. Zhao has been constantly supportive during my Ph.D. study at VCU by providing research insights and resources to implement many of the works in this study. Her kind guidance and support have provided me to deeply understand the concepts and shape my studies. Beyond her duties as an advisor, Dr. Zhao has been a continuous source of inspiration, as well as an excellent model of a dedicated mentor. Dr. Kuzlu has guided me with valuable expertise and effort, which helped me to develop skills to be an eligible Ph.D. He has offered me the opportunity to reach my full potential. I feel lucky to work with both of my advisors: Dr. Zhao and Dr. Kuzlu. Without their helpful support, effort, and time, this dissertation and many of my achievements during this time would not be possible.

Besides my advisors, Dr. Ozgur Guler has been a true example of being an academician and an entrepreneur. He has been resourceful in every aspect, and he never hesitates to share them to improve my skills. He has fostered my interest in chronic wounds by sharing his research and required materials with me. I have tremendous admiration for the way he thoughtfully approaches every discussion we have.

I would also like to extend my gratitude to the other members of my dissertation committee: Dr. Ruixin Niu, Dr. Sherif Abdelwahed, and Dr. Changqing Luo, for their valuable suggestions and practical advice regarding this dissertation. I feel fortunate to have each of them on my committee, as their insightful inputs have helped me to reach the current form of this dissertation.

I am also grateful to Dr. Can Korman and Dr. David Nagel from GWU for advising me during my M.Sc. I have benefited from their vision, their generous help, and valuable discussions on my research.

Abstract of Dissertation

Artificial Intelligence-Powered Chronic Wound Management System: Towards
Human Digital Twins

Artificial Intelligence (AI) has witnessed increased application and widespread
adoption over the past decade. AI applications to medical images have the potential
to assist caregivers in deciding on a proper chronic wound treatment plan by helping
them to understand wound and tissue classification and border segmentation, as well
as visual image synthesis.

This dissertation explores chronic wound management using AI methods, such as
Generative Adversarial Networks (GAN) and Explainable AI (XAI) techniques. The
wound images are collected, grouped, and processed. One primary objective of this
research is to develop a series of AI models, not only to present the potential of AI in
wound management but also to develop the building blocks of human digital twins.

First of all, motivations, contributions, and the dissertation outline are summa-
rized to introduce the aim and scope of the dissertation. The first contribution of
this study is to build a chronic wound classification and its explanation utilizing XAI.
This model also benefits from a transfer learning methodology to improve perfor-
mance. Then a novel model is developed that achieves wound border segmentation
and tissue classification tasks simultaneously. A Deep Learning (DL) architecture,
i.e., the GAN, is proposed to realize these tasks. Another novel model is developed
for creating lifelike wounds. The output of the previously proposed model is used
as an input for this model, which generates new chronic wound images. Any tissue
distribution could be converted to lifelike wounds, preserving the shape of the original
wound.

The aforementioned research is extended to build a digital twin for chronic wound
management. Chronic wounds, enabling technologies for wound care digital twins,

are examined, and a general framework for chronic wound management using the digital twin concept is investigated. The last contribution of this dissertation includes a chronic wound healing prediction model using DL techniques. It utilizes the previously developed AI models to build a chronic wound management framework using the digital twin concept. Lastly, the overall conclusions are drawn. Future challenges and further developments in chronic wound management are discussed by utilizing emerging technologies.

Table of Contents

xiv

List of Tables

1   Introduction

After alternating between periods of great passion and setback [1], Artificial Intelligence (AI) has found its place as a critical component of growth in a variety of applications [2]. These applications range from diagnostic decision assistants in healthcare to safety-critical systems in autonomous vehicles and to long-term financial investment planning. These applications benefit from AI breakthroughs to solve complex problems [3].

Recent advancements in AI and its applications in computer vision have paved the way for systems with more human-like performance. AI applications can now detect malicious cancer cells from medical images and can diagnose diseases exceeding radiologists' performance. One of the promising visual applications of AI is wound management in healthcare.

Wounds or injuries form as a result of disruptions in the normal architecture of any body tissue, especially on the skin [4]. Wounds could be classified into two groups: (1) acute wounds, which follow an orderly healing process, and (2) chronic wounds that do not progress in an orderly manner. Wound management could be defined as organizing a comprehensive care plan for a wound by reviewing all the factors that affect its development and healing. One of the recent studies indicates that mortality rates of chronic wounds are also on par with cancer for some patients [5]. That is why proper and continuous care is critical for hard-to-heal or non-healing wounds.

The Use of AI will enhance chronic wound management in many ways. AI-powered wound management could be beneficial for identifying the chronic wound, classifying both the wound and the tissue type, segmenting the wound and its tissues, creating chronic wound images using its segmentation, and predicting the healing process of the wound [6, 7, 8, 9, 10, 11]. From wound classification to the wound healing process, AI is perceived as an invaluable asset in the improvement of wound management. With the use of emerging technologies such as AI and cloud computing, wound management

could be tracked dynamically by extending research to the digital twin concept.

1.1   Motivation

Chronic wound healing is a highly complex and dedicated process that requires constant and planned care [12]. Whereas an acute wound follows sequential healing progress with anatomic and functional restoration in a timely manner, chronic wounds lack routine restoration due to various physiological impairments [13]. Without proper attention and care, chronic wound treatment becomes a heavy burden that costs over US$25 billion yearly in the US only [14]. Furthermore, a single diabetic ulcer treatment case could reach nearly US$50,000 [15]. Even spending on wound care products outstretches to the US$30 billion per year [16]. Wound care gets the highest share among any other skin diseases [17].

Chronic wounds are snowballing with an increasing rate among the elderly and people with diabetes, foot ulcers, and pressure ulcers [13]. Therefore, wound healing draws enormous attention where actual healing progress is tracked with a simple visual inspection [18, 19]. Continuous development in healthcare methods opens new opportunities with the broad acceptance of AI.

The importance of developing a healthcare technique using AI has long been investigated, both by academia and the industry [20, 21]. However, numerous attempts are still being plagued by privacy concerns and cost overruns. Specifically, patient privacy prevents the collection of a large dataset to train AI models. Without an extensive collection of data in wound care (e.g., thousands of structured wound images), an AI model cannot be trained well enough to be used in healthcare due to the high risks associated with the health of the patient. A literature review shows that past studies have primarily been focused on the classification of a particular type of chronic wound, such as diabetic or pressure injury wounds [22]. Limited progress has been made in identifying, classifying, and segmenting various wounds and their

tissues in a comprehensive manner. Prediction of chronic wound healing has not been studied extensively. Overall, the following challenges exist for precise and complete wound care.

First of all, wound management systems are still lagging technologically, and most caregivers depend only on imprecise optical assessment [23], which can cause many complications like infection risks and inaccurate measurements which, in turn, cause wrong assessment of the wound [24]. In addition to this complexity, nearly 7 million people in the US have received needed care for chronic wounds [14, 25]. A human-centered approach and its high cost complicate the healing of the wound in time. However, this labor-centric financial burden could be reduced with advanced AI-assisted computer vision models.

Secondly, every chronic wound is unique and has its own characteristics and sets of properties that affect its healing process. As a result, the healing process of a chronic wound is subject to numerous constraints that limit the track of the healing progress, which invariably poses a significant negative impact on chronic wound management. Moreover, without proper practice and experience, the care for the wound takes a long time, which further increases both the financial burden and patient discomfort. Therefore, building an AI-assisted model for chronic wound care is a fundamental and critical task in AI research in healthcare.

Thirdly, it is challenging to predict the healing progress without the involvement of an experienced practitioner [26, 27]. The determination of the right chronic wound treatment plan plays a crucial role in a systematic healing process. A number of parameters should be considered for an accurate prediction, whereas today's wound healing prediction is only done by optical assessments. Forecasting using a multi-parameter has been a major challenge to the AI research community. However, AI-based models such as computer vision could enhance continuous and accurate wound healing monitoring at a lower cost.

In addition, the education and training of clinicians are still done by manual optical assessments. The labor-intensive nature of the training takes a long time, as well. Since the amount of available data is so scarce due to both privacy concerns and collection difficulty, additional generative models such as GANs [28], and autoencoders [29] could be used to tackle this problem.

Moreover, a data-driven model could be built using emerging technologies. The digital twin concept has a promising potential to overcome many of the challenges in chronic wound management. It will allow continuous track of chronic wound development and foresee the healing status of the wound in near real-time. Health concerns could be detected beforehand so that required medical intervention could be provided promptly.

1.2 Contributions

The goal of this research is to develop a formalized chronic wound management system using AI. As defined previously, wound management consists of a series of challenging processes, including identifying the chronic wound, classifying and segmenting the wound and tissue type, creating chronic wound images using this segmentation, and predicting the healing process of the wound as well. The objective of this study is to provide a comprehensive review of the literature and to develop state-of-the-art techniques for chronic wound management. This research is also extended to digital twin use in wound management. In order to address the above-mentioned challenges in chronic wound management using AI, in this dissertation, we conduct studies on designing new network architectures utilizing state-of-the-art ML techniques such as GAN, transfer learning, and Explainable AI (XAI) methods by hyperparameter optimization and crafting learning objectives and metrics.

The work presented in this study has been published in three journals and has been submitted to a journal and a book chapter. The first journal paper includes

a study on chronic wound classification using XAI techniques [30]. Another journal paper proposes a hybrid approach for the identification of wound borders and the classification of wound tissues [31]. The research about the creation of wound images from tissue combinations appeared in the last published journal paper [32]. Research about the digital twin in chronic wound management and its enabling technologies is accepted as a book chapter and will be published in 2022 [33]. Proposed digital twin framework and wound healing prediction research have also been submitted to a journal, currently under review [34].

Contributions made in this study are summarized as follows:

(i) A new chronic wound classifier is built using transfer learning methodology and XAI techniques to support caregivers. Previous works in wound classification are restricted to a single wound type and do not utilize any explanation. The proposed model classifies chronic wounds through transfer learning and fully connected layers. Classified chronic wound images serve as input to the XAI model for an explanation. Interpretable results can help shed new perspectives for clinicians during the diagnostic phase. The proposed method successfully provides chronic wound classification and its associated explanation to extract additional knowledge that can also be interpreted by non-data-science experts such as medical scientists and physicians. This hybrid approach is shown to aid in interpreting and understanding the AI decision-making processes.

(ii) A novel GAN model of medical image synthesis is built to assist caregivers in deciding on a proper chronic wound treatment plan by helping them visually understand the border segmentation and the wound tissue classification. This study proposes a hybrid wound border segmentation and tissue classification method utilizing conditional GAN, which can mimic actual data without expert knowledge. We trained the network on chronic wound datasets of different sizes. The performance of the GAN algorithm is evaluated through the Mean

Squared Error (MSE), Dice coefficient metrics, and visual inspection of generated images. This study also analyzes the optimum number of training images and the number of epochs using GAN for wound border segmentation and tissue classification. The results show that the proposed GAN model performs efficiently for wound border segmentation and tissue classification tasks with 2000 images at 200 epochs.

(iii) A medical image synthesis model is developed and presented, which shows great potential in assisting clinician training. This work proposes a synthetic wound image generation model based on GAN architecture to increase the quality of clinical training. The proposed model is trained on chronic wound datasets of various sizes taken from natural hospital environments. Hyperparameters such as epoch count and dataset size for training tasks are also studied to find the optimum training conditions. The performance of the developed model is evaluated through the MSE metric to determine the similarity between generated and actual wounds. Visual inspection is performed to examine generated wound images. The results show that the proposed synthetic wound image generation (WG2AN) model has great potential to be used in medical training. It performs well in producing synthetic wound images with a 1000-image training dataset and 200 epochs of training.

(iv) The digital twin is one of the emerging technologies that promise personalized and predictive healthcare. We have proposed the use of digital twin in wound care management by utilizing AI models developed throughout this dissertation. The potential of the digital twin in chronic wound management is examined. After reviewing the concepts and approaches in wound care, several enabling technologies such as cloud computing, AI, and advances in communication are discussed. The digital twin in chronic wound care will shed light on the ability

to provide both optimal treatment pathways and better interpretation of the treatment.

(v) A data-driven wound healing prediction framework has been developed that could effectively guide chronic wound treatment. This model will lead to improved treatment outcomes by utilizing digital twins in chronic wounds proposed in the previous chapter. Digital correspondence of the actual wounds will simulate and imitate the healing progress. The early identification of non-healing wounds also is possible, which will help arrange and adjust chronic wound treatment effectively. By building a digital twin in healthcare, tailored treatments will play an essential role in identifying problems beforehand.

1.3   Dissertation Outline

This dissertation first reviews various types of chronic wounds and their tissues. Based on this understanding, a classification and segmentation method is developed to categorize the wounds and their tissues. In the second stage of this study, a wound generation architecture is developed to map tissue segmentation to a lifelike wound image. The third stage contains the development of a wound healing prediction model. Finally, an overall chronic wound management system is outlined.

The remainder of this dissertation is organized as follows. First, an AI-based chronic wound classification and its explanation using XAI are introduced in Chapter 2. Second, a border segmentation and tissue classification model using AI is proposed in Chapter 3. The AI-based synthetic chronic wound generation is presented in Chapter 4. Chronic wound management and the digital twin concept are examined in Chapter 5. A chronic wound management framework utilizing the digital twin concept is proposed in Chapter 6. Chapter 7 summarizes the general conclusions and points out some future directions.

2 AI-based Chronic Wound Classification

AI has seen increased applications and widespread adoption over the past decade despite, at times, offering a limited understanding of its inner working. AI algorithms are, in large part, built on weights, and these weights are calculated as a result of large matrix multiplications. Computationally intensive processes are typically harder to interpret. XAI aims to solve this black-box approach through the use of various techniques and tools.

In this chapter, we present chronic wound classification. We discuss and propose a model for wound classification and its explanation. Section 2.1 discusses wound classification. In Section 2.2, extended discussions on transfer learning, XAI, and the model pipeline are presented. One of our primary contributions in this area of research is a novel explanation method. Section 2.3 introduces data collection, pre-processing, the environment, and validation subsections of the study. Section 2.4 presents the outputs of the implementation of transfer learning, and XAI approaches to wound classification. Section 2.5 provides results and related discussions. Section 2.6 concludes the chapter.

2.1 Introduction

AI is capable of analyzing complex data and exploiting non-intuitive approaches to derive meaningful relationships [35]. Healthcare applications based on AI are utilized in early detection, diagnosis, treatment, as well as outcome prediction and prognosis evaluation [36]. The barrier that stands in the way of AI applications is sourced from the lack of transparency and black-box nature that cannot be explained directly [37]. The black-box nature of AI systems could be explained as follows. When an AI model learns and gives an output, it processes the data and deciphers the processed information immediately instead of storing the learned data as a clear digital memory [38]. This is why an explainable and understandable glass-box approach should be

8

taken to enable transparent, trustable, and re-traceable AI applications [39]. Chronic wound management, which is one of the critical fields in healthcare, also requires XAI models. In this chapter, AI techniques are applied to the classification of chronic wounds, i.e., diabetic ulcers, lymphovascular, surgical, and pressure injury.

The XAI term is coined to provide transparency, and guided inference in understanding the decision-making processes of the AI system [40]. The study in [41] provides a comprehensive review of XAI in terms of concepts, taxonomies, opportunities, and challenges, as well as a discussion on adopting XAI techniques to image processing. The study in [42] summarizes the recent developments in XAI and its connection with artificial general intelligence, as well as identified trust-related problems of AI applications. The study in [43] examines the state of AI-based FDA-approved medical devices and algorithms. Although millions of dollars funded medical AI research in 2019, only ten (10) medical devices have been approved by the FDA. The authors in [44] present a comparative analysis of approved AI and Machine Learning (ML) medical devices. The approved devices are being used mainly in radiology, and a few are qualified as high-risk devices. The acceptance of AI is still low amongst medical practitioners with various matters related to trustworthiness and reliability [45]. Authors in [46] identified nuances, challenges, and requirements for the design of interpretable and explainable ML models and systems in healthcare and described how to choose the right interpretable ML algorithm. Conventional black-box AI systems are turned into glass-box systems with the help of XAI techniques which provide data about the intermediate steps of the inference process [47, 48]. An example of this would be a computer-aided diagnosis system that not only outputs a prediction but also shows where it looked during the decision-making process by overlaying a heat map on top of an X-ray image. The study in [48] presents the Grad-CAM technique by utilizing the gradients that are taken from the convolution layer to generate a highlighted localization map. Grad-CAM benefits the convolutions, whereas our proposed

9

method calculates the most effective features by tweaking the input and perceiving its effect on classification. Authors in [49] presented classification tasks using Local Interpretable Model-Agnostic Explanations (LIME) to explain predictions of DL models, to be able to make these complex models partly understandable.

In [50], the authors proposed a classification technique where they combined the Genetic Algorithm (GA) and Adaptive Neural Fuzzy Inference System (ANFIS) to predict heart attack through XAI at satisfactory rates. Authors in [51] developed an assisted and incremental medical diagnosis system using XAI, which allows the interaction between the physician (i.e., human agent) and the AI agent. Authors in [52] investigated the problem of explainability in AI in the medical domain where wrong system decisions can be very harmful and proposed two approaches to explain predictions of DL models, (i) computes sensitivity of the prediction with respect to changes in input, and (ii) decomposes decision in terms of the input variables. Authors in [53] investigated how to increase the trust in computer vision through XAI and how to implement XAI to better understand AI in a critical area such as disease detection. This chapter presents a highly transparent XAI tool for the classification of chronic wounds, i.e., diabetic ulcer, lymphovascular, surgical, and pressure injury.

The objectives of the study are:

(i) Build a wound type classification model using DL and transfer learning methods.

(ii) Showcase an approach to make common AI models more transparent and explainable to understand the results and gain trust in the AI model.

(iii) Utilize readily available AI neural networks to show that more transparency or explainability can be introduced to a variety of commonly available models, such as transfer learning.

(iv) Apply XAI methods to convert complex black-box AI systems to more understandable glass box AI systems that aim to provide a look into the internal

decision-making mechanics to give the user the ability to follow the reasoning behind the AI models' prediction.

(v) Provide insights into the complex decision-making processes of an AI system in the field of healthcare applications, especially chronic wound type classification.

2.2 Methodology

This section discusses the methodology of transfer learning for the wound type classification and XAI for providing transparency to the classification task as well as the overall model pipeline.

2.2.1 Transfer Learning

Predictions on new data utilizing data distributions and statistical properties of a previously trained model are called transfer learning [54]. The same distribution of the training and the testing dataset is needed for traditional ML models [55]. However, transfer learning provides flexibility and capability of training on a smaller dataset by transfer of learned features from an old model to the new model.

The transfer learning application comprises two steps, (i) feature extraction and (ii) fine-tuning. The pre-trained network will extract meaningful features from new data samples, with a final classifier added on top of the pre-trained network to do classification tasks in the target domain. The pre-trained network masters feature extraction tasks with Convolutional (Conv) layers. The second step is fine-tuning through freezing and unfreezing some of the top layers from the pre-trained model to train for higher performance jointly. ResNet [56], EfficientNet [57], and VGG16 (Very Deep Convolutional Neural Networks for Large-Scale Image Recognition) [58] networks are a few of the successful DL models for classification tasks. In this study, transfer learning is utilized with VGG16 architecture in order to utilize its object

11

detection capabilities. Its architecture is shown in Figure 2.1, which gives the flexibility and best score among other DL models. VGG16 consists of roughly 138 million parameters and is trained over 14 million images on the ImageNet [59] database. The network is initialized with random weights before the training [60]. The pre-trained Conv layers of the VGG16 architecture are kept frozen, and only fully connected output layers after Conv layers are trained in the first phase of the transfer learning, where Conv layers' weights are not updated. In the second phase, the Conv layers are kept frozen, but the last Conv layer is kept unfrozen. The last Conv layer and fully connected layers are trained together to fine-tune the model, i.e., Deep Neural Networks (DNN). The weights of Conv layers from the VGG16 are transferred to utilize their feature extraction skills. The training of the last Conv layer provides the fine-tuning necessary to obtain better classification results.
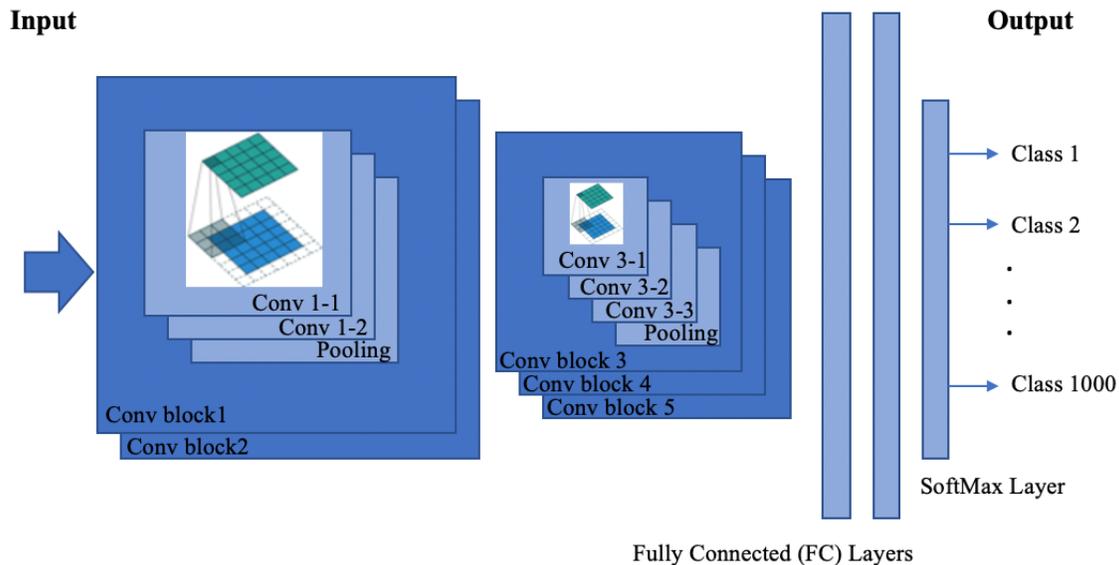
Figure 2.1: VGG16 architecture.

2.2.2 Explainable AI (XAI)

AI provides tremendous benefits in various sectors, but its adoption is limited due to the non-intuitive, opaque nature of ML models [61]. The internal working of an AI

model is complicated and requires a solid mathematical background to understand. This can be a significant barrier to entry [62]. There are two kinds of approaches to explaining an AI model; (i) the comprehensible and (ii) the interpretable model. Comprehensible models are explained with posthoc explainability approaches. Classical ML methods (e.g., regression models and decision trees) are interpretable models as these reveal greater transparency when compared to Conv networks [63]. The inner workings of ML models might be complicated and hard to interpret, yet their efficiency and accuracy are higher than human performance in many cases [64]. This improved efficiency and accuracy are the main reasons why we need to comprehend the inner workings of ML models.

Generalized Linear Models (GLM) provide meaningful, clear, and accessible feature importance that indicates the relative importance of each feature when making a prediction for the regression models. Outputs of regression models are a linear combination of features with different weights depending on the significance of features [65]. Tree-based models have individually meaningful features, with tabular-style datasets used in these models. The connection of tree-based models to the training data results in greater interpretability with local explanations in comparison to linear regression models [66].

DL is a relatively new research field compared to classical ML models. The sheer number of parameters and non-linear structure of DL prevents linking inputs to the model prediction. Therefore, a post-hoc explainability approach is taken. Gradient and attention-based methods are developed and used in the context of the image and text-based models, respectively. The gradient-based method brings attention to important regions in the input image in the backward pass. The attention-based method trains attention weights, which determine how much each of the elements is in the final output [67].

Generalized XAI methods are designed to treat any ML model as a black-box

13

with inputs and some outputs [68]. One of these methods is LIME [69]. It finds the statistical connection between input and model prediction by training local surrogate models on perturbed inputs instead of training them globally [70]. It provides both an explanation of an instance by an interpretable representation as well as visualization. This study provides the explainability and transparency of chronic wound classification using transfer learning implementation with Keras and XAI methods.

2.2.3 Model Pipeline

The proposed model architecture consists of two main parts, i.e., classification and explanation. In the first part of the process, the chronic wound images are classified into four categories, i.e., diabetic, lymphovascular, pressure injury, and surgical. This part of the model employs a pre-trained VGG16 network, i.e., transfer learning, which is capable of extracting features using 13 Conv layers. These layers are already loaded with pre-trained weights using the ImageNet dataset that is publicly available. The last three Fully Connected (FC) layers and the softmax layer is trained with the chronic wound dataset from the ground up to provide weights for the classification of chronic wounds. After training the classification part of the model with these steps, images are fed to the XAI part of the model, where the LIME XAI tool and heatmap are utilized for the explanation. The process of classification and explanation of chronic wound images is illustrated in Figure 2.2. The input wound image is simply classified by the model consisting of transfer learning and DNN and then explained with an XAI tool, i.e., LIME and heatmap, to provide transparency to the classification.

2.3 Data Collection, Pre-processing, Environment, and Validation

This section discusses data collection, data pre-processing, and the test environment. Details about the dataset are given in the data collection section. Forming

14
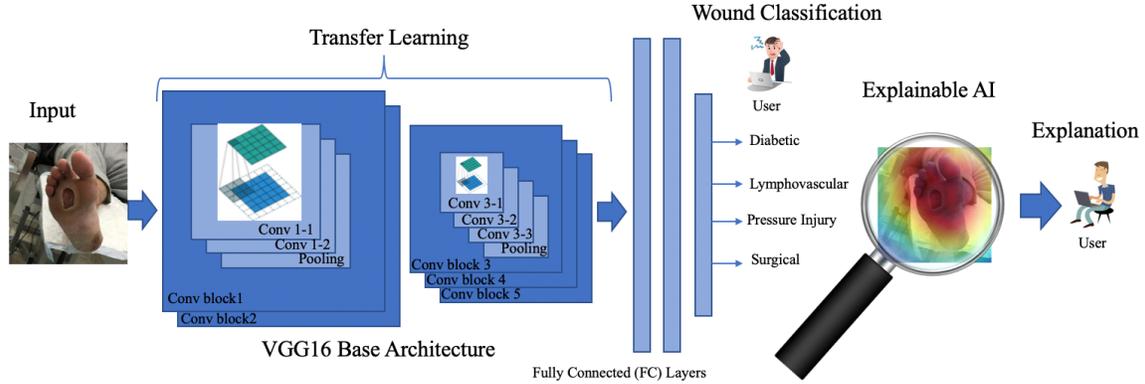
Figure 2.2: Wound classification model with transfer learning, DNN, and XAI tool.

a ground truth for classification and the environment that the model runs on is explained in the data pre-processing and environment sections, respectively.

2.3.1 Data Collection

The chronic wound data repository, which includes diabetic, lymphovascular, pressure injury, and surgical wound types, are collected from the eKare Inc. data repository and was anonymized for patient privacy. eKare Inc. specializes in wound management, with its services used by many hospitals and wound clinics for patient/wound management. A total of 8690 wound images were chosen by an MD specialized in wound care to represent the aforementioned wound types. The dataset comprises 1811 diabetic, 2934 lymphovascular, 2299 pressure injuries, and 1646 surgical wound images. The proposed model uses wound images to predict wound etiology utilizing transfer learning, data augmentation, and DNNs.

2.3.2 Pre-processing

The dataset was reviewed by a trained MD to ensure the correct classification of underlying chronic wound etiology. This validated classification serves as the clinical ground truth. Wound images are then hand-labeled for wound type classification. The distribution of the dataset is not even, as the dataset is fine-tuned for a correct representation of chronic wound classes. Data augmentation techniques such as mir-

15

roring, rotation, and horizontal flip are used to increase dataset size and maintain class balance. The dataset, 8690 images in total, was split into training and test sets comprising 6520 and 2170 images, respectively. The collected data was pre-processed to increase data quality. This includes formatting, rescaling, and normalization of the images. Images were scaled to 224×224 pixels and normalized for a faster learning process.

2.3.3 Environment

The proposed model was implemented using the Keras DL framework with Python version 3.6. We used a workstation to run our model, which has an Intel® Core ™ i7 -8700X CPU @3.20 GHz with 32 GB memory, NVIDIA GeForce GTX 1080 GPU with 8 GB dedicated and 16 GB shared memory. We trained the model for 1000 epochs where the model has warmed up 250 epochs with only training FC layers, then an additional 750 epochs with the training of FC layers, and the final set of the Conv layers. The total training of the model took around 8 hours. We used a constant learning rate of 0.001 for the "RMSprop" optimizer for the training.

2.3.4 Validation

Validation was done using the confusion matrix shown in Table 2.1. Precision gives the ratio of correctly classified wound types over total positive wound type predictions. The recall is a measure of how many of the positive wounds are correctly classified. This metric checks predictions in the eye of true labels. A high recall value relates to the identification of more true positive and, therefore, fewer incorrectly classified samples. Interestingly, both of these metrics could be high, yet the model could still underperform. This is why a third metric is utilized to characterize the model performance. F1-score is a hybrid measurement that brings together both precision and recall for a better evaluation.

Performance measures are given in Equations 2.1 - 2.3 below.

16

Table 2.1: Confusion matrix.

		Prediction	
		y'=0	y'=1
True Label	y=0	True Negative	False Positive
	y=1	False Negative	True Positive

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{2.1}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{2.2}$$

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{2.3}$$

The Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) are also used as performance measures and shown in Figure 2.3. Higher AUC values indicate the classification capability of the proposed model. The X-axis of the ROC curve is recall, and Y-axis is the false positive rate (FPR) which is given in Equation 2.4 below.

$$FPR = \frac{FalsePositive}{TrueNegative + FalsePositive} \tag{2.4}$$

17

2.4 Implementation of Transfer Learning and XAI Approaches on Wound Classification

The objective of this subsection is to explore and apply XAI methods to chronic wound classification to expand knowledge about the opaque "black-box" structure of the ML model. The test dataset comprised 25% of the data, while the remaining 75% was used as training data. Data augmentation techniques such as mirroring, rotation, and horizontal flip are used to avoid overfitting and to increase the dataset for better training performance. Test data is indexed for generalization of the model and proper comparison. Transfer learning is realized in two steps: first, a warm-up phase, and second, a fine-tuning phase. This study, using transfer learning, provided satisfying results according to performance metrics, which are F1-score, recall, and precision (features extracted from the confusion matrix). Precision, recall, and F1-scores of each wound type, and their averages, are compared in Table 2.2.

Table 2.2: Classification performance evaluation of the proposed model.

Model	Precision	Recall	F1-Score
Diabetic	0.85	1.00	0.92
Lymphovascular	0.95	0.98	0.96
Pressure Injury	1.00	0.86	0.92
Surgical	1.00	0.91	0.95
Average	0.95	0.94	0.94

Higher precision values of lymphovascular, surgical, and pressure injury wound types indicate the model performed very well with these wound types. In contrast, pressure injuries were harder to diagnose (low recall score for pressure injury wounds). This means that some pressure injury wounds are not learned or are similar to another wound type and misclassified by the model. Lymphovascular wounds have one of the highest recall scores among all wound types, which reveals that the proposed method is capable of diagnosing lymphovascular wounds. The F1-score on the performance of lymphovascular wounds is high, and pressure injury is low. Surgical wounds have

fair precision and F1-scores but have low recall scores. Hence our model is likely to classify a surgical wound as diabetic. The recall of diabetic wound types is pretty high, and it has one of the lowest F1-scores, which is a result of low precision. The ROC curve and AUC results are depicted in Figure 2.3. Lymphovascular and surgical wounds have the highest AUC values, whereas diabetic and pressure injury suffers from low precision (diabetic) and recall (pressure injury).
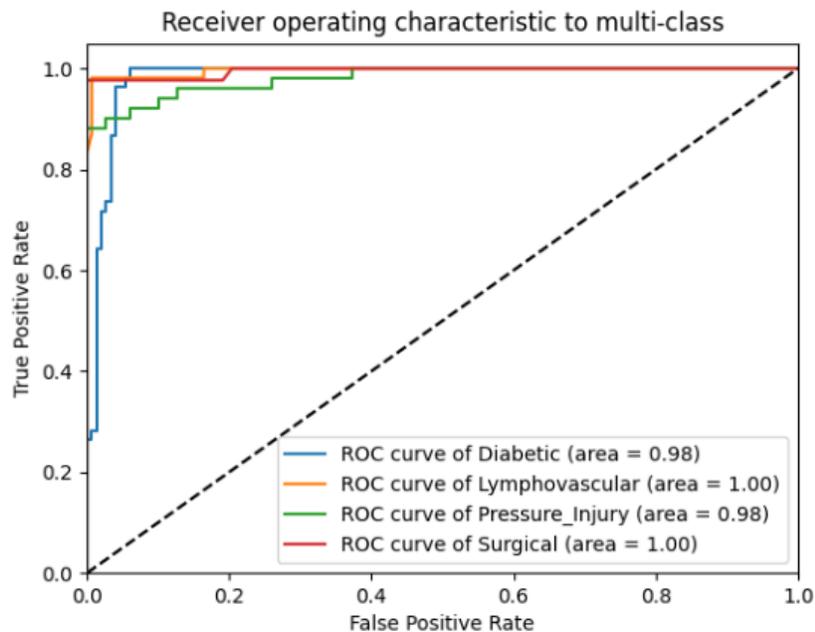
Figure 2.3: ROC curve of the wound classification.

As AI-based products provide efficiency and automation, AI has become very popular in low-risk fields, such as agriculture, customer services, and manufacturing. However, applications of AI remain limited in high-risk domains, such as health care, as trust is critical in medical practice [45]. Reliability issues concerning patients and medical practitioners, as well as regulations, hinder the adoption of AI-based systems [43]. Understanding the rationale behind model predictions would undoubtedly help users decide when to trust or not to trust their predictions.

A DNN using the transfer learning technique was trained using chronic wound

images to predict the wound type. Accurate wound type designation helps a clinician to classify the wound type, which serves to better steer the treatment approach. Prediction of the image classification is then explained by an "explainer" that points to visual features of the image that are the most important to the model. With this information related to the model rationale, the clinician can decide whether to trust the model or not. Model outputs include an understandable qualitative link between inputs and predictions, which is an essential part of the explainability aspect [71]. The rich model feature-set is too numerous and difficult to interpret directly, yet by facilitating a guided qualitative approach, human reasoning can be augmented with additional model data [72]. Another significant property that a reliable explainer should have is local faithfulness. Local faithfulness is achieved by characterizing the response of a local function with a range of adjacent inputs [73].

In this study, the DNN model with transfer learning and extended XAI technique is used to provide explainability and transparency for wound image classifiers by visually indicating what particular class is estimated for various model regions. The proposed model forms a hybrid XAI framework through the use of LIME and heatmap proposals. LIME architecture using superpixels is implemented similar to the study in [71]. LIME provides a set of correlated and connected pixels, which are used as input to the heatmap method. The proposed model provides a focus on the classification task through the use of a heatmap. Medical practitioners often conceptualize the clinical problem based on their knowledge acquired in medical school, as well as clinical experience. The heatmap approach is a fairly naive method of raising focus to different image regions based on the model. The basic intuition with the use of the heatmap is that by drawing focus to certain image regions, practitioners will narrow their attention to regions where the heatmap data correlates with their medical intuition. Warmer colors indicate the more critical areas of the wound in the importance map.

<div align="center">(a) Original Wound Image   (b) Importance Map</div>
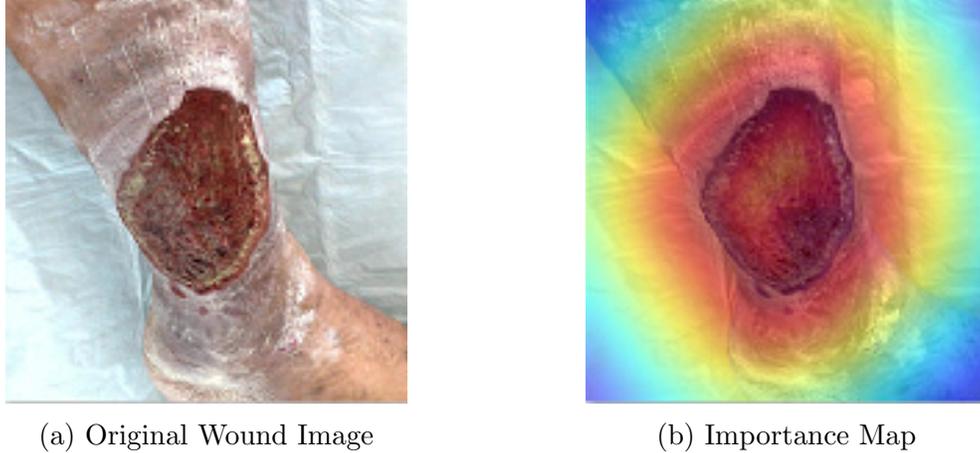
Figure 2.4: Original lymphovascular wound image and its explanation using heatmap.

The proposed model classifies a chronic wound as a lymphovascular wound with a probability of 99.9%, shown in Figure 2.4a. Figure 2.4b highlights the model's focused area for classification tasks in the wound image with an importance map as an explanation.

Figures 2.5-2.8 show images of diabetic, lymphovascular, pressure injury, and surgical wounds. Each wound type has a respective heatmap highlighting the focused area that affects the model to choose the proper wound type. Diabetic wound is correctly predicted at 95.36% (Pressure injury: 4.07%, lymphovascular: 0.01%, surgical: 0.56%) and lymphovascular wound is predicted at 100% (Diabetic: 0%, pressure injury: 0%, surgical: 0%) in Figures 2.5 and 2.6, respectively. The low diabetic wound classification probability can be increased with additional data to amplify feature extraction of diabetic wounds during training.

Probabilities of wound classification are very high for Figure 2.7, i.e., pressure injury wound at 100% (Lymphovascular: 0%, surgical: 0%, diabetic: 0%), and for Figure 2.8, i.e., surgical wound at 99.91% (Diabetic: 0.05%, pressure injury: 0.03%, lymphovascular: 0.01%).

Figures 2.5a and 2.5b show explanations of the most important features that contribute to the prediction. Like Figure 2.5a and 2.5b, Figure 2.6a and 2.6b shows

<div align="center">21</div>

(a) Original Wound Image (b) Importance Map

Figure 2.5: The probabilities of wound types: Diabetic: 95.36%, pressure injury: 4.07%, lymphovascular: 0.01%, surgical: 0.56% (Diabetic).
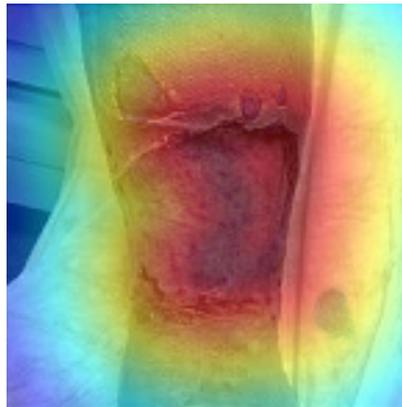
(a) Original Wound Image (b) Importance Map

Figure 2.6: The probabilities of wound types: Lymphovascular: 100%, diabetic: 0%, pressure Injury: 0%, surgical: 0% (Lymphovascular).

explanations and map features with the highest contribution to prediction for lymphovascular classification. Both figures provide insights into why the wound type was predicted to be diabetic or lymphovascular. Focus on the diabetic wound includes surrounding wound tissues and toes, with the shape of the ulcer and its proximity to toes as the explanations for the diabetic foot ulcer.

The lymphovascular wound, as seen in Figure 2.6a, is explained with a focus on deeper damaged tissue. This kind of explanation enhances trust in the wound classifier and helps caregivers make a decision and support their decision with a visual
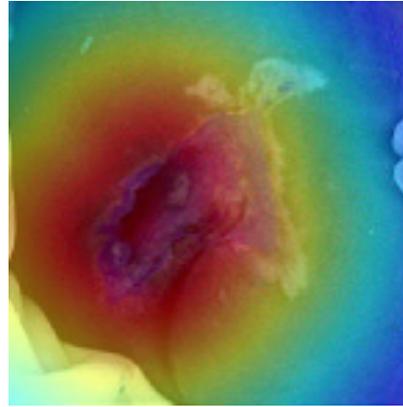
(a) Original Wound Image   (b) Importance Map

Figure 2.7: The probabilities of wound types: Pressure Injury: 100%, lymphovascular: 0%, surgical: 0%, diabetic: 0% (Pressure Injury).
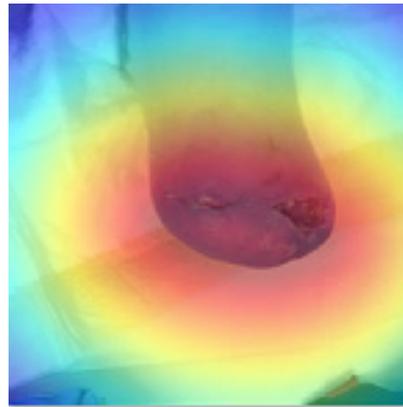
(a) Original Wound Image   (b) Importance Map

Figure 2.8: The probabilities of wound types: Surgical: 99.91%, diabetic: 0.05%, pressure injury: 0.03%, lymphovascular: 0.01% (Surgical).

explanation. The pressure injury wound explainer focuses on the wounded area and indicates the correct placement of the wound, shown in Figure 2.7b. In Figure 2.8, a surgical wound image is explained with a scar pattern and the shape of the wound. The explainer identifies the scar of the wound as the highest feature, and the wound area is highlighted by the proposed model with an importance map.

The proposed method explains diabetic wounds with respect to wound tissue and ulcer location. Diabetic ulcers mainly occur under the foot and follow a similar pattern. A different diabetic wound occurs just below the ankle in Figure 2.9, which

is misclassified as a lymphovascular wound. This kind of ulcer is hard to differentiate from lymphovascular wounds because of its location, as lymphovascular wounds frequently occur at the ankle. Misclassification of a diabetic wound can also be the result of a large wound area; wherein lymphovascular wounds typically cover larger areas than diabetic ulcers.

Lymphovascular wounds are detected with a high probability. There is a slightly lower probability of a lymphovascular wound in Figure 2.10. The spread of the wound forms a line that looks like a surgical wound's scar. The darker part of the wound also looks like a diabetic ulcer. That's why the proposed model gives about a seven percent probability to each wound. Nonetheless, the proposed method highlights the vital areas for the lymphovascular wound correctly.

It is assumed that the pressure injury wound in Figure 2.11 is misclassified due to the size and the shape of the wound area. Pressure injury typically has a large wound area with surrounding damaged skin. As shown in Figure 2.11, the wound occurs under the foot, which is a common diabetic wound area, and also, the wound area is smaller in comparison to the regular pressure injury wounds. These comprise the reasons why the proposed model misclassified the image of pressure injury wounds.

Figure 2.12 depicts a surgical wound, which is correctly classified with a probability of 63.4%. This surgical wound might be the result of a previous pressure injury that covered a larger area. The vast spread of the wound causes this conclusion for the model. In addition to this, the model is confused with the edge of the white cloth, which causes a larger highlighted area. The darker and deeper wound in the middle might be the reason for the high diabetic wound percentage. On the other hand, surgical wounds tend to take a longer time to heal and may convert to diabetic ulcers in diabetic patients. Model classification performance could be increased by collecting more data as this will strengthen the extraction of wound features in the training phase.
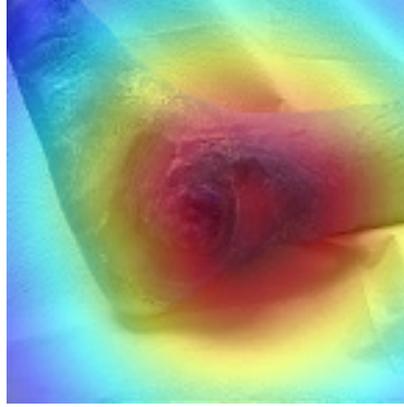
(a) Original Wound Image   (b) Importance Map

Figure 2.9: The probabilities of wound types. Diabetic: 29%, pressure injury: 14%, lymphovascular: 56%, surgical: 1% (Diabetic).
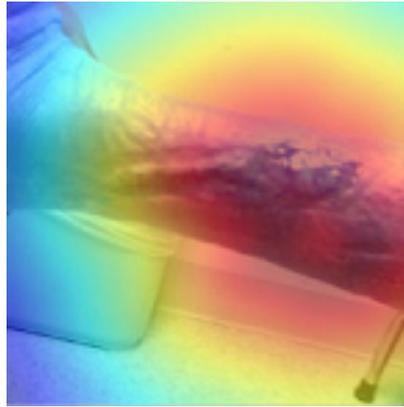
(a) Original Wound Image   (b) Importance Map

Figure 2.10: The probabilities of wound types: Lymphovascular: 80.8%, diabetic: 7.4%, pressure injury: 4.5%, surgical: 7.3% (Lymphovascular).

2.5 Results and Discussion

The proposed model extracts features with Conv networks from a pre-trained VGG16 network. The use of transfer learning accelerates training and produces efficient results, as shown in Figures 2.4 - 2.8. Performance metric evaluation of the model on diabetic wounds (with a precision of 0.85, recall of 1.00, and F1-score of 0.92) indicate that the model has limitations with feature identification for this wound type. This is especially evident with sparse datasets. Surgical wounds have a fair performance on the evaluation metrics where precision, recall, and F1-scores are
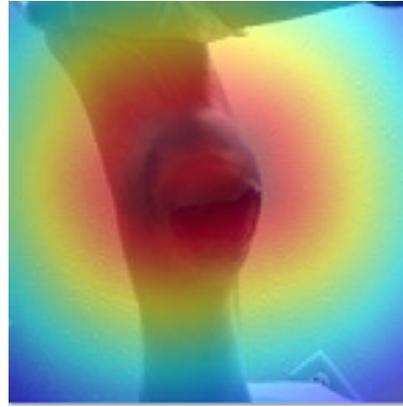
(a) Original Wound Image   (b) Importance Map

Figure 2.11: The probabilities of wound types: Pressure injury: 27.3%, lymphovascular: 12%, surgical: 2.5%, diabetic: 58.2% (Pressure injury).
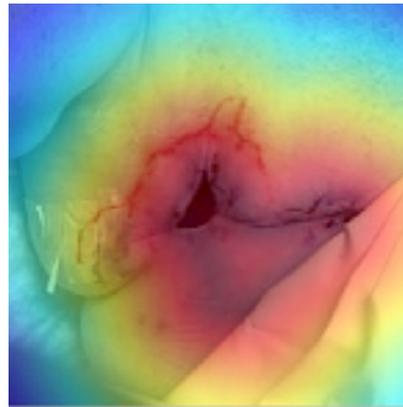
(a) Original Wound Image   (b) Importance Map

Figure 2.12: The probabilities of wound types: Surgical: 63.4%, diabetic: 19.4%, pressure injury: 15%, lymphovascular: 2.2% (Surgical).

1.00, 0.91, and 0.95, respectively. Precision, recall, and F1-scores of lymphovascular wounds are 0.95, 0.98, and 0.96, respectively. Pressure injury wound type has one of the highest precisions, 1.00, a low recall score, 0.86, and an F1-score of 0.92. Surgical and pressure injury wounds have good precision and low recall scores. The recall score of pressure injury wounds is low, which is an indicator that the proposed model has some difficulty in learning the features of pressure injury wounds. The proposed model has the average precision at 0.95, the recall at 0.94, and the F1-score at 0.94. The ROC curve and the AUC provide a visualization related to the performance of

the model on the classification task. The performance of the model could be improved with a more extensive training dataset [31] and fine-tuning the hyperparameters [32].

The second part of the model is specialized in explaining why the model gives a specific output with a hybrid structure. This part extends the LIME technique using a heatmap model. Heatmap is used as a tool to draw focus to image regions based on work done, with the intuition being that practitioners will take less time under guidance. The explainer of the proposed model is successful, while the classification part of the hybrid model could be further improved with additional data (a common problem in data-hungry DL models). The explainer provides visual cues through the use of a heatmap overlaid on wound images to indicate image regions identified by the AI model.

A clinician may eliminate certain wound types for consideration based on the location of the wound. For example, in the case of a plantar foot ulcer, a doctor will likely eliminate sacral pressure injury wounds from the possible wound type list. This is why wound location is essential, and an explanation of a wound type should also indicate location information for a complete understanding. Diabetic wound type is explained via the corresponding deeper and darker damaged tissue size and location on toes. These features are stressed and shown in Figure 2.5. Lymphovascular wound features are highlighted and shown in Figure 2.6, where the size and texture of the damaged tissue are essential indicators. Explanation of the lymphovascular wound type is unexpected; its focus is on the border of the lesion and the adjacent areas instead of the whole lesion. This is another case whereby DL utilizes a non-intuitive search space that provides essential information. Pressure injury wounds are explained via wound tissue and the surrounding wound area, as seen in Figure 2.7. Pressure injury wounds often have a surrounding region of newly healed or damaged skin immediately adjacent to the larger wound. A surgical wound has more specific features to explain, such as postoperative scar and stitches.

Observations deduced from the results of the proposed model are summarized below:

Observation 1: AI applications with XAI have a high potential to improve explainability and transparency in high-risk industries, such as healthcare, where trust is critical.

Observation 2: Limitation in the classification task is carried to the explanation part of the model.

Observation 3: The list of possible wound types is decreased significantly based on wound location.

Observation 4: Explainer has different approaches for each class, yet it uses a qualitative method to explain decisions.

Observation 5: Qualitative methods may explain AI models better to non-subject experts as model parameters and inputs alone are too numerous to be meaningful to non-experts.

Observation 6: Given hardships in understanding quantitative methods, human reasoning can be augmented through qualitative methods.

Observation 7: XAI has excellent potential to improve overall model performance by analyzing the effect and importance of features.

Observation 8: Non-expert users are often able to intuitively grasp the rationale behind class decisions made by the model.

Observation 9: AI decision-making processes might be unanticipated, yet they can provide insights and improve how we handle specific tasks through a bottom-up approach.

2.6 Conclusion

This chapter presents a use case of wound type classification in the healthcare domain using an XAI model. The proposed model is used to augment decision-making

through clinician guidance. Moreover, the proposed method reveals the underlying reason for a particular output by analyzing the relationship between input and output. This study intends to showcase an approach to make common AI models more transparent and explainable to understand the results and gain trust in the AI model. By utilizing readily available AI neural networks, it can be shown that more transparency or explainability can be introduced to a variety of commonly available models, such as transfer learning.

DNN using the transfer learning technique is utilized to predict the classification of four wound types: diabetic, lymphovascular, pressure injury, and surgical. The model accepts an image as input and predicts the etiology of a chronic wound as output. It is discussed that trust is crucial for effective human interaction with ML systems and that explaining individual predictions is vital in assessing trust. We used XAI techniques identified here in a healthcare application to faithfully explain predictions of wound-type classifications in an interpretable manner through the use of heatmaps. The proposed model extends the LIME technique with a heatmap method for better explainability. XAI techniques allow AI systems to cooperate with non-expert end-users. The AI and end-user give each other feedback to arrive at a decision together by guiding a human, e.g., researcher or caregiver, during a classification task. It can also explain how a decision was made, tracing back to the inner workings of the AI system. Transparency is crucial in developing caregiver confidence and improving wound treatment.

This study demonstrated that explanations are helpful for wound type classification in the healthcare domain, when assessing trust, to develop new approaches to wound classification and prediction insights. The proposed hybrid model performs well on both chronic wound classification and explanation tasks. Collecting additional data will increase classification performance further. Interpretation of the results obtained from the XAI module provides adequate information about the chosen wound

29

type. Application of other XAI techniques such as Taylor Decomposition, Grad-CAM, and sensitivity analysis will enhance the overall trustworthiness of the model as well.

It is expected that this work can benefit researchers and caregivers who work in the chronic wound management field in healthcare by providing insights into the XAI potential and availability in healthcare applications.

3 Wound Border Segmentation and Tissue Classification using AI

Medical image processing has the potential to assist caregivers in deciding on a proper chronic wound treatment plan by understanding the border segmentation and the wound tissue classification visually.

In this chapter, we propose a hybrid wound border segmentation and tissue classification method utilizing conditional GAN, which can mimic real data without expert knowledge. Section 3.1 lays down the background for the use of AI in wound and tissue detection and localization. In Section 3.2, the proposed GAN-based DL model is examined thoroughly. One of our primary contributions in this area of research is a novel simultaneous wound and tissue segmentation. Section 3.3 introduces data collection, pre-processing, environment, and validation of the study. The performance of the GAN algorithm is evaluated through MSE, Dice Coefficient metrics, and visual inspection of generated images. Section 3.4 presents the model outputs and also analyses the optimum number of training images, as well as the number of epochs using GAN for wound border segmentation and tissue classification. Section 3.5 concludes the chapter.

3.1 Introduction

Wound management technologies are an essential part of the treatment of chronic wounds, which affect around 6.5 million patients at the cost of $25 billion yearly in the U.S. [14]. However, they are lagging technologically, and most caregivers only depend on imprecise optical assessment [23], which brings some complications, such as infection risks, inaccurate measurements, and discomfort to patients [24]. Advanced computer vision methods assist in the accurate monitoring of wound healing [74]. Image processing and ML automate the evaluation of medical images [75]. The computer vision paired with AI would provide caregivers with continuous and accurate

31

wound healing monitoring at a lower cost. Familiarity with wound tissue types and their sizes play an important role in determining the right chronic wound treatment plan. One of the goals of this dissertation is to contribute to the development of such a system for wound border segmentation and tissue classification utilizing the conditional GAN algorithm in a hybrid way.

Yann LeCun, an AI expert in neural networks, called adversarial training "the most important idea in the last two decades in Machine Learning" [76]. The GAN algorithm of DL techniques has been used successfully in many applications, such as style transfer, image synthesizing, and the famous DeepFake synthetic media creator. The power of the GAN algorithm comes from learning directly from data without human knowledge [77]. That means that GAN does not require a human to select features to predict; it extracts from the data itself. On the other hand, the GAN is challenging to train as the complicated loss functions are hard to interpret [78]. Tweaking hyperparameters such as the number of images and epochs for training in neural networks is still a subject of research and is being done empirically [79]. Finding the correct hyperparameters is like black art where there is no absolute path to follow [80].

The data-driven GAN algorithm provides automatic feature generation, which saves time and labor, but it needs a higher number of images as a trade-off [81]. That is why the number of images is the critical parameter to achieving good approximation in GAN-based models. At the same time, data collection and management processes cost tens of millions of dollars in healthcare, such as clinical trials [82]. There are also significant concerns over privacy, confidentiality, and control of the data [83], which makes it difficult to obtain data in healthcare. Collecting data in healthcare is not an easy task, but the GAN algorithm could generate synthetic images that have no cost and could be used without hesitation.

The number of epochs for training is another critical parameter that requires many

32

trials to find the optimum amount and expertise in the healthcare field [80]. High performance in minimal epoch is needed to achieve significant time and labor savings [27]. The question of how many images, and epochs you need to train a GAN has not been answered. This study also provides a rule of thumb for choosing the correct number of training images and epochs for GAN algorithms in healthcare applications.

Prior efforts for wound tissue classification and segmentation include the development of an image analysis algorithm that is capable of wound area assessment, segmentation, and extraction of wound color without correlating to wound tissue from wound images using smartphone cameras [84]. The study [85] proposes the use of the K-means clustering algorithm, which requires feature engineering, for the wound border segmentation and tissue classification using 113 images. Multispectral imaging is utilized by Thatcher et al. [86]. This study examines the tissue characteristics of burn wounds in the light of medical imaging without segmentation of the wound area and tissue. Authors in [87] explored the feasibility of RGB-D (Red Green Blue - Depth) cameras in wound detection, segmentation, and chronic wound area measurement in 3D. However, the use of special RGB-D cameras increases the cost and model complexity of wound management systems. Also, tissue segmentation is not included in this study [87]. K-Nearest Neighbours (KNN), Decision Tree (DT), and Linear Discriminant Analysis (LDA) are used for the burn tissue classification. This study is limited to burn wounds, whereas the proposed method in this chapter covers a variety of wound types. The study in [88] proposes a DL and data augmentation model for wound-region segmentation. The model used in the study [88] segments each wound tissue separately. Also, detection and segmentation tasks are done using different ML models. Authors in [89] propose an automatic skin ulcer region assessment framework using Convolutional Neural Network (CNN) and encoder/decoder DNN. The study in [89] achieves overall wound segmentation, but the segmentation of different wound tissues is not studied. The study in [90] proposes a CNN-based model for the segmen-

tation of wound tissue types. Authors in [90] provide tissue segmentation of pressure injury wounds with the help of manual pre-processing steps, including external mask application and flashlight removal. The study in [91] describes chronic wound status monitoring with wound tissue segmentation using LDA, DT, random forests, and naïve Bayesian. This study only segments the wound into two tissue types, where our proposed method gives more details. The authors in [92] proposed a model that utilizes color correction and a CNN for wound region segmentation. A two-step pre-processing pipeline is discussed in [92] to segment the overall wound without tissue segmentation. The authors in [22] propose a model that segments solely diabetic wounds using CNN and the removal of artifacts with probability maps after a pre-processing step. The study in [93] investigated CNN with different architectures, i.e., U-Net [94], Segnet [95], FCN8 [96], and FSN32 for the wound tissue segmentation. The study in [97] proposes a wound segmentation model using both traditional and DL methods. In [93], the authors added a pre-processing step that includes detection of the wound and a post-processing step that segments solely the overall wound area. The model also could not be trained end-to-end because of the model complexity. Authors in [98] propose a model for automatic wound region segmentation and wound condition analysis with infection detection and healing progress prediction. This study [98] utilizes traditional pre-processing and post-processing steps to improve segmentation performance and does not have tissue classification. Authors in [99] provide a tool for segmenting and locating chronic wounds to facilitate bioprinting treatment using edge detection and segmentation algorithms. In [99], authors utilize semi-automatic overall wound segmentation on a limited number of wound images. Pre-processing and feature extraction steps are used to improve the performance of the segmentation task. The study in [100] proposed a framework for tissue classification based on the appearance and texture of the current and prior visual appearance of chronic wounds. Pre-processing and feature extraction steps are used

for the segmentation task.

In this chapter, the state-of-the-art GAN algorithm [28, 101] is utilized to develop a model that can classify and segment different wound tissue types simultaneously. Unlike previous studies that require pre-processing or post-processing steps, which increases the model complexity, the proposed method provides wound detection and segmentation without implementing such additional steps. Hence, end-to-end training is possible. Furthermore, while many of the previous studies lack the segmentation of different wound tissues, this study provides segmentation of wound tissues; hence, important information related to wound healing status can be recognized. Additionally, the proposed novel approach could be applied to various wound types such as diabetes, pressure injury, and burn despite the prior studies. The medical image synthesis using GAN for hybrid wound border segmentation and tissue classification has not been done previously. These two tasks are realized individually by the previous studies with a focus on one type of wound.

The main contributions of this chapter include:

(i) The development of a hybrid GAN algorithm to perform wound border segmentation and wound tissue classification in one step on different wound types,

(ii) Providing guidance to healthcare researchers with respect to how many images and epochs are needed to perform successful medical image synthesis with GANs for various applications.

3.2   Methodology

A GAN model comprises two neural networks, which are the generator (G), and the discriminator (D). Both generator and discriminator are concurrently trained with real data to capture the data distribution. A random uniform or a Gaussian noise (z) is fed to the generator network to produce fake images(y), G: z $\rightarrow$ y [102]. This makes the output of the generator unique. This newly created fake image is then fed

to the discriminator network [102]. The discriminator network aims to determine if the generated image is from the training set or not, D: y → [0,1]. The generated images are labeled as fake or real, depending on the training data distribution. The generator is to deceive the discriminator network that generated images are from the training set [103]. Figure 3.1 shows the basic structure of the GAN model.
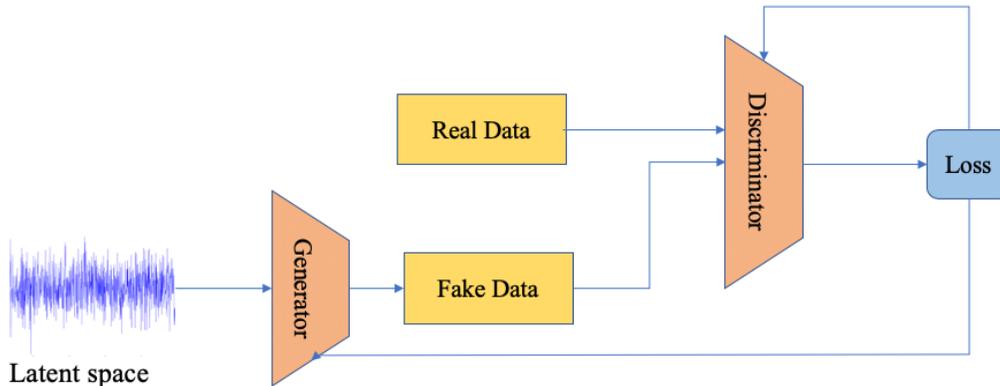
Figure 3.1: The basic structure of the GAN model.

Different versions of the GAN model are developed for different applications, such as conditional GAN (cGAN) [104], cycle-consistent GAN (CycleGAN) [105] , Gaussian-Poisson GAN (GP-GAN) [106], and super-resolution GAN (SRGAN) [107]. Results of a cGAN-based model are evaluated in the scope of this study. The CycleGAN-based model was also examined in this study. However, after initial trials, this model did not yield good results and suffered from the mode collapse issue, a well-known problem in the GAN field, which causes the generation of a particular output image regardless of different inputs [105], for border segmentation and tissue classification tasks. Since this approach failed to produce output results, it was discarded from this study. Other DL-based segmentation methods [108], to our knowledge, have no evidence of their use to simultaneously perform wound segmentation and classification. Existing research firstly performs the segmentation step and consecutively the classification step. Hence, the proposed algorithm could not be compared to other work as a whole because of its novelty. On the other hand, the proposed

novel model in this chapter successfully accomplishes the border segmentation and tissue classification tasks simultaneously by utilizing end-to-end training. In addition to this, the border segmentation task performance is compared with the five different DL models, i.e., VGG16, Segnet, U-Net, Mask R-CNN [109], and MobileNetV2 [110], using the Dice Coefficient metric.

The cGAN architecture has additional properties over the regular GAN architecture, which is also called vanilla GAN. cGAN gets an image as an input (x) in addition to the random noise (z) and generates an output (y) conditioned on that input image, G: x, z→y. The generated image carries similar features to the input image while maintaining the data distribution of the training set, consisting of paired and aligned images. The mapping of input to the output images is learned by the generator network, where the discriminator network learns a loss function to train this mapping, D: x, y→ [0,1] [111]. The objective of both generator and discriminator networks is the same as the vanilla GAN algorithm, with the difference that discriminator and generator observe the input image [104]. Figure 3.2 shows the general architecture of the cGAN model.
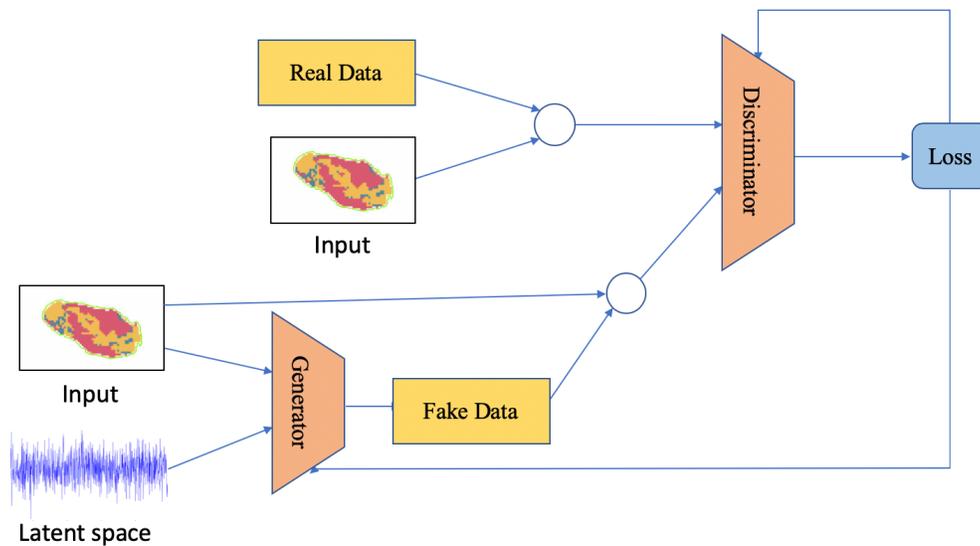
Figure 3.2: The cGAN model overview.

The cGAN model encapsulates two networks and four loss functions to generate plausible fake data. The discriminator networks are updated directly, but generator networks are trained by the feedback coming from the discriminator model while updating the loss function. Trained by the second model, the generator network lacks an objective function, which is the primary reason for GANs' hardship to train [112].

For the generator network, "U-Net" encoder-decoder with skip connections architecture is utilized to get high resolution. The skip connection is a widely used method to keep the original data between the layers. The input is downsampled and flows through many layers, which concludes the input to be a bottleneck. On the other hand, for the image translation, there should be some shared common features needed to have. That's why the cGAN is trained over paired and aligned data, which helps to predict the conditioned output. We used a 70x70 patch-wise comparison of images by discriminator network to classify the generated image as fake or real.

The discriminator network learns to classify real images and fake images with binary cross-entropy loss. There are two loss functions to update the discriminator for real and fake samples, namely D_{real} and D_{fake}. The generator network also has two different losses to provide plausible generated images. The weights of the generator model are then updated with adversarial loss (G_{GAN}) via the discriminator network and L1 loss (G_{L1}). L1 loss is calculated by comparing the generated images with the real image. Adversarial loss and L1 loss scores are combined to obtain the loss of the generator network, shown as:

$$\mathcal{L}_{\text{Gen}} = \mathcal{L}_{\text{Adv}} + (\mathcal{L}_{\text{L1}}) \cdot \lambda \tag{3.1}$$

Where:

$\mathcal{L}_{\text{Gen}}$: Generator network loss

$\mathcal{L}_{\text{Adv}}$: Adversarial loss from the discriminator network

$\mathcal{L}_{\text{L1}}$: L1 loss

λ: Regularizing hyperparameter

L1 loss serves as a regularizing term in the generator network loss with a hyperparameter lambda, $\lambda = 100$. The objective of adversarial loss of cGAN architecture can be depicted as:

$$\mathcal{L}_{\text{Adv}}(G, D) = \mathbb{E}_{\text{x,y}}[log D(x, y)] + \mathbb{E}_{\text{x,z}}[log(1 - D(x, G(x, z)))] \qquad (3.2)$$

Where the generator (G) competes with the discriminator while G is trying to minimize this objective, and D is trying to maximize it [112]. The final objective function can be expressed as shown in equation (3).

$$G^* = argmin_{\text{G}} max_{\text{D}} \mathcal{L}_{\text{Adv}}(G, D) + \lambda \mathcal{L}_{\text{L1}}(G) \qquad (3.3)$$

3.3 Data Collection, Pre-processing, Environment, And Validation

This section discusses data collection, data pre-processing, the simulation environment, and model validation.

3.3.1 Data Collection and Preparation

The chronic wound data repository is provided by eKare Inc., which provides professional wound imaging and analysis services. Images are taken with commercially available cameras by regular users in a natural hospital environment on a normal wound assessment process at the clinic. The chronic wound images, including burn, pressure injury, and diabetic wounds, are semi-automatic segmented for training and testing purposes. The wound tissues are classified as necrotic, sloughy, and gran-

ulation, which are represented in blue, yellow, and red colors, respectively, in the segmentation task. The variety of wounds improves the applicability of the algorithm implemented in this chapter. In this study, anonymized wound images were rescaled to 512x512 pixels. To test the effect of the number of images by the GAN algorithm, we created a set of 100, 500, 1000, 2000, and 4000 images from 13,000 images containing different wound types. The test set was fixed to the same 100 images., and data augmentation, i.e., flipping, was used. Some of the images used in this study can be seen below in the result section.

The number of publicly available chronic wound images is minimal and not sufficient for the comparison of a training dataset of DL-based wound border segmentation and tissue classification tasks. Additionally, it is very challenging or impossible to find chronic wound images with ground truths. Another issue is related to the quality of the images. Medetec wound database [113] is a publicly available dataset that suffers degraded image quality because of the presence of mold growth on the original 35mm transparencies. This will further decrease the resolution of generated images as well. In contrast, the unique eKare Inc. chronic wound image repository provides us with a sufficient number of images, higher quality, and above all, ground truth data in order to sustain high-quality training.

3.3.2 Environment

We implemented the wound border segmentation and tissue classification model using the PyTorch DL framework on the Anaconda platform with Python version 3.6. Our implementations ran on Intel® Core ™ i7 -7800X CPU @3.50 GHz with 16 GB RAM and NVIDIA GeForce GTX 1080 GPU with 8 GB dedicated and 8 GB shared memory. We trained our model 2000 epochs using 100, 500, 1000, 2000, and 4000 images, which took 4 hours, 9 hours, 20 hours, 42 hours, and 76 hours, respectively. The batch size is chosen as 64 to increase the benefit from the GPU. We used a constant learning rate of 0.0002 and an "adam" optimizer for the first half of the

training. The rest of the training was done with a linearly decaying learning rate to zero until convergence.

3.3.3 Validation

Validation was done using MSE and Dice Coefficient metrics for the evaluation of (generated) fake image quality. MSE, which is a pixel-wise loss function, was used to measure the quality of the generated images in addition to losses of GAN. Minimizing the pixel-wise error measurement provides converging results in contrast to GAN loss. Generated segmented images are expected to be very similar to the actual segmented images. In addition, segmented images consist of a combination of three colors, which makes them easy to compare. That's why the MSE metric fits appropriately for the evaluation of this similarity. MSE score was calculated by comparison of real and fake images on pixel level in three color channels. MSE metric can be written as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[(Y_R - Y'_R)^2 + (Y_G - Y'_G)^2 + (Y_B - Y'_B)^2] \qquad (3.4)$$

Where:

n: Number of pixels

Y_R, Y_G, Y_B: RGB (Red Green Blue) pixel values of the real images

Y'_R, Y'_G, Y'_B: RGB pixel values of generated images.

The Dice Coefficient is used to evaluate the performance of the proposed method in addition to the MSE metric. The harmonic means of recall and precision provide a Dice Coefficient which is also known as the F1-score, which is calculated as follows:

$$DiceCoefficient = \frac{2\mid A \cap B \mid}{\mid A \mid + \mid B \mid} \qquad (3.5)$$

<center>(a)   (b)   (c)</center>
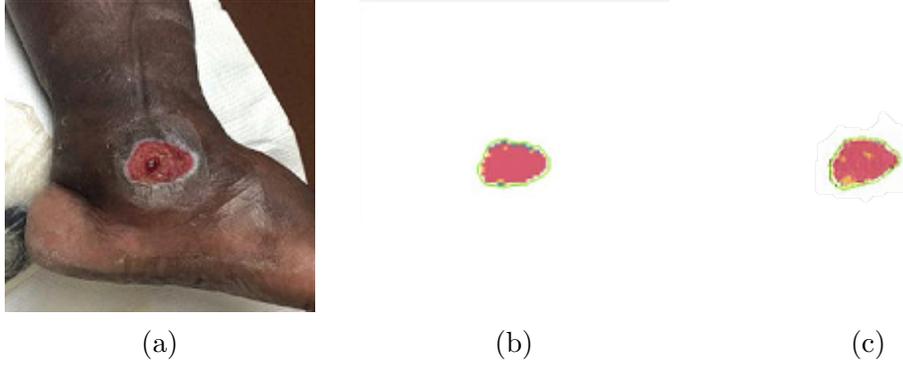
Figure 3.3: cGAN model achieves good results with 2,000 images at 200 epochs: (a) Original image, (b) Segmentation ground truth and (c) Generated segmented wound.

Where A and B are the ground truth, and model output respectively. Dice scores range from 0 to 1 where a score of 1 indicates a perfect segmentation.

3.4   Results and Discussion

This section discusses the output of the model, loss graphs, the effect of epoch on border segmentation and tissue classification, as well as the optimum training conditions of the model.

3.4.1   Model Output

The output of the proposed method was compared with the ground truth. A successful result from the model is given in Figure 3.3, which indicates a proper border segmentation and tissue classification of the wound by training with 2,000 images and 200 epochs.

As shown in Figure 3.3, the proposed model successfully segments the wound border and classifies the wound tissue concurrently. The model learned the wound area in Figure 3.3, where there are paled areas around the heel and the side of the foot. The model is insensitive to color changes and could identify the wound in a crowded environment. The background is discarded as well.

<center>42</center>

3.4.2 Effect of Number of Images on Model Loss

The loss curves of cGAN are depicted in Figure 3.4 and Figure 3.5, when trained with 100, 500, 1,000, 2,000, and 4,000 images, respectively. The G_{L1} loss has the most meaningful loss for the generated image quality. G_{L1}, G_{GAN}, D_{real}, and D_{fake} losses oscillate because the GAN model moves from one type of sample generation to another type of generation before reaching a balance [114]. Training two opposing neural networks concurrently in zero-sum game results in a non-converging problem [114]. G_{L1} represents the generator loss only, and it lacks the contribution of the adversarial loss. G_{L1} loss could be used only to determine the learning capability of the proposed model with respect to dataset size. However, the training progress is unpredictable from the loss alone. That is why an additional technique is needed to predict the progress of training and the quality of generated images.

(a) Loss curve for 100 images.

(b) Loss curve for 500 images.

(c) Loss curve for 1000 images.

(d) Loss curve for 2000 images.
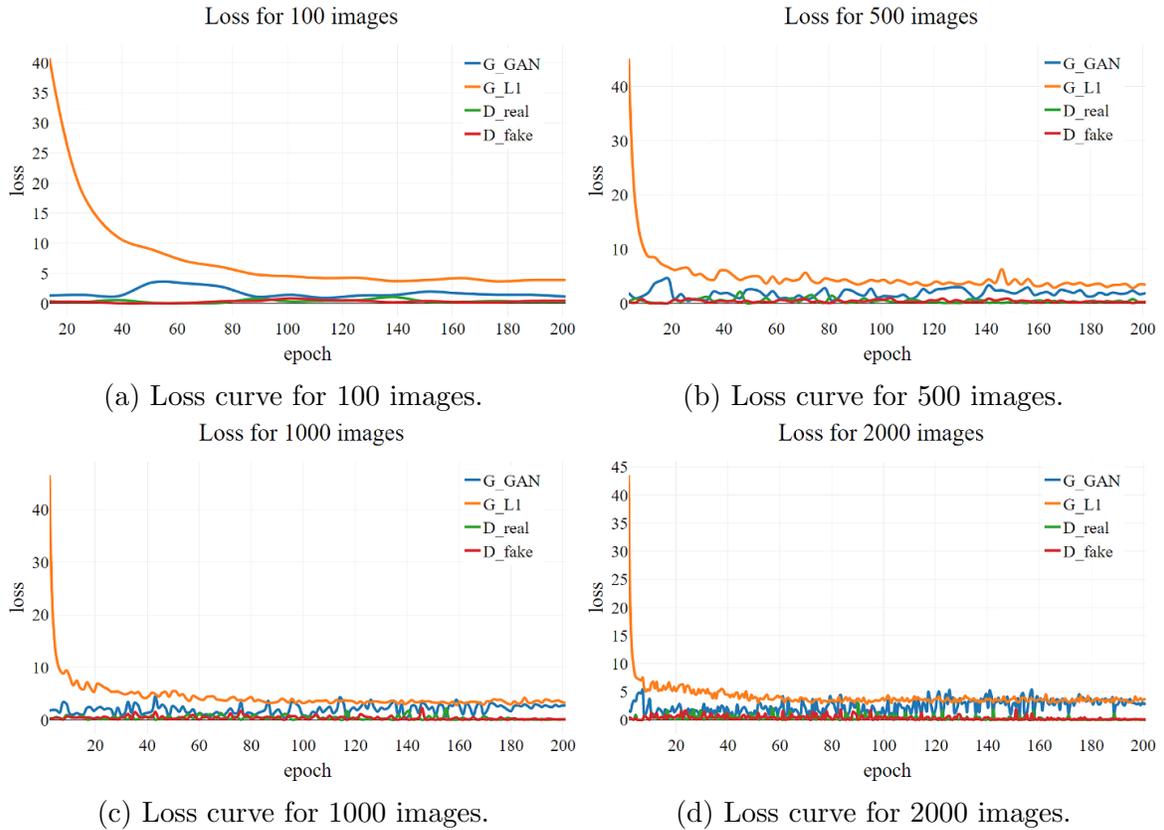
Figure 3.4: Model's loss graphs using a different number of images.

The loss curves for the model with 4000 training images are shown in Figure 3.5. A comparison of loss graphs in Figure 3.4 and Figure 3.5 reveals the drop rate of the

Loss for 4000 images

Figure 3.5: Loss curve for the model with 4000 training images.

G_{L1} loss increases with an increasing number of images for the training. The G_{L1} loss drops to 10 at around the 40th epoch and stays stable under five (5) around the 100th epoch with a training set of 100 images, see Figure 3.4a. The drop rate of G_{L1} loss increases in Figure 3.4b, which is the model loss with a dataset of 500 images. The loss of the proposed model reaches 10 at the 10th epoch and stabilizes under five (5) at the 75th epoch. These are four times and one-and-a-half times improvements over the model with 100 images, respectively. A five-fold increase in the training set provides a similar increase in the G_{L1} drop rate. The drop rate further increases two and four times, respectively, in the models trained with 1,000 and 2,000 images, shown in Figures 3.4c and 3.4d. The results of the model trained with 4000 images are consistent with Figure 3.4b - 3.4d. G_{L1} loss converges to 10 in the first couple of epochs and becomes stable under five (5) in 15 epochs. The results indicate that our proposed model learns the data distribution faster with a larger dataset.

44

(a) Original Image   (b) Importance Map

Figure 3.6: (a) Original image, and (b) Ground truth for border segmentation and wound tissue classification tasks.

(a) Epoch=5   (b) Epoch=75   (c) Epoch=200
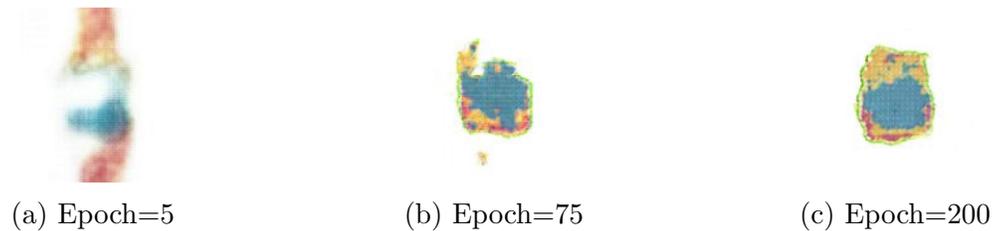
Figure 3.7: Effect of the epoch count on border segmentation and tissue classification tasks.

3.4.3 Effect of Number of Epochs on Border Segmentation and Tissue Classification

Figure 3.6a - 3.6b depicts the original wound image and the ground truth wound tissue classification. By varying the number of epochs, Figure 3.7a - 3.7c shows the results of the model with a dataset of 2000 images, which is trained using 5, 75, and 200 epochs, respectively. With five epochs, as shown in Figure 3.7a, the result does not represent the segmented wound, although it carries similar tissue features. Training the model further to 75 epochs in Figure 3.7b provides a better representation. Adequate generation of the wound segmentation is achieved after 200 epochs of training, shown in Figure 3.7c.

Figure 3.8 depicts the original wound image and the ground truth for wound tissue classification. Figure 3.9 illustrates the output of the model with a dataset of 100

45
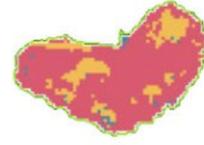
(a) Original image   (b) Ground truth

Figure 3.8: (a) Original image, (b) Ground truth for border segmentation and tissue classification tasks.

images after it was trained for 5, 500, 1000, and 2000 epochs. Training the model for five (5) epochs produces a similar shape but a blurry result with Figure 3.8a, which indicates that the model could not get the data distribution yet. An increase in the epoch count generates better-segmented wound images, but these results could not catch the wound shape as a result of inadequate training images.

The results were also analyzed using MSE scores, as summarized in Table 3.1, which shows the MSE values for a different number of epochs and training images. MSE score is a good indicator of the model's learning ability to mimic the real image data distribution. The MSE score of the model trained for five epochs with 100 images is the highest and improves with the increase in the number of training images and epochs.

Table 3.1: MSE scores for the cGAN model.

Epochs / Images	5	200	500	1000	2000
100	40737	4094	3911	3695	3329
500	11843	2847	2730	2528	2333
1000	6983	3319	3251	3293	3144
2000	3565	2027	2024	2061	2056
4000	3907	1958	1955	1968	2019

MSE values of the different numbers of images are shown in Figure 3.10. The model trained with 100 images dataset did not yield efficient results and was omitted

46

(a) Generated image at the 5th epoch.
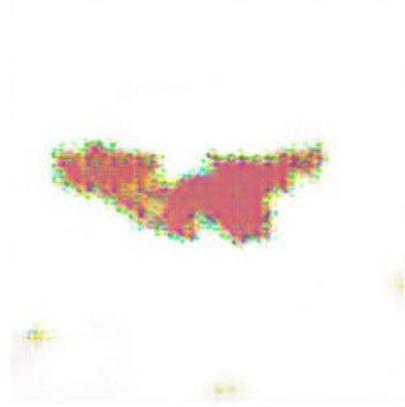
(b) Generated image at the 500th epoch.
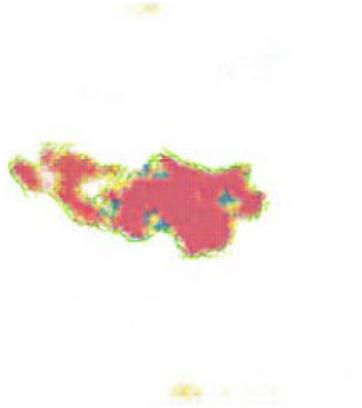
(c) Generated image at the 1000th epoch.
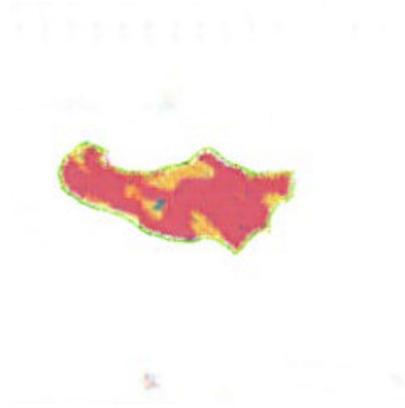
(d) Generated image at the 2000th epoch.

Figure 3.9: The model trained with 100 images is not converging to synthesize border segmentation and tissue classification.

for simplicity in Figure 3.10. The decrease in the MSE score in the first 200 epochs is the highest for all dataset sizes. The dramatic decline in the first 200 epochs indicates that the proposed method successfully learns to segment the wound and classify the tissue type at 200 epochs. Results confirm that increasing epoch count results in a better MSE value for the first 200 epochs, and training for more epochs decrease MSE values slightly. This is a good indicator of optimum training parameter selection that 200 epochs are the optimum epoch number for training the proposed method.

Note that: The model trained with 500 and 1000 images has an equilibrium around 3,000 MSE score, and the model trained with 500 images keeps decreasing, which could be a result of a limited number of images representing a few samples

47
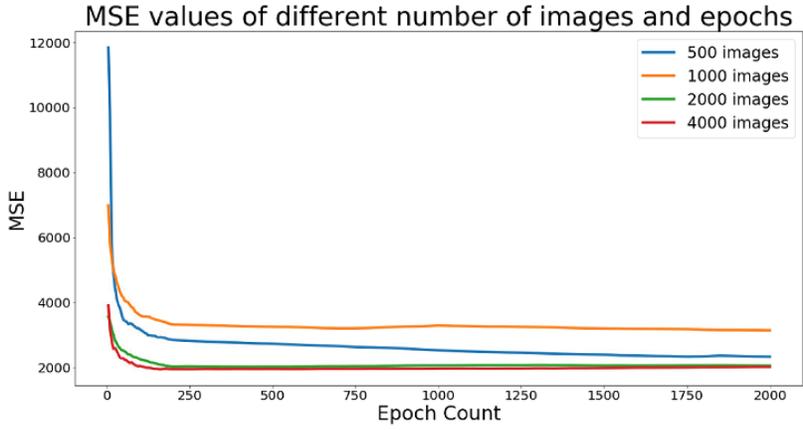
Figure 3.10: MSE vs epoch for the different number of images.

and overfitting that data. There may be a potential overfitting problem. The model trained with 2,000 and 4,000 images share similar MSE values of around 2,000. The outcome in Figure 3.10 indicates that increasing the number of images for training produces lower MSE values, which is a good sign that the proposed model works as expected.

The changes in MSE values with 5-200 and 200-2000 epochs are compared for different datasets, i.e., 100-4000 training images, as shown in Table 3.2. It appears that the MSE value improves significantly during the first 200 epochs of training. Training from 200 to 2000 epochs improves MSE slightly. The changes in MSE score with a smaller number of training images, i.e., 100 and 500, are higher than those with a higher number of training images, i.e., > 1000. This is because the model with a larger dataset converges faster during the first couple of epochs. Negative values imply the increase in the MSE metric, which results from the deformation of generated images and noises in input images.

The Dice Coefficient is also calculated for each epoch and image combination. The Dice scores are shown in Table 3.3 for further analyses. The correlation between the MSE score and Dice Coefficient indicates that the model with 2000 images and 200 epochs is the best performing model, requiring a lower number of images and epochs. The differences between MSE scores and the Dice scores are sourced from the calcu-

48

Table 3.2: The decrease in the MSE score between different numbers of epochs.

Epochs / Images	5 to 200 Epochs	200 to 2000 Epochs
100	90.0 %	18.7%
500	76.0%	18.1%
1000	52.5%	5.3%
2000	43.1%	-1.4%
4000	49.9%	-3.1%

lation of both metrics. The MSE metric considers both overall wound segmentation and the segmentation of the wound tissues. It provides more information about the segmentation performance.

Table 3.3: Dice Coefficients of the proposed model.

Epochs / Images	5	200	500	1000	2000
100	0.09	0.77	0.62	0.84	0.73
500	0.09	0.77	0.89	0.88	0.86
1000	0.18	0.74	0.83	0.76	0.82
2000	0.79	0.90	0.88	0.85	0.78
4000	0.42	0.93	0.93	0.91	0.79

The Dice Coefficient metric provides a measurement of wound area segmentation performance regardless of the wound tissue. The models with a lower number of training dataset images, i.e., 100 and 500, do not provide higher scores as expected. The Dice Coefficient of the model with 1000 training images increases with an increasing number of epochs. The 2000-image model's Dice score is also in line with its MSE score. The 4000-image model has the highest performance metrics, whereas the required number of images doubles compared to the 2000-image model.

The comparison of the proposed model with the previous works [115] is shown in Table 3.4. Five different previous models are compared with the proposed model.

The comparison indicates that the proposed model has similar performance to other highest-performing models. In addition to wound segmentation, the proposed

Table 3.4: The comparison of Dice Coefficients of the proposed model and other models.

Model	VGG16	Segnet	U-Net	Mask-RCNN	MobileNetV2	Proposed Model
Dice Score	0.81	0.85	0.90	0.90	0.90	0.90

(a) Original image   (b) Ground truth

Figure 3.11: (a) Original and (b) Ground truth of tissue classification images.

model provides tissue classification and respective segmentation of tissues as well. That's why the proposed model has not only good segmentation performance but also tissue classification capability as well.

3.4.4 Effect of Number of Images on Border Segmentation and Tissue Classification

Figure 3.11 depicts the original wound image and the ground truth for wound tissue classification. The fixed number of epochs at 200, the results of the models with different training datasets are shown in Figure 3.12. Input datasets that have fewer than 500 images give poor performance. Therefore, they are excluded from Figure 3.12. As shown, the proposed method provides efficient segmentation and tissue classification on a dataset consisting of around 2,000 images or more. It is a significant conclusion that having at least 2,000 images at hand results in efficient training for GAN to generate qualified images in this study. Smaller datasets face difficulties in mimicking the data distribution, or these models overfit the training images, which is the case for the model with a training set of 500 images or less,

(a) 500 images   (b) 1000 images.

(c) 2000 images.   (d) 4000 images.
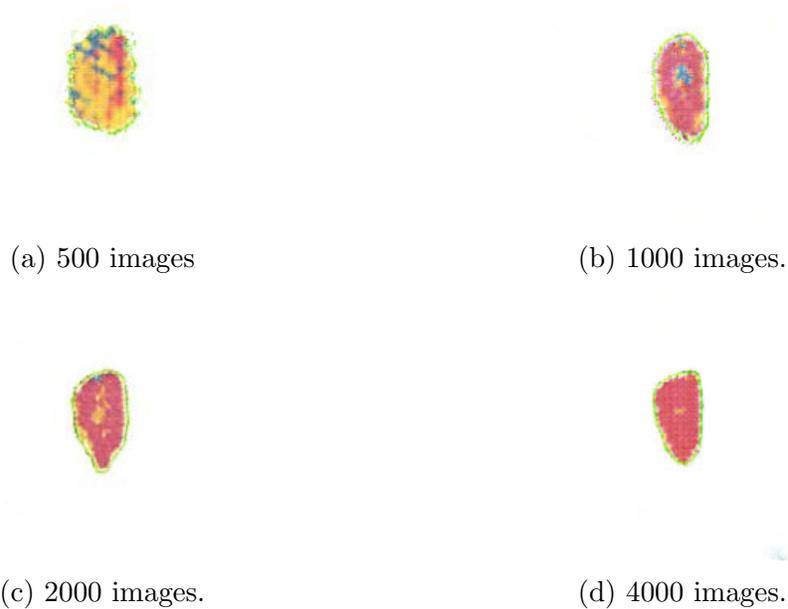
Figure 3.12: Effect of training dataset size at 200 epochs.

whereas datasets with higher than 2,000 images generate plausible images.

3.4.5 Discussion

Based on the study results, the following observations regarding the application of the GAN algorithm could be made.

Observation 1: The proposed method can perform both wound border segmentation and tissue type classification in one step.

Observation 2: cGAN has a high potential of producing close to real synthetic images for wound tissue segmentation and classification.

Observation 3: The quality of the generated images is in line with the image count. 2,000 image count is the threshold for a valid generated image as the results of our study.

Observation 4: The epoch count has a significant impact on the generated image quality, but after surpassing the 200-epoch threshold, the model reaches its convergence, and additional training has a marginal effect on the quality of the generated

image.

3.5   Conclusion

This chapter presents that the cGAN algorithm can achieve chronic wound border segmentation and tissue classification efficiently. The wound border segmentation and the wound tissue type classification using GANs were performed for the first time. Results from different numbers of dataset sizes and epoch counts are evaluated through the MSE metric and visual inspection of generated images. MSE metric provides valuable information in interpreting the quality of the generated segmentation and classification tasks due to the simplicity of the generated images. The optimum training dataset size and epoch count are determined at 2000 images and 200 epochs. This study confirms that the generated image quality increases significantly by increasing the dataset size to 2,000 images. After that threshold, the image quality improves marginally. Currently, data collection in healthcare is an expensive task and process; this study introduces the optimum dataset size for related healthcare applications utilizing GAN. The proposed method achieves border segmentation and tissue classification simultaneously without additional processing steps and expertise. The MSE score decreases, and the Dice Coefficient increases with the increase in generated segmented image quality. The proposed model is in line with these conditions, which are explained in the validation section. The ability to perform end-to-end training and testing ability simplifies the application of the proposed model in healthcare for broader adoption. However, the healthcare industry requires robust and explainable models that will require adopted models to be transparent. The proposed method and DL models in general, lack transparency and behave as a black-box.

The scope of this study includes detection of the various wound types such as burn, lymphovascular, pressure injury, and classification of three different wound tissues, namely necrotic, slough, and granulation. Some limitations in this study could be further addressed in future work. Firstly, the image quality of the overall model

could be further improved. The image quality selected for this study is to provide a fast and straightforward implementation which is the case for many algorithms in the object detection and segmentation field. This is also due to the resolution of available data sets. Since images were collected by various cameras with different settings, it is necessary to format them to a common size for further processing. Secondly, due to the non-converging nature of the GAN algorithm, the loss curves of our model also have limitations providing the relationship between the training and the generated image quality. That's why the hyperparameter optimization was performed by observing both the generated images and the loss curves together.

Possible future work may include the modification of the algorithm to generate high-resolution images. The structure of the proposed algorithm resizes the images to 512x512 pixels. With model modification, the generated image quality may increase to 2048x2048 pixels. Another future research direction can be the consideration of an additional class of tissues, i.e., bones or foreign objects such as metal fixations in the wound. This will enhance and increase the use cases of this model. The next iteration of this model may identify wound etiology, such as diabetic, lymphovascular, pressure injury, and surgical. Identification of wound type will enhance wound management further by determining the right wound care plan.

It is expected that this study will help caregivers in deciding the wound treatment plan by understanding the wound tissue classification visually, as well as assist researchers in providing an insight into the wound border segmentation and tissue classification through advanced ML methods.

4 AI-based Synthetic Wound Image Generation

Medical image synthesis is a diverse field that has great potential to assist clinician training. In this chapter, we propose a synthetic wound image generation model based on GAN architecture to increase the quality of clinical training. We also discuss and examine the use of this model for patient privacy in Section 4.1. In Section 4.2, details about the synthetic wound generation model are presented. One of our primary contributions in this area of research is a novel synthetic wound generation using DL methods. Section 4.4 shares the results of the experiments done with chronic wound datasets of various sizes taken from real hospital environments. Hyperparameters such as epoch count and dataset size for training tasks are presented to find optimum training conditions as well. Section 4.5 provides related discussions and Section 4.6 concludes the chapter.

4.1 Introduction

Computer-human interaction has an important impact on facilitating knowledge generation, dissemination, access, and utilization [116]. The study of this interaction has evolved in broader terms into the active research domain of ML or AI over the past 60 years [117].

A key element of successful DL depends on the availability of massive amounts of data [118]. DL applications in healthcare are found to be lagging, in large part, due to the high cost of accurate and tagged data collecting. The total cost of data management in healthcare could rise to millions of dollars [119], which affects medical training as well. Medical students' education costs are high, with many instances of not enough data to showcase the pathologies, wounds, and diseases. In addition, class imbalances hinder the performance of AI models, such as decision trees, neural networks, and support vector machines [120]. Limited generalization ability and overfitting problems are too common without reliable training datasets utilizing neural

networks with millions of parameters [121]. One could say that data augmentation techniques could be used to prevent overfitting and improve imbalanced class problems [122]. However, augmented data still resembles the original image intuitively [123]. This is why the quantity of training data is crucial to train and validate a model utilizing ML techniques [124]. Determining optimal hyperparameters is critical to AI models. However, the process is largely considered black art [80]. Training dataset and the number of epochs play an essential role while building new types of AI models, i.e., data-hungry and DL models. The number of epochs to train a model could be determined through loss functions in traditional ML models, whereas a GAN algorithm provides complicated loss functions. Hence, finding the optimum number of epochs is also a critical parameter that requires labor, time, and expertise in the healthcare field [26, 27]. In the scope of this study, the performance of the synthetic wound generation was examined extensively regarding the number of training epochs and the training dataset size. Six (6) different wound datasets were created to measure the effect of dataset size, and the model was trained for 2000 epochs. Results of every five epochs till the 2000th epoch were compared to analyze the effect of epoch count in detail.

In the literature, there are limited studies using the GAN algorithm in the healthcare domain. A conditional Wasserstein GAN framework was introduced for Electroencephalography (EEG) data augmentation to improve emotion recognition [125]. Magnetic Resonance (MR) image generation using Wasserstein GAN was presented for data augmentation, and physician training [126]. In the study by Antoniou et al. [127], a data augmentation GAN model was developed to generate broader data augmentation techniques such as mirroring and rotation. Synthetic data was generated through GAN models to augment the dataset with synthetic images to maximize the performance of the classifier [128] and to improve the classification task on blood cells [129]. Authors in [130] propose balancing GAN (BAGAN) to balance class dis-

tribution while training the GAN with all available images. It generates synthetic underrepresented class images utilizing conditioning in the latent space. A study by Wang et al. introduces the generation of belittled class plankton images while training the GAN with belittled class images and classifying plankton types using shared Conv layers with a discriminator network [131]. The authors in [132] propose a training scheme that implements classical data augmentation techniques to enlarge the Computed Tomography (CT) images of the liver and then enlarge the dataset a second time through synthetic image generation with GAN using classically augmented data for training. The study in [133] investigates different GAN architectures, such as SRGAN and DualGAN [134], to generate CT images from real Magnetic Resonance Imaging (MRI). The authors in [135] propose an autoencoder combined GAN to synthesize jellyfish using a small dataset compared to other GANs. The study in [136] investigates the generation of synthetic fundus images of Age-related Macular Degeneration (AMD), which are indistinguishable from real ones using Progressively grown GANs (ProGANs) [137]. In this chapter, a GAN architecture is utilized to create a novel model that generates synthetic wound images, which was not done previously. This can help the training of medical students and clinicians on wound type and wound healing prediction more accurately. This study also presents a criterion for the optimum number of training images and epochs for generating wound images using GAN algorithms.

The objectives of this study are:

(i) Development of a model using the GAN algorithm in a conditional setting to generate synthetic wound images.

(ii) Use of MSE metric to compare similarity between generated and actual wound images.

(iii) Validate and evaluate model performance visually.

(iv) Develop a hyperparameter selection guideline that can be utilized by healthcare researchers while training medical AI applications.

4.2 Methodology

GAN has the ability to mimic any data distribution with the help of adversarial generator and discriminator networks. The minimax optimization between the adversarial generator (G) and discriminator networks (D) is at the heart of the GAN architecture. Both networks are trained simultaneously with real data to learn the data distribution evenly. The generator network gets uniform or Gaussian noise (z) as an input and produces fake images (y), which are fed to the discriminator network so as to be classified as real or fake, D: $y \rightarrow [0,1]$. The generator and the discriminator networks are trained once with the training set. Following this, the generator network is used to generate fake images that are not differentiable from the training set. Figure 4.1 shows the basic structure of the GAN model.
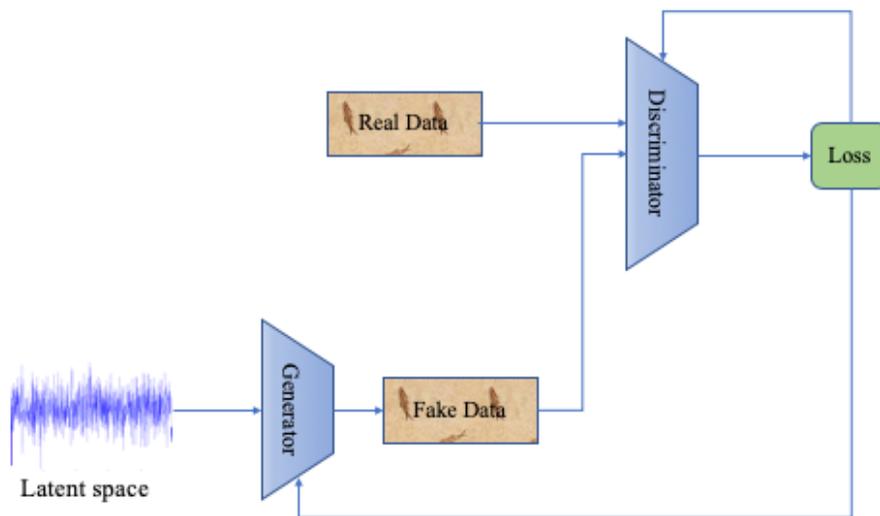
Figure 4.1: The basic structure of the GAN model.

There are various versions of the GAN algorithm for different applications. For our study, Vanilla GAN, DCGAN [138], CycleGAN, and cGAN-based versions were

examined. The Vanilla GAN and DCGAN-based models for synthetic wound generation were not suitable for our study as generated wound images were not compatible with the tissue segmentations. On the other hand, the CycleGAN-based model was promising due to its ability for unpaired image-to-image translation. The initial trials with the CycleGAN model revealed that the generated synthetic wound images experience a mode collapse problem, a well-known complication in the GAN field, which causes the synthesized wound images to look similar despite different tissue segmentation inputs [105]. As these models failed to generate proper wound images coherent to the given wound tissue segmentation images, these different GAN approaches, namely Vanilla, DCGAN, and CycleGAN-based models, are discarded from this study due to mentioned issues above. A cGAN-based architecture is used for conditioning the output (generated image), for specific data distributions, with an input image (x) and the addition of a noise (z) factor. In order to condition the output of the generator, the input is given to the generator. The input image is the segmented image ground truth that is gathered from eKare LLC. The discriminator is also fed with the segmented input image and the real data to distinguish whether the generator performed meaningful image synthesis or not.

The generator network learns a mapping from the input domain to the output domain, G: x, z→ y. The discriminator network learns a loss function to train this mapping [111] and tries to differentiate the fake image (y) from the real one, D: x, y→ [0,1]. This architecture is used to generate outputs that are similar to the input. This architecture gets paired and aligned images as an input to generate look-alike images. Both networks, the discriminator and generator, observe the input. Figure 4.2 shows the general architecture of the model that is fed with noise and input data. Loss coming from discriminator is fed to generator.

The discriminator network has two losses, namely D_{real} and D_{fake}, which indicate the ability of the discriminator network to differentiate the real and the fake images.

58
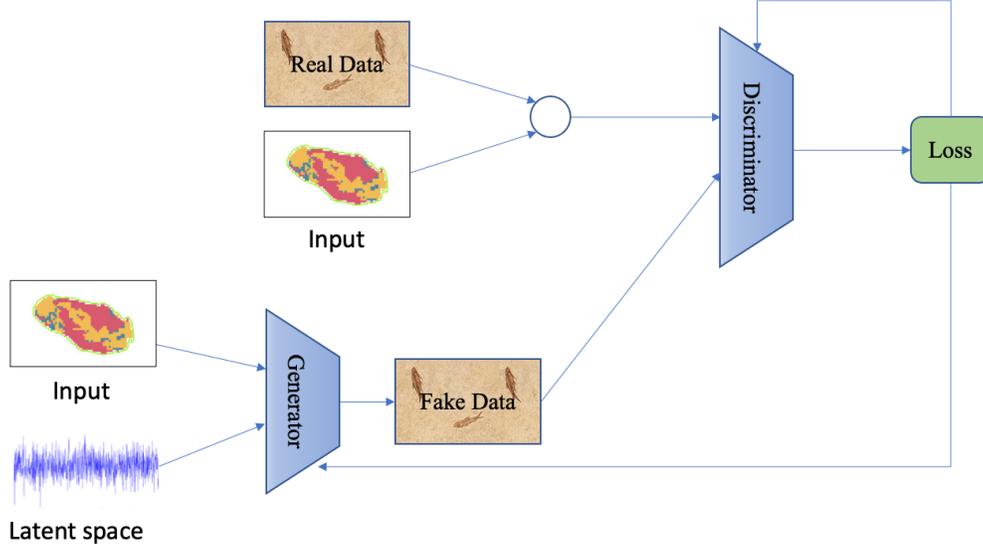
Figure 4.2: The general framework of the model architecture.

Another loss used for the generator loss is G_{L1}, which compares the generated fake image and the real image to generate more plausible fake images. Adversarial loss ($\mathcal{L}_{Adv}$) coming from the discriminator network together with L1 loss ($\mathcal{L}_{L1}$) of the generator network (G_{L1}) used for generator training causes complicated objective functions for the generator network.

$$\mathcal{L}_{Adv}(G, D) = \mathbb{E}_{x,y}[logD(x,y)] + \mathbb{E}_{x,z}[log(1 - D(x, G(x, z)))] \tag{4.1}$$

This is why the process of training a GAN architecture is hard to interpret. L1 loss behaves as a regularizing term with a hyperparameter, lambda (λ), which is chosen as 100. The final objective function is an analogy of a minimax game where the generator tries to minimize, and the discriminator maximizes an adversarial objective. This relationship is depicted as:

$$G^* = argmin_G max_D \mathcal{L}_{Adv}(G, D) + \lambda\mathcal{L}_{L1}(G) \tag{4.2}$$

The hyperparameter problem was examined through various experiments. To measure the effect of dataset size, six (6) training datasets were created. For each of these datasets, a model was trained for 2000 epochs. The performance of the model has been measured every five (5) epochs through the MSE metric for extensive analyses of the hyperparameter selection. The significance of the hyperparameter is examined via visual inspection as well.

4.3 Data Collection, Pre-processing, Environment, and Validation

This section discusses data collection, data pre-processing, the simulation environment, and the validation methods.

4.3.1 Data Collection

Chronic wound images were provided by eKare, Inc (Fairfax, VA). Chronic wounds of various types, i.e., diabetic, burn, lymphovascular, pressure injury, and surgical, are anonymized and semi-automatically labeled for training and testing. The diversity of chronic wound types increases the applicability of the model. This dataset includes around 4100 wound images ranging in size from 1224x1224 to 2160x2160 pixel dimensions (depending on the camera used to acquire measurement). Wound images were acquired under standard ambient room lighting by clinical users with commercially available 3D wound cameras [139].

4.3.2 Pre-processing

Chronic wound images are rescaled to 512x512, with the original wound and its labeled pair concatenated to form a 512x1024 pixel image. Before concatenation, the background of the labeled image is cleaned and made white. One hundred wound images out of the 4100 are chosen for the test dataset. The training dataset constitutes the rest of the dataset. To study the effectiveness of the dataset size, training datasets

with sample sizes of 500, 1000, 2000, and 4000 were created. Normalization and formatting were done to ensure high performance during training.

4.3.3 Environment

The proposed synthetic wound image generation using the GAN (WG2AN) model is implemented using the Pytorch DL framework with Python version 3.7. The model runs on an Intel® Core ™ i7 -8700X CPU @3.20 GHz with 32 GB memory and NVIDIA GeForce GTX 1080 GPU with 8 GB dedicated and 16 GB shared memory. The model training for 2000 epochs took around 4 hours, 6 hours, 10 hours, 16 hours, 29 hours, and 58 hours for 100, 250, 500, 1000, 2000, and 4000 images, respectively. The batch size is 64 for maximum GPU performance.

The Adam optimizer with a learning rate of 0.0002 is utilized for the first half of the training, while the rest is done with a linearly decaying learning rate to zero. The Adam optimizer is used as it converges faster, and also lower loss values are generated for the generator network in comparison to other optimizers in GAN applications [140]. The learning rate setting is chosen in line with the previous works done in image synthesis using GAN algorithms [104].

4.3.4 Validation

Validation of the proposed method is done with the MSE metric, which is a pixel-wise error measurement. Three color channels (RGB) are considered while calculating the MSE scores where Y_R, Y_G, Y_B are used for the color channels of real images, and Y'_R, Y'_G, Y'_B are used for the color channels of generated images. Additionally, n denotes the number of pixels.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} [(Y_R - Y'_R)^2 + (Y_G - Y'_G)^2 + (Y_B - Y'_B)^2] \qquad (4.3)$$

The MSE metric is applied to each pixel in the fake and real images. Each pixel

has three channels, namely red, green, and blue. Each corresponding pixel in fake and real images is compared, and the MSE score is calculated. The MSE metric provides better quality scores [141]. Converging and meaningful results for the synthetic wound generation task are obtained using the MSE metric. The synthetic wound results are expected to be similar to actual wound images, where the MSE metric could evaluate such similarity properly.

4.4 Experiments

Model and loss function outputs, the effect of dataset size, and epoch count on synthetic wound generation are discussed with MSE score and visual inspection in this section. Additionally, the optimum training conditions are determined with hyperparameter tuning.

4.4.1 Model Output

The proposed model gets wound tissue segmentation and produces synthetic wound images. Wound segmentation is done with respect to wound tissue types, namely necrotic (blue color), granulation (red color), and slough (yellow color). This classification is of foremost importance to the caregiver while assessing a wound and deciding on a treatment plan. A successful result from the WG2AN model is depicted in Figure 4.3, which indicates proper synthetic wound generation by training with 4000 images and 2000 epochs. Actual wound, Figure 4.3b, and synthetic wound image, Figure 4.3c, were presented without background for a better comparison with the segmentation input, Figure 4.3a.

In addition to the original wound image, a combination of the synthetic wound and original wound background is shown in Figure 4.4a - 4.4b. The generated image

(a) Segmentation   (b) Actual wound   (c) Synthetic wound
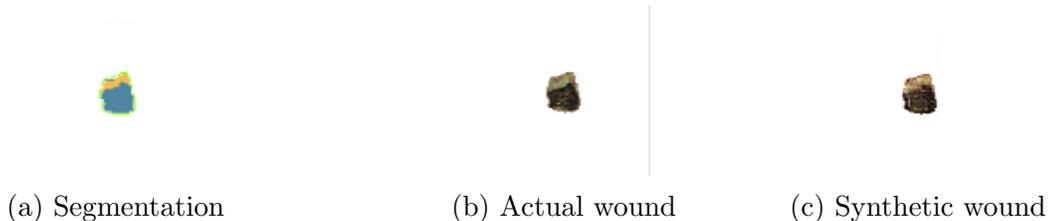
Figure 4.3: WG2AN model results trained with 4000 images and at 2000 epochs.

(a) Original wound   (b) Combined wound

Figure 4.4: WG2AN model results applied on a real limb that is trained with 4000 images and at 2000 epochs.

has characteristics of a real wound.

4.4.2   Effect of Number of Images and Epochs on Model Loss

The chronic wound dataset is pre-processed and arranged as six (6) subgroups, which have 100, 250, 500, 1000, 2000, and 4000 images. The proposed model with WG2AN architecture to synthesize wound images is trained for 2000 epochs. The effects of dataset size and epoch count are examined to find the optimum training dataset.

The G_{L1} loss curves of the model with different training dataset sizes are depicted in Figure 4.5. The adversarial and the discriminator network losses, i.e., D_{fake} and D_{real}, do not provide useful information, and they are not included in Figure 4.5. The proposed model could not settle down in the first 75th epochs as the losses increase, which is a common problem because of the randomized weights in the neural networks. The loss curves also share a similar fluctuating but decreasing pattern. The general zigzag behavior of the loss curves is a result of the alternating wound image samples

during training. The proposed model moves along the different training samples without reaching an equilibrium, which causes this oscillation [114].

The loss curve of the model with 100-images has the lowest loss. The model with 250-images has the second-lowest loss values. Other curves follow this characteristic as well. It can be deduced that models with fewer training images produce lower loss values. The loss graphs are compared to evaluate the learning ability of the model in Figure 4.5. The model with 100-images goes below the loss of 25 at the 180th epoch and stabilizes under the loss of 15 at the 600th epoch. 250-image model's loss decreases to 20 at the 200th epoch and becomes stable under 15 at the 700th epoch. The models with 100 and 250 images move closer curves as their training dataset size difference is small, i.e., 150.

The loss of the model with 500 images reduces to 25 at the 300th epoch and balances under 15 at the 950th epoch. The 1000-image model drops to 25 at the 350th epoch and flattens under 15 at the 1100th epoch. The 2000-image model decreases to the 275th epoch and stays stable under 15 at the 950th epoch. The model with 4000 images goes down to 25 at the 250th epoch and balanced around 15 at the 1000th epoch.

The 4000-image model has been trained for 4000 epochs; however, it does not stabilize under 15 and oscillates. The drop rate of the G_{L1} loss indicates that our proposed model learns the data distribution faster with a smaller dataset. The models with a larger training set have more complicated data distributions that take more epochs to mimic.

As stated in 4.2, the generator and the discriminator networks compete over a min-max game, which results in a nonconverging problem [111]. Although G_{L1} loss has a meaningful curve that could be used to determine if the model is learning the training data distribution, it lacks the contribution of the adversarial loss. That's why an additional evaluation metric is needed, i.e., MSE. MSE score provides a better
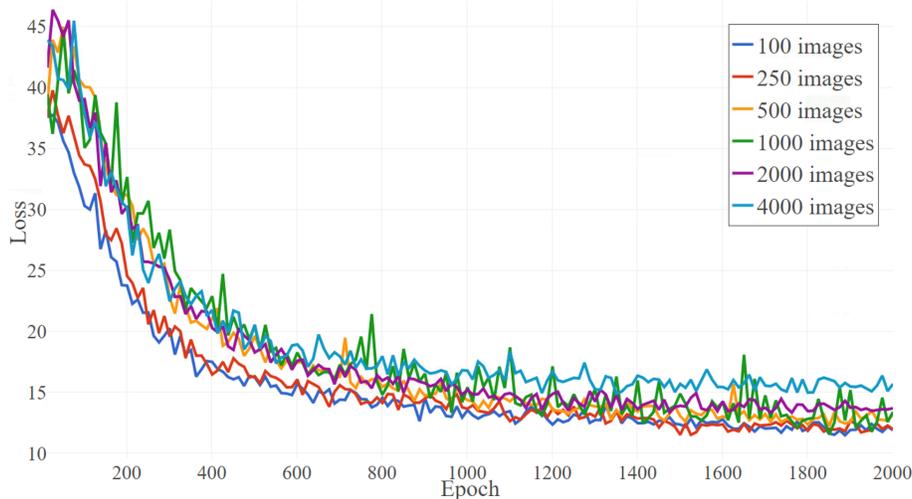
Figure 4.5: Model's loss graphs using a different number of images.

understanding of the generation of synthetic wound images by the proposed WG2AN model, whether they look like real wounds or not.

4.4.3 Effect of Number of Images and Epochs on Synthetic Wound Generation

The effect of training parameters, i.e., dataset size and epoch count, are discussed in this section. The models with different training dataset sizes (100, 250, 500, 1000, 2000, and 4000) are evaluated visually and using the MSE score. Figure 4.6 indicates the actual wound and the segmented wound, which is the input of the model. The outputs of four different models are compared in Figures 4.7 - 4.12 with respect to epoch counts (200, 500, 1000, and 2000).

The model with the 4000-image dataset is trained for 4000 epochs, but the effect is negligible. That's why we cut training at the 2000th epoch for each model. The generated images before the 75th epoch lack the details and texture of a wound in general and suffer from the checkerboard effect significantly. The up-sampling layer in the generation pipeline of a GAN model, which produces high-resolution images from low-resolution ones, causes checkerboard artifacts [142]. The checkerboard pattern emerges when deconvolution has uneven overlap [143].

65

(a) Original wound   (b) Segmentation

Figure 4.6: (a) Original wound and (b) wound segmentation.

The models with smaller training datasets tend to be biased towards a particular data distribution, which is the case for 250 and 500-image models, which produced darker wounds. After further training of the same model, dataset limitation recovered at the 500-image model at 2000th epochs, as seen in Figure 4.9d. Training the models longer provides detailed texture and balance in the distribution of the tissue (also visible in Figure 4.9d).

The 100-image model generated insufficient wounds for the 500 epochs and the lower epoch counts. However, increasing the training to 1000 epochs improved the generated images significantly. Results for the 100-image model in Figure 4.7c - 4.7d have close to real composition but generated samples in Figure 4.7a - 4.7b are far from a real wound nature.

The 250-image model generated primitive wounds at 200 epochs in Figure 4.8a. Generated wounds have more balanced wound tissue characteristics with some limitations, such as overall darker colors at 500, 1000, and 2000 epochs in Figure 4.8b - 4.8d. The 200 epochs of training is not enough for a life-like wound image generation, as seen in Figure 4.8a.

The model with the 500 images produced more life-like wounds at 200 and more epochs of training, shown in Figure 4.9c - 4.9d. The result of 500 images at 200 epochs in Figure 4.9a does not represent a real wound.

The model with 1000 images exhibits better performance than previous models.

(a) 200 epochs

(b) 500 epochs

(c) 1000 epochs

(d) 2000 epochs

Figure 4.7: WG2AN output using 100 images for different epochs.

(a) 200 epochs

(b) 500 epochs

(c) 1000 epochs

(d) 2000 epochs

Figure 4.8: WG2AN output using 250 images for different epochs.

(a) 200 epochs

(b) 500 epochs

(c) 1000 epochs
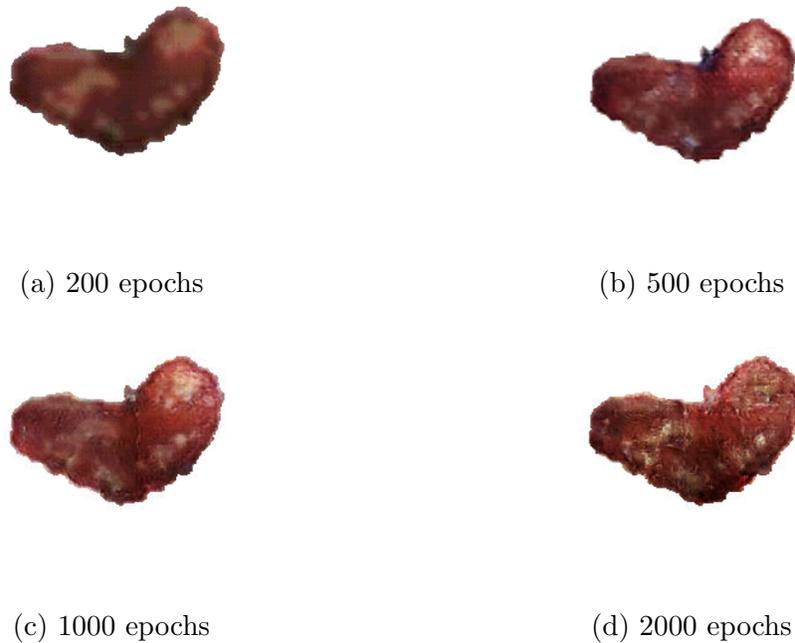
(d) 2000 epochs

Figure 4.9: WG2AN output using 500 images for different epochs.

At 200 epochs of training in Figure 4.10a, it shows a reasonable output where previous models just produce a crude output from fed segmentation data. The results of 500, 1000, and 2000 epochs of training produce better-wound images, as seen in Figure 4.10b - 4.10d.

The 2000 image model produced life-like images at 500, 1000, and 2000 epochs in Figure 4.11b - 4.11d. The 200 epochs of training in Figure 4.11a also form a wound with close to wound tissue characteristics.

The model with 4000 images generates life-like wounds at 200 epochs to 2000 epochs in Figure 4.12a - 6.3d. Every generated sample carries characteristics of the wound and well-balanced tissue distribution.

By varying the number of epochs and dataset size, generated image quality has improved significantly. It can be inferred that the lack of data could be substituted by further training the model.

(a) 200 epochs

(b) 500 epochs

(c) 1000 epochs

(d) 2000 epochs

Figure 4.10: WG2AN output using 1000 images for different epochs.

(a) 200 epochs

(b) 500 epochs

(c) 1000 epochs

(d) 2000 epochs

Figure 4.11: WG2AN output using 2000 images for different epochs.

(a) 200 epochs

(b) 500 epochs

(c) 1000 epochs

(d) 2000 epochs
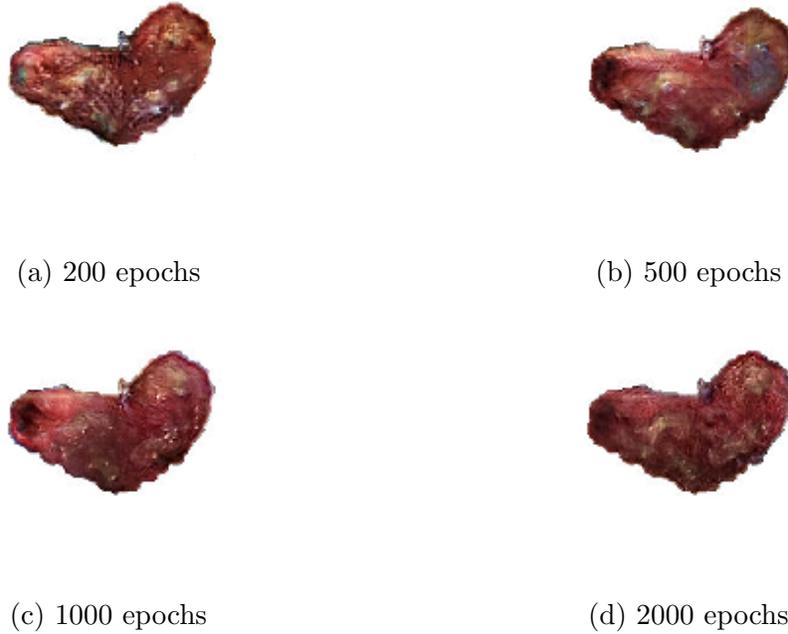
Figure 4.12: WG2AN output using 4000 images for different epochs.

4.4.4 Evaluation of the Model with MSE Metric

The model performance is evaluated by the MSE metric, as summarized in Table 4.1, which indicates the MSE scores of the models with different training images and epochs. A comparison of the MSE scores can be examined in Figure 4.13. MSE scores are calculated on individual images, after which the median average is calculated. A comparison of the MSE scores provides guidance for the generated image quality. A lower MSE score means better-generated image quality. The overall trend of the MSE scores in Table 4.1 is decreasing with expanding dataset size and further training of the models. The first 75 -100 epochs of training are necessary for the model to settle. The models' MSE scores for 100 epochs are also included in Table 4.1. The MSE score of the 100-image model was the highest for every epoch. It starts with the MSE score of 745, decreases to 732, increases to 747, and then keeps increasing with additional epochs of training.

The 250-image model has an MSE score of 727 at the 100th epoch and drops to

Table 4.1: MSE scores for the WG2AN model.

Epochs Images	100	200	500	1000	2000
100	745	732	747	759	785
250	727	711	720	721	721
500	627	707	707	701	691
1000	736	703	703	704	709
2000	729	709	709	706	706
4000	721	697	697	705	710

711, and has higher MSE scores with the increase in training. The model with 500 images does not settle at the 100th epoch, where the MSE score is lower than any other MSE score in the table. It becomes 707 and stays stable at that MSE score. There is a drop in the 2000th epoch of the 500-image model, which is a result of an overfitting problem.
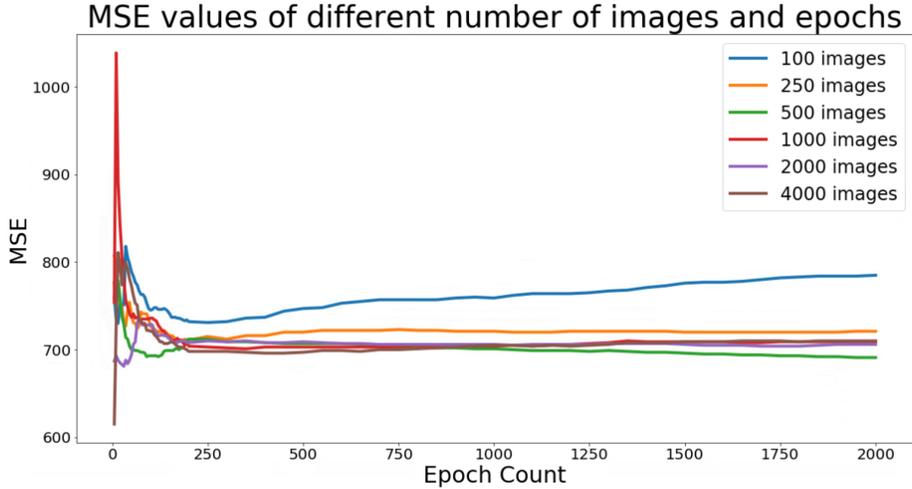
Figure 4.13: MSE scores (Median) of the models with the different training dataset.

The 1000-image model has an MSE score of 736 at the 100th epoch and reduces to 703 at the 200th epoch. It stays stable till the 1000th epoch. It rises slightly to 704 at the 1000th epoch and 709 at the 2000th epoch. The smaller increase means that there is a deformation in the generated images. This can be observed by comparing Figure 4.10c and 4.10d. The slough area gets smaller in the former one, which causes

71

a slight MSE increase. Models with a higher number of training images, i.e., 2000 and 4000, follow a similar pattern to the 1000-image model.

MSE score of the fifth epoch results is negligible as generated images are not good enough to mimic wound texture and detail. Training the model with a smaller dataset for a more extended number of epochs results in overfitting, i.e., an increase in MSE score. This can be seen in Figure 4.13, especially after the 500th and 200th epochs. It is confirmed that increasing dataset size and epoch count until the 500th epoch gives a better MSE score. For the first 200 epochs of training, the decrease in MSE score is so significant that later increase of training has minimal effect on the score. Therefore, 200-epoch is the optimum epoch number for training the proposed method. The amount of dataset size to generate plausible images is 1000 because the 1000-image model generated most life-like wounds, as seen in Figure 4.10a. That's why we conclude that the 1000 images and 200 epochs of training are the optimum training hyperparameters for plausible wound image generation.

4.5 Discussion

The proposed method generates synthetic wound images that have an input of segmented tissue outline, shown in Figure 4.2. The output is compared with the ground truth wound images. Two views of a successful result from the model are presented in Figure 4.3 and Figure 4.4, indicating optimal synthetic wound generation from the segmentation of a wound by training with 4000 images and 2000 epochs. The generated wound image is combined with the rest of the limb. The generated wound tissue has a life-like structure, proper color, and detailed texture. The colors of the wound tissues are well-matched and conformable. Figure 4.5 shows the loss graphs of the L1 loss of models with the different training datasets. The loss curves have similar characteristics. The model with fewer training images learns the data distribution faster and stabilizes loss earlier than the model with more training images.

The generated image comparison with respect to epoch count and training dataset size is shown in Figures 4.7 - 4.12. An increase in the training dataset size helps to generate more plausible wound images. Blurry images become more realistic with an increase in dataset size given the same conditions, i.e., epoch count. These results show that a lack of training can be overcome by increasing the training dataset size. The effect of epoch count is also examined in Figures 4.7 - 4.12, where better wound images are produced with increased training. Figure 4.9a - 4.9d prove that the impact of epoch count overcomes the deficiency of the training dataset. Further training yields better-wound images.

These extensive investigations suggest that the optimal performance for synthetic wound generation could be achieved using 1000 images and 200 epochs of training.

On the study results, the following observations regarding the application of the WG2AN model could be made.

Observation 1: The proposed model can perform synthetic wound generation from provided wound segmentation.

Observation 2: WG2AN has a high potential of producing close to real synthetic images.

Observation 3: The quality of the generated images is in line with the image count. 1,000 image count is the threshold for a valid generated image as the results of our study.

Observation 4: The epoch count has a significant impact on the generated image quality. Yet, after surpassing a 200-epoch threshold, the model reaches its convergence, and additional training has a marginal effect on the quality of the generated image.

Observation 5: The WG2AN model can generate detailed tissue texture.

Observation 6: Lack of training can be solved by increasing the training dataset.

Observation 7: Scarcity of training images can be mitigated to some extent by

further training the model.

Observation 8: Generated wound images can be combined with any part of the body to demonstrate the wound characteristics at that body location.

4.6   Conclusion

This chapter compiles a synthetic wound image generation using the proposed WG2AN architecture. Synthetic wound generation using GAN is implemented for the first time in the literature. The wounds are segmented with respect to tissue types using semi-automatic ML techniques. The proposed model is then fed with segmented wound images to generate synthetic wounds.

With respect to patient privacy and the lack of enough datasets for wound images in healthcare, the generation of synthetic anonymous wound images could enable further studies in AI, which could improve the adoption of AI in clinician training. L1 loss is also examined in order to understand the impact of training dataset size and epoch count, where the loss curve does not reveal much information other than models' learning behavior. The generated images are examined and compared visually to evaluate their resemblance to a real wound. It can be concluded that the hardship of finding adequate training images in healthcare can be mitigated by additional model training. The effect of dataset size is also evaluated visually for further analyses. An increase in the training dataset brings in more life-like wound images. The results of different dataset sizes and epoch counts are evaluated through the MSE metric to compare the generated image. The generated images are expected to be very similar to actual wound images. This is why the MSE metric gives acceptable guidelines when comparing actual and synthetic images. The proposed model confirms that the 1000-image model with 200 epochs of training yields optimum results. Increasing these parameters could provide better synthetic images.

As future work, the segmentation of the wound could be done in more detail. In

74

addition, the use of the proposed model in an education package could bolster the performance and practice of medical training.

Balancing underrepresented classes with synthetic image generation could help with the adoption of AI in the healthcare industry, where sourcing context-specific data is expensive. It is expected that this study can provide a handy clinician tool for generating and interacting with live wound models.

5   Chronic Wound Care and Digital Twin

Throughout this dissertation, we have proposed many AI models to enhance chronic wound management with the idea of generalized care for any wound. Developed models form a framework in the wound healing continuum that could successfully provide end-to-end care using wound images. What we proposed and developed in this study's scope aligns with the digital twin concept to wound care. That is why we bring up the digital twin concept to wound care.

There have been many attempts to provide optimum healthcare services with advanced computing and communication technologies in the last decades. The digital twin is an emerging technology that promises personalized and predictive healthcare. The digital twin concept has been broadly accepted in many fields, such as manufacturing, construction, smart building, smart grids, smart cities, and more. However, its application in healthcare is still in its infancy. The digital twin in chronic wound management will shed light on providing optimal treatment pathways and better interpretation of the treatment.

This chapter will thoroughly examine the potential of the digital twin in chronic wound management. Section 5.1 lays the scene for digital twin and chronic wound management. Section 5.2 visits the potential and use of digital twins in healthcare. One of our primary contributions in this area is to research the applicability of the digital twin concept to chronic wound care management. Section 5.3 reviews chronic wound management and its components. Section 5.4 assesses the enabling technologies that lead to real-time digital health twin. Section 5.5 and 5.6 put forward the digital twin concept in chronic wound management. Section 5.7 concludes the chapter.

5.1   Introduction

Wound management has been studied for many years since its healing requires

dedicated care [12]. This chapter will extend these studies to the digital twin concept, where a digital clone of the wound would be built to ease the plan and track the wounds. Besides diagnosis and management, prognosis and decision support systems for chronic wounds could be developed using AI and the recently adopted digital twin concept in healthcare. Digital twins are essentially virtual replications of physical objects and processes. They use the Internet of Things, AI, and complex data in models that create insights and support (real-time) decision-making. It is perhaps healthcare that holds tremendous potential for digital twins.

Healing chronic wounds take weeks or months, requiring periodic examinations and continuing care. Digital twin application in wound care is feasible as longer and slower healing pace of hard-to-heal wounds will provide flexibility to handle surfacing health conditions. This chapter examines chronic wounds and the digital twin concept to develop a digital twin framework for chronic wound management. Contributions of this study two folds:

(i) Investigation of chronic wounds and digital twin technologies for healthcare.

(ii) Development of digital twin framework for chronic wound management and human digital twins.

5.2   Digital Twin in Healthcare

The digital twin concept in healthcare is still in its infancy due to privacy concerns and the vulnerable nature of healthcare. The authors study the security aspect of the digital twin usage in healthcare in [144]. Blockchain-based secure digital twin framework is proposed, and a case study covering the recent COVID-19 pandemic is discussed to provide security of the shared data. Another study that emphasizes vulnerability detection for cyber resilience in healthcare digital twin is explored by the authors in [145]. One of the prior works proposes cloud-based digital twin healthcare (CloudDTH) for elderly patients [146]. The cloud-based healthcare service platform

is proposed for real-time monitoring, crisis warning, medication reminder, and disease diagnosis. One of the case studies that utilize digital twin is proposed for an electrocardiogram classification framework to diagnose heart disease and detect heart problems using ML techniques [147]. An Ischemic Heart Disease (IHD) recognition digital twin architecture is also proposed by the authors in [148]. The presented digital twin model works on the edge and utilizes CNN to classify myocardial conditions with an accuracy of 85.77%. Authors in [149] integrate digital twin with multi-agent systems as mirror worlds in a case study. This case study digitalizes to support the process of trauma management. An augmented digital twin is proposed to lay the foundations of the complete life cycle digital twin of a human being [150]. In addition to these, the potential of the digital twin technology in health is reviewed by Erol et al. [151].

These studies indicate that the utilization of the digital twin concept has great potential. More case studies should be done to unleash this potential since the needed technology is available. New methods and technologies should also be adopted to overcome privacy and security issues in the digital twin.

5.3 Review of Chronic Wound Management

Chronic wound management requires special care in order to heal [12]. Treatments and the lengthy healing course are tracked with long-established visual methods. In this section, the tissue types, treatment methods, and evaluation metrics are examined for the construction of digital twins.

5.3.1 Types of Wounds and Tissues

There have been various diseases and incidents that could harm and break the integrity of the skin. Wounds could be classified into two groups regarding their healing process, i.e., acute and chronic. Acute wounds follow an orderly path during the healing and tend to heal in a short period of time. In comparison, chronic wounds

distinguish by their complicated and challenging healing process. These kinds of wounds are also called hard-to-heal wounds. Along with hardship in healing, some could pose severe risks like loss of limbs and mortality.

Some of the hard-to-heal wounds are pressure injury, burn, venous, diabetic, and surgical. Pressure injury wounds are caused by stress or force on the skin's surface, resulting from limited mobility. Prolonged inpatient stay and lack of movement at hospitals also pose a greater risk of pressure ulcers [152, 153]. Diabetes is one of the largest epidemics in this century; around half a billion people suffer from it [154]. One of the complications of diabetes is nerve damage, and neuropathy [155]. Studies have shown that patients with diabetes have around 20% risk of developing an ulcer that costs more than US$10 billion in the US only [156, 157]. Another primary wound type is vascular or arterial wounds, caused by poor blood flow below the knee, affecting both legs. A similar wound type, venous ulcers, are developed as a result of damaged veins from high blood pressure [158]. These wounds share similar characteristics, but different approaches are required to cure them. Burn wounds are also prevalent and caused by heat which damages the tissues and underlying structure of the body [159]. These are some of the notable wound types.

The wound tissues are tracked for centuries as visual inspection of the wounds plays utmost importance in determining the status of the wounds. There is a Red-yellow-black (RYB) tissue classification methodology that was introduced by Cuzzell in 1988 [160]. This tissue classification provided a more straightforward and universally accepted system that red areas can be granulation, yellow areas can be slough, and black areas contain eschar (necrotic) tissue [161]. Eschar (Black) and slough (Yellow) are tissues that are not ready to heal. These tissues will be removed to accumulate a swift healing process.

5.3.2 Continuum of Wound Healing

After the incident, the injury breaks the blood circulation and causes bleeding.

Granulation tissue (Red tissue) plays an essential role in healing. There are four phases for the healing of wounds. The hemostasis phase is the first step of healing. In this phase, the wound is sealed with various molecular binding agents. The inflammatory phase will cause swelling due to the production of a transudate. With this phase, infections are prevented. The proliferative phase will form new tissues and blood vessels for the circulation of enough oxygen and nutrients. The remodeling or maturation phase will provide the required materials to heal a wound completely. These phases follow a planned healing procedure that chronic wounds do not follow any order.

5.3.3 Treatment Methods and Evaluation

The healing endeavor has been discussed and practiced for many centuries since wounds are easily recognized visually, and their track depends on visual assessment in the first place. Cleansing the wound and removal of the dead tissue, also called debridement, is critical in some cases. After the application of debridement, readily healed tissues have emerged, and healing is sped up with the living tissue. Ultrasound, laser surgery, or irrigation are some debridement methods used in current wound management. Another essential method to treat and heal wounds is dressing, which could be dry or wet [162].

Each wound type has its own characteristics and requires distinct treatment methods. Pressure injury wounds are treated using various dressing hydrogels (water-based gels), hydrocolloid dressing, and foam dressings. Diabetic ulcers are treated with a silver ion foam dressing. Besides this, this type of wound care requires a care team to manage the wound appropriately. Improving blood circulation is one of the essential methods to cure arterial or vascular wounds. Compression stockings could be used to prevent blood pressure to cure venous ulcers. Burn wounds could benefit from dressings as well. Severe burns could require skin grafting from other healthy parts of the body.

Various ointments and dressings could be used to fight infection and oral medicines. The techniques mentioned above are some of the traditional treatment methods that can cure wounds. There are also dressing-free therapies and skin grafting, and 3D bioprinting techniques. Emerging therapy solutions are investigated and proposed continuously to improve healing procedures and patients' quality of life.

The wound evaluation is carried out by surface area measurement, which is a simple and least expensive method. Some of the previously utilized methods are rulers, mathematical models, manual planimetry, digital planimetry, stereophotogrammetry, and digital imaging methods [163]. The latter method has been adopted as it is noninvasive and provides better results. There are also near-infrared methods such as angiography, laser speckle contrast imaging, and optical coherence tomography [164].

5.4 Enabling Technologies

5.4.1 High-Performance Computing

High-performance computing (HPC) is essential for developing the digital twin in healthcare [165]. The digital twin requires efficient data management, processing, and analysis. These tasks are demanding in terms of speed and accuracy. For instance, the digital twin requires a large amount of data to model the patient, which is impossible with a single computer. Therefore, HPC is required to manage and analyze this data. In addition, HPC will help to take advantage of the data collected from various sources, such as Electronic Health Records (EHRs), clinical trials, sensors, and imaging devices [166]. HPC can be used to develop new treatments, diagnostic tools, personalized medicine, and clinical trials in the healthcare field.

5.4.2 Internet of Things

Internet of Things (IoT) is a network of physical devices, vehicles, buildings, and other items embedded with electronics, software, sensors, actuators, and connectivity, enabling these objects to connect, collect, and exchange data.

IoT in healthcare can provide many benefits, such as improving patient care, reducing costs, and improving clinical outcomes. In addition, IoT can be used for remote patient monitoring, patient engagement, and disease management. IoT can monitor patients' vital signs and wound healing progress in chronic wound management. For instance, IoT can collect data from sensors placed on the patient's skin to monitor the wound's healing progress.

In addition, IoT can be used for remote patient monitoring. For instance, patients with chronic wounds can be monitored remotely using IoT devices. These devices can collect data from patients and send it to healthcare providers. This data can monitor the patients' health and provide them with the necessary care. That is why IoT systems are essential parts of the digital twin in healthcare.

5.4.3 Artificial Intelligence (AI)

AI is the ability of a computer system to perform tasks that require intelligence. AI can be used for various tasks, such as decision making, pattern recognition, and natural language processing. AI has been used in many fields, such as manufacturing, automotive, and healthcare. AI in healthcare can provide many benefits, such as improving patient care, reducing costs, and improving clinical outcomes.

In chronic wound management, AI can be used for various tasks, such as wound classification, segmentation, and healing prediction. AI can be used to classify wounds into different categories, such as burns, ulcers, and cuts. AI can also be used to segment wounds into different parts, such as the edges, center, and surrounding tissue. AI can also be used for wound healing prediction to predict the healing progress of wounds.

5.4.4 Edge / Fog Computing

Edge computing is a distributed computing paradigm that brings computation and data storage closer to the location needed to improve response times and save

bandwidth. Edge computing is used in many applications, such as video streaming, virtual reality, and autonomous vehicles. In healthcare, edge computing can be used for digital twin tasks, such as patient monitoring, disease management, and clinical decision support.

5.4.5 Cloud Computing

Cloud computing is a type of computing that provides computing resources over the Internet. Cloud computing can be utilized for digital twin tasks, such as storage, data analysis, and ML. Cloud computing can be used for multiple tasks in healthcare, such as patient data management, disease management, and clinical decision support. Cloud computing has many benefits, such as scalability, flexibility, and cost-effectiveness.

5.4.6 Privacy Enhanced Techniques for Electronic Health Records

A massive amount of data is generated by chronic wound patients, which is confidential and must be protected. Data privacy concerns and privacy-enhanced techniques should be adopted to protect the data used in digital twins. In particular, data privacy concerns and privacy-enhanced techniques are used to protect the data in EHRs. There are two different approaches to protecting the privacy of sensitive health records.

The first approach is to encrypt the EHRs using homomorphic encryption techniques before storing them in the cloud. The second approach uses federated learning techniques to train the ML models on the EHRs without sharing the EHRs with the central server. Homomorphic encryption is a type of encryption that allows mathematical operations to be performed on ciphertexts, which results in an encrypted result that, when decrypted, matches the result of the operations as if they had been performed on the plaintext. There are three subcategories of homomorphic encryption: partially homomorphic encryption, somewhat homomorphic encryption, and

fully homomorphic encryption.

The second approach to protecting sensitive EHR privacy is federated learning. Federated learning is an ML method that allows multiple devices to train an ML model without sharing their data or with a central server. The devices train the ML model on their local data and send the model parameters to the central server. The central server then aggregates the model parameters from all the devices and updates the global model. The devices then download the updated global model and continue training the model on their local data. This process is repeated until the global model converges. The main advantage of federated learning is that it allows the ML model to be trained on the data of multiple devices without sharing the data or with a central server.

5.4.7 Blockchain

Distributed Ledger Technology (DLT) refers to one of the emerging Industry 4.0 technologies implemented as Bitcoin in 2009. Blockchain Technology is a subcategory of DLT and a more popular terminology. Thus DLT is used interchangeably in the industry and academy. DLT is offering new opportunities for the challenging nature of the digital twins in healthcare. Collected data in digital twin applications in healthcare require utmost security to eliminate any breach or modification. Corrupted data could bear a high risk for the digital twins as it will lead to wrong decisions [167]. Many functions in the digital twin could also be realized using smart contracts. After a triggering event happens, a smart contract could initiate a default behavior [168]. Having secure, immutable, and robust characteristics, DLT enables reliable and automated digital twins. Figure 5.1 characterizes the use of Blockchain in the digital twin.

5.4.8 Communication Networks

As communications technology evolves rapidly from 4G to 5G and beyond net-

84
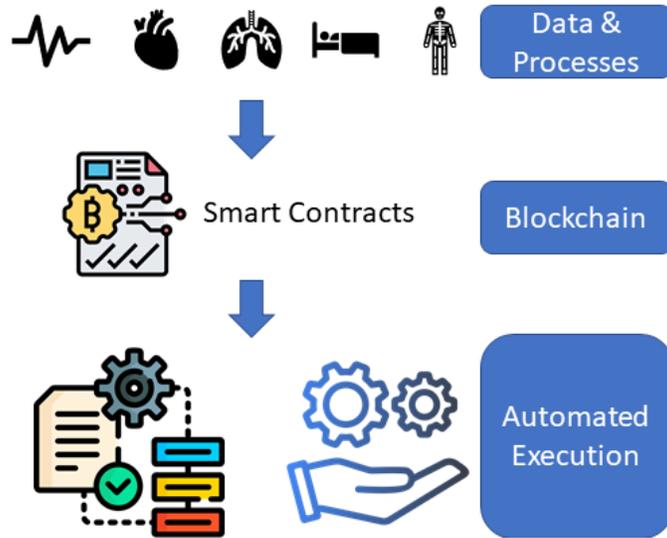
Figure 5.1: Blockchain implementation on digital twin.

works, a large number of technologies are expected to benefit from these networks in terms of high speed, latency, accuracy, and security. The complexity of the services such as the digital twin is not an incremental increase but an exponential leap. In addition, the implementation of digital twin applications often demands always-on service and near real-time solutions. These factors pose grand challenges in developing digital twin models utilizing 5G and beyond networks.

To overcome these challenges, advancement in communications is a promising and potential solution that can facilitate the Data Acquisition (DAQ) for digital twins with less complexity, less costly, and more timely. From data collection to its storage, network technologies such as Bluetooth, Wi-Fi, NextG, LoRa, 5G, and beyond will be employed for digital twins.

5.5 Chronic Wound Management System using Digital Twin

Wounds generally heal between 4 and 6 weeks and show signs of healing within this time frame. However, the chronic wound fails to heal through the typical phases of wound healing in an orderly and timely manner. Although there are many studies and efforts to develop therapeutic ways to treat chronic wounds effectively, there is

limited clinical success in chronic wound healing. One of the main reasons is the lack of effective chronic wound management systems or their limitations in terms of technology integration and adaptation [169]. Fortunately, the advanced data communication and computing technologies along with new concepts, e.g., digital twin, can overcome all these issues and provide a perfect wound management system beyond expectations. A digital twin is defined as a digital replica of anything from people and processes to systems. A proposed general system architecture is shown in Figure 5.2 for human digital twin. The proposed architecture consists of four main components, namely data collection, data management, analysis model management, and digital twin. Each component is briefly explained below.
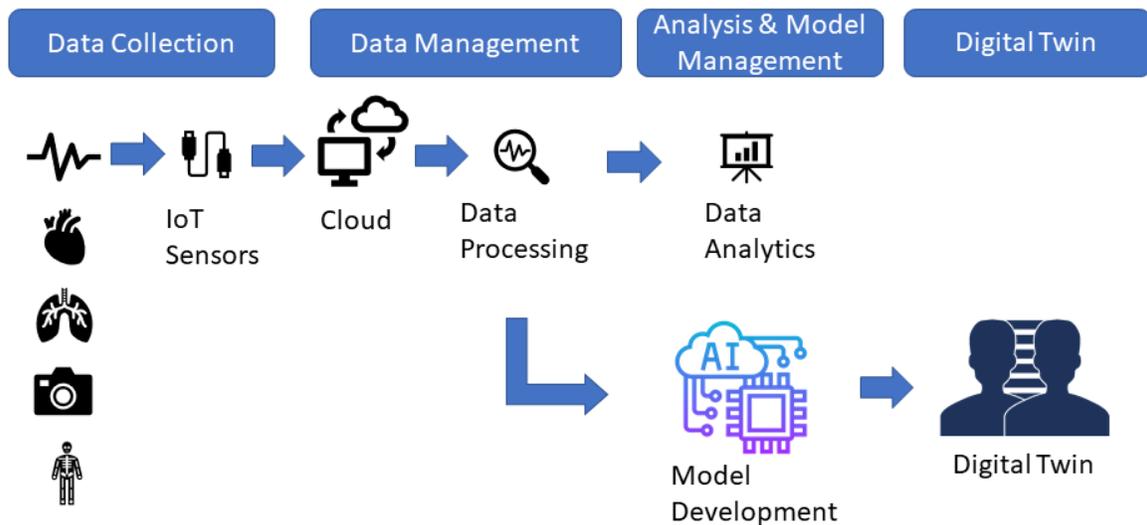
Figure 5.2: Proposed system architecture for human digital twin.

5.5.1 Data Collection

The data collection is the starting point of the proposed platform by collecting data from several resources, such as IoT based-devices, i.e., the Internet of Medical Things (IoMT), and healthcare systems. The primary IoMT sensors are pressure, temperature, blood oxygen, image, and flow sensor. This list can be extended as well. IoMT based-devices enable healthcare systems to be more interactive and con-

nect with the patients. This type of patient care leverages connected devices with IoT sensors to offer providers a continuous stream of real-time health data such as heart rate, blood pressure, and glucose monitoring. Data collected from IoMT devices can also help to find out the best treatment process for patients. IoMT-based applications are also promising in the healthcare sector. These IoMT-based applications include remote patient monitoring, glucose monitoring, heart-rate monitoring, hand hygiene monitoring, depression and mood monitoring, Parkinson's disease monitoring, connected inhalers, ingestible sensors, connected contact lenses, and robotic surgery, and many more [170]. In addition to the data collected from IoMT sensors, patient information is also very crucial in healthcare applications, especially wound management, such as age, gender, smoking, any drug, any previous condition, chronic wound type, diabetes, etc.

5.5.2 Data Management

The second component of the proposed architecture is data management using cloud sources. This component consists of two steps, namely cloud integration and data processing. Cloud-based systems provide enhanced interoperability and consolidation of the data. Without a uniform data management system, data transmission, processing, and model development tasks suffer from asynchronous and complex data flow. Cloud services provide real-time data collection and monitoring, which pave the way for the digital twin concept. Without cloud sources, the required infrastructure for data processing and model deployment will cause a burden to many health organizations [146]. Monitoring patients with chronic wounds at home is possible by using cloud services. The use of the cloud will increase the treatment outcomes by providing real-time data.

The collected data also needs processing, i.e., data processing, to produce meaningful information for robust analysis and decision and digital twins. This process includes removing or filling in the missing values, detecting outliers, and normalizing

the data.

5.5.3 Analysis and Model Management

The third component of the proposed architecture is the analysis and model management. The proposed architecture offers a prescriptive analytic, where AI and big data combine to help predict outcomes and identify what actions to take. In the proposed architecture, prescriptive analytics allows one to take a deeper look into the data and answer "what" and "why" questions for wound data, such as healing stages, wound size details, etc.

Model management is a sub-component of the proposed architecture. A model should be consistent and accurate in terms of performance metrics. Therefore, a logical, easy-to-follow policy for model management is crucial for model management. The primary purpose of model management is to provide a system for the development, training, versioning, and deployment of models. It is expected that the models to be developed will be AI-based models. The model management makes it easier to manage the model life-cycle from creation, configuration, experimentation, and tracking of the different experiments, all the way to model deployment. Under the model management, models are also monitored, trained, or retrained with different deployment deployment strategies. Tracking, comparing, and deploying a model without model management would be challenging.

5.5.4 Digital Twin

The last component of the proposed architecture is the digital twin. In wound management systems, digital twins are utilized to build digital representations of those wound data through computer models. Digital twin technology can be used to generate a virtual twin of a wound to review healing stages, size, and type details, to identify the improvement and challenges. Collected tabular data from both sensors and patient information could be utilized for forecasting tasks such as wound closure

or healing times. By estimating patient discharge time, used resources could be arranged more effectively. In addition to this, required treatments or operations could also be forecasted using various information gathered from patients' health history and current status with the help of AI models.

The digital twin subpart in the proposed architecture provides a software-as-a-medical service. The digital twin of the wound is generated from the developed model using collected data. It is also expected that the proposed systems embedded in digital twins can help optimize the software in medical devices as well as caregivers capture and find information shared across physicians and multiple specialists. A proposed image-based digital twin system for chronic wound management is depicted in Figure 5.3. Images of the wound are used to construct the digital twin of the chronic wound. The wound image is segmented to understand underlying tissue distribution then a wound healing prediction model forecasts the healing progress of the current chronic wound using AI models trained on similar cases. A lifelike wound is generated to visualize the wound itself. The status of the actual wound is assessed, and an appropriate treatment method is chosen accordingly.

5.6 Discussion

The application of the digital twin concept into chronic wound management will ease the overwhelming burden that cumbered the healthcare system. Required continuous clinic visits and care will be reassessed since real-time monitoring and virtual twin of the wound could detect and flag any imminent health issues. The more digitalized wound care could bring many benefits and challenges, such as:

Observation 1: Digital twins can significantly enhance the track and care of wounds by increased monitoring and forecasting capabilities.

Observation 2: Personalized care could be achieved with data-driven approaches that could lead to more efficient and effective care and improved outcomes.
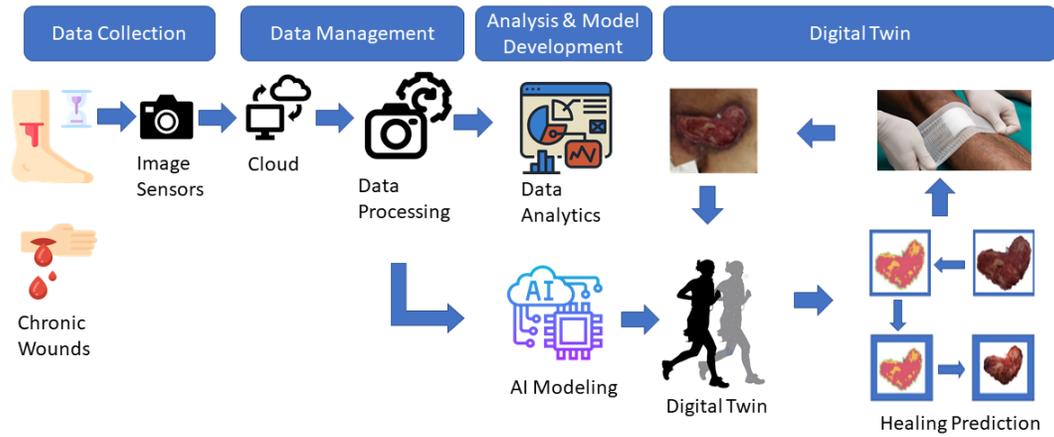
Figure 5.3: Proposed framework for chronic wound management using a digital twin.

Observation 3: Digital twins will provide more information to the patients and encourage them to take more actions for their wellbeing.

Observation 4: Digital twins will increase the use of telehealth applications as a result of synchronous track of the wounds. Any medical issue could be detected, and appropriate medical attention could be sought right away.

Observation 5: Newly emerging technologies such as augmented reality and metaverse could be utilized together with the digital twin concept.

5.7   Summary and Conclusion

Chronic wounds or hard-to-heal wounds require special care as their healing course is out of order and takes longer times in comparison with acute wounds. In order to alleviate the high cost of constant care and track, new technologies should be adopted. The digital twin is one of the critical technologies that could lift this burden with the use of improved data management and ML techniques.

In this chapter, the adoption of the digital twin in healthcare, i.e., chronic wound

management, is examined thoroughly to reach optimal treatment and management of wounds. Pieces of the digital twin are reviewed to increase the opportunities and awareness further. Chronic wounds and enabling technologies of the digital twin are explored. A framework for the proposed chronic wound management system is detailed using the digital twin concept. The proposed image-based model gathers images of wounds and processes them to feed into an ML model. ML model then predicts the healed version of the current wound to enhance wound management. In addition to an image-based system, a hybrid model that utilizes both tabular data coming from IoMT devices and visual data could realize a more comprehensive digital twin formations.

This chapter sheds light on the digital twin concept adaptation into chronic wound management. It is expected that this chapter will provide comprehensive knowledge that can enhance the researchers, engineers, and institutions' vision to reshape their approaches for more efficient and effective wound management systems.

6 Digital Twin Framework for Chronic Wound Management

Digital twin technology has been in use for almost fifty years, but its adoption in healthcare is relatively new. Providing treatments to chronic wounds could be enhanced by using data-driven models, i.e., digital twin, which could improve the outcomes. Tailored treatments by building a digital twin in healthcare will play an essential role in identifying problems beforehand. The progress of the wound could be foreseen with the application of different treatments with the help of a digital twin in advance, which is not possible on the real wound. So possible treatment methods could be assessed before applying to the real wound. The early identification of non-healing wounds could also be accomplished, which will help arrange and adjust chronic wound treatment effectively.

In this chapter, we analyze and design a chronic wound management framework using state-of-the-art AI and digital twin technology. Digital correspondence of the actual wounds will simulate and imitate the healing progress. Section 6.2 and Section 6.3 review the previous works to tackle chronic wound management problems and the use of digital twins in healthcare. The proposed chronic wound management framework using digital twin is presented in Section 6.4. Phases of wound care with a digital twin are explained in Subsection 6.4.3. Details of the developed model for wound healing prediction are presented in Section 6.5. Section 6.6 share and discuss the results. Section 6.7 overview the opportunities and challenges while adopting the digital twin concept. Section 6.8 conclude the chapter.

6.1 Introduction

Classic wound assessment methods such as optical are still in use since the best way to describe a wound's progress is explained visually [171, 172]. The change in tissue distribution and growth is tracked by most caregivers, where proportions of the

tissues provide helpful information to analyze the state of the wound and its healing status [173]. In addition to this, there is still room for the advancement of imprecise optical techniques with the help of image processing and AI [174, 175, 176]. Non-healing or slow-healing wounds could also be identified early with the use of advanced computer vision techniques to provide medical care promptly [177].

ML techniques, especially DL, provided an unprecedented leap in image-based medical tasks such as detection, classification, and segmentation [178]. The use of these models in healthcare, i.e., wound care, is also realized by many works. These prior studies provide image analysis models capable of either measuring wound area or extracting the wound tissue's characteristics individually. The classification of chronic wounds is also studied. However, AI techniques could be incorporated together to complete the modeling of human physiology. In the previous chapter, digital twin use in wound management is introduced.

The digital twin concept will allow us to track and model health issues successfully. A digital twin for chronic wounds, which is essentially a virtual replica of a chronic wound, could provide many benefits, such as automated and enhanced medical care with the use of AI techniques that could improve patients' quality of life. Furthermore, telehealth applications will also harness recent advances in the digital twin. Even after the recent pandemic, telehealth applications boomed at an unpredictable rate where management of wounds through less-trained resources (patient or family) became possible [179]. Increased use of smartphones and wearable devices allows patients to be more proactive with any health concerns, from identifying to tracking health conditions. By incorporating these technologies and processes, digital twins of chronic wounds could revolutionize the management of chronic wounds.

Wound care requires frequent visits to a clinic on a regular basis for wounds to be evaluated by the caregiver, which is time-consuming and causes high costs [180]. Moreover, tracking manual visual examination in these continuous visits requires a

laborious and complicated procedure that lacks quantifiable healing parameters. In this chapter, a digital twin-based chronic wound management framework is developed, and a new model that can predict wound healing is proposed. The main contributions of this study two folds:

(i) Development of the digital twin concept and its components for chronic wound care.

(ii) Development of a novel chronic wound healing prediction model that could forecast the progress of wounds and identify the non-healing wounds.

6.2 Related Works

Many previous works utilize the ML methods to tackle problems in chronic wound management. For instance, in [181], the author discussed the potential role of ML in wound care. It is proposed that standardization and specialization of wound care will be achieved with the use of data and DL techniques. The study in [21] systematically reviews the image-based AI's current state in wound assessment. Results suggest that AI-based wound management platforms will deliver data-driven care efficiently. The authors in [182] developed a prototype with a sunburst diagram to evaluate the clinical decision support system to gain insights regarding its functionality. Their results are affected by the insufficient data, while the sunburst diagram for diagnosis was found to be beneficial. In another study, Howell et al. in [183] proposed an AI-based wound assessment tool to evaluate the percentage of wound tissues quantitatively and qualitatively. Around 1200 wound photographs were collected from two wound care centers, and 199 images were selected for this study in total because of the quality of the images. Results validated that AI models could be in good use to determine wound features to obtain accurate wound assessment. Recent advances in wearable bioelectronics are examined in [184]. Principles and current trends in wound biosensors are discussed with the challenges and future perspectives.

Also, advanced techniques based on AI have been proposed and employed to assess wounds remotely in recent studies. In [185], the authors explored the telehealth application of CNN to detect wounds. Four CNN architectures, namely SegNet, LinkNet, U-Net, and pre-trained U-Net, are prepared to evaluate their performances. Results show that U-Net-based architectures demonstrated better performances where the base U-Net model achieves a specificity of 0.943 and sensitivity of 0.993. In another study [186], wounds of 150 patients are evaluated using an AI-powered medical device for clinical validation. Results show that medical devices reached 97% accuracy against wound bed preparation classification and tissue segmentation analysis in comparison with a physician. In [20], data-driven specialization of wound care and up-skilling through the use of AI-based models are discussed in the context of digital health. A smartphone-based wound assessment system is proposed by Wang et al. in [187] to enable the active participation of patients with diabetes in daily wound care. After capturing them with a smartphone, they segment the wound images using an accelerated mean-shift algorithm. Authors determine the healing status of the wound by the widely accepted RYB wound tissue classification system [160]. Experiments have been performed on two categories of wound images. The first dataset consists of 30 simulated wound images, whereas the second dataset consists of 34 images of actual patients. Simulated wounds are relatively simple than the actual ones. Results show that their model accurately classifies tissues of simulated wounds. In contrast, the results of the analysis of clinical images are not sufficient due to the complex nature of the skin color and texture. The location and variable illumination conditions affected their model as well. It could be inferred that the number of data and their quality affects the performance of the models. Another study that uses the RYB system is conducted by Chitra et al. to investigate the performance of the Random Forest (RF) algorithm on the classification of chronic wound tissues [188]. In general, existing approaches attempt solely to solve wound detection, segmentation, and tissue

classification problems. However, complete wound care could be accomplished with wound healing prediction and ultimately forming a digital twin of the wound.

6.3   Digital Twin in Healthcare

As mentioned previously, the digital twins can be defined as digital replications of physical systems and processes with maintaining synchronization [189]. Digital twins are a relatively new concept in healthcare, while their first use dates back to Apollo 13 shuttle in 1970 [190]. Its incredible potential to revolutionize healthcare brought colossal attention. The developments in IoT systems such as Wireless Body Area Networks (WBAN) and underskin implants enable digital twin technology by the real-time accumulation of personal vitals more accurately [191].

Transferring the idea of a digital twin to healthcare has been discussed and studied in recent studies [146, 147, 151, 192, 193, 194, 195, 196]. These studies examine the digital twin concept to orchestrate an ecosystem of data for the healthcare domain. Many studies, such as [146, 147] proposed digital twin architectures based on IoT and cloud systems. These critical enabling technologies are investigated, and details regarding system workflows are assessed in detail. These studies provide illustrative solutions to digital twins' implementations, focusing on patients and the data flow. In a position paper by Corral-Acero et al. [197], the applications of the digital twin concept to cardiovascular medicine are reviewed. This study suggests that mechanistic and statistical models need to be coordinated to tackle challenges for the deployment in the treatment and prevention of cardiovascular diseases.

Another potential use of digital twins is to personalize the medicine. Current healthcare utilizes only a small number of biomarkers, whereas increased data and processing capabilities yield countless opportunities thanks to the availability of IoT devices and advanced computing power. Customized medicines and vaccines have already been proven to fight against diseases, i.e., recent COVID19 mRNA vaccines.

In [198], the authors investigated the potential use of digital twins for personalized medicine. The drug with the significant effect could be selected for each patient with the use of a digital twin concept. The studies laid down here are only some of the recent applications and reviews that have been published in recent years. Thousands of new digital twin studies in healthcare are reported at an increasing rate each year.

6.4 The Proposed Chronic Wound Management Framework Using Digital Twin

The proposed chronic wound management framework consists of many elements, including data collection, data processing, analysis, model development, and forming a digital twin. After digital twin formation, processing and prediction with the use of AI tools are made. Then wound treatment method is chosen considering the outcomes of the previous step. The status is reassessed, and the digital twin will be synchronized with regard to feedback from the examination. A simplified overview of the proposed framework is shown in Figure 6.1.

6.4.1 Collection, Processing, and Analysis of the Data

Data collection and its security are essential parts of the digital twin concept in healthcare. Wound data could be collected via many sources such as prescreening questionnaires, heart sensors, glucose levels, and intelligent devices such as watches, phones, and IoT gadgets. The CIA triad [199] should be implemented to assure information security to provide confidentiality, integrity, and availability of the data. Data encryption and authentication methods should be adopted to address privacy concerns. Deployment of blockchain technologies will be an alternative to the current methods to enhance data storage and availability safely.

There are different kinds of data available in a digital twin concept. Without a doubt, tabular data is one of them, and it could provide information about patients' demographics such as gender, medical history, smoking status, and age. Besides pa-

97
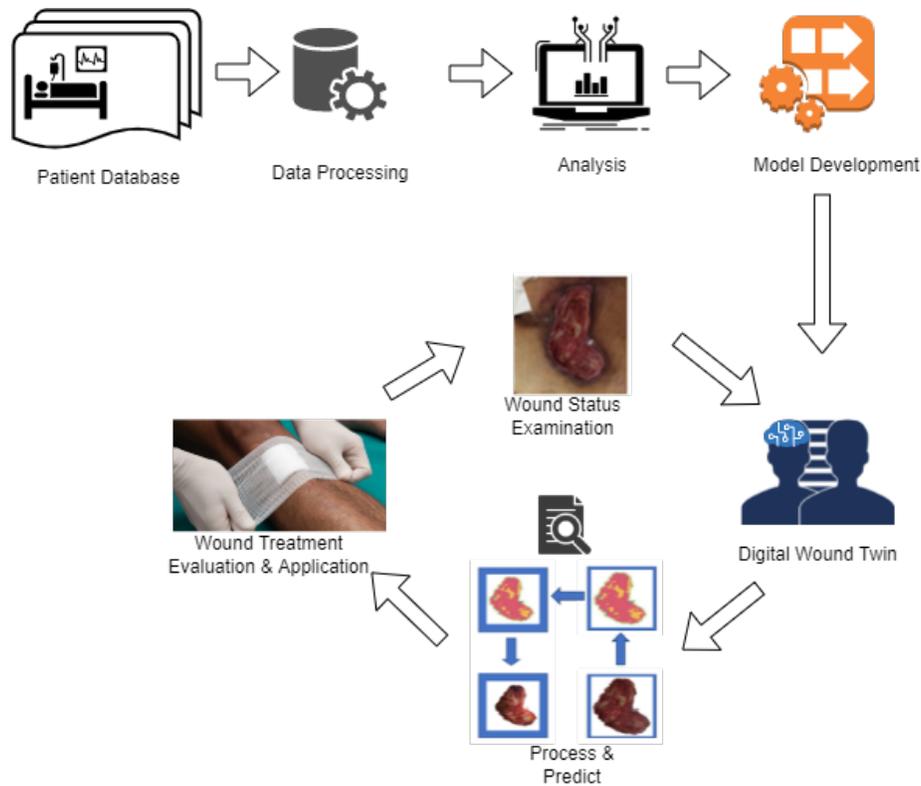
Figure 6.1: Chronic wound management framework using digital twin.

tient demographics, wound characteristics such as wound area, type, and location are also recorded in wound care facilities. However, varying recordings of this information include inaccuracies, inconsistencies, and duplications [200]. Depending the low quality of the data causes unreliable models, which hinders the applicability of the models. One of the critical studies by Fife et al. [201] compared the healing rates of the publicly reported wound outcomes. This study finds that the healing rates of the wounds posted by the providers appear impossible in such a short time with an acceptable outcome. A detailed criterion needs to be developed and adopted for the honest recording of wound outcomes. The proposed model in section five of this study utilizes image data to form a digital wound. After collecting both tabular and image data, preprocessing will be done with the use of python or R libraries such as OpenCV, NumPy, and pandas to deal with data properties such as data shape, scale, and missing data. Feature engineering techniques are also great tools in tabular data

in order to find the underlying trends. Then an analysis of the data will be utilized to gather insights to develop analytic tools and models.

Our proposed model in section five uses an RYB tissue classification system where blue is chosen instead of black for better visuality. The strength of this classification technique is well studied and universally accepted in the wound care field. Traditionally, the wound healing process continuum consists of three phases: reaction, regeneration, and remodeling [161], which could be analyzed by the RYB system as well. Granulation, slough, and eschar are represented by red, yellow, and blue colors, respectively. This method offers a dependable, universal, comprehensible, and comprehensive representation of wounds for assessing the wounds. By using this system, appropriate interventions have been designed and applied in wound care facilities for many years for better treatments. That is why we have decided to use the red-yellow-blue system to evaluate and form a chronic wound management framework.

6.4.2 Model Development and Digital Wound

Model development and its adoption in healthcare are still in their infancy due to high risk and privacy issues. Two different kinds of datasets could be used in data-driven digital wound management. The first one utilizes tabular data to forecast the wound status as well as the wound healing rate [202]. The classification and regression methods are used in these tasks. These ML architectures will reveal the relations between the predictor variables [203]. Forecasting methods will be used to predict the outcome of the wound using the wound area measurements.

Image-based ML models are utilized for most standard computer vision tasks in wound care, i.e., detection, classification, and segmentation. As previously explained in related works, current studies focus on assessing the wound status by either classifying the wound or segmenting it. The measurement of the wound area is also calculated with image-based methods. Wound detection and classification models are

developed using CNNs as well as utilizing transfer learning methods. Transfer learning benefits from previously learned features and transfer this information to new tasks. The most used transfer learning models are ResNet, VGG16, and EfficientNet. These networks are pre-trained using different datasets such as ImageNet and CIFAR-10. These datasets consist of millions of images with hundreds of classes. Learned feature extraction capabilities and then re-trained to classify chronic wounds. In addition to classification models, XAI techniques are developed to make the classification results more transparent since ML models, especially DL models, are black-boxes to the users [30, 204, 205].

Segmentation of the wound and its tissues is another task that will be used to form a digital twin of the wound as well. Segmentation of wounds is simply a classification of each pixel to a class. If only the wound is segmented, there will be two classes, i.e., wound and rest. If there are more classes, such as three different tissue classes, this will be a multi-class segmentation. This task provides the necessary information to understand the current status of the wound. However, segmenting the wound will not be sufficient for forming a digital wound. The segmentation of tissues has great importance as it will allow for further processing and forecasting of the wound by gathering tissue information. Tissue segmentation will allow tracking the tissue distribution and wound healing status in consecutive images. Tissue distribution provides essential information regarding the wound status since different tissues indicate unalike healing paths.

Other tasks that could be developed using wound images are detecting non-healing wounds and predicting wound healing. These tasks will predict the wounds' development in a certain period. To forecast the wounds' succeeding status requires the knowledge of current wound properties. As tissue types affect the wound appearance and distribution of these tissues is used to disclose the status of the wound, tissue segmentation could be used to identify the successive wound status.

6.4.3 Phases of Wound Care with Digital Twin

Digital representation of the wound consists of many elements that require synchronization periodically. The digital twin framework for chronic wounds has four steps, i.e., digital twin of the wound, process and predict, evaluation and treatment, and status examination, that makes a loop through the life of the wound. These stages enhance the wound status and improve treatment method choice.

(i) Digital Twin: Construction of a digital twin framework for chronic wound management is achieved using the data collection from similar cases, model building, and storing this information in a database. Models and their properties are designed and implemented to illustrate the physical wound on a digital platform in this phase. Data analysis tools and ML techniques are used to represent the features of the wound.

(ii) Processes and Predictions: After constructing the digital twin with various data-driven models, these models are applied to the chronic wound at hand by utilizing the sensor data to analyze and predict the outcomes using AI models. AI models previously trained on similar cases are used to classify, localize, and segment the wounds as a first step. These processes will help understand and quantify the underlying wound features such as tissue distribution and wound area. By using these and other tabular data coming from both sensors and prescreening questionnaires, additional models will be run to analyze the development of the wound. The wound healing prediction model is also proposed to identify non-healing wounds in this study. The proposed model provides essential information for the early detection of severe conditions that might lead to limb losses.

(iii) Evaluation and Treatment: In this stage, the interaction between the digital twin and the physical twin is realized by evaluating the previous stage's results.

101

After examining the wound with its digital, proper wound care could be planned and implemented by the intervention of professional caregivers. The digital twin is more than just a decision support system for the treatment methods; it could furnish the Key Performance Indicators (KPIs) and increase the situational awareness related to the status of the wound and, ultimately, the patient's status.

(iv) Wound Status Examination: After choosing and implementing the customized treatment, the obtained outcome of the medical administration is monitored and assessed. Furthermore, the resulting wound status will be utilized to update the digital twin framework using the sensors' data. Physical and digital twins are synchronized by enhancing the whole framework with current wound data.

6.5   Development of an AI Tool for Digital Twin

Tracking chronic wounds is an essential part of wound care as chronic wounds do not heal timely and in a proper way. Identification of non-healing wounds has utmost importance since non-healing status poses serious outcomes such as illnesses and loss of limbs. The proposed model is developed to forecast the progress of wound healing using wound images. To project the progress of a chronic wound using images following tasks should be implemented:

(i) Classification of the wound type.

(ii) Segmentation and the classification of wound tissues.

(iii) Prediction of the wound tissue distribution.

(iv) Generation of lifelike wound.

These tasks will provide essential information to form a digital twin for wound care. The flow of the proposed model and tasks are visualized in Figure 6.2. An

image of the wound will be taken and preprocessed to remove the background from the image. The tissue segmentation will be done to comprehend and extract the wound's features. The wound healing prediction model will forecast the distribution of the wound tissues after four weeks of treatment. Another model will map this tissue distribution to lifelike wound images in order to be easily understood by the caregiver. Four weeks of the treatment are chosen as this time period is used to evaluate wound status, and it is found that 50% of the wound area will be closed in four weeks [172, 206, 207, 208, 209].
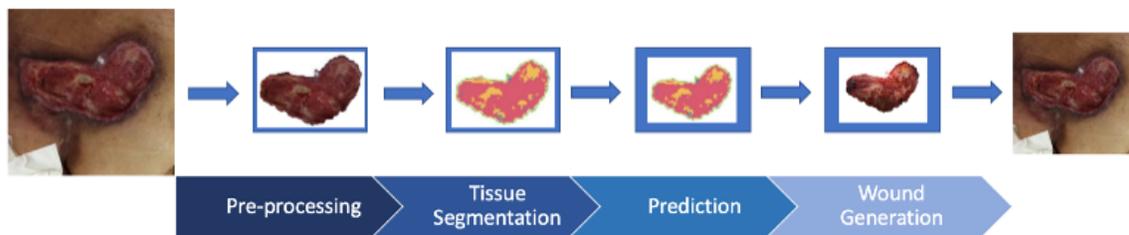
Figure 6.2: Wound healing projection model.

Perfect prediction of a wound's progress is almost impossible as every wound heals differently. The healing progress prediction could be beneficial significantly while detecting a non-healing wound as this condition is an essential indicator of treatment outcome. The tasks mentioned above are studied in our prior works [30, 32, 31] except wound healing. This section develops the wound healing prediction and identification of non-healing wound tasks using the state-of-the-art DL model, i.e., GANs [28].

6.5.1 Data Collection, Preprocessing, Environment, and Validation

Image data used in this study is collected from the chronic wound data repository provided by eKare Inc. It provides 3D imaging and AI solutions to analyze wound healing and optimize clinical workflow in hospitals and clinics worldwide. Provided images are collected using commercially available cameras by caregivers in a natural hospital environment during regular patient visits. Various chronic wound types are included in the repository, such as pressure injury, diabetes, lymphovascular, and

surgical wounds. The diversity of wound types increases the performance of the model. The proposed model gets tissue segmented wound images and its healed tissue segmented images after four weeks of conventional treatment, shown in Figure 6.3.

(a) Wound Image

(b) Tissue Segmentation of (a)

(c) Wound Image

(d) Tissue Segmentation of (c)
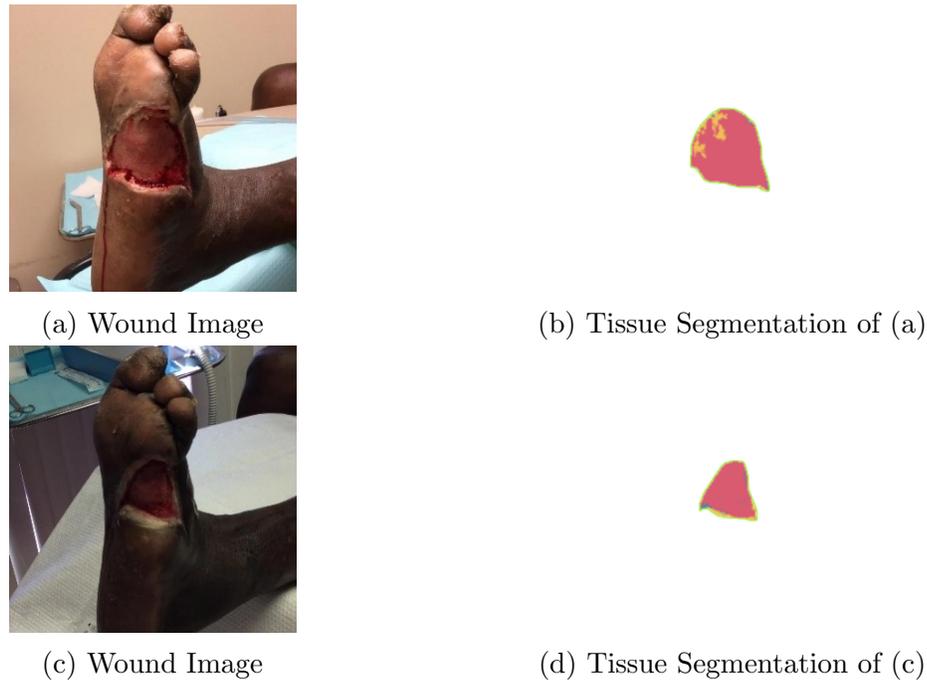
Figure 6.3: Samples from the dataset (a) and (b) a wound image and its segmentation prior to treatment, (c) and (d) same wound and its segmentation after four weeks of treatment.

Tissue distribution is one of the critical features in chronic wounds, which unfold the status of the wound successfully. Therefore, tissue segmentation could be seen as a feature extraction method for wound images. There are 700 wound couples, and %10 of them are used for testing, and the rest is used for training.

A broadly accepted RYB tissue classification technique is adopted in segmenting the images. Instead of black, blue is chosen for better visibility. The wounds are semi-automatically segmented for the ground truth, and the resulting segmentation is reviewed by the MD specializing in wound care. Wound images are anonymized and rescaled to 512x512 pixels. Data augmentation techniques such as flipping, mirroring,

and rotation are utilized to improve the model's performance and robustness.

The proposed model is trained using the PyTorch DL framework on the VS Code with Python version 3.6. Our implementations ran on Intel® Core ™ i7 -7800X CPU @3.50 GHz with 16 GB RAM and NVIDIA GeForce GTX 1080 GPU with 8 GB dedicated and 8 GB shared memory. The proposed model is trained for 200 epochs, which took around 7 hours. It has a batch size of 64 and a learning rate of 0.0002.

Validation of the study is challenging due to the complexity of the problem. Since each wound heals differently and depends on many conditions such as diets and activities, producing an exact healed version of a wound is almost impossible. That is why we have used percentages of wound tissues and wound area as a validation method in our study. Also, a comparison of these values will give sufficient information for identifying non-healing wounds, which is one of the goals of this study. To compare the percentages of the tissues', the MSE metric is utilized. MSE metric can be written as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}[(P_{\mathrm{R}} - P'_{\mathrm{R}})^2 + (P_{\mathrm{Y}} - P'_{\mathrm{Y}})^2 + (P_{\mathrm{B}} - P'_{\mathrm{B}})^2] \tag{6.1}$$

Where:

n: Number of pixels

P_{R}, P_{Y}, P_{B}: RYB pixel percentages of the actual images

P'_{R}, P'_{Y}, P'_{B}: RYB pixel percentages of generated images.

Area measurement error is also compared with the following equation 6.2:

$$PE = 100 \cdot \frac{Actual - Predicted}{Actual} \tag{6.2}$$

Where:

PE: Percentage Error

Actual: Real Image Area Measurement

Predicted: Predicted Image Area Measurement

6.5.2 Implementation Using GAN

GANs have been used in many applications such as DeepFake and style transfer. It is a powerful tool to generate new images with the help of the generator (G) and discriminator (D) couple. The overview of the model is depicted in Figure 6.4. The generator network is fed with the random Gaussian noise (z) and wound images before the treatment (x), then it generates the healed version of the wound after four weeks of treatment (y), G: x, z→ y. This type of GAN implementation is also described as conditioning the generated image with an input image. Input images behave like labels, and the outputs are generated with the same input image domain structure. That is why these implementations are also referred to as cGAN [104]. Feeding the generator network with noise produces unique output generation. The generated images are unique with this formation while maintaining the training dataset's aligned and paired data distribution. The Discriminator network (D) is trained by the training dataset simultaneously with the generator network to learn the data distribution of output images. Discriminator network classifies newly generated images whether they are from the training set or not D: x, y→ [0,1]. Each network has different loss functions; the discriminator network is updated directly by classifying actual and generated images' sigmoid cross-entropy losses. At the same time, generator loss is a sigmoid cross-entropy loss with the L1 loss, which depends on the discriminator network. This concurrent training of both networks results in non-converging loss results in training [112]. After maturing the generator network and training with the
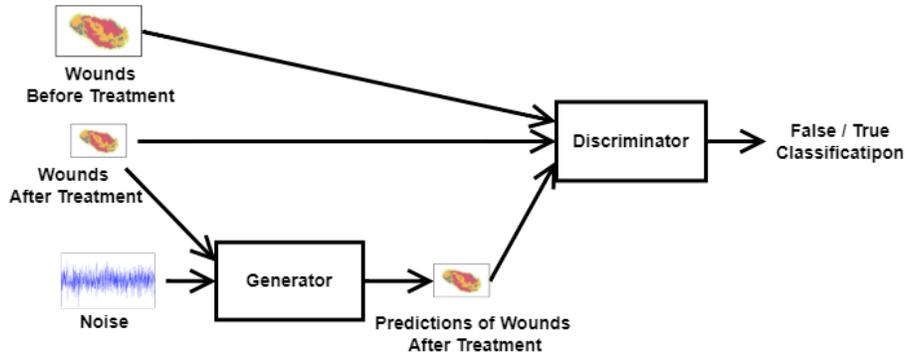
Figure 6.4: The overview of the model.

(a) Original wound segmentation before the treatment

(b) Ground truth segmentation after treatment

(c) Predicted wound segmentation after treatment
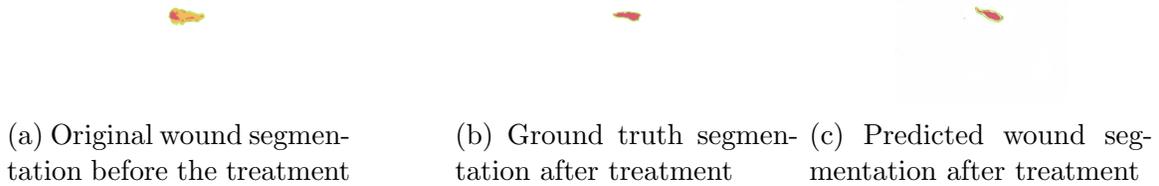
Figure 6.5: The proposed model outputs successful results.

discriminator network, the generator network is employed to generate test images.

6.6 Results and Discussion

After running the model on the test set, results show that the proposed model successfully predicts the healing progress of a chronic wound so that it can be adapted to identify non-healing wounds correctly. A sample output of the model is conveyed in Figure 6.5c.

A sample wound progress and the output of the proposed model are shown in Figure 6.5. The development of the natural wound is mimicked by the model successfully. The wound prior to the treatment has slough and granulation as well as a small amount of eschar. After four weeks of treatment, the look of the wound indicates that previously granulated areas are healed thoroughly. Slough tissues healed; moreover, the granulation tissue appeared in the same sites. The necrotic tissue also falls off, and the granulation tissue grows in these sites. A similar healing path can be tracked

in the predicted model output depicted in Figure 6.5c. The tissue distribution and area measurements show a similar development. The shape of the wound is also similar to the actual wound. The orientation of the predicted wound is not in line with the actual ones ideally since data augmentation techniques are implemented, and every wound heals differently. By considering these features, it could be said that the proposed model successfully predicts the healing progress of a chronic wound. The healing prediction is evaluated using area percentage (predicted vs. actual) and the change in tissue distribution (predicted vs. actual) since the healing wounds share similar tissue distribution.

6.6.1 Comparison of Area

Area comparison is one of the KPIs used in wound care. In this section, the comparison between natural wound healing and predicted wound healing is being made to determine the performance of the proposed method. The MSE metric and the difference in wound area percentages will be calculated to discuss the results further.

Figure 6.6 depicts the progress of a sample wound. It can be deduced that the progress of the wound is not always healing. After an excellent decreasing wound area trend is broken, the wound area increases significantly. This change can result from a treatment method or another secondary health problem complication. The early identification of this non-healing condition is vital to prevent further health concerns such as limb loss. Also, there is a wrong recording which leads to a substantial decrease.

Wound area decreases or increases during the whole continuum of wound care. The decreasing trend is targeted in most cases to heal the wound completely. Wound area information of the dataset is given in Table 6.1, including mean, max, and min

Figure 6.6: A sample wound healing with respect to area.

Table 6.1: Change in wound area after four weeks of treatment (Real vs. Predicted).

Change / Area	Original wound change (%)	Predicted wound change (%)
Min Area	-37.8	46.2
Mean Area	49.4	51.2
Max Area	82.2	56.8

area.

The results indicated that the area increase could happen in chronic wounds that need to be detected before the severe outcomes. After four weeks of treatment, min area change is found to be -27.8, indicating that the wound is worsening. Figure 6.7 depicts an example of a worsening wound and the output of the model's prediction.

The depicted wound in Figure 6.7 is a non-healing wound, and it is not healing with the applied treatment; however, our model indicates the expected healing progress, which will be used to warn caregivers and patients. Looking at our model output caregiver or the patient could detect the non-healing wound successfully.

109

(a) Wound segmentation prior to treatment

(b) Segmentation after four weeks of treatment

(c) Predicted wound segmentation after four weeks of treatment.
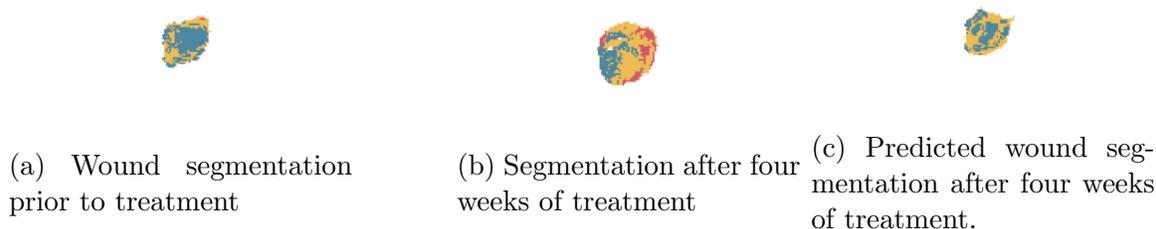
Figure 6.7: A sample of worsening wounds and its prediction by the model.

6.6.2 Comparison of Tissue Distributions

The wound tissues are of utmost importance as the wounds heal with changing tissue phases. Change in tissue distributions is given in Table 6.2. Areas with eschar or necrotic tissue are non-viable due to the reduced blood supply, and it will fall off during the healing. Areas with slough are also dead tissue caused by increased cell deaths and white blood cells. Moreover, areas with granulation are an indicator of healing [210]. That is why a healing wound should have more granulation concerning other types of tissues. The proportion of the eschar should be decreased during the healing process. The absence of granulation indicates a lack of adequate blood flow and impedes healing.

Table 6.2: Change in wound tissue distributions.

Stage vs Tissue Change	Granulation (Red) Tissue (%)	Slough (Yellow) Ratio	Eschar (Blue) Tissue (%)
Before Treatment	58.7	32.9	8.4
After Treatment	69.9	24.5	5.6
Predicted	64.7	31.4	3.9

The results are also analyzed using the MSE metric. MSE value of 25.84 is calculated using equation 6.2. MSE score indicates that around 74 percent of the tissues are predicted correctly. Results indicate that the proposed model could predict similar tissue distribution. Another result is conveyed in Figure 6.8. Original and healed wound segmentations are shown in Figure 6.8a and 6.8b, respectively. Figure 6.8c

(a) Wound segmentation prior to treatment.

(b) Segmentation after the four weeks of treatment.

(c) Predicted wound segmentation after four weeks of treatment.
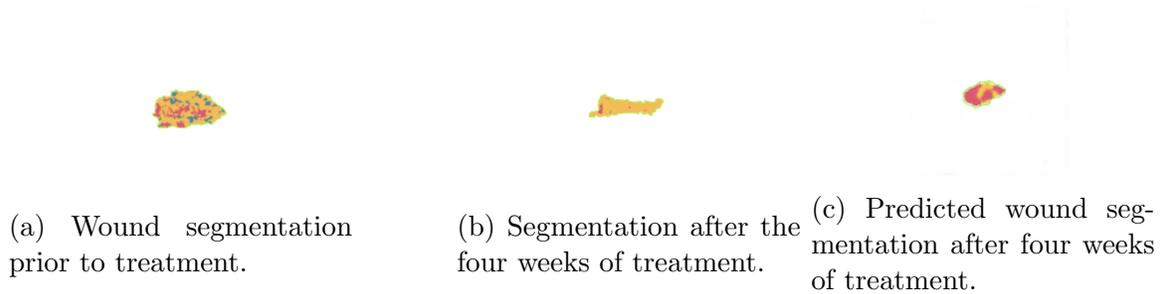
Figure 6.8: A sample of worsening wounds and its prediction by the model.

indicates the predicted wound segmentation after four weeks of treatment.

6.7   Opportunities and Challenges

With the adoption of the digital twin concept, digitalized healthcare will provide many benefits. The digital transformation which revolutionized many industries could be carried to healthcare with the help of the digital twin concept. Its benefits could be listed as:

1. Personalized wound care will be the first benefit of the digital twin concept as it presents a data-driven and synchronous approach to the wound at hand. Imminent health issues will be detected, and caregivers will be warned in real-time. Personalized treatments will improve the outcomes.

2. Lower costs and higher efficiency are other benefits of the digital twin for wound management. Frequent visits to clinics will be minimized, saving time and money for both patients and clinicians. Furthermore, with the improved tele-health approaches that could use digital twin, patients and their families could play an active role. Treatments and medicines tailored to patients' needs will also result in efficient resource management.

3. Improved and increased self-aware decision support systems could be developed, and treatments and medicines will be better analyzed for each patient. Digital twin for wound care management conveys previous experiences to the new pa-

tients that will improve the performance of the applied treatment. Optimized healthcare will be given to the patients with detailed risk assessments.

4. Advanced analysis by gathering unseen correlations and insights will be made using the digital twin for wound care. AI is capable of completing many tasks with high accuracy. The underlying correlations could be detected effectively by the AI systems. This is the result of using data-driven models and depends on the quality of the data. The insights given by the analysis of AI systems will be helpful in choosing and adjusting the treatment parameters.

Despite having enormous potential in digital twin for chronic wound management, there are still challenges that need to be addressed. Following problems persist in broad acceptance of the digital twin concept in wound care.

1. Technical difficulties such as multiple sourcing caused the complexity of the processes that need to be tackled. Integration of IoT devices will also be practical and effective in providing real-time data flow. Besides, simultaneous processing of this data is essential to uncover the interactions of the accumulated information. Furthermore, Cloud-based services, IoT and AI will play a significant role in overcoming these difficulties.

2. Data security and privacy are other fields that need attention for adequately handling sensitive information by the digital twin frameworks. The CIA triad will be put to use with data encryption authentication techniques. Besides these conventional methods, blockchain technologies have significant potential to be applied in digital twin implementations for chronic wound management and other data-sensitive areas.

3. Interoperability and standardization are the key elements in digital twin systems to gather and process valuable information. Differences and inconsistencies in

record-keeping hinder the data quality, resulting in inaccurate models. Interoperable systems and standard data and processing structures should be developed for a feasible and scalable digital twin framework.

4. Resource management and network requirements in a healthcare facility are other domains that need enhancement. The increased need for reliable, concurrent, and secure communication between digital and physical twins demands better quality of service. Display and update the real-time data, and their analysis will generate a massive memory load. New cloud-based technologies, fog, and edge computing will be utilized to address these requirements.

5. Collaborative environment creation of digital twin in healthcare is still in its infancy. Many major companies provide digital twin solutions for manufacturing industries. However, the healthcare industry still cries out for reliable, efficient, functional digital twin systems.

6.8   Conclusion

Chronic wounds cause a decrease in the quality of a patient's life with the loss of independent movement and altered lifestyle, including daily activities, emotions, sleeping, and eating habits. These wounds could be non-healing, primarily resulting from incompetence, misdiagnosis, inappropriate treatment strategies, or neglect [207], and cause pain and reduced quality of life.

This study presents a framework for using the digital twin concept for chronic wound care management to improve the outcomes. The elements of the digital twin are examined, and the phases of digital twin use are explained. A new AI tool to predict the wound healing status is proposed to enhance the use of AI in digital twin for chronic wound care. The proposed model uses images of chronic wounds before treatment and the same wounds' images after a four-week treatment. The first step is to segment images to extract their tissue features. The proposed model

is trained to learn the tissue segmentation of the same wound after four weeks of treatment. With this study, wound healing prediction was performed for the first time using tissue segmentation. The proposed model provides straightforward, fast, and accurate implementation without requiring labor-intensive preprocessing.

Visual and quantitative analyses are conducted to validate the study. MSE results indicate that the model performs successfully, but further improvement will be implemented for finetuning. The scope of this study includes the prediction of various wound outcomes after four weeks of treatment. The progress of each wound type differs notably. However, tissue healing follows similar phases. In order to overcome the limitation of finding structured data, tissue segmentation is a practical alternative. With increased and structured data, the results will be further enhanced. Another future study will be the extension of the study to enhance the proposed model by hyperparameter optimization and the use of high-quality structured images. We should admit that wounds and their care have a colossal complexity that requires professional involvement. This study is expected to benefit the wound care community and researchers working on digitalized healthcare by extending the concept of a digital twin to chronic wound care. The telehealth applications using the digital twin concept will also provide valuable insights regarding the active participation of patients and their families in wound care.

7 Conclusions and Further Developments

The preceding chapters have provided several aspects of chronic wound management, including model and framework developments. In this chapter, a summary of the major contributions is presented in Section 7.1. Further research directions are suggested in Section 7.2.

7.1 Conclusions

The main contributions presented in this dissertation can be summarized in five folds:

(i) A novel chronic wound classifier that can categorize chronic wounds and explain why they belong to that category. The transfer learning technique is leveraged to extract features of the wound images which will shorten the time for classification and improve the classifier performance. In addition to transfer learning, a novel explanation method is developed and proposed to highlight the wound features that affect its classification. Utilizing this proposed XAI technique to explain wound classification can help shed new perspectives on clinicians and physicians during the diagnostic phase.

(ii) A new AI-based computer vision model is developed to localize and segment the wound as well as its tissues. This hybrid approach is developed using a state-of-the-art DL model, i.e., GAN, which provides a straightforward implementation advantage and catches the data distributions accurately. The proposed model will assist caregivers and clinicians in determining a proper wound treatment plan by assessing the wound location, area, and tissue distribution. Assessing these features is key to well-planned chronic wound care management. Hyper-parameter optimization is also studied to find optimal conditions.

(iii) Scarcity of medical images hinders the development of computer vision models

due to the privacy issues related to patients and hospitals. A new approach has been developed to overcome privacy issues as the proposed model could generate lifelike images of wounds without breaching the privacy of the patients and hospitals. The proposed model is built conditionally using the GAN model. This model could also be suitable for clinician training in medical schools to improve chronic wound care. Hyperparameter optimization is also realized to find optimal conditions for the DL-based model.

(iv) The digital twin concept is one of the critical technologies that transform many industries, but its adoption in healthcare is still in its infancy. The use of digital twin in chronic wound management is proposed by using AI methods. The digital twin concept for health has great potential to fulfill personalized and predictive healthcare. Enabling technologies of a digital twin for wound management are examined after revising concepts and approaches to chronic wound care management.

(v) We have proposed a data-driven wound healing prediction framework to improve wound healing tracking with the assistance of DL. Instead of complex and time-consuming ruler-based measuring systems, a data-driven approach using wound images of successive weeks is used to predict the wound status after a certain period of time. A digital version of the wound is built using the digital twin concept, which could imitate the healing progress of the actual wound. By using this framework, tailored treatments will be planned for optimal treatment strategy for better outcomes.

7.2 Further Developments

Chronic wound care management is a critical healthcare field that requires continued care. Many new healthcare approaches arise from applications in AI and the

digital twin. They provide new venues and motivations for future works. These directions include:

(i) The aforementioned chronic wound care management approaches utilize the wound images, whereas tabular data is not used. Although there are different approaches and inconsistencies for keeping tabular data, it is required to collect tabular data for a complete treatment plan. Tabular data includes patient demographic information and other health data such as smoking status and previous health conditions. Forecasting of wound closure rate and healing time could be realized with the help of tabular data.

(ii) Most of the studies in the chronic wound care field utilize the RGB images. However, the technological advancement in hyperspectral imaging could open new approaches to chronic wound management by accessing beyond the RGB spectrum. Although current literature has some implementations using hyperspectral imaging [211], these applications are limited by the number of patients and wounds. New AI models could be developed to incorporate the features detected by hyperspectral imaging, such as blood flow, temperature, and oxygenation.

(iii) Many healthcare fields utilize visual assessments in many forms, such as regular images, X-rays, and CT scans. The previously developed AI models could be utilized for different healthcare problems like eye and oral diseases. Human digital twin will also be possible by utilizing these AI models for personalized healthcare.

References

[1] C. Villani, Y. Bonnet, B. Rondepierre *et al.*, *For a meaningful artificial intelligence: Towards a French and European strategy.* Conseil national du numérique, 2018.

[2] H. Lu, Y. Li, M. Chen, H. Kim, and S. Serikawa, "Brain intelligence: go beyond artificial intelligence," *Mobile Networks and Applications*, vol. 23, no. 2, pp. 368–375, 2018.

[3] I. Peña-López *et al.*, "Artificial intelligence in society," 2019.

[4] B. C. Nwomeh, D. R. Yager, and I. K. Cohen, "Physiology of the chronic wound," *Clinics in plastic surgery*, vol. 25, no. 3, pp. 341–356, 1998.

[5] D. G. Armstrong, J. Wrobel, and J. M. Robbins, "Guest editorial: are diabetes-related wounds and amputations worse than cancer?" pp. 286–287, 2007.

[6] M. Dabas, D. Schwartz, D. Beeckman, and A. Gefen, "Application of artificial intelligence methodologies to chronic wound care and management: A scoping review," *Advances in wound care*, no. ja, 2022.

[7] D. Anisuzzaman, Y. Patel, B. Rostami, J. Niezgoda, S. Gopalakrishnan, and Z. Yu, "Multi-modal wound classification using wound image and location by deep neural network," *arXiv preprint arXiv:2109.06969*, 2021.

[8] C. Aguirre Nilsson and M. Velic, "Classification of ulcer images using convolutional neural networks," Master's thesis, 2018.

[9] H. Wannous, Y. Lucas, and S. Treuillet, "Enhanced assessment of the wound-healing process by accurate multiview tissue classification," *IEEE transactions on Medical Imaging*, vol. 30, no. 2, pp. 315–326, 2010.

[10] R. Niri, H. Douzi, Y. Lucas, and S. Treuillet, "A superpixel-wise fully convolutional neural network approach for diabetic foot ulcer tissue classification," in *International Conference on Pattern Recognition.* Springer, 2021, pp. 308–320.

[11] N. Rania, H. Douzi, L. Yves, and T. Sylvie, "Semantic segmentation of diabetic foot ulcer images: dealing with small dataset in dl approaches," in *International Conference on Image and Signal Processing.* Springer, 2020, pp. 162–169.

[12] G. Han and R. Ceilley, "Chronic wound healing: a review of current management and treatments," *Advances in therapy*, vol. 34, no. 3, pp. 599–610, 2017.

[13] H. Brem, M. Tomic-Canic, A. Tarnovskaya, H. P. Ehrlich, E. Baskin-Bey, K. Gill, M. Carasa, S. Weinberger, H. Entero, and B. Vladeck, "Healing of elderly patients with diabetic foot ulcers, venous stasis ulcers, and pressure ulcers," *Surgical technology international*, vol. 11, pp. 161–167, 2001.

[14] D. Stern and H. Cui, "Crafting polymeric and peptidic hydrogels for improved wound healing," *Advanced Healthcare Materials*, vol. 8, no. 9, p. 1900104, 2019.

[15] D. Paquette and V. Falanga, "Leg ulcers," *Clinics in geriatric medicine*, vol. 18, no. 1, pp. 77–88, 2002.

[16] C. K. Sen, G. M. Gordillo, S. Roy, R. Kirsner, L. Lambert, T. K. Hunt, F. Gottrup, G. C. Gurtner, and M. T. Longaker, "Human skin wounds: a major and snowballing threat to public health and the economy," *Wound repair and regeneration*, vol. 17, no. 6, pp. 763–771, 2009.

[17] D. R. Bickers, H. W. Lim, D. Margolis, M. A. Weinstock, C. Goodman, E. Faulkner, C. Gould, E. Gemmen, and T. Dall, "The burden of skin diseases: 2004: A joint project of the american academy of dermatology association and the society for investigative dermatology," *Journal of the American Academy of Dermatology*, vol. 55, no. 3, pp. 490–500, 2006.

[18] M. Flanagan, "The characteristics and formation of granulation tissue," *Journal of wound care*, vol. 7, no. 10, pp. 508–510, 1998.

[19] W. McGuiness, S. V. Dunn, T. Neild, and M. J. Jones, "Developing an accurate system of measuring colour in a venous leg ulcer in order to assess healing," *Journal of wound care*, vol. 14, no. 6, pp. 249–254, 2005.

[20] D. Queen and K. Harding, "Data-driven specialisation of wound care through artificial intelligence," *International Wound Journal*, vol. 16, no. 4, p. 879, 2019.

[21] D. Anisuzzaman, C. Wang, B. Rostami, S. Gopalakrishnan, J. Niezgoda, and Z. Yu, "Image-based artificial intelligence in wound assessment: A systematic review," *Advances in Wound Care*, 2021.

[22] C. Cui, K. Thurnhofer-Hemsi, R. Soroushmehr, A. Mishra, J. Gryak, E. Domínguez, K. Najarian, and E. López-Rubio, "Diabetic wound segmentation using convolutional neural networks," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 1002–1005.

[23] K. L. Wu, K. Maselli, A. Howell, D. Gerber, E. Wilson, P. Cheng, P. C. Kim, and O. Guler, "Feasibility of 3d structure sensing for accurate assessment of chronic wound dimensions," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 10, no. 1, pp. 13–14, 2015.

[24] S. Zahia, M. B. G. Zapirain, X. Sevillano, A. González, P. J. Kim, and A. El-maghraby, "Pressure injury image analysis with machine learning techniques: A systematic review on previous and possible future methods," *Artificial intelligence in medicine*, vol. 102, p. 101742, 2020.

[25] H. Brem, O. Stojadinovic, R. F. Diegelmann, H. Entero, B. Lee, I. Pastar, M. Golinko, H. Rosenberg, and M. Tomic-Canic, "Molecular markers in pa-

tients with chronic wounds to guide surgical debridement," *Molecular medicine*, vol. 13, no. 1, pp. 30–39, 2007.

[26] S. Hwangbo, S. I. Kim, U. Cho, Y.-S. Song, and T. Park, "Identification of hyperparameters with high effects on performance of deep neural networks: application to clinicopathological data of ovarian cancer," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2019, pp. 1982–1987.

[27] L. N. Smith, "A disciplined approach to neural network hyper-parameters: Part 1–learning rate, batch size, momentum, and weight decay," *arXiv preprint arXiv:1803.09820*, 2018.

[28] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[29] W. H. L. Pinaya, S. Vieira, R. Garcia-Dias, and A. Mechelli, "Autoencoders," in *Machine learning*. Elsevier, 2020, pp. 193–208.

[30] S. Sarp, M. Kuzlu, E. Wilson, U. Cali, and O. Guler, "The enlightening role of explainable artificial intelligence in chronic wound classification," *Electronics*, vol. 10, no. 12, p. 1406, 2021.

[31] S. Sarp, M. Kuzlu, M. Pipattanasomporn, and O. Guler, "Simultaneous wound border segmentation and tissue classification using a conditional generative adversarial network," *Journal of Engineering*, vol. 2021, no. 3, 2021.

[32] S. Sarp, M. Kuzlu, E. Wilson, and O. Guler, "Wg2an: Synthetic wound image generation using generative adversarial network," *The Journal of Engineering*, vol. 2021, no. 5, pp. 286–294, 2021.

[33] S. Sarp, M. Kuzlu, Y. Zhao, F. O. Catak, U. Cali, and O. Guler, "Chronic wound management in digital twin," in *Digital Twin Driven Intelligent Systems.* Springer.

[34] S. Sarp, M. Kuzlu, Y. Zhao, and O. Guler, "Digital twin in healthcare: A study for chronic wound management," *IEEE Journal of Biomedical and Health Informatics.*

[35] A. Ramesh, C. Kambhampati, J. R. Monson, and P. Drew, "Artificial intelligence in medicine." *Annals of the Royal College of Surgeons of England*, vol. 86, no. 5, p. 334, 2004.

[36] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, and Y. Wang, "Artificial intelligence in healthcare: past, present and future," *Stroke and vascular neurology*, vol. 2, no. 4, 2017.

[37] A. Adadi and M. Berrada, "Peeking inside the black-box: a survey on explainable artificial intelligence (xai)," *IEEE access*, vol. 6, pp. 52 138–52 160, 2018.

[38] D. Castelvecchi, "Can we open the black box of ai?" *Nature News*, vol. 538, no. 7623, p. 20, 2016.

[39] A. Holzinger, C. Biemann, C. S. Pattichis, and D. B. Kell, "What do we need to build explainable ai systems for the medical domain?" *arXiv preprint arXiv:1712.09923*, 2017.

[40] K. Gade, S. C. Geyik, K. Kenthapadi, V. Mithal, and A. Taly, "Explainable ai in industry," in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2019, pp. 3203–3204.

[41] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins *et al.*, "Explainable artificial

intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information fusion*, vol. 58, pp. 82–115, 2020.

[42] F. K. Došilović, M. Brčić, and N. Hlupić, "Explainable artificial intelligence: A survey," in *2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO)*. IEEE, 2018, pp. 0210–0215.

[43] S. Benjamens, P. Dhunnoo, and B. Meskó, "The state of artificial intelligence-based fda-approved medical devices and algorithms: an online database," *NPJ digital medicine*, vol. 3, no. 1, pp. 1–8, 2020.

[44] U. J. Muehlematter, P. Daniore, and K. N. Vokinger, "Approval of artificial intelligence and machine learning-based medical devices in the usa and europe (2015–20): a comparative analysis," *The Lancet Digital Health*, vol. 3, no. 3, pp. e195–e203, 2021.

[45] C.-W. Park, S. W. Seo, N. Kang, B. Ko, B. W. Choi, C. M. Park, D. K. Chang, H. Kim, H. Kim, H. Lee *et al.*, "Artificial intelligence in health care: Current applications and issues," *Journal of Korean medical science*, vol. 35, no. 42, 2020.

[46] M. A. Ahmad, C. Eckert, and A. Teredesai, "Interpretable machine learning in healthcare," in *Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics*, 2018, pp. 559–560.

[47] R. Schmelzer, "Understanding explainable ai'. forbes," 2019.

[48] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based

localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

[49] S. M. Mathews, "Explainable artificial intelligence applications in nlp, biomedical, and malware classification: a literature review," in *Intelligent computing-proceedings of the computing conference.* Springer, 2019, pp. 1269–1292.

[50] M. Aghamohammadi, M. Madan, J. K. Hong, and I. Watson, "Predicting heart attack through explainable artificial intelligence," in *International Conference on Computational Science.* Springer, 2019, pp. 633–645.

[51] I. Monteath and R. Sheh, "Assisted and incremental medical diagnosis using explainable artificial intelligence," in *Proceedings of the IJCAI/ECAI Workshop on Explainable Artificial Intelligence (XAI)*, 2018, pp. 104–108.

[52] W. Samek, T. Wiegand, and K.-R. Müller, "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models," *arXiv preprint arXiv:1708.08296*, 2017.

[53] C. Meske and E. Bunde, "Using explainable artificial intelligence to increase trust in computer vision," *arXiv preprint cs.HC/2002.01543*, 2020.

[54] S. Pan and Q. Yang, "A survey on transfer learning. ieee transaction on knowledge discovery and data engineering, 22 (10)," 2010.

[55] W. Dai, Y. Chen, G.-R. Xue, Q. Yang, and Y. Yu, "Translated learning: Transfer learning across different feature spaces," *Advances in neural information processing systems*, vol. 21, 2008.

[56] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[57] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.

[58] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[59] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[60] K. He, Y. Wang, and J. Hopcroft, "A powerful generative model using random weights for the deep image representation," *Advances in Neural Information Processing Systems*, vol. 29, 2016.

[61] S. Sarp, M. Kuzlu, U. Cali, O. Elma, and O. Guler, "An interpretable solar photovoltaic power generation forecasting approach using an explainable artificial intelligence tool," in *2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, 2021, pp. 1–5.

[62] R. Goebel, A. Chander, K. Holzinger, F. Lecue, Z. Akata, S. Stumpf, P. Kieseberg, and A. Holzinger, "Explainable ai: the new 42?" in *International cross-domain conference for machine learning and knowledge extraction*. Springer, 2018, pp. 295–303.

[63] D. Doran, S. Schulz, and T. R. Besold, "What does explainable ai really mean? a new conceptualization of perspectives," *arXiv preprint arXiv:1710.00794*, 2017.

[64] S. M. McKinney, M. Sieniek, V. Godbole, J. Godwin, N. Antropova, H. Ashrafian, T. Back, M. Chesus, G. S. Corrado, A. Darzi *et al.*, "Interna-

tional evaluation of an ai system for breast cancer screening," *Nature*, vol. 577, no. 7788, pp. 89–94, 2020.

[65] T. De, P. Giri, A. Mevawala, R. Nemani, and A. Deo, "Explainable ai: a hybrid approach to generate human-interpretable explanation for deep learning prediction," *Procedia Computer Science*, vol. 168, pp. 40–48, 2020.

[66] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, "Explainable ai for trees: From local explanations to global understanding," *arXiv preprint arXiv:1905.04610*, 2019.

[67] L. Hulstaert, "Interpreting machine learning models," Feb 2018. [Online]. Available: https://towardsdatascience.com/interpretability-in-machine-learning-70c30694a05f

[68] G. Vilone and L. Longo, "Explainable artificial intelligence: a systematic review," *arXiv preprint arXiv:2006.00093*, 2020.

[69] M. T. Ribeiro, S. Singh, and C. Guestrin, "Model-agnostic interpretability of machine learning," *arXiv preprint arXiv:1606.05386*, 2016.

[70] S. Saha, "Local interpretable model-agnostic explanations (lime)-the eli5 way," Aug 2019. [Online]. Available: https://medium.com/intel-student-ambassadors/local-interpretable-model-agnostic-explanations-lime-the-eli5-way-b4fd61363a5e

[71] M. T. Ribeiro, S. Singh, and C. Guestrin, ""why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135–1144.

[72] P. Struss, "Model-based and qualitative reasoning: An introduction," *Annals of Mathematics and Artificial Intelligence*, vol. 19, no. 3, pp. 355–381, 1997.

[73] R. C. Fong and A. Vedaldi, "Interpretable explanations of black boxes by meaningful perturbation," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 3429–3437.

[74] A. A. Mehl, B. Schneider Jr, F. K. Schneider, and B. H. K. D. Carvalho, "Measurement of wound area for early analysis of the scar predictive factor," *Revista Latino-Americana de Enfermagem*, vol. 28, 2020.

[75] K. Yu, M. N. Syed, E. Bernardis, and J. M. Gelfand, "Machine learning applications in the evaluation and management of psoriasis: A systematic review," *Journal of psoriasis and psoriatic arthritis*, vol. 5, no. 4, pp. 147–159, 2020.

[76] M. Salvaris, D. Dean, and W. H. Tok, "Deep learning with azure," *Building and Deploying Artificial Intelligence Solutions on Microsoft AI Platform, Apress*, 2018.

[77] T. Pinetz, J. Ruisz, and D. Soukup, "Actual impact of gan augmentation on cnn classification performance." in *ICPRAM*, 2019, pp. 15–23.

[78] L. Mescheder, A. Geiger, and S. Nowozin, "Which training methods for gans do actually converge?" in *International conference on machine learning.* PMLR, 2018, pp. 3481–3490.

[79] S. R. Young, D. C. Rose, T. P. Karnowski, S.-H. Lim, and R. M. Patton, "Optimizing deep learning hyper-parameters through an evolutionary algorithm," in *Proceedings of the workshop on machine learning in high-performance computing environments*, 2015, pp. 1–5.

[80] N. Reimers and I. Gurevych, "Optimal hyperparameters for deep lstm-networks for sequence labeling tasks," *arXiv preprint arXiv:1707.06799*, 2017.

[81] Y. Roh, G. Heo, and S. E. Whang, "A survey on data collection for machine learning: a big data-ai integration perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 4, pp. 1328–1347, 2019.

[82] L. Martin, M. Hutchens, C. Hawkins, and A. Radnov, "How much do clinical trials cost," *Nat Rev Drug Discov*, vol. 16, no. 6, pp. 381–382, 2017.

[83] P. Kostkova, H. Brewer, S. De Lusignan, E. Fottrell, B. Goldacre, G. Hart, P. Koczan, P. Knight, C. Marsolier, R. A. McKendry *et al.*, "Who owns the data? open data for healthcare," *Frontiers in public health*, vol. 4, p. 7, 2016.

[84] T. W. K. Poon and M. R. Friesen, "Algorithms for size and color detection of smartphone images of chronic wounds for healthcare applications," *IEEE Access*, vol. 3, pp. 1799–1808, 2015.

[85] F. J. Veredas, R. M. Luque-Baena, F. J. Martín-Santos, J. C. Morilla-Herrera, and L. Morente, "Wound image evaluation with machine learning," *Neurocomputing*, vol. 164, pp. 112–122, 2015.

[86] J. E. Thatcher, J. J. Squiers, S. C. Kanick, D. R. King, Y. Lu, Y. Wang, R. Mohan, E. W. Sellke, and J. M. DiMaio, "Imaging techniques for clinical burn assessment with a focus on multispectral imaging," *Advances in wound care*, vol. 5, no. 8, pp. 360–378, 2016.

[87] D. Filko, R. Cupec, and E. K. Nyarko, "Wound measurement by rgb-d camera," *Machine Vision and Applications*, vol. 29, no. 4, pp. 633–654, 2018.

[88] N. PHOLBERDEE *et al.*, "Wound-region segmentation from image by using deep learning and various data augmentation methods," Ph.D. dissertation, Silpakorn University, 2019.

[89] D. Y. Chino, L. C. Scabora, M. T. Cazzolato, A. E. Jorge, C. Traina-Jr, and A. J. Traina, "Segmenting skin ulcers and measuring the wound area using deep convolutional networks," *Computer methods and programs in biomedicine*, vol. 191, p. 105376, 2020.

[90] S. Zahia, D. Sierra-Sosa, B. Garcia-Zapirain, and A. Elmaghraby, "Tissue classification and segmentation of pressure injuries using convolutional neural networks," *Computer methods and programs in biomedicine*, vol. 159, pp. 51–58, 2018.

[91] C. Chakraborty, "Computational approach for chronic wound tissue characterization," *Informatics in Medicine Unlocked*, vol. 17, p. 100162, 2019.

[92] H. Lu, B. Li, J. Zhu, Y. Li, Y. Li, X. Xu, L. He, X. Li, J. Li, and S. Serikawa, "Wound intensity correction and segmentation with convolutional neural networks," *Concurrency and computation: practice and experience*, vol. 29, no. 6, p. e3927, 2017.

[93] V. Godeiro, J. S. Neto, B. Carvalho, B. Santana, J. Ferraz, and R. Gama, "Chronic wound tissue classification using convolutional networks and color space reduction," in *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*.   IEEE, 2018, pp. 1–6.

[94] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*.   Springer, 2015, pp. 234–241.

[95] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[96] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[97] F. Li, C. Wang, X. Liu, Y. Peng, and S. Jin, "A composite model of wound segmentation based on traditional methods and deep neural networks," *Computational intelligence and neuroscience*, vol. 2018, 2018.

[98] F. Li, C. Wang, Y. Peng, Y. Yuan, and S. Jin, "Wound segmentation network based on location information enhancement," *IEEE Access*, vol. 7, pp. 87 223–87 232, 2019.

[99] P. Gholami, M. A. Ahmadi-Pajouh, N. Abolftahi, G. Hamarneh, and M. Kayvanrad, "Segmentation and measurement of chronic wounds for bioprinting," *IEEE journal of biomedical and health informatics*, vol. 22, no. 4, pp. 1269–1277, 2017.

[100] A. Khalil, M. Elmogy, M. Ghazal, C. Burns, and A. El-Baz, "Chronic wound healing assessment system based on different features modalities and non-negative matrix factorization (nmf) feature reduction," *IEEE Access*, vol. 7, pp. 80 110–80 121, 2019.

[101] J. Zhao, M. Mathieu, and Y. LeCun, "Energy-based generative adversarial network," *arXiv preprint arXiv:1609.03126*, 2016.

[102] H. Huang, P. S. Yu, and C. Wang, "An introduction to image synthesis with generative adversarial nets," *arXiv preprint arXiv:1803.04469*, 2018.

[103] A. Dosovitskiy and T. Brox, "Generating images with perceptual similarity metrics based on deep networks," *Advances in neural information processing systems*, vol. 29, 2016.

[104] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.

[105] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[106] H. Wu, S. Zheng, J. Zhang, and K. Huang, "Gp-gan: Towards realistic high-resolution image blending," in *Proceedings of the 27th ACM international conference on multimedia*, 2019, pp. 2487–2495.

[107] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[108] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE transactions on pattern analysis and machine intelligence*, 2021.

[109] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.

[110] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[111] M. C. Bakkay, H. A. Rashwan, H. Salmane, L. Khoudour, D. Puig, and Y. Ruichek, "Bscgan: Deep background subtraction with conditional generative adversarial networks," in *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018, pp. 4018–4022.

[112] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," *Advances in neural information processing systems*, vol. 29, 2016.

[113] S. Thomas, "Medetec wound database: Stock pictures of wounds," Feb 2014. [Online]. Available: http://medetec.co.uk/files/medetec-image-databases.html

[114] I. Goodfellow, "Nips 2016 tutorial: Generative adversarial networks," *arXiv preprint arXiv:1701.00160*, 2016.

[115] C. Wang, D. Anisuzzaman, V. Williamson, M. K. Dhar, B. Rostami, J. Niezgoda, S. Gopalakrishnan, and Z. Yu, "Fully automatic wound segmentation with deep convolutional neural networks," *Scientific reports*, vol. 10, no. 1, pp. 1–9, 2020.

[116] B. R. Gaines, "Knowledge acquisition: Past, present and future," *International Journal of Human-Computer Studies*, vol. 71, no. 2, pp. 135–156, 2013.

[117] E. S. Brunette, R. C. Flemmer, and C. L. Flemmer, "A review of artificial intelligence," in *2009 4th International Conference on Autonomous Robots and Agents*. Ieee, 2009, pp. 385–392.

[118] A. Jiménez-Sánchez, S. Albarqouni, and D. Mateus, "Capsule networks against medical imaging data challenges," in *Intravascular imaging and computer assisted stenting and large-scale annotation of biomedical data and expert label synthesis*. Springer, 2018, pp. 150–160.

[119] I. Pavlović, T. Kern, and D. Miklavčič, "Comparison of paper-based and electronic data collection process in clinical trials: costs simulation study," *Contemporary clinical trials*, vol. 30, no. 4, pp. 300–316, 2009.

[120] N. Japkowicz and S. Stephen, "The class imbalance problem: A systematic study," *Intelligent data analysis*, vol. 6, no. 5, pp. 429–449, 2002.

[121] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *2018 international interdisciplinary PhD workshop (IIPhDW)*. IEEE, 2018, pp. 117–122.

[122] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: when to warp?" in *2016 international conference on digital image computing: techniques and applications (DICTA)*. IEEE, 2016, pp. 1–6.

[123] H. Hosseini, B. Xiao, M. Jaiswal, and R. Poovendran, "On the limitation of convolutional neural networks in recognizing negative images," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 2017, pp. 352–358.

[124] C. Grajeda, F. Breitinger, and I. Baggili, "Availability of datasets for digital forensics–and what is missing," *Digital Investigation*, vol. 22, pp. S94–S105, 2017.

[125] Y. Luo and B.-L. Lu, "Eeg data augmentation for emotion recognition using a conditional wasserstein gan," in *2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2018, pp. 2535–2538.

[126] C. Han, H. Hayashi, L. Rundo, R. Araki, W. Shimoda, S. Muramatsu, Y. Furukawa, G. Mauri, and H. Nakayama, "Gan-based synthetic brain mr image

generation," in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018, pp. 734–738.

[127] A. Antoniou, A. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," *arXiv preprint arXiv:1711.04340*, 2017.

[128] C. S. Nielsen, "Improving image classification through generative data augmentation," Master's thesis, Schulich School of Engineering, 2019.

[129] O. Klang and M. Carlberg, "Blood cell data augmentation using deep learning methods," *Master's Theses in Mathematical Sciences*, 2020.

[130] G. Mariani, F. Scheidegger, R. Istrate, C. Bekas, and C. Malossi, "Bagan: Data augmentation with balancing gan," *arXiv preprint arXiv:1803.09655*, 2018.

[131] C. Wang, Z. Yu, H. Zheng, N. Wang, and B. Zheng, "Cgan-plankton: Towards large-scale imbalanced class generation and fine-grained classification," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 855–859.

[132] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using gan for improved liver lesion classification," in *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE, 2018, pp. 289–293.

[133] D. Prokopenko, J. V. Stadelmann, H. Schulz, S. Renisch, and D. V. Dylov, "Unpaired synthetic image generation in radiology using gans," in *Workshop on Artificial Intelligence in Radiation Therapy*. Springer, 2019, pp. 94–101.

[134] Z. Yi, H. Zhang, P. Tan, and M. Gong, "Dualgan: Unsupervised dual learning for image-to-image translation," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2849–2857.

[135] K. Kim and H. Myung, "Autoencoder-combined generative adversarial networks for synthetic image data generation and detection of jellyfish swarm," *IEEE Access*, vol. 6, pp. 54 207–54 214, 2018.

[136] P. M. Burlina, N. Joshi, K. D. Pacheco, T. A. Liu, and N. M. Bressler, "Assessment of deep generative models for high-resolution synthetic retinal image generation of age-related macular degeneration," *JAMA ophthalmology*, vol. 137, no. 3, pp. 258–264, 2019.

[137] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv preprint arXiv:1710.10196*, 2017.

[138] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[139] J. D. Bills, S. J. Berriman, D. L. Noble, L. A. Lavery, and K. E. Davis, "Pilot study to evaluate a novel three-dimensional wound measurement device," *International Wound Journal*, vol. 13, no. 6, pp. 1372–1377, 2016.

[140] M. Srivastava, S. Pallavi, S. Chandra, and G. Geetha, "Comparison of optimizers implemented in generative adversarial network (gan)," *International Journal of Pure and Applied Mathematics*, vol. 119, no. 12, pp. 16 831–16 835, 2018.

[141] J. Löhdefink, F. Hüger, P. Schlicht, and T. Fingscheidt, "Scalar and vector quantization for learned image compression: A study on the effects of mse and gan loss in various spaces," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*.   IEEE, 2020, pp. 1–8.

[142] X. Zhang, S. Karaman, and S.-F. Chang, "Detecting and simulating artifacts in gan fake images," in *2019 IEEE International Workshop on Information Forensics and Security (WIFS)*. IEEE, 2019, pp. 1–6.

[143] J. Gauthier, "Conditional generative adversarial nets for convolutional face generation," *Class project for Stanford CS231N: convolutional neural networks for visual recognition, Winter semester*, vol. 2014, no. 5, p. 2, 2014.

[144] A. EL Azzaoui, T. W. Kim, V. Loia, and J. H. Park, "Blockchain-based secure digital twin framework for smart healthy city," in *Advanced Multimedia and Ubiquitous Engineering*. Springer, 2021, pp. 107–113.

[145] J. Zhang, L. Li, G. Lin, D. Fang, Y. Tai, and J. Huang, "Cyber resilience in healthcare digital twin on lung cancer," *IEEE Access*, vol. 8, pp. 201 900– 201 913, 2020.

[146] Y. Liu, L. Zhang, Y. Yang, L. Zhou, L. Ren, F. Wang, R. Liu, Z. Pang, and M. J. Deen, "A novel cloud-based framework for the elderly healthcare services using digital twin," *IEEE access*, vol. 7, pp. 49 088–49 101, 2019.

[147] H. Elayan, M. Aloqaily, and M. Guizani, "Digital twin for intelligent context-aware iot healthcare systems," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16 749–16 757, 2021.

[148] R. Martinez-Velazquez, R. Gamez, and A. El Saddik, "Cardio twin: A digital twin of the human heart running on the edge," in *2019 IEEE international symposium on medical measurements and applications (MeMeA)*. IEEE, 2019, pp. 1–6.

[149] A. Croatti, M. Gabellini, S. Montagna, and A. Ricci, "On the integration of agents and digital twins in healthcare," *Journal of Medical Systems*, vol. 44, no. 9, pp. 1–8, 2020.

[150] W. Shengli, "Is human digital twin possible?" *Computer Methods and Programs in Biomedicine Update*, vol. 1, p. 100014, 2021.

[151] T. Erol, A. F. Mendi, and D. Doğan, "The digital twin revolution in healthcare," in *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2020, pp. 1–7.

[152] N. Graves, F. Birrell, and M. Whitby, "Effect of pressure ulcers on length of hospital stay," *Infection Control & Hospital Epidemiology*, vol. 26, no. 3, pp. 293–297, 2005.

[153] L. A. Borojeny, A. N. Albatineh, A. H. Dehkordi, and R. G. Gheshlagh, "The incidence of pressure ulcers and its associations in different wards of the hospital: A systematic review and meta-analysis," *International Journal of Preventive Medicine*, vol. 11, 2020.

[154] S. A. Tabish, "Is diabetes becoming the biggest epidemic of the twenty-first century?" *International Journal of health sciences*, vol. 1, no. 2, p. V, 2007.

[155] E. L. Feldman, B. C. Callaghan, R. Pop-Busui, D. W. Zochodne, D. E. Wright, D. L. Bennett, V. Bril, J. W. Russell, and V. Viswanathan, "Diabetic neuropathy," *Nature reviews Disease primers*, vol. 5, no. 1, pp. 1–18, 2019.

[156] P. R. Cavanagh, B. A. Lipsky, A. W. Bradbury, and G. Botek, "Treatment for diabetic foot ulcers," *The Lancet*, vol. 366, no. 9498, pp. 1725–1735, 2005.

[157] A. Gordois, P. Scuffham, A. Shearer, A. Oglesby, and J. A. Tobian, "The health care costs of diabetic peripheral neuropathy in the us," *Diabetes care*, vol. 26, no. 6, pp. 1790–1795, 2003.

[158] N. Greer, N. A. Foman, R. MacDonald, J. Dorrian, P. Fitzgerald, I. Rutks, and T. J. Wilt, "Advanced wound care therapies for nonhealing diabetic, venous,

and arterial ulcers: a systematic review," *Annals of internal medicine*, vol. 159, no. 8, pp. 532–542, 2013.

[159] A. Oryan, E. Alemzadeh, and A. Moshiri, "Burn wound healing: present concepts, treatment strategies and future directions," *Journal of wound care*, vol. 26, no. 1, pp. 5–19, 2017.

[160] J. Z. Cuzzell, "Wound care forum the new ryb color code," *AJN The American Journal of Nursing*, vol. 88, no. 10, pp. 1342–1346, 1988.

[161] D. Krasner, "Wound care how to use the red-yellow-black system," *The American journal of nursing*, vol. 95, no. 5, pp. 44–47, 1995.

[162] C. Shi, C. Wang, H. Liu, Q. Li, R. Li, Y. Zhang, Y. Liu, Y. Shao, and J. Wang, "Selection of appropriate wound dressing for various wounds," *Frontiers in bioengineering and biotechnology*, vol. 8, p. 182, 2020.

[163] L. B. Jørgensen, J. A. Sørensen, G. B. Jemec, and K. B. Yderstræde, "Methods to assess area and volume of wounds–a systematic review," *International wound journal*, vol. 13, no. 4, pp. 540–553, 2016.

[164] M. G. Sowa, W.-C. Kuo, A. C. Ko, and D. G. Armstrong, "Review of near-infrared methods for wound assessment," *Journal of biomedical optics*, vol. 21, no. 9, p. 091304, 2016.

[165] Y. Zhang, Y. Sun, R. Jin, K. Lin, and W. Liu, "High-performance isolation computing technology for smart iot healthcare in cloud environments," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16 872–16 879, 2021.

[166] R. Zhang, G. Cavallaro, and J. Jitsev, "Super-resolution of large volumes of sentinel-2 images with high performance distributed deep learning," in *IGARSS*

2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, 2020, pp. 617–620.

[167] P. Raj, "Empowering digital twins with blockchain," in *Advances in Computers*. Elsevier, 2021, vol. 121, pp. 267–283.

[168] H. R. Hasan, K. Salah, R. Jayaraman, M. Omar, I. Yaqoob, S. Pesic, T. Taylor, and D. Boscovic, "A blockchain-based approach for the creation of digital twins," *IEEE Access*, vol. 8, pp. 34 113–34 126, 2020.

[169] M. Mahmoudi and L. J. Gould, "Opportunities and challenges of the management of chronic wounds: A multidisciplinary viewpoint," *Chronic Wound Care Management and Research*, vol. 7, p. 27, 2020.

[170] "MS Windows NT kernel description," http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm, accessed: 2010-09-30.

[171] P. Foltynski, J. Wojcicki, P. Ladyzynski, and S. Sabalinska, "A comparison of three techniques for wound area measurement," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013*. Springer, 2014, pp. 1071–1074.

[172] P. Sheehan, P. Jones, A. Caselli, J. M. Giurini, and A. Veves, "Percent change in wound area of diabetic foot ulcers over a 4-week period is a robust predictor of complete healing in a 12-week prospective trial," *Diabetes care*, vol. 26, no. 6, pp. 1879–1882, 2003.

[173] A. Prodan, M. Rusu, R. Câmpean, and R. Prodan, "Artificial intelligence for wound image understanding." in *ICEIS (2)*, 2008, pp. 213–218.

[174] M. Flanagan, "Improving accuracy of wound measurement in clinical practice." *Ostomy/wound management*, vol. 49, no. 10, pp. 28–40, 2003.

[175] H. Vermeulen, D. T. Ubbink, S. M. Schreuder, and M. J. Lubbers, "Inter-and intra-observer (dis) agreement among nurses and doctors to classify colour and exudation of open surgical wounds according to the red-yellow-black scheme," *Journal of Clinical Nursing*, vol. 16, no. 7, pp. 1270–1277, 2007.

[176] F. Buntinx, H. Beckers, A. Briers, G. De Keyser, M. Flour, G. Nissen, T. Raskin, and H. De Vet, "Inter-observer variation in the assessment of skin ulceration," *Journal of Wound Care*, vol. 5, no. 4, pp. 166–170, 1996.

[177] K. Jung, S. Covington, C. K. Sen, M. Januszyk, R. S. Kirsner, G. C. Gurtner, and N. H. Shah, "Rapid identification of slow healing wounds," *Wound Repair and Regeneration*, vol. 24, no. 1, pp. 181–188, 2016.

[178] J. K. Jopling, B. C. Pridgen, and S. Yeung, "Setting assessment standards for artificial intelligence computer vision wound annotations," *JAMA Network Open*, vol. 4, no. 5, pp. e217 851–e217 851, 2021.

[179] K. G. Harding, "Trends in wound care–the development of a specialty," *International Wound Journal*, vol. 3, no. 3, p. 147, 2006.

[180] K. Järbrink, G. Ni, H. Sönnergren, A. Schmidtchen, C. Pang, R. Bajpai, and J. Car, "The humanistic and economic burden of chronic wounds: a protocol for a systematic review," *Systematic reviews*, vol. 6, no. 1, pp. 1–7, 2017.

[181] D. Queen, "Artificial intelligence and machine learning in wound care—the wounded machine!" *International Wound Journal*, vol. 16, no. 2, p. 311, 2019.

[182] S. Vogel, J. Richter, S. Wache, K. Pischek-Koch, S. Auchter, S. Zebbities, K. Güttler, U. Hübner, M. Pryzsucha, J. Hüsers *et al.*, "Evaluation of a clinical

decision support system in the domain of chronic wound management," *Studies in Health Technology and Informatics*, vol. 281, pp. 535–539, 2021.

[183] R. S. Howell, H. H. Liu, A. A. Khan, J. S. Woods, L. J. Lin, M. Saxena, H. Saxena, M. Castellano, P. Petrone, E. Slone *et al.*, "Development of a method for clinical evaluation of artificial intelligence–based digital wound assessment tools," *JAMA network open*, vol. 4, no. 5, pp. e217 234–e217 234, 2021.

[184] C. Wang, E. Shirzaei Sani, and W. Gao, "Wearable bioelectronics for chronic wound management," *Advanced Functional Materials*, vol. 32, no. 17, p. 2111022, 2022.

[185] N. Ohura, R. Mitsuno, M. Sakisaka, Y. Terabe, Y. Morishige, A. Uchiyama, T. Okoshi, I. Shinji, and A. Takushima, "Convolutional neural networks for wound detection: the role of artificial intelligence in wound care," *Journal of wound care*, vol. 28, no. Sup10, pp. S13–S24, 2019.

[186] G. Zoppo, F. Marrone, M. Pittarello, M. Farina, A. Uberti, D. Demarchi, J. Secco, F. Corinto, and E. Ricci, "Ai technology for remote clinical assessment and monitoring," *Journal of wound care*, vol. 29, no. 12, pp. 692–706, 2020.

[187] L. Wang, P. C. Pedersen, D. M. Strong, B. Tulu, E. Agu, and R. Ignotz, "Smartphone-based wound assessment system for patients with diabetes," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 477–488, 2014.

[188] T. Chitra, C. Sundar, and S. GOPALAKRISHNAN, "Investigation and classification of chronic wound tissue images using random forest algorithm (rf)," *International Journal of Nonlinear Analysis and Applications*, vol. 13, no. 1, pp. 643–651, 2022.

[189] M. Mashaly, "Connecting the twins: A review on digital twin technology & its networking requirements," *Procedia Computer Science*, vol. 184, pp. 299–305, 2021.

[190] H. Aydemir, U. Zengin, and U. Durak, "The digital twin paradigm for aircraft review and outlook," in *AIAA Scitech 2020 Forum*, 2020, p. 0553.

[191] J. I. Jimenez, H. Jahankhani, and S. Kendzierskyj, "Health care in the cyberspace: Medical cyber-physical system and digital twin challenges," in *Digital twin technologies and smart cities*. Springer, 2020, pp. 79–92.

[192] C. Angulo, J. A. Ortega, and L. Gonzalez-Abril, "Towards a healthcare digital twin 1," in *Artificial Intelligence Research and Development*. IOS Press, 2019, pp. 312–315.

[193] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges and open research," *IEEE access*, vol. 8, pp. 108 952–108 971, 2020.

[194] Y. Feng, X. Chen, and J. Zhao, "Create the individualized digital twin for noninvasive precise pulmonary healthcare," *Significances Bioengineering & Biosciences*, vol. 1, no. 2, 2018.

[195] C. Patrone, G. Galli, and R. Revetria, "A state of the art of digital twin and simulation supported by data mining in the healthcare sector," in *Advancing Technology Industrialization Through Intelligent Software Methodologies, Tools and Techniques*. IOS Press, 2019, pp. 605–615.

[196] C. Patrone, M. Lattuada, G. Galli, and R. Revetria, "The role of internet of things and digital twin in healthcare digitalization process," in *The World Congress on Engineering and Computer Science*. Springer, 2018, pp. 30–37.

[197] J. Corral-Acero, F. Margara, M. Marciniak, C. Rodero, F. Loncaric, Y. Feng, A. Gilbert, J. F. Fernandes, H. A. Bukhari, A. Wajdan *et al.*, "The 'digital twin'to enable the vision of precision cardiology," *European heart journal*, vol. 41, no. 48, pp. 4556–4564, 2020.

[198] B. Björnsson, C. Borrebaeck, N. Elander, T. Gasslander, D. R. Gawel, M. Gustafsson, R. Jörnsten, E. J. Lee, X. Li, S. Lilja *et al.*, "Digital twins to personalize medicine," *Genome medicine*, vol. 12, no. 1, pp. 1–4, 2020.

[199] B. Barnes and T. Daim, "Information security maturity model for healthcare organizations in the united states," *IEEE Transactions on Engineering Management*, 2022.

[200] J. H. Brice, K. D. Friend, and T. R. Delbridge, "Accuracy of ems-recorded patient demographic data," *Prehospital Emergency Care*, vol. 12, no. 2, pp. 187–191, 2008.

[201] C. E. Fife, K. A. Eckert, and M. J. Carter, "Publicly reported wound healing rates: the fantasy and the reality," *Advances in wound care*, vol. 7, no. 3, pp. 77–94, 2018.

[202] S. K. Cho, S. Mattke, H. Gordon, M. Sheridan, and W. Ennis, "Development of a model to predict healing of chronic wounds within 12 weeks," *Advances in wound care*, vol. 9, no. 9, pp. 516–524, 2020.

[203] D. Culter, "What statisticians should know about machine learning," in *Proceedings of 2017 SAS Global Forum*, 2017.

[204] Y. Zhang, Y. Weng, and J. Lund, "Applications of explainable artificial intelligence in diagnosis and surgery," *Diagnostics*, vol. 12, no. 2, p. 237, 2022.

[205] S. Sarp, M. Kuzlu, E. Wilson, U. Cali, and O. Guler, "A highly transparent and explainable artificial intelligence tool for chronic wound classification: Xai-cwc," 2021.

[206] R. Avrahami, J. Rosenblum, M. Gazes, S. Rosenblum, and L. Litman, "The effect of combined ultrasound and electric field stimulation on wound healing in chronic ulcerations." *Wounds: a compendium of clinical research and practice*, vol. 27, no. 7, pp. 199–208, 2015.

[207] M. Abdelmalek and J. Spencer, "Retinoids and wound healing," *Dermatologic surgery*, vol. 32, no. 10, pp. 1219–1230, 2006.

[208] M. Cardinal, D. E. Eisenbud, T. Phillips, and K. Harding, "Early healing rates and wound area measurements are reliable predictors of later complete wound closure," *Wound repair and regeneration*, vol. 16, no. 1, pp. 19–22, 2008.

[209] E. Rodriguez-Merchan, "Surgical wound healing in bleeding disorders," *Haemophilia*, vol. 18, no. 4, pp. 487–490, 2012.

[210] J. E. Grey, S. Enoch, and K. G. Harding, "Wound assessment," *Bmj*, vol. 332, no. 7536, pp. 285–288, 2006.

[211] G. Saiko, P. Lombardi, Y. Au, D. Queen, D. Armstrong, and K. Harding, "Hyperspectral imaging in wound care: a systematic review," *International Wound Journal*, vol. 17, no. 6, pp. 1840–1856, 2020.

Vita

Salih Sarp

Research Interests

Deep Learning, Explainable AI, Forecasting, AI Education, Image Processing, AI Use in Healthcare and Digital Twin, Embedded Systems, Application of Generative Machine Learning Models, Use of AI in Wireless Networks.

Education

- Virginia Commonwealth University, , Richmond, VA, USA,　　　August 2022
  Ph.D. in Electrical and Computer Engineering

- George Washington University , Washington, DC , USA,　　　May 2018
  M.Sc. in Electrical and Computer Engineering

- Dogus University, Istanbul, Turkey,　　　February 2014
  B.S. in Electronics and Communications Engineering

Publications

Journal Publications:

[J1] S. Sarp, M. Kuzlu, Y. Zhao, F. O. Catak, U. Cali, and O. Guler, "Chronic wound management in digital twin," in Digital Twin Driven Intelligent Systems. Springer, 2022. (Under review).

[J2] S. Sarp, M. Kuzlu, Y. Zhao, and O. Guler, "Digital twin in healthcare: A study for chronic wound management," IEEE Journal of Biomedical and Health Informatics, 2022. (Under review).

[J3] F. Wibawa, F. O. Catak, S. Sarp, and M. Kuzlu, "Bfv-based homomorphic encryption for privacy-preserving cnn models," Cryptography, vol. 6, no. 3, p. 34, 2022.

[J4] M. Kuzlu, S. Sarp, F. O. Catak, U. Cali, Y. Zhao, O. Elma, and O. Guler, "Analysis of deceptive data attacks with adversarial machine learning for solar photovoltaic power generation forecasting," Electrical Engineering, pp. 1–9, 2022.

[J5] S. Sarp, M. Kuzlu, E. Wilson, U. Cali, and O. Guler, "The enlightening role of explainable artificial intelligence in chronic wound classification," Electronics, vol. 10, no. 12, p. 1406, 2021.

[J6] S. Sarp, M. Kuzlu, M. Pipattanasomporn, and O. Guler, "Simultaneous wound border segmentation and tissue classification using a conditional generative adversarial network," Journal of Engineering, vol. 2021, no. 3, 2021.

[J7] S. Sarp, M. Kuzlu, E. Wilson, and O. Guler, "Wg2an: Synthetic wound image generation using generative adversarial network," The Journal of Engineering, vol. 2021, no. 5, pp. 286–294, 2021.

[J8] S. Sarp, M. Kuzlu, Y. Zhao, M. Cetin, and O. Guler, "A comparison of deep learning algorithms on image data for detecting floodwater on roadways," Computer Science and Information Systems, pp. 58–58, 2021

Peer-reviewed Conference Papers:

[C1] H. Tang, S. Sarp, Y. Zhao, W. Wang, C. Xin, "Security and Threats of Intelligent Reflecting Surface Assisted Wireless Communications," International Conference on Computer Communications and Networks (ICCCN), 2022.

[C2] F. Wibawa, F. O. Catak, M. Kuzlu, S. Sarp, and U. Cali, "Homomorphic encryption and federated learning based privacy-preserving cnn training: Covid-19 detection use-case," in EICC 2022: Proccedings of the European Interdisciplinary Cybersecurity Conference, 2022, pp. 85–90. 148

[C3] S. Balci, G. M. Demirci, H. Demirhan, and S. Sarp, "Sentiment analysis using state of the art machine learning techniques," in Conference on Multimedia, Interaction, Design and Innovation. Springer, 2022, pp. 34–42.

[C4] S. Sarp, H. Tang, and Y. Zhao, "Use of intelligent reflecting surfaces for and against wireless communication security," in 2021 IEEE 4th 5G World Forum (5GWF). IEEE, 2021, pp. 374–377.

[C5] S. Sarp, M. Kuzlu, U. Cali, O. Elma, and O. Guler, "An interpretable solar photovoltaic power generation forecasting approach using an explainable artificial intelligence tool," in 2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT). IEEE, 2021, pp. 1–5.

[C6] S. Sarp, H. Demirhan, A. Akca, F. Balki, and S. Ceylan, "Work in progress: Activating computational thinking by engineering and coding activities through distance education," in 2021 ASEE Virtual Annual Conference Content Access, 2021.

[C7] M. Kuzlu, S. Sarp, M. Pipattanasomporn, and U. Cali, "Realizing the potential of blockchain technology in smart grid applications," in 2020 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT). IEEE, 2020, pp. 1–5.

[C8] S. Sarp, M. Kuzlu, M. Cetin, C. Sazara, and O. Guler, "Detecting floodwater on roadways from image data using mask-r-cnn," in 2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA). IEEE, 2020, pp. 1–6.

[C9] V. M. Jovanovic, M. Kuzlu, O. Popescu, A. R. Badawi, D. K. Marshall, S. Sarp, S. Tsouganatou, P. J. Katsioloudis, L. Vahala, and H. Wu, "An initial look into the computer science and cybersecurity pathways project for career and technical education curricula," 2020.

[C10] F. Balki, H. Demirhan, and S. Sarp, "Neural machine translation for Turkish to English using deep learning," in Conference on Multimedia, Interaction, Design and Innovation. Springer, 2020, pp. 3–9.

Honors And Awards

VCU College of Engineering, Outstanding Graduate Student Service Award   2022
VCU ECE Department, Outstanding Graduate Research Assistant Award   2022
Turkish Ministry of Education Scholarship to pursue M.Sc. and Ph.D.   2014
Turkish Ministry of Education Scholarship to pursue B.Sc.   2007