

POLITECNICO DI TORINO



**Politecnico
di Torino**

Master Degree course in Biomedical Engineering

Master Degree Thesis

**Neuromorphic methods for detection,
classification and analysis of chronic
wounds**

Supervisors

Prof. Jacopo SECCO

Eng. Elisabetta SPINAZZOLA

Dr. Filippo BEGARANI (Corporate Supervisor)

Candidate

Sara BECCHI

ACADEMIC YEAR 2023-2024

Acknowledgements

I would like to thank Prof. Jacopo Secco and Eng. Elisabetta Spinazzola for their help and support during these months, which have allowed me to grow as a student and reach an important milestone. I would also like to express my gratitude to Dr. Filippo Begarani for the opportunity and trust given to me.

Abstract

Vulnology is a field focused on treating wounds, an often underestimated health issue. When a chronic ulcer worsens without monitoring it can become chronic in 30 % of cases sometimes leading to non healing wounds and, in severe instances, even death. Hence consistent follow up care is essential to prevent complications. It's crucial to establish protocols in this area and incorporate telemedicine tools to aid healthcare providers. With the advancements in deep learning technology for image analysis neuromorphic systems are now being integrated into devices to assist clinicians during examinations. These systems can automatically analyze wound images using deep learning algorithms reducing the need for contact and lowering the risk of exacerbating the wound condition. It is crucial to assess the capability of these algorithms to identify and classify wounds, and simultaneously determine which medical images are best suited for extracting useful features. After examining the field and considering the methods used, it was decided to evaluate the capabilities of the algorithm through the use of thermal images. This method represents a new approach within the framework of automated assessments to wound management. By accomplishing this task we can offer vulnologists an AI powered tool that streamlines their work process, enhances examination quality and follow up care and fosters efficient healing outcomes for patients.

Table of Contents

1	Introduction	3
1.1	The field of vulnology	3
1.2	Impact of chronic wounds on the health system	4
1.3	Analysis of wounds, scale bars to asses the healing process	6
1.3.1	Wound Bed Preparation [WBP]	6
1.3.2	Photographic Wound Assessment Tool	7
1.3.3	Diabetic ulcers classification	9
1.3.4	General protocol of approaching ulcers	9
2	Imaging and measurment in wound care	13
2.1	Types of imaging techniques	13
2.2	Ultraviolet radiation for bacterial fluorecence	17
2.3	Thermography	18
3	Integration of deep learning in digital imaging techniques	25
3.1	Basic operating principle of a neural network	26
3.2	Different neural network architectures	27
3.3	Applications of deep learning in medical imaging	28
3.3.1	Segmentation networks	28
3.3.2	Detection and classification networks	30
3.3.3	Deep Learning and wound care	33
4	Analysis of thermal images to assess infection status	35
4.1	Dataset	35
4.2	Ground truth generation and preprocessing	36
4.2.1	Construction and Test set	38
4.3	Segmentation and classification of wounds	40
4.3.1	U-Net	40
4.3.2	PSP-ResNet50	42
4.3.3	YOLOv8-seg	43
4.4	Classification of wound status	45
4.5	Metrics for evaluation of performances	46
4.6	Results for segmentation task	47
4.7	Results for classification task	48
4.8	Discussion	51
A	Clinical trial protocol	55

Chapter 1

Introduction

Neuromorphic methods are techniques based on artificial intelligence and thus on neural networks that allow us to automate and thus speed up several processes in healthcare. Neural networks can in fact be used not only for detection tasks but also for segmentation and classification, supporting the physician in clinical decisions when integrated on medical devices. It emerges that the vulnological field can benefit from this type of device as there is a clear need for standardisation and sharing of clinical data. In order to be able to analyse something we need to see it, which is why the subject of imaging in wound care is also explored. The type of medical images used is indeed essential in deciding the type of feature to be extracted by AI, in order to provide the physician with the most useful information for diagnosis and treatment and to digitise this area of medicine that is often left aside.

1.1 The field of vulnology

Vulnology is the field of medicine that deals with wounds and lacerations of the skin, derived in fact from the Latin word 'vulnus' meaning wound. Skin lacerations, contrary to popular belief, are clinically widespread in all age groups with a marked prevalence from the age of 60 onwards. Under this branch of study we find skin ulcers, a condition with many possible causes that if left untreated can lead to amputation of the affected body part or in extreme cases death from infection. Skin ulcers can be linked to neoplasms, they can be of diabetic origin or linked to a sedentary condition where they take the name of pressure or decubitus ulcers. Then there are vascular ulcers, caused by a poor outflow capacity of the blood vessels affecting the area where the lesion forms. In some cases a chronic wound may also originate post-operatively, due to errors during surgery or factors that slow down complete healing such as bacterial infections. Chronic wounds are divided into different types, depending also on the causes of the pathology. One of the complication of diabete is the diabetic foot ulceration which is a kind of lesion that expands more underneath the skin in respect of the surface. Because of the the type of lesion's structure there is a major risk of infection and chronic vascular changes in the wound area. These factors can compromise the healing process.

Vascular alterations can, in the case of venous recirculation dysfunction, lead to the formation of venous ulcers, which cause oedema and problems with normal venous return. Poor perfusion to a certain district, often in the limbs, leads to the formation of arterial ulcers, wounds that very often do not heal and cause permanent infections with the need for subsequent amputation. Another determining factor is sedentariness or being bedridden, which causes bed sores. During the covid period, this type of chronic wounds has seen a dramatic increase, where body areas with protruding bones such as the heels and sacral area are exposed to prolonged and unnatural body

weight pressure with the consequent formation of these types of sores. In a hospital environment it is not unusual to be exposed to non-sterile environments with viruses or bacteria, which if they come into contact with an already present wound can trigger a local infection. This type of wound has pus or abscesses that must be drained to prevent the infection from becoming permanent and spreading through the circulatory system. Post-surgical wounds are also exposed to this danger and may take a long time to heal or reopen.

Incidence of types of ulcers

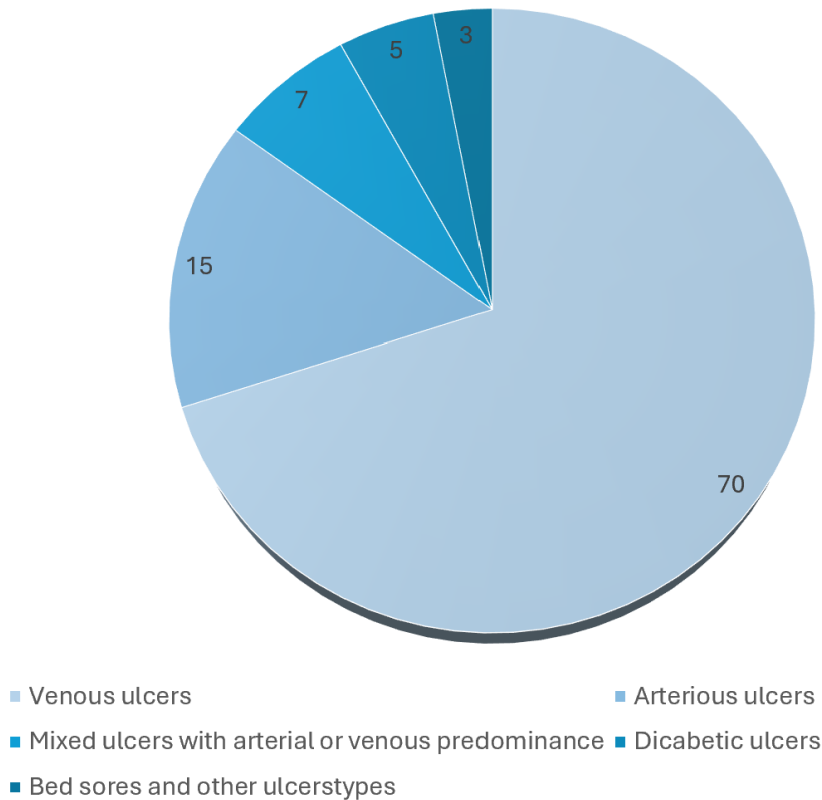


Figure 1.1. Incidence of varois types of ulcers of the lower limbs among the population [1]

1.2 Impact of chronic wounds on the health system

From data available from the literature, it is estimated that 2% of the world's population will develop at least one skin ulcer-related disease during their lifetime [2]. Unfortunately, a healthy lifestyle cannot always prevent the onset of such complications, but there is evidence that intervening early on in an acute phase injury can prevent it from becoming chronic. Total Medicare spending estimates for all wound types ranged from 28.1 to 96.8 billion. Including infection costs, the most expensive estimates were for surgical wounds (11.7, 13.1, and 38.3 billion), followed by

diabetic foot ulcers (6.2, 6.9, and 18.7 billion,). The highest cost estimates in regard to site of service were for hospital outpatients (9.9–35.8 billion), followed by hospital inpatients (5.0–24.3 billion). The data refer to the US federal health insurance Medicare established in 1965 [3]

In Europe alone, The costs in the European population were 2% of the European health budget, which results in a life time cost of 4 billion to European healthcare system [4]. Patients suffering from chronic injury-related diseases number more than 10 million in Europe. This economic problem does not only concern Europe, recent studies worldwide have shown that the costs of hospitalisation and treatment for patients with chronic wounds are not small.

A recent analysis of costs in Canada for the conditions of pressure sores, diabetic foot and lower limb ulcers reported that the average life time cost for individuals hospitalised due to these types of wounds is approximately one billion dollars[5]. Often either because of lack of beds or because the patient's condition is not critical enough to require hospitalisation, home follow-up is opted for. It is essential that the wound is monitored and dressed periodically; it has been shown that early detection of a critical wound can be a determining factor in healing and preventing it from becoming chronic. After 4 weeks there is a 30% greater chance of the injury never healing, a 50% greater chance of potentially losing the limb, and a 50% greater chance of death over the next 5 years[6].

The choice of home care introduces quite a few problems from a logistical point of view: there are few specialists in the field compared to the demand, and home visits may be prevented by external causes (covid). It is therefore necessary to make the processes of home care delivery more efficient, so as to make effective use of resources. One of the solutions to this problem is the introduction of telemedicine, understood as telecare and telemonitoring. It allows patients fewer visits to the hospital, optimising documentation and examination times, and encourages the periodic follow-up that acute injuries need to heal. Telemedicine services are based on the introduction of technological devices to support the clinician's activity for data storage, diagnosis or detection of a given pathology. Data archiving also allows the creation of large databases used to extrapolate knowledge for medical or research purposes.

Managing the large volume of data in medicine has required the implementation of AI (Artificial Intelligence) based technologies. Archiving is not the only use of this type of technology, detection and characterisation can also be implemented after training the algorithms on appropriate datasets to provide support for the physician's work. An automation of certain practices is achieved, resulting in considerable time savings. Another problem sought to be solved by these devices is the variability of inter- and intra-operator judgement: the degree of infection or other parameters may in fact be judged differently between different specialists or even by the same doctor at different times due to situations that generate stress or other concomitant causes.

At the AIUC national congress in Turin in October a survey was carried out by questioning the attending nursing specialists to understand the position of healthcare professionals regarding the introduction of new technologies. The need for hospital-wide centralisation of clinical data on individual patients to keep track of wound trends and applied dressings to ensure continuous treatment of the individual and to improve communication and cooperation between practitioners also emerged among nursing specialists. Respondents were then asked for their opinion on the usefulness of telemedicine in supporting skin ulcer care processes, with a positive response in 100% of the interviews. The aspects found most interesting in the introduction of new technologies to support the diagnosis and treatment of skin ulcers were:

- Objective assessment of the skin ulcer
- Remote monitoring and direct connection with home patients
- Shortening of diagnosis and treatment time Wound evolution history and continuity of care
- Facilitation of communication between clinicians and formation of a multidisciplinary team

- Easy access and sharing of patient data

1.3 Analysis of wounds, scale bars to assess the healing process

In order to outline the best treatment for a skin ulcer, one initially proceeds with the patient's anamnesis, in order to ascertain whether any pathologies are present as a result of which the wound may have been created or worsened.

The next step is the wound visual inspection and analysis of wound-specific parameters

- localisation
- the multiplicity or uniqueness
- the size and morphology of the wound
- the staging of the wound floor
- the appearance of the edges and periulcerous skin (which maintains in relation to the ulcer relations of close interdependence)
- the depth with particular attention to the presence of subdominations, sinuous formations sinuous and fistulous pathways or pus collections

Characterising the wound serves to distinguish what type of injury it is and to understand its status and criticality. Collecting data correlated with this information allows a preliminary assessment of the lesion, which is then contextualized with present complications, such as infection and edema or specific risk factors, such as family history, thrombophlebitis, surgery undergone, and ischemia present in the patient's medical history. There are also several scales in the literature to classify the state of the wound with respect to different aspects.

1.3.1 Wound Bed Preparation [WBP]

Cleansing and searching for infection of wounds is fundamental to the treatment and preparation of the wound bed in order to achieve faster healing. In the context of the WBP we identify the TIME scale, a framework designed to identify 4 clinical areas that should be considered for wound bed preparation: tissue, infection or inflammation, moisture imbalance, epidermal margin.

For each scale parameter, what to look at is suggested and thus what aspects are of interest for proper diagnosis, how to intervene on the items to be corrected, and the expected results for each field. For the T parameter, attention is paid to the amount of necrotic tissue present, for the I parameter the heat of the cure, edema and fever, for the M parameter we note whether excess exudate or dryness is present to quantify the fluid imbalance present in the lesion, and for the E parameter we focus on the wound margins, hence undermining and hypertrophy. According to the WBP standard for facilitated healing, the outcomes that should be aimed for with treatment are:

- restoration of the bed of the lesion
- control of inflammation and infection
- balancing of fluid supply to the affected area
- proper proliferation of the wound margins

Table 1.1: Table showing wound classification according to the percentage of granulation tissue [7]

CLASS	Specifications
A	Wound is composed of 100% granulation tissue.
B	The wound consists of more than 50% but less than 100% granulation tissue. Absence of eschar or necrotic tissue.
C	The wound is composed of less than 50% granulation tissue. Neither eschar nor necrotic tissue is present.
D	Present eschar or necrotic tissue

1.3.2 Photographic Wound Assessment Tool

The Wound Photographic Assessment Tool (PWTA) provide a sub-classification according to wound characteristics. PWTA takes into account more parameters than WBP, also characterising the type of granulation and necrotic tissue present in the wound. Basically, the difference with the WBP scale is a more qualitative than quantitative assessment of the tissue present in the wound area.

Table 1.3: Wound assessment table

Item	Description	Assessment
1. Size	It is used to describe the size of the wound, if the wound is larger the scores are higher	0 = skin intact ($< 0.3 \text{ cm}^2$) 1 = $0.5 - 2.0 \text{ cm}^2$ 2 = $2.0 - 10.0 \text{ cm}^2$ 3 = $10.0 - 20.0 \text{ cm}^2$ 4 = $> 20.0 \text{ cm}^2$
2. Depth	Is used to quantify damaged layers of tissue	0 = wound is healed 1 = Full thickness
3. Necrotic type	Characteristics of the necrotic tissue present in the lesion	0 = Non visible 1 = necrotic tissue is thin grey or yellow 2 = necrotic tissue is thick, majority white yellow or fibrin 3 = necrotic tissue is devitalized, eschar, white or grey 4 = majority of tissue is hard with black eschar
4. Necrotic amount	Percentage of necrotic tissue present in the lesion	0 = non visible or closed lesion 1 = characterize 25% of lesion

Continued on next page

Table 1.3 – Continued from previous page

Item	Description	Assessment
		2 = characterize 25% to 50% of lesion 3 = characterize > 50% but < 75% of lesion 4 = characterize 75% to 100% of lesion
5. Granulation Type	Features of the granulation tissue	0 = wound is closed 1 = > 50% of the lesion is characterized by healthy tissue 2 = most of the granulation tissue is unhealthy 3 = most of the tissue is damaged 4 = the tissue present in the wound is all necrotic
6. Granulation Amount	Quantity of granulation tissue in the wound	0 = wound is closed 1 = 75% to 100% of the lesion is composed by granulation tissue 2 = > 50% and < 75% of the wound is composed by granulation tissue 3 = 25% to 50% of the wound is composed by granulation tissue 4 = < 25% of wound bed is covered with granulation tissue
7. Edges	Characteristics of wound edges	0 = wound is closed 1 = > 50% of the border has epithelial tissue 2 = most of edges are not advancing 3 = most of edges are undermined 4 = most of edges are fibrotic
8. SKin	diseas associated with skin near lesion	number of collateral diseas affecting the area 0 = None 1 = One 2 = Two or Three 3 = Four or Five 4 = more than Five

1.3.3 Diabetic ulcers classification

Diabetic ulcers, particularly diabetic flail ulcers if the foot area is affected, are very difficult wounds to treat with a high risk of amputation, so again proper follow-up and classification is essential so as to prevent limb loss. The most common classification methods found in the literature are shown below.

Meggitt-Wagner system

It consists of six grades of classification, which focuses on the depth of the lesion and the presence of necrotic tissue. One of the weaknesses of this method is that it cannot distinguish between infection and ischemic injury.

Table 1.4: Table showing DFU classification according to Meggitt-Wagner system [8]

GRADE	Ulcer depth
0	Pre-ulcerative area without open lesion
1	Superficial ulcer (partial/full thickness)
2	Ulcer creep to tendon, capsule, bone
3	Stage 2 with abscess, osteomyelitis, or joint sepsis
4	Localized gangrene
5	Global foot gangrene

Texas classification system

The method proposed by the University of Texas focuses on the depth, infection and ischemic state of the wound. It divides patients into 4 categories with respect to the 3 parameters mentioned above.

Table 1.5: Table showing DFU classification according to Texas scale [8]

GRADE	DESCRIPTION	STATUS	STATUS DESCRIPTION
0	Skin without lesions	A	Neither infection nor ischemia is present
1	Superficial ulcer	B	Infection presence
2	Deep ulcer	C	Infection presence
3	Ulcer with bone involvement	D	Presence of both infection and ischemia

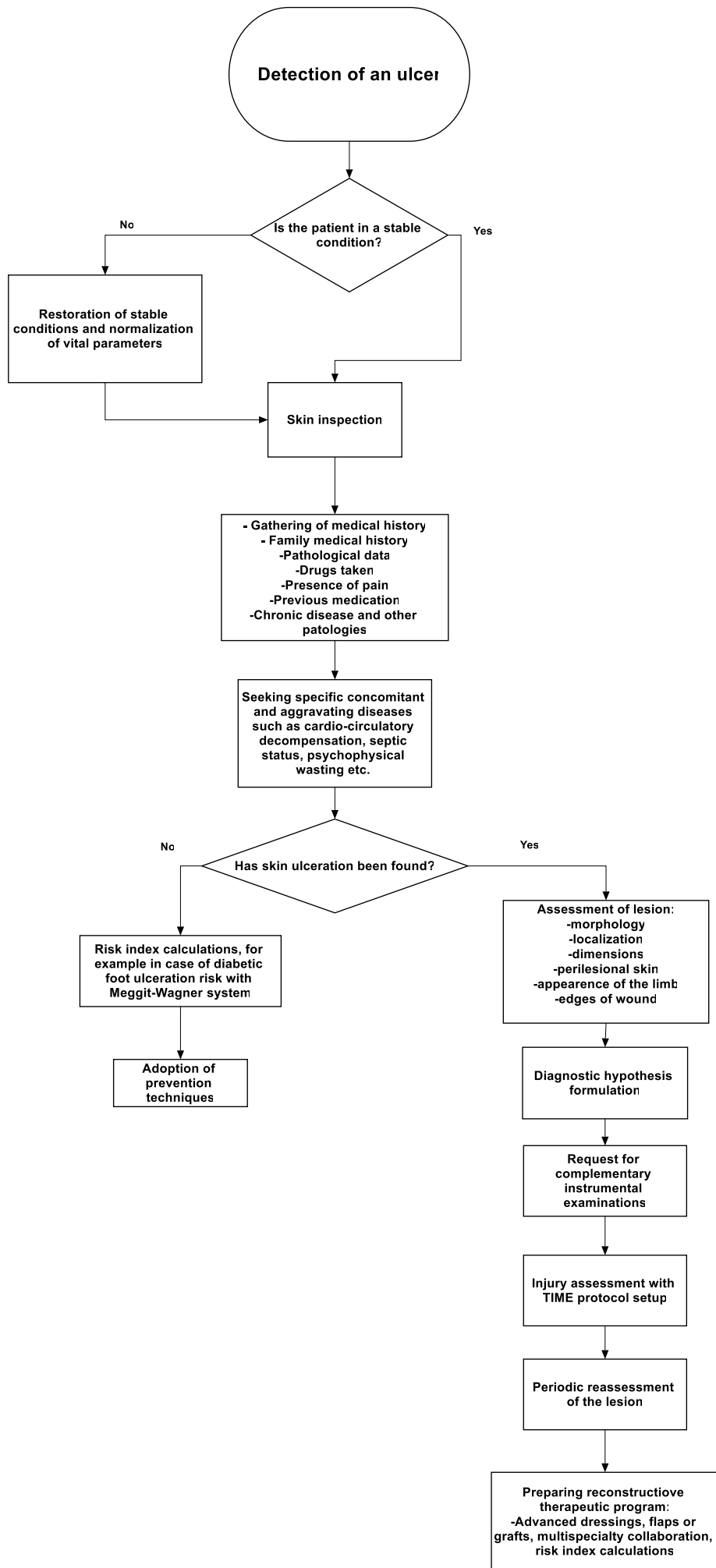
1.3.4 General protocol of approaching ulcers

As far as there are no defined clinical pathways for the treatment of ulcers, an attempt has been made to summarize what is usually done by physicians and nurses, confirmed by data taken at the AIUC national exhibition.

It is wrong to treat the lesion without an adaguate family and patient history. In fact, it is noted that often the formation of the chronic lesion is a consequence of another disease, which if left untreated leads to failure to heal the ulcer itself. The Medical History and General Examination should be conducted simultaneously upon the patient’s intake to gather all necessary data. This aims to stabilize the patient’s general condition and gather information regarding any medical conditions that may have contributed to the development of chronic skin lesions, thus providing insights into the potential etiology of the ulcer. Following this, we can proceed with the physical

examination.

Shown in the figure is what may be a process of analyzing a lesion, where history, other concomitant diseases, and the final follow-up of the ulcer are taken into account. We assess not only the presence of the lesion but also the possible formation based on determinant risk factors **Figure??**



Chapter 2

Imaging and measurement in wound care

The first thing that is done by the experienced vulnologist is to observe the wound so as to make a preliminary assessment of its condition. The ulcer is also measured, in some cases manually through the use of calibrated instruments such as a ruler. This way of performing the measurement can bring pain to the patient since it includes direct contact with the instrument; it also creates the risk of worsening infection if the instruments are not used in a sterile environment such as an apartment. For these reasons, digital imaging used for measurement can be a great help to the clinician bringing benefit to both him and the patient. By photographing the wound, measurements can be automated, or digital floor plans can be obtained, which superimpose a grid over the image to aid the physician in estimating the area. The size and depth of the lesion are not the only parameters that interest us in assessing the healing status of the ulcer, of which a complete evaluation can only be obtained by integrating microbiological and biochemical analyses as well. Commonly, swab tests are practiced where with a metal instrument the bacteria present are collected by direct contact with the wound for later analysis, or a biopsy of the ulcer is performed directly. These two practices again present the problem of direct contact and can also further damage tissue that was already struggling to heal. An integral aspect of wound care involves the removal of devitalized or bacteria-laden tissue through debridement, which acts as a barrier to healing and can compromise the efficacy of topical antimicrobial agents. Determining the extent of debridement required can be challenging and may sometimes lead to overly aggressive measures to ensure complete removal. Therefore, the use of digital imaging can be helpful in overcoming these limitations, until the advent of imaging, immediate information regarding the presence and location of bacteria was largely based on conjecture. Imaging provides specific target for debridement but not only, is useful in order to monitor the wound over time, conduct tele-visits, seek second opinions. It is therefore clear that it may be useful to keep simple RGB photos of ulcers on file for comparison and to find increasingly effective methods for analysis, characterisation and monitoring. Monitoring in this type of disease seems to be the key element to be included in treatment protocols to avoid serious complications.[9]

2.1 Types of imaging techniques

Different types of medical imaging could be used, beyond visible light imaging. Two main types of imaging used in vulnology can be distinguished, optical and nonoptical imaging. Optical-type imaging uses electromagnetic waves(light) at different wavelengths to generate an

image. The most common technology attributable in this category of imaging is based on X-rays. Although this type of technology is suitable in many practices, in wound care it is preferred to use 'EDI' type techniques that stands for Enhanced Digital Imaging, which work in the visible light range. One of the advantages of these techniques is the standardization of color [6]. The light incident on the sensor is identified through the 3 main components of color: red, green and blue. This feature can be used for the analysis of wounds, distinguishing their condition by color, something not available with any type of imaging. Another type of analysis that falls under the classification of optical-type imaging is multispectral imaging (HI), which is coupled with EDI-type techniques where emitters are installed. Light is emitted at different light frequencies, the reflected photons are detected by specific sensors in order to create the image based on their intensity and energy. HI techniques differ from each other in the wavelength used during the procedure, as different values of this parameter change the type of analysis and experimental set-up.

NIRS, near infrared spectroscopy, is a noninvasive optical imaging technique that is used to measure blood flow to a specific district. This technique is often used in neurology but also in wound care. The procedure is performed with at least 2 wavelengths used to quantify the chromophores of interest by absorption and scattering phenomena. Quantification of these chromophores provides insight into the amount of oxygenated and deoxygenated hemoglobin flux in the district of interest. We take advantage of the different wavelengths at which the phenomenon of absorbance occurs in the infrared spectrum, which differs for the two states of hemoglobin. Thus there is an emitter and a receiver, photons are emitted from the emitter, when they meet the body surface they are deflected and travel a curved path(banana shape) to the receiver. The distance between emitter and receiver has to be chosen according to the type of analysis and as a trade-off between the optical path that the photons travel, which the longer it gets the more absorbance phenomena increase along with scattering phenomena: these two things can compromise the spatial resolution and sensitivity of the analysis done by NIRS. The temporal resolution of the technique turns out to be relatively good(order of a few seconds) while the spatial resolution is relatively low (of the order of cm).

Then we can identify imaging techniques based on a non-optical system that do not use to create images light reflected or emitted but other physical phenomena like magnetization of water molecules in Magnetic Resonance Imaging (MRI). MRI and ultrasonography are used in the field of wound care to analyze hemodynamic. Ultrasonography is based on sound waves sended and detected after the reflection: areas with small variation will be seen as dark, instead the areas with wide variations are seen as bright.

However, there remains the problem of the absence of a standardized protocol, observation of the ulcer is subject to the opinion and evaluation of the physician, with no objective assessment. Many imaging techniques used are very expensive, such as MRI, PET, CT, so the need to find cheaper alternatives is highlighted. All techniques used in wound care are shown in the **Table2.1** with main advantages and limitations.

Table 2.1: Comparison of Various Imaging Techniques Used in Wound Infection Diagnosis

[10]

Imaging Technique	Application in Wound Infection Diagnosis	Pros and Cons	Sensitivity	Specificity
Plain radiography	Initial examination of soft-tissue infections.	Simple operation, low cost, wide availability; radioactive; reveals inflammatory changes. Can distinguish swelling due to infection from fractures; nonspecific findings. Misled by other conditions.	Low	Low
Computed tomography	Diagnosis of soft-tissue infection and intra-abdominal abscesses; evaluation of deeper structures and the extent of surrounding inflammations; identification of small infected collections.	Wide availability; fast scanning speed; high spatial resolution; multiplanar reformatting capabilities; high penetration depth. Radioactive, sometimes requires contrast agent.	High	High
Magnetic resonance imaging	Diagnosis of soft-tissue infection; Can provide anatomic and pathophysiologic information about the extent of infection within both soft tissue and the underlying bone.	High spatial/contrast resolution; nonradioactive. Costly; low availability; requires special training/-facility; hard to distinguish foreign bodies from adjacent structures within superficial wounds.	High	High
Ultrasound imaging	Diagnosis of skin and soft tissue infections. Can evaluate suspected radiolucent foreign bodies.	Fast; Accurate; cost-effective; portable; available in many clinics; no ionizing radiation. Interference from air; low penetration depth; relies on operator's skill.	Medium	Medium
PET	Diagnose and predict remission of antibiotic treatment for diabetic foot infections by developing 3D images from accumulation of radioisotopes.	High-resolution 3D imaging; high penetration depth. Costly; Short tracer half-life; requires special training/facility; Possible misdiagnosis resulting from sterile inflammation.	High	High

Continued on next page

Table 2.1 – continued from previous page

Imaging Technique	Application in Wound Infection Diagnosis	Pros and Cons	Sensitivity	Specificity
SFDI	Identify burn wound infection by quantifying volume fraction of tissue chromophores.	Noncontact; distinguishes infected and noninfected burn wounds. Limited scanning area and wound types.	High	High
Thermography	Use an infrared camera to measure infrared radiation emitted from the wound tissue. Smart phone-based thermography has been developed for diabetic foot ulcer detection and wound healing prediction.	Simple; portable; cost-effective; real-time imaging; noninvasive; remote diagnosis. Limited accuracy and specificity.	Low	Medium
Luminescence imaging	Portable imager distinguishes infected wounds from uninfected ones in a pig model based on intensity and distribution of visible photons emitted by Chernow radiation.	Simple; portable; cost-effective; real-time imaging; noninvasive; remote diagnosis. Limited specificity; requires more research.	Medium	Medium
Autofluorescence imaging	A handheld portable device diagnoses bacterial infection in diabetic foot ulcers in real time by detecting autofluorescence due to the light absorbing properties of endogenously produced bacterial porphyrins.	Simple; portable; cost-effective; real-time imaging; noninvasive; sensitive; remote diagnosis. Low specificity; Cannot determine microbial species.	High	Low
Microwave-microfluidic biosensor	A microwave-microfluidic biosensor for quantification of Escherichia coli within medium solutions to increase the efficacy of clinical wound infection assessment; the growth of bacteria can be monitored over time.	Small; cost-effective; rapid; contactless; real-time measurement; noninvasive; sensitive. Limited detectable bacterial species.	High	N/A

In particular, we find in the state of the art techniques that also allow a microbiological assessment of the lesion to be performed. In addition, when wound swabs are subjected to microbiological analysis, they can only detect microorganisms present on the surface of the wound or within the depth of the tissue biopsy obtained. Consequently, deeper-seated pockets of infection might go unnoticed during these evaluations. Additionally, it's important to acknowledge that while laboratory and microbiological assessments offer objectivity, they might fail to detect local changes that haven't yet triggered a systemic inflammatory response, and their effectiveness can be influenced by the operator's proficiency in sample collection. It generally takes 36 to 48 hours for the outcomes of wound cultures to be available. In cases of significantly mixed cultures, the task of distinguishing between different bacterial strains can introduce further hold-ups. Such delays in receiving culture outcomes are hardly beneficial for surgeons who are in urgent need of making decisions about whether a wound should be closed or not. As a result of these postponements, the bacterial load present in the wound can significantly alter by the time the culture results are received [11]. Hence, these limitations highlight significant gaps that could be addressed through the concurrent utilization of these imaging techniques[12].

2.2 Ultraviolet radiation for bacterial fluorescence

The discovery of fluorescence was initially made by Stokes in 1852. However, the realization that tissue fluorescence could have diagnostic applications was acknowledged considerably later, in 1911 by Stubel [13]. The detection of principles of infection is critical for wound healing, and with fluorescence uv signals we are able to obtain useful information for this purpose without going for a biopsy. Recently this innovative technology has been integrated on medical devices such as the Moleculight[14]. Tissue is excited and emits a spectrum of fluorescent signals that are emitted by different elements in the wound. Many proteins that contain amino acids with aromatic structure are in fact self-fluorescent. Structural proteins such as collagen and elastin are also affected by this phenomenon [15]. Fluorescence imaging offers clinicians immediate diagnostic information and feedback on treatment efficacy by revealing the presence and extent of bacterial infection in wounds. One can then evaluate the different components of the wound based on the color of the fluorescent signal emitted as shown in **Figure 2.2**.

Table 2.2. Characteristic fluorescence signals from wound tissues

Color	Characteristic fluorescence signals from wound tissues
Green	High collagen content (i.e., emitted by a tendon) Pseudomonas aeruginosa High fibrin content
Brown	Blood or high vascular tissue (i.e., granulation tissue)
Pink	Fluorescence from bacteria
Red	Endogenous porphyrins from bacteria
Cyan	Pyoverdines, specific to Pseudomonas aeruginosa

To identify regions of bacterial fluorescence, the following should be analyzed and distinct: tissue and fluid fluorescence signals, bacterial fluorescence signals, and artifacts that can be created [16]. The result is a mapping of the bacteria in the wound that can be used by the physician for targeted cleaning of the wound or to make decisions about what type of therapy to pursue. No contrast agents are used because bacterial autofluorescence properties are exploited. The information captured in these images can facilitate more targeted and thorough wound debridement,

support clinical decision-making, aid in dressing selection, and help determine the necessity for antimicrobial therapy, ultimately improving clinical outcomes. The technique brings with it some

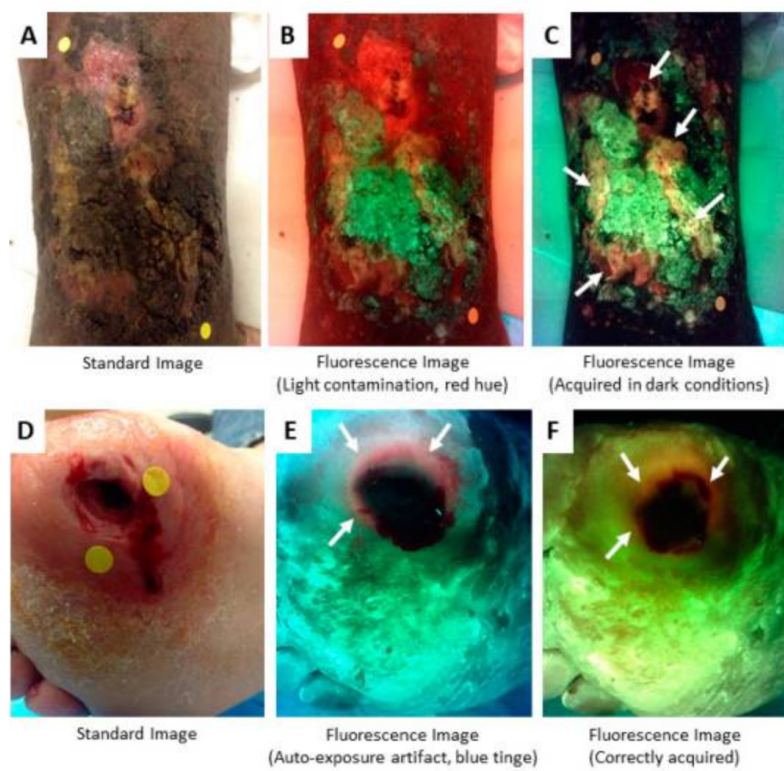


Figure 2.1. Limitations of bacterial fluorescence are highlighted[16]

issues, including the need to perform the analysis in a totally dark environment to avoid artifacts as shown in the figure. As we can see from the image **Figure 2.1**, if the illumination conditions are not right, the image is contaminated with blue fluorescence that has no clinical significance. Unfortunately, many creams and gels and some contrast agents also exhibit the autofluorescence phenomenon that contaminates such analyses. Other causes of artifacts can be tattoos, cotton products such as hospital bed sheets, dust in the environment and on wounds, collagen-based bandages and ointments used to treat wounds with necrotic or eschar tissue. Devices such as MolecuLight present no small innovation in the field of wound care, as under certain conditions it is possible to obtain feedback on the degree of infection affecting the wound, without performing biopsies and thus without contact and harm to the patient.

2.3 Thermography

Thermography was developed in the 1950s for military use to allow soldiers with night vision capabilities. Beyond enhancing night vision, it has been utilized to improve the visibility for military vehicle operators in conditions obscured by smoke, dust, light fog, and rain, and has also been applied in target identification, and tracking. In the last years the scope of thermal imaging has broadened significantly due to ongoing advancements in technology, making it a valuable tool in various sectors including healthcare and veterinary science.

Thermography is an imaging technique based on radiation emitted in the infrared band, captured by an infrared camera. Thanks to this types of cameras we are able to detect the temperature of the object of interest, in this case a wound. Thermography has been used in wound care in several applications: Thomas et al. [17] evaluated the use of infrared technology in laser skin treatment of vascular lesion, Mason et al. [18] concluded that for burn injury depth thermographic imaging was more accurate than clinical based pathways. Mercer et al. have used Infrared for monitoring the improvement of wound healing in human subjects with chronic venous stasis ulcers [19]. Through the heat detected we can tell what the state of blood perfusion is in the area observed, which is crucial information for understanding the state of a wound. The correlation between The heat detected at the skin surface and the surrounding blood flow is expressed by the Pennes equation of bio-heat [20]

$$k\Delta^2T - c_b w_b (T - T_a) + q_m = 0 \quad (2.1)$$

Several studies suggested a correlation between surface's heat and infection of wounds [21]. Physicians are often tasked with differentiating between infected and non-infected wounds, a process that can vary significantly among specialists due to the inherent challenges in assessing a wound's condition. The manual methods usually used by the physician to assess warming of the affected area are manual palpation or direct contact with thermometers, practices that can again bring discomfort to the patient with all the problems that direct contact with the wound causes. It is important to keep track of the temperature of the ulcerated area because when tissue undergoes cooling, it poses a heightened risk of infection due to several physiological responses. Cooling may result from several factors and may also be an effect of the change in dressing of the lesion. These include vasoconstriction, which reduces blood flow and oxygen delivery to the area. Consequently, there's a decreased oxygen availability for neutrophils, key cells in the immune response against infection. Additionally, as temperature drops, the activity of important cells such as neutrophils, fibroblasts, and epithelial cells diminishes. Hypothermia also adversely affects platelet activation and the ability of neutrophils to kill pathogens through oxidative mechanisms. Moreover, collagen deposition, crucial for wound strength, declines under hypothermic conditions [22].

Utilizing thermographic images of the lesion can provide valuable support in these situations. Given that clinical inspection alone has been shown to accurately identify infected wounds less than 60% of the time, there is a crucial need for additional diagnostic tools to enhance detection rates. Thermal imaging has found applications across various medical fields, notably in predicting ulcer formation in diabetic patients' feet. This utility arises from its capacity to discern and quantify minute variations in skin temperature associated with pathological transformations, including the inflammation of soft tissues, the ensuing breakdown, and the infection of ulcers. These methods have also been used in individuals at risk of developing wounds by comparing the skin temperature distribution of both feet of participants, termed as asymmetry analysis: the foot with the higher temperature is considered to be at risk of ulceration. The heat given off by the skin may depend on the blood supply to the ulcerated area, so a lower temperature may indicate impaired cellular metabolism. This method is resilient to natural variations in ambient conditions, scaling and rotation, and does not require contact with the wound of the individual. Furthermore, computerised analysis of textural features is a time efficient and cost-effective method to identify delayed healing of VLU and further research to assess generalisability and to reference the methods is warranted.

The measurement can be affected by the external environmental factors, such as the temperature of the environment in which the analysis is performed. To mitigate these effects, the patient is made to lie on his or her side for a few seconds to dissipate heat from the area that may have accumulated in the supine position. A reference area is then defined, which is subjected to the same environmental heat conditions. The reference area must have a temperature change of no more than 1 degree and no bone protrusions [23]. Clinically, temperature differences between a wound and a control area can provide significant information on the presence or absence (as well

as degree) of inflammation and/or infection in a wound. It can also provide an indication of the underlying perfusion to the area. That being said, skin temperature changes resulting from a pathologic process can be a matter of degrees. The interpretation of very similar temperatures may be normal or pathologic, depending on the clinician and the type of wound. In this way, the relative measured temperature can be associated with physiological or pathological conditions independently of other factors[24]. In addition to temperature information and thermal image observation, recent research proposes the calculation of new parameters that can be associated with these types of images. The Wound Inflammatory Index (WII)[25] is obtained through a process of thermographic image analysis. Initially, the anatomical surfaces of the foot are examined to identify any hot or cold spots where inflammation or circulatory leakage is occurring, respectively. Next, thermographic and visual images are examined to determine the shape, area, curvature, and eccentricity characteristics of the wound. This process helps to identify the shape of the wound and describe its base in terms of granular, fibrotic, or necrotic tissue. Once these details have been identified, we proceed to calculate the Thermal Index (TI), which is a dimensionless unit derived from the thermal characteristics of the wound. TI is calculated using the following formula:

$$TI = \frac{DT \times a}{A}$$

DT is the temperature difference between the wound and the average temperature of the foot a is the area of the isotherm (highest or lowest temperature) in the wound area, A is the area of the wound bed. Area is calculated in terms of pixels for this analysis. The choice of the highest or lowest isotherm must be made at the beginning of the analysis and followed consistently. Currently, these features are manually extracted from the thermographic image. Once TI is obtained, it can be used as an indicator of wound inflammation and the healing process over time. Analysis of thermographic images and calculation of TI can then provide an objective indication of the health of the tissue and the progress of wound healing over time.

Some injury assessment criteria include temperature among the parameters such as the STONES criteria. The STONEES mnemonic consists of seven clinical signs: Size enlargement, Temperature increase of 3° F or more versus the opposite limb mirror image temperature, Os (bone exposed or direct probing), New areas of break down on the wound margin, Exudate increase, Erythema and/or Edema (usually indicates cellulitis), and Smell. Any three of the seven STONEES criteria are diagnostic of deep and surrounding infection (with a sensitivity of 90% and specificity of 69.4%).[26] Khalad Maliyar et al.[26] found that noncontact infrared thermometry comparing maximum temperatures of mirror image of a wound on the opposite extremity, when combined with two or more other STONEES criteria, is a significant indicator of deep and surrounding infection. Thus, a correlation emerges from the literature between lesion temperature and the infectious and microbiological processes occurring in situ. This correlation indicates that similar information can be obtained from thermal images to that which comes from images obtained by fluorescence, without the problems of light and reflection that are encountered using ultraviolet radiation. Another problem encountered for ultraviolet radiation is that results can be compromised by the presence of dust, which can distort the analysis, whereas infrared technology is not affected by dust, water droplets or salt particles since the absorption contribution made by these elements is largely negligible [27]. Thermal imaging is also portable and much cheaper than MRI or other very precise technique, also it should be noted that it is not harmful to the patient in contrast to many imaging techniques now more widely used, such as X-rays. Thermography, although a relatively new technology in the field of wound care imaging, is gaining recognition in the scientific literature for its significant success. This technique is establishing itself as a cost-effective solution, providing a non-invasive, rapid, and reliable method for assessing and monitoring the status of wounds. Thanks to its ability to detect subtle temperature variations associated with different stages of wound healing, thermography is positioned among the techniques that could be both economically viable and highly predictive of wound conditions.

Process of acquisition

In a thermal imaging camera we can find 3 main components: focal plane array module, control and digital processing module, imaging module. The signals from IR sensors are converted from analog to digital by the analog digital converter so that they can be processed by the image processing unit. As in any electronic system, each added module causes greater resource consumption, so it is necessary to understand the basic steps to process the acquired image. The main task performed by the image processing unit is the control of the focal plane array so as to read the input values from each sensor detector, correct the non-uniformities in the image, and generate the necessary data for the display unit that will allow the acquired image to be displayed[28]. Basically, the electromagnetic radiation from the objects passing through the focal plane array is analyzed by the device so as to visualize the already preprocessed image. The infrared (IR) sensor

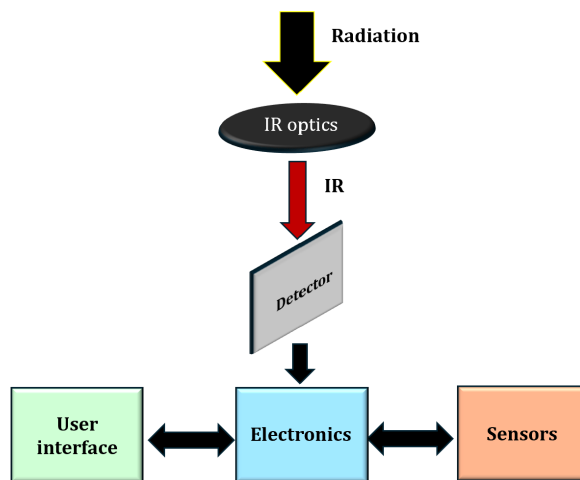


Figure 2.2. Thermal camera operating logic

is an electronic device that plays a crucial role in detecting specific features in its environment by either emitting or sensing infrared radiation. These sensors are capable of detecting or measuring an object's heat and motion. Such sensors operate similarly to human vision in sensing obstacles, though they do not rely on visible light. Typically, in the infrared spectrum, all objects emit some level of radiation, which, while invisible to the human eye, can be detected by IR sensors. All objects at temperatures above absolute zero emit IR radiation and the amount of the emitted radiation increases with temperature.

The radiation emitted in the infrared band is approximately between 0.3 and 1000 micrometers. This type of sensor has an aperture that defines the field of view (FOV), where a lens is inserted to allow signals in the infrared region to pass through it. The FOV of the sensor is critical determines the field I want to analyze, within which I must place the objects whose temperature I want to measure. A very narrow field of view indicates point measurements, less thermal power will be collected but it will be very sensitive. The signal of interest is focused and penetrates the device, the sensor heats up according to the temperature of the object being measured. We then find a sensing element that can absorb infrared, heat up, and produce a signal proportional to the heat absorbed. The sensing element itself has a temperature, so the voltage signal produced is proportional to the difference between the temperature of the sensing element and the temperature of the container. By placing a reference sensor that measures the temperature of the container, we

are able to finalize the measurement. This is the basis working principle of an IR sensor. IR sensors are governed by three fundamental principles of physics: Planck's Radiation Law, the Stefan-Boltzmann Law, and Wien's Displacement Law. Planck's Radiation Law states that any object with a temperature above absolute zero emits radiation. The Stefan-Boltzmann Law relates the total energy emitted across all wavelengths by a black body to its absolute temperature. Wien's Displacement Law suggests that the peak wavelength of emission from objects shifts inversely with temperature, meaning hotter objects emit radiation at shorter wavelengths.



Figure 2.3. FLIR A300 thermal camera [29]

The raw frames captured by the sensor initially do not offer any discernible details. To convert these into a visually interpretable format that highlights thermal attributes of the scene, the Thermal Image Processing pipeline performs a sequence of operations on the raw image. These operations include tone mapping to adjust brightness and contrast, correcting defects such as bad pixels, compensating for the camera's non-uniform response across the image, and enhancing the image digitally using various filters. Following these initial steps, the pipeline employs overlays and assigns color palettes to distinguish different temperature ranges within the scene. The process

results is shown in the referenced figures, illustrating the transformation from non-thermal to thermal visualization. The output of this process is a thermal image that is easily comprehensible to humans, resembling the example depicted in the image below **Figure 2.4** **Figure 2.5**. We can see some details from this image [Figure 2.5] with thermal staining, we have a red area, so warmer on the leg, which does not reach the wound. This indicates a strong circulation problem affecting the limb, we also notice cooler areas inside the wound, which may indicate principles of necrosis.

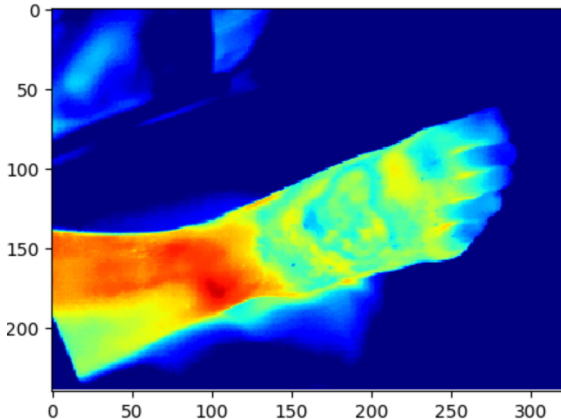
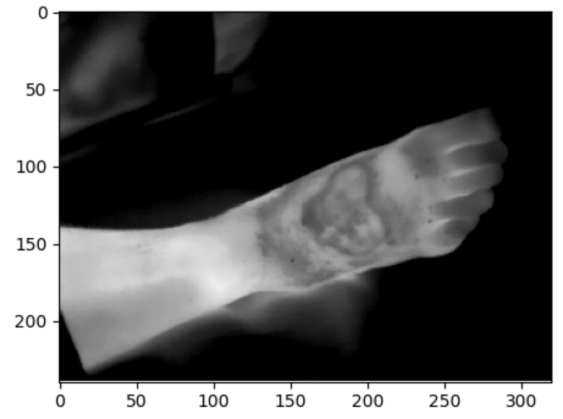


Figure 2.4. Thermal camera output image

Figure 2.5. Thermal camera output image with colourmap applied

Chapter 3

Integration of deep learning in digital imaging techniques

Artificial intelligence is a set of techniques that mimic human behavior, mainly the techniques are divided into machine learning and deep learning. Machine learning techniques are based on supervised learning, i.e., algorithms are first trained on categorized and structured data extracted from the data to be used. In deep learning techniques, a subgroup of machine learning, there is no need to provide input data features because they are automatically extracted by the algorithm. The structures of these techniques are the multilayer neural networks and are used when the features that are decisive for proper classification cannot be extracted a priori or it is not known which features may be decisive for classification.

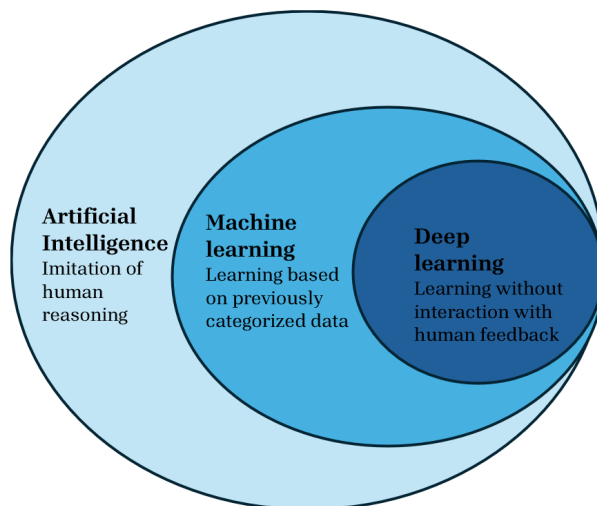


Figure 3.1. Relationship between machine learning and deep learning techniques

3.1 Basic operating principle of a neural network

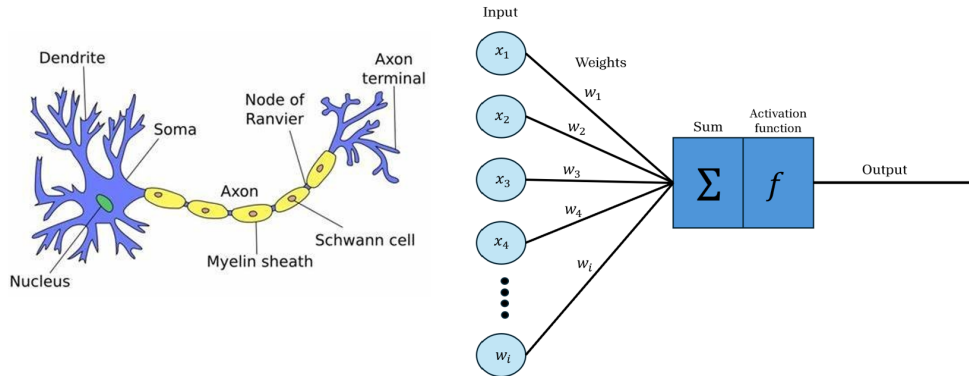


Figure 3.2. Illustration of perceptron

All artificial neural networks share several characteristics: a set of interconnected nodes, several layers and one or more activation functions. The network receives different inputs and associates different weights with the information, what it tries to do is to change the weights until the output class is equal to the input class. The final output of a neuron is obtained by applying the activation function used, which can be for example sigmoid, tangentoid, ReLu or SiLu. These functions inject non-linear elements into the architecture, allowing the network to capture intricate patterns within the dataset. Regarding the layers that can make up a neural network there are 3 main categories: the input layers, where there are neurons that receive information from outside, the hidden layers where the information is processed, and finally the output layer. All the neurons in the intermediate layers are connected to the other layers to approximate precisely the interconnections of human neurons.

The simplest mathematical approximation of the human neuron is shown in the **Figure3.2**, known as *perceptron*, introduced by Frank Roseblatt in 1958. It consists of an input layer, no hidden layers, and the output. The artificial neuron is the basic element of a Neural Network. Artificial neuron is mentioned because the idea is to create a system that simulates the human brain, or at least the human way of learning. In fact, neural networks are based on a learning algorithm that after exposure to a sufficient number of example samples are able to "learn," for example, to classify the item being proposed into a certain class.

The type of learning can be mainly of two types-supervised and unsupervised. One can speak of supervised learning when the neural network needs to receive as input a set of already correctly classified items in order to be able to understand the classification task being asked. On the other hand, one speaks of unsupervised learning when it is not necessary to provide the model with examples, the neural network will try on the basis of the characteristics of the elements provided to it to identify groups (clusters) of homogeneous elements. One of the most popular learning algorithms is the backpropagation algorithm. Its structure can be summarized in 4 steps [30]:

- feed-forward computation
- back propagation to the output layer
- back propagation to the hidden layer

- weight updates

An error function is created that represents the difference between the input and the output. The algorithm stops when the value of the error function is small enough. A limitation of the perceptron model is that classification can only be done for two classes at a time. By complicating the model, multiclass classification can be achieved with the Multilayer Perceptron.

Multilayer Perceptron is composed of several neurons connected to each other and organized in different layers in order to perform tasks more complicated than the single perceptron. The structure consists of an input layer that receives the information, several hidden layers by which the network is able to represent not only continuous functions such as the perceptron but also discontinuous ones so as to perform complex classification tasks. There is finally an output layer, where the number of neurons present will depend on the number of input classes.

As previously mentioned there is also unsupervised learning, where input data need not be labeled. What is called "competitive learning" is used: during the training phase, neurons in the layers become active only under certain conditions, there is then a "winner" neuron. During the subsequent learning phase only the weights of the winning neuron will be updated. An example of neural networks using this learning technique are the Kohonen self-organizing maps (SOM) [31].

3.2 Different neural network architectures

There are several neural network architectures, which differ in how the output is obtained, how the data is processed, and what operations are performed to extract the features of interest for classification or other tasks. Feedforward Neural Networks (FNNs) are the foundational architecture of artificial neural networks where data flows linearly from input to output. FNNs are layered structures, each of which is composed of nodes where calculations are performed and output is transmitted to subsequent layers. They are structured so as not to have loops where the output of one layer is carried back to a previous layer: the information only moves in one direction, hence the name 'feedforward'[32]. Instead, there are types of neural network architectures such as the recurrent neural network where neurons through cycles and loops send feedback back to each other. The idea behind this architecture is to simulate at a more complex level human brain activity where the various neuronal areas interact with each other with feedback signals[33]. This architecture is ideal for analyzing sequential or time-series data, finding its place in applications ranging from speech recognition to natural language understanding. To overcome the limitations of RNNs, more complex models have been proposed based on two units in particular: long short-term memory unit (LSTM) and gated recurrent unit (GRU). LSTM was proposed by Hochreiter and Schmidhuber in 1997. In contrast to simple network recurrence units, LSTMs maintain memory for a certain period of time t . When the weights are updated, the memory cell is partially updated: new information is added and some of the old is deleted. It can be modulated via the input parameters how much the memory is updated each time. As for the GRU, invented by Cho et al [34], its mechanism also involves a partial update of the memory state of the unit by a linear interpolation of the state before and after activation. The difference is that the degree of memory update cannot be modulated [35]. These modifications allow the networks to capture long-term dependencies more effectively, enhancing performance in complex sequential tasks like language translation and speech-to-text conversion. Convolutional Neural Networks (CNNs) excel in processing structured grid data such as images.(Multilayer Perceptrons in Machine Learning: A Comprehensive Guide) Another type of neural network are autoencoders, designed for unsupervised learning, which are designed to decode input data to obtain a determinant representation and decode the extracted features to obtain an output similar to the input [36]. They are often used for a better representation of input data e.g. images in order to decrease noise. This approach is useful for reducing dimensionality, denoising data, and enabling generative models. It is

also important to mention a type of network capable of generating realistic data that can also be used to expand an image dataset: this architecture is known as generative adversarial networks (GAN). The configuration of this type of model involves the training of two separate networks, one taking the role of generator and the other that of discriminator. The generator must try to reproduce data increasingly similar to the input data in order to confuse the discriminator [37]. When the discriminator can no longer tell whether the data provided by the generator is real or synthetic, the training is over. This framework has proven that can generate realistic synthetic images, video and tokens

3.3 Applications of deep learning in medical imaging

In medicine one may be interested in the classification of an element of an image, for oncological purposes for example one is interested in the nature of a lesion that may be benign or malignant. In the present case one may be interested in classifying the wound according to different scales such as the WBP, to understand its status and the decisions to be made regarding treatment. In addition to the classification task, as also evident in chronic wound assessments one may be interested in the area and thus the segmentation of a given image element so as to have an automatic and precise measurement. Alternatively or in conjunction with segmentation some networks are used for detection tasks and thus detection of a suspicious element or simply an element of interest in an image. There are complex neural networks, created from scratch or obtained as a combination of different architectures so that all tasks are achieved together. Alternatively, different simpler, non-combined models are used if one is interested in achieving certain performance for a specific task, or different, non-combined models can be used at different times to lighten the computational cost if one is interested in multiple tasks.

3.3.1 Segmentation networks

Networks for segmenting objects or areas of interest in the image have a common basic structure that consists of two parts: encoder and decoder. The encoder part extracts features useful for classifying each pixel by convolution and pooling operations from the input image by reducing its size. Once the initial dimensions have been reduced, the decoder part builds the output of the network through an upsampling mechanism to bring the image back to the output dimensions, assigning to each pixel the probability of belonging to the class it is interested in segmenting. The basic structure of this type of network is likened to a U, hence the name of the U-net neural network **Figure3.3**. Another type of network used for segmentation is the FastFCN. FCNs prior to the introduction of the improvements due to the studies of Huikai Wu et al [39]. were used much less due to model heaviness and high memory consumption The core of FastFCN is the introduction of the Joint Pyramid Upsampling (JPU) module, which replaces the dilated convolutions commonly used in FCN networks to increase the resolution of the final feature map without losing the receptive field. While dilated convolutions are effective in maintaining spatial resolution, they lead to a significant increase in computational complexity and memory usage.

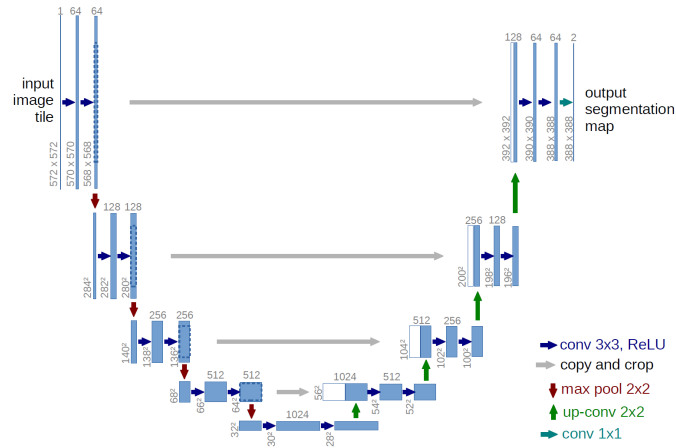


Figure 3.3. U-net structure [38]

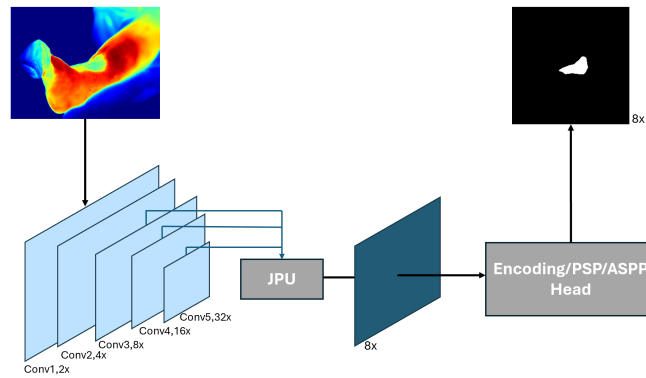


Figure 3.4. FastFCN structure

In the context of discussing Segmentation Networks, it's pivotal to mention the utilization of Deep Convolutional Neural Networks (DCNNs). DeepLab is an advanced semantic image segmentation system that leverages the power of DCNNs to segment objects within an image, i.e., to label each pixel of the image with a class representing the object to which it belongs. Semantic segmentation is a dense prediction task that demands high localization accuracy as well as semantic understanding of the image [40].

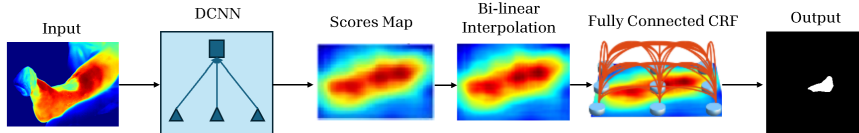


Figure 3.5. Deep Convolutional neural network structure

3.3.2 Detection and classification networks

One of the most widely used architectures for image classification and object detection is the convolutional neural network. The basic structure of a CNN consists of layers practising the convolution operation, layers practising pooling operations, fully connected layers and various functions for activating between the different layers and for extracting the final output as shown in the **Figure3.6**. Convolutional layers allow the network to learn patterns within the image,

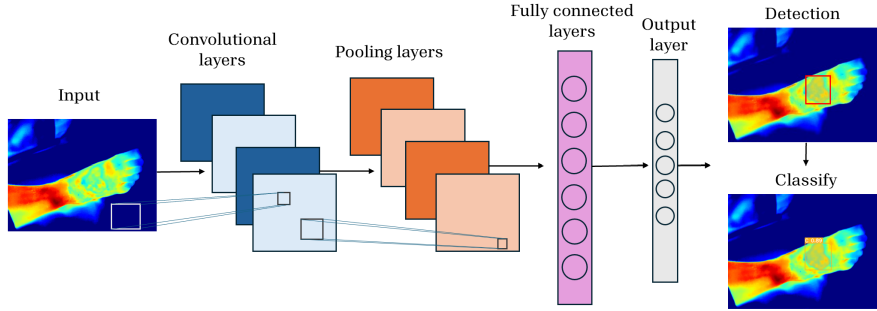


Figure 3.6. Typical structure of a convolutional neural network

using filters (kernels) of different sizes to extract the most interesting spatial features. The convolution operation is precisely defined by the size of the kernel used, which turns out to be a matrix of dimensions defined a priori when building the model. The matrix is moved over the image, and the result will be the sum of the multiplication of each kernel pixel value by the image pixels. The output will be a matrix equal to the size of the kernel used for the convolution operation as shown in **Figure3.7**. The edges of the image are a problem for this operation, however, since when the

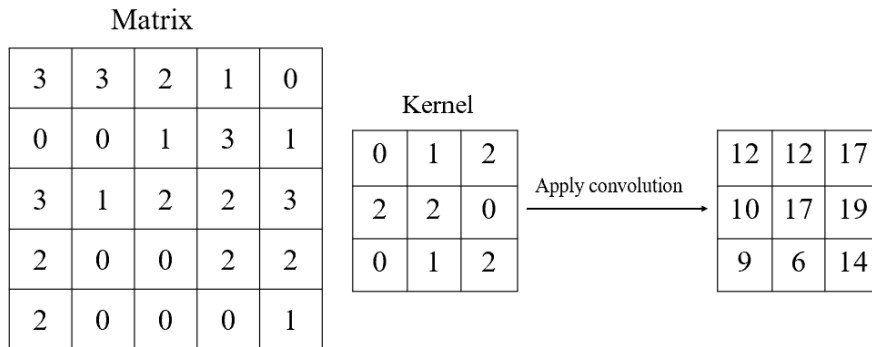


Figure 3.7. Convolution operation between two matrices

kernel is centred on an edge pixel it will fall out of the range of the image, a solution may be to use zero padding so that we do not have to perform an impossible operation but simply multiply by zero. Decreasing the size of the initial matrix is also a problem, first of all we don't want the size to be reduced every time otherwise we would not be able to reconstruct the output: padding

presents a solution to this problem as well [3.7](#). The kernel in addition to size can be set with several parameters among which we find the stride. The parameter refers to the displacement in number of pixels that the kernel makes at each step. This is useful for computational cost since increasing the stride, increases the movement of the kernel in number of pixels and thus the operation becomes inherently faster. In addition, a high value stride can be useful for the convolution to consider different parts of the image together so that new patterns can be found. Pooling layers are another key component of cnn, where the spatial dimension of the input is reduced again. There are several ways of performing the pooling operation, the two most common being max pooling and average. After the size of the image submatrix on which to perform pooling is initialized, max pooling involves retaining only the largest value present in the submatrix, while average mode averages over the submatrix of interest and that value is retained as representative. Due to this type of architecture, this network model has become very popular and as already mentioned can be used for detection and subsequent classification of detected objects.

Object detection task can usually be summarized in 3 main steps the process of object detection performed by a CNN. First, the entire image is analyzed and resized so as to begin predicting bounding boxes of objects of interest. The candidate regions are then further analyzed to find characteristic features. The last part of the procedure involves the final classification of what is found in each bounding box.

Object detection methods can be distinguished based on anchor free or anchor based methodology settings. Anchor-based methods are a cornerstone technique in the field of object detection, playing a crucial role in identifying objects within images. These methods rely on the concept of anchor boxes, which are predefined bounding boxes of various sizes and aspect ratios that serve as references for predicting the presence and location of objects in images. The anchor boxes are strategically designed to cover different shapes and sizes, mirroring the diversity of objects that the model might encounter in the dataset. When an image is processed, these anchor boxes are tiled across the image in a dense grid, ensuring comprehensive coverage. Each anchor box is evaluated by the detection model to predict whether it closely aligns with an actual object within the image. The model does not directly predict the bounding box from scratch but adjusts the size, position, and aspect ratio of the anchor boxes based on the objects present. This adjustment is typically done by predicting offsets and applying these to the anchor boxes, refining them to tightly encompass the detected objects. The use of anchor boxes allows the model to efficiently search for objects across different scales and aspect ratios, significantly improving the detection of varied object sizes and shapes. It simplifies the object detection problem by transforming it into a classification and regression problem, where the model classifies each anchor box as containing an object (or not) and regresses the box dimensions to best fit the objects in the image. Two-stage methods for object detection, such as the well-known Faster R-CNN, employ a sequential approach to identify and classify objects within an image. Initially, these methods focus on generating region proposals. This entails scanning the image to pinpoint areas that are likely to contain objects, essentially pre-selecting regions of interest. Following this, each proposed region undergoes classification and bounding box refinement. This two-phased approach allows for a more meticulous examination of potential objects, which is why two-stage methods are generally praised for their accuracy. They excel in scenarios where precision in detecting and localizing objects is paramount, thanks to the dedicated step of generating high-quality region proposals before proceeding to the actual object detection task. However, this meticulousness comes at a cost. The inherent nature of the two-stage process, requiring region proposal generation followed by classification and localization, makes these methods slower in terms of inference time. This slower processing rate renders two-stage methods less ideal for real-time applications where speed is crucial. The additional step of generating region proposals before moving on to detection introduces a delay that can be a significant drawback in scenarios demanding quick responses. Thus, while two-stage methods stand out for their accuracy, their practicality is limited in contexts

where rapid object detection is necessary.

One-stage detectors such as YOLO (You Only Look Once) and SSD (Single Shot MultiBox Detector) mark a significant shift in the approach to object detection by eliminating the region proposal step seen in two-stage methods. These models operate under a streamlined principle, aiming to identify and classify objects across the entire image in a singular, consolidated pass. This direct approach is designed to enhance processing speed, allowing these models to classify and localize objects without the need for a preliminary step to highlight areas of interest. The advantage of this method is its efficiency, making one-stage detectors particularly well-suited for real-time applications where speed is of the essence. They strike an appealing balance between speed and accuracy, enabling rapid object detection without the procedural delay introduced by generating and processing region proposals. However, the simplification that bolsters their speed also introduces certain drawbacks. Primarily, one-stage methods tend to lag behind their two-stage counterparts in terms of detection accuracy. The root of this issue lies in their direct approach, which, while eliminating delays, also increases the likelihood of false positives. Predicting over the entire image without first isolating regions of interest can lead to a broader margin of error, as the model may mistakenly identify irrelevant parts of the image as objects. This can be particularly challenging in complex scenes with multiple objects or in situations where objects blend closely with their background. Despite these limitations, the appeal of one-stage detectors remains strong in scenarios where the immediacy of detection is paramount, offering a practical solution that balances the need for speed with reasonable accuracy [41].

Anchor-free detectors represent a significant evolution in object detection technology, diverging from the traditional anchor-based approach by eliminating the need for predefined anchor boxes. These models directly detect objects within an image without relying on a set group of bounding box shapes and sizes, offering a more flexible and intuitive method for locating and classifying objects. YOLOv1 was one of the early models to adopt an anchor-free approach, simplifying object detection into a single process. By dividing the image into a grid, YOLOv1 directly predicts bounding boxes and class probabilities for each grid cell, streamlining the detection process and enabling real-time predictions. This method significantly reduces the complexity involved in object detection, offering a more efficient and faster approach. Following YOLOv1, several innovative anchor-free models emerged, each introducing unique strategies for object detection. CornerNet utilizes keypoints to identify the corners of objects, avoiding traditional anchors and employing a novel corner pooling layer for more efficient localization. ExtremeNet takes a bottom-up approach, detecting the outermost points of objects to define the bounding box, emphasizing geometric relations within the object's structure. RepPoints represents objects as sets of adaptable sample points, allowing for a more detailed and flexible object representation than traditional bounding boxes. The FSAF model addresses issues in anchor-based detectors by employing a dynamic, online feature selection strategy that optimizes detection across different feature levels without predefined anchors. The FCOS model further simplifies the detection process by treating it similarly to semantic segmentation, predicting objects per-pixel and removing the need for complex anchor and proposal mechanisms. ATSS and OTA introduce advanced sampling and label assignment techniques, respectively, enhancing the precision of anchor-free detection by leveraging statistical characteristics and optimization theory. DSLA improves the transition between positive and negative samples, offering a refined approach to classification and localization. YOLOv8, building on the legacy of YOLO, presents a state-of-the-art model that combines speed, accuracy, and ease of use, incorporating advancements in architecture and feature detection. In summary, anchor-free detectors have revolutionized object detection by offering models that are not only faster and simpler but also capable of achieving impressive accuracy. These models avoid the complexities and limitations associated with anchor boxes, opting instead for direct detection methods that can more intuitively adapt to the varied shapes and sizes of objects within an image [41].

3.3.3 Deep Learning and wound care

The application of deep learning methodologies in vulnology has been a growing trend in recent years, not only has there been an increase in research in the literature, but actual devices with this type of integration have sprung up to support the physician in decision-making processes. A prime example is the WoundViewer, designed to capture images of injuries and automatically classify them according to the WBP, Texas scale in the case of diabetic feet (DFU). The device consists of a CMOS camera and 16 IR sensors. The IR sensors are essential for measuring distance, area and wound depth. These measurements are possible thanks to the deeplearning algorithm integrated in the device's software. The algorithm consists of two parts, each relating to a different neural network: the first part has the task of extracting the regions of interest (ROI) from the input image, its structure is that of a CNN with several layers performing the convolution operation connected to each other. The second part is composed by a Discrete Time Cellular Non-linear Network (DT-CNN): here the output image from the first neural network is analysed and divided into 4 colour macro-groups (white, yellow, black and red) to identify the wound bed and then proceed to identify the area and classify the lesion according to WBP score. Some examples are provided in **Figure 3.8**. The wound ROI is then extracted from the first network so that the




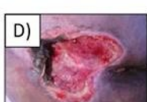
	Physician's Blind Evaluation (WBP)	WV Evaluation (WBP)	Clinical Characteristics	WV Colour Analysis
A) 	A	A	Completely composed of granular tissue	<ul style="list-style-type: none"> • 100% Red • 0% Yellow • 0% White • 0% Black
B) 	B	B	<50% of slough, no necrosis	<ul style="list-style-type: none"> • 82% Red • 11% Yellow • 7% White • 0% Black
C) 	C	C	>50% of slough, no necrosis	<ul style="list-style-type: none"> • 45% Red • 32% Yellow • 23% White • 0% Black
D) 	D	D	Presence of necrosis	<ul style="list-style-type: none"> • 23% Red • 36% Yellow • 3% White • 38% Black

Figure 3.8. Wound Bed Preparation scale and Wound Viewer algorithm classification [42]

second one can operate directly on the roi, resulting in a computationally less demanding model as only essential image information is presented to the second analysis block [43]. The results obtained with the WoundViewer algorithm were then compared with the opinion of an expert in order to calculate the static parameters necessary for the validation of the device. In the test phase all 4 wound types identified by the WBP scale were identified with more than 90% accuracy. In addition to the classification, information was obtained on the percentage of granulation tissue present in the wound, where the data reported by the device were consistent with the respective identified wound classifications. Correct classification is essential to the function of the physician support device, but other features are also included such as the ability to manually enter various parameters that can be shared among specialists creating a digital patient record and a strong tool for data sharing among physicians. This device achieve only 6% of error measuring wound's area and 99% of precision in calculating granulation percentage, for this reason it ranks among the candidates to be considered a gold standard in the field of vulnology. Other automatic area

measurement algorithms are known from the literature, such as Chino et al.[44] who use a deep neural network to obtain an error of 14% or Spinczyk et al. [45] who implement the measurement using a triangulation technique with an error of 11%. Regarding the estimation of granulation tissue, algorithms composed of deep neural networks were proposed by Nejati et al.[46] with an accuracy of 86.4 % and by Hsu et al.[47] with an accuracy of 83.54 %. Results comparable with those of the WoundViewer were achieved by Maity et al.[48] with 99% accuracy in classifying the amount of granular tissue in the wound.

H. Nejati et al. [46] proposed a pipeline for the classification of chronic wounds not into 4 subgroups but rather 7, expanding the classification of B, C and D wounds according to wbp even more specifically. The proposed tool consists of training by transfer learning using the AlexNet from which features are extracted for subsequent classification using support vector machine. The results obtained using this method performed better than using the single AlexNet to classify all classes. The network individually obtained statistically significant results only by dividing the dataset into three classes. Another important application of machine learning in this area is the tracking of how the wound evolves over time. Changan et al. [49] proposed a unified framework that not only includes the calculation of wound area but also allows for a binary classification on the presence of infection and a prediction on the probability of wound healing based on ulcer behaviour over time. The prediction is made on the basis of the decrease or increase in area relative to the initial area of the wound. The model estimates are promising in that the results show a 95% confidence interval on the healing predictions of the wound time series used as a test set.

From these data, evidence emerges that the integration of deep learning in this field can bring enormous benefits both clinically and logistically. A lot of research and applications focus on RGB images but they are not the only type of images that can be useful. A new frontier of study are images acquired by thermal imaging, which can supplement the information that RGB images can provide. Innovative research on this type of image applied to wounds is emerging, Fletcher and Niyigena [50] developed two neural network models for surgical site infections analysis using thermal images. Their aim is to emphasise the importance of accurate prediction of infection, which is essential for wound healing. The ASEPSIS score was used to calculate the probability of a possible infection, under the supervision of a trained physician. This type of scale was invented by Wilson et al in 1986, possible additional treatments, serous discharge, erythema, purulent exudate, separation of deep tissues, presence of bacteria, impaired mobility for more than 14 days are analysed [51]. The two models used in Fletcher's research are a basic CNN with 3 convolutional layers where the activation function implemented is ReLu. The second model is more complex, with a ResNet50 architecture optimised for application via Transfer Learning. The results obtained with the ResNet50 model have a classification accuracy of 90%. The limitation of this research is that the classification is binary, i.e. 'infected wound' or 'non-infected wound'. Beau et al. [52] have used both thermal and RGB images to provide a unique model able to generalise predictions for both type of images. Two models were generated: one trained with only RGB images and the other one use RGB+Thermal. Predictions were made on 3 classes, as 'wound', 'skin but not wound', 'non-skin'. The results obtained are excellent, but no significant improvement of the model is noted with the addition of thermal images, indicating that thermal features are not particularly used by proposed models for classification.

Chapter 4

Analysis of thermal images to assess infection status

Images acquired by thermal imaging contain features that can bring significant advances in the clinical treatment of skin ulcers, however, it is interesting to understand whether information typically extracted from RGB images can also be obtained through the automatic analysis of thermal images. These types of images have been used for a short time in wound care, so it was decided to use them to see if the features extracted from deep learning models are informative enough to create generalisable and complete models. The deep learning models most often found in the literature are then applied to the dataset of images acquired with a thermal imaging camera.

4.1 Dataset

The images used are from the European database of chronic wounds[29]. This database contains wound images of 188 cases from 79 patients. This dataset was chosen because it includes both RGB images and the respective thermal images of the wounds. Thermal images are acquired with a FLIR A300 thermal camera: images are captured 2D with 320x240 resolution. The camera has a field of view (FOV) of 25° x 18.8° and a thermal sensitivity of 0.05°C. The detection of IR is made thanks to a focal plane array (FPA).

Were also acquired stereo-images with MicronTracker Hx40, depth images with depth camera, SwissRanges SR4000. The calibration phase was done with a checkerboard image, to find the spatial correlations between the 4 cameras and the global room reference. For each patient, all the above-mentioned types of images were acquired. Both before and after calibration data are available. In the following analysis, the images after calibration were used so that the thermal and RGB images were superimposable. This factor was crucial and was decisive in the choice of dataset, as the algorithm used for ground truth works with images of an RGB nature. In **Figure 4.20** and **Figure 4.21** an example of 2 images from the dataset is provided.

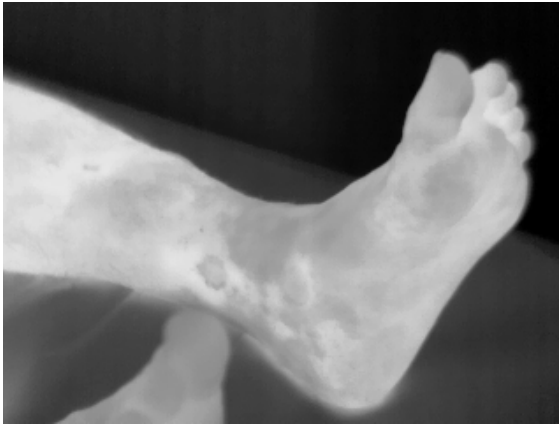


Figure 4.1. Calibrated thermal image



Figure 4.2. Calibrated RGB image

4.2 Ground truth generation and preprocessing

The WoundViewer algorithm was used to obtain a reliable ground truth, in order to compare performance with a method validated in clinical practice. RGB images were utilised for this part of the work, of which the outline is segmented **Figure 4.3**. In order to obtain a mask of binary

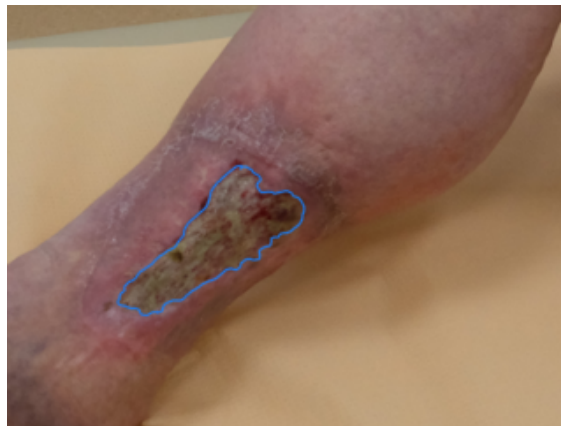


Figure 4.3. Contour identified by the algorithm

type, the image was further processed. The image is converted into HSV **Figure 4.4**, the pixel value with which the segmentation was performed is identified.

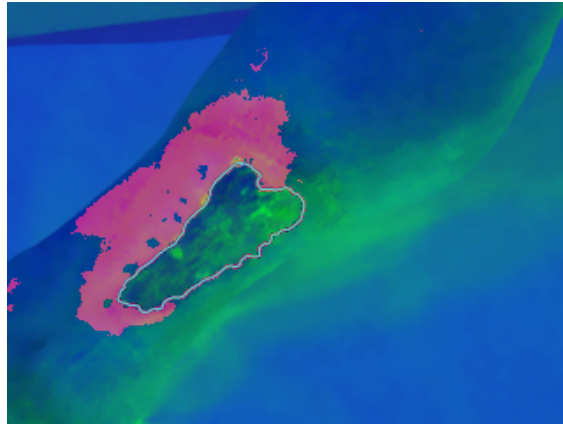


Figure 4.4. HSV Transformation

A thresholding in the specific blue range was performed with the following settings so that only the wound contour was obtained.

$$\text{lower_blue} = [110, 50, 50] \quad \text{upper_blue} = [130, 255, 255]$$

After the thresholding a binary image is obtained, but the contours are not closed and defined **Figure 4.5**.

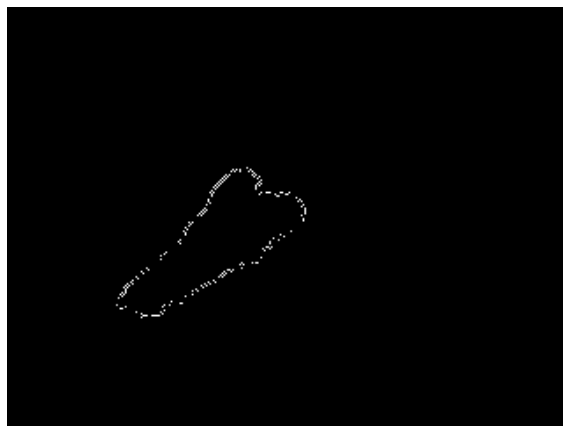


Figure 4.5. Thresholding result

A morphological operation is performed to close the contours. The closing operation is equivalent to two cascading morphological operations with a kernel of equal size. First a dilation and then an erosion is performed. The kernel chosen to achieve the best result is a 12x12 elliptical shape. The result is shown in **Figure 4.6**.



Figure 4.6. Final binary mask

The images were then classified according to WBP scale using WoundViewer algorithm.

4.2.1 Construction and Test set

The dataset makes available 250 images acquired by thermal imaging camera. The data and they are divided into 4 classes (A,B,C,D) according to the classification of the algorithm used as ground truth. In order to ensure representation of all classes in Construction and Test set, thirty per cent of the images are then randomly drawn for each class to be used as Test sets using a *stratified sampling* approach 4.1. The classes are unbalanced, particularly class C and D. Class D is essential to be represented as it indicates the most severe state of the ulcer. Data augmentation techniques are performed so that the classes are equally represented 4.2.

Class	Construction Set	Test Set
A	61 images	25 images
B	54 images	22 images
C	35 images	16 images
D	25 images	12 images

Table 4.1. Split dataset in construction and test set for each class

Class	Training set	Training set + augmentation	Validation set
A	36	108	25
B	32	96	22
C	21	126	14
D	15	135	10

Table 4.2. Split construction set

To increase the number of class D images of the Construction set, vertical and horizontal translations of random value are done. For the more numerous class C images, only horizontal

translations are performed, as shown in **Figure 4.10**. Translations are applied equally to the corresponding masks. A colourmap jet is then applied to make the temperature differences and thermal information represented in grey tones more visible. It was decided not to perform image cropping on the ulcer with a possible clinical model in mind where, due to logistics, images with high accuracy cannot be acquired. Furthermore, it was intended to test the ability of the various models to generalise even with a different, non-homogeneous background.

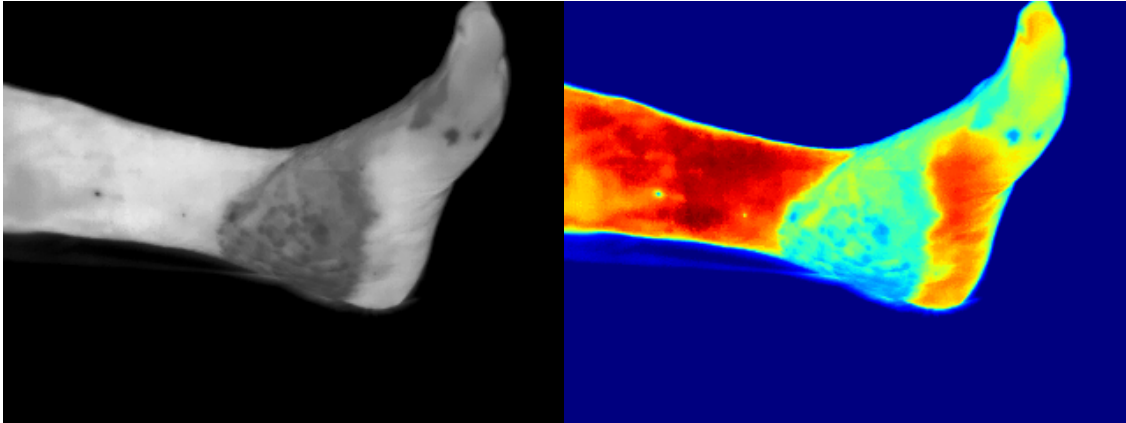


Figure 4.7. Original image

Figure 4.8. Image with applied colourmap

The numerosity of Construction set is further increased by performing horizontal and vertical flips of all images. 60% of the images in the Construction Set are used as Training Sets, the remaining images are used as Validation Sets.

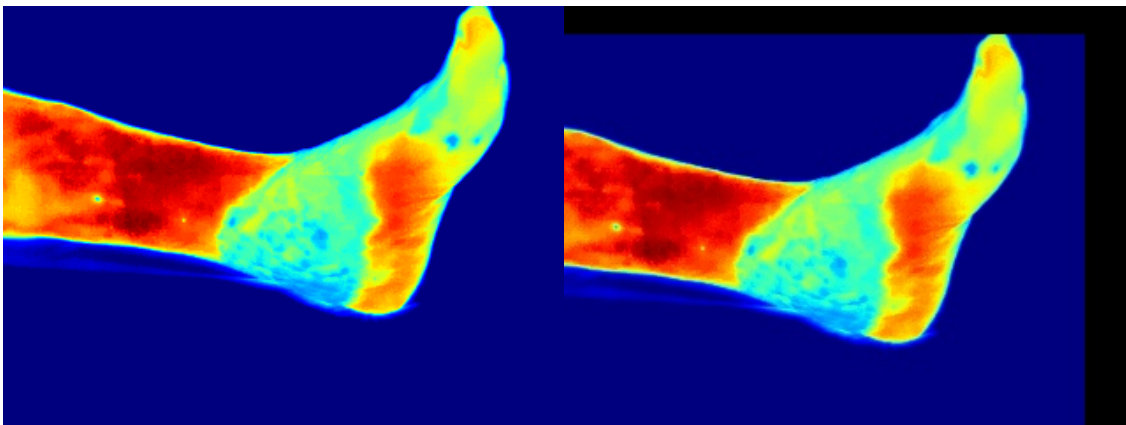


Figure 4.9. Original image

Figure 4.10. Image with random translation

4.3 Segmentation and classification of wounds

Segmenting and measuring the area correctly in wound care is a very important task. The area of the wound over time gives fundamental information such as the status of the healing progress. Three networks were chosen based on literature sources:

- *U-net*[38] is a very popular structure for segmentation, often used and validated in wound care for wound segmentation [53].
- *PSP-Resnet50* this network has achieved good segmentation performance in several searches [54], it also has a Resnet backbone which is one of the best networks for extracting features from wound images[53].
- *YOLOv8*[55], represents the gold standard of object detection in real time [56] [53] , but it is also capable of classification and segmentation. This model has been also used in wound care to classify various state of DFU ulcers [57]. *YOLOv8*[55] was selected to evaluate the classification of ulcers obtainable with thermal imaging. Wound classification is equally important so that the patient receives the right medical attention and avoids infection and amputation.

The training of PSP-Resnet and YOLOv8 was carried out using a *Transfer Learning* approach. Respectively, the training starts from the weights of the ImageNet dataset and the COCO (commun object in context) dataset. It was shown that transfer learning enables better model convergence in the case of a small dataset and prevents overfitting. Furthermore, several studies show that the images of the initial training dataset are not decisive for the final performance on the medical dataset.[58] Models were implemented in Python using Google Colab T4 GPU.

4.3.1 U-Net

The architecture has 3 main steps which are the downsampling path,made of 2 convolutional layers, each followed by the ReLU 4.11 activation function.

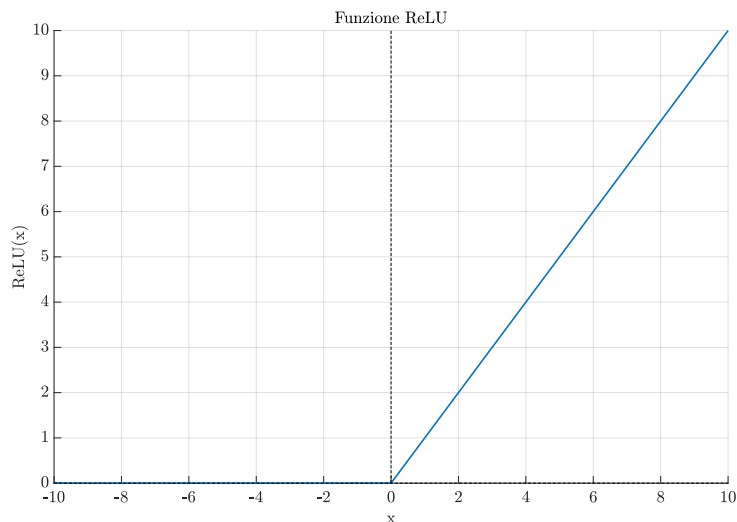


Figure 4.11. Graph showing the trend of ReLu loss function

The second part is the bottleneck, followed by the upsampling path. This last part is made of upsampling operations of the feature map followed by convolution layers **Figure 3.3**. The code used is the one provided by Caffe implementation [38]. The images and their respective masks were normalised in the range of 0 to 1 and resized with final dimension of 256x256. The number of epochs chosen for training was 40 based on the improvement in loss function after the following epoch showed an increase without improvement in performance: this parameter was monitored through early stopping. Early stopping is a method to prevent overfitting which allows to choose a *patience*, representing the maximum number of epochs of training without increasing performance. Initially, 100 epochs were set [59] in order to figure out the best number of epochs to use. The loss function used is the binary cross entropy 4.1 [60]. The function takes a value of 0 when y_m and x_m are equal, in other cases the value of the function is positive.

$$J_{bce} = -\frac{1}{M} \sum_{m=1}^M [y_m \times \log(h_{\theta}(x_m)) + (1 - y_m) \times \log(1 - h_{\theta}(x_m))] \quad (4.1)$$

In 4.1[61] M is the number of training samples, y_m is the target label for training images, x_m is the input image of the training, h is the network model with weights. Binary cross entropy loss is commonly used in binary classifications. It measures the discrepancy between the actual labels and the model's predictions. The decrease in the loss function indicates the level of learning of the model. When the model stops learning, the loss value stops decreasing, based on this principle the epochs for training are chosen to prevent overfitting. Values of validation loss function relating different epochs is shown in **Figure 4.12**

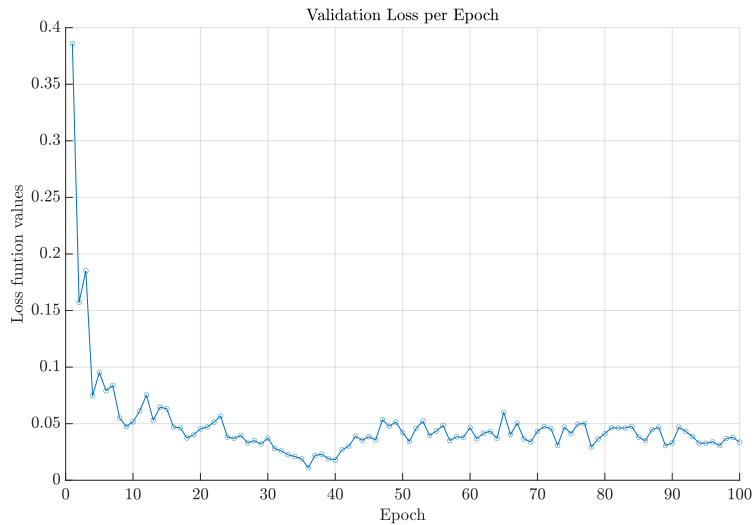


Figure 4.12. Values of loss function for each epoch

The first layer of the network was modified with the input image dimensions. The output of the network is the feature map, which is thresholded to obtain the output binary mask. The output image values are related with the probability of the network prediction 4.13, threshold was evaluated with ROC curve and set at 0.7 in order to obtain the final mask 4.14.



Figure 4.13. Output feature map

Figure 4.14. Final prediction of the model

Figure 4.15. Ground truth mask

4.3.2 PSP-ResNet50

This type of network consists of a Pyramid scene parsing network (PSP) [62] decoder with ResNet50 backbone encoder 4.16. PSPNet is a model capable of semantic segmentation that thanks to a pyramid parsing module analyses global context information by how the features are aggregated in the different parts of the input image. PSPNet combined with a pretrained CNN is able to extract the feature map [62].

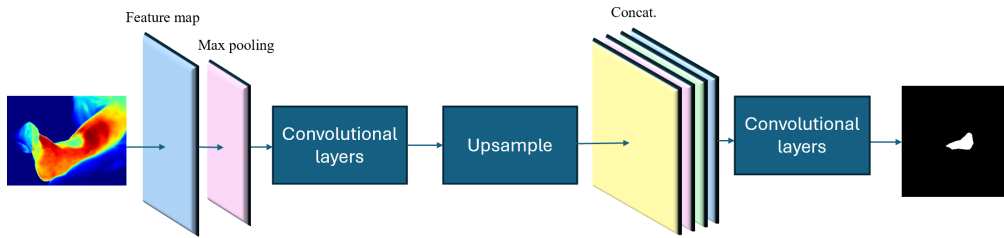


Figure 4.16. Diagram of PSP-Resnet structure

The ResNet model was devised by He et al.[63] to solve the convergence and degradation problems of deep neural networks, the concept of residual learning is introduced. A residual network is characterised by shortcut connections that connect the convolutional layers to each other [63]. The training begins with weights from mmsegmentation[64] pre-train weights library[65]. Since the model is pre-trained, the input images must be normalised. Another parameter to be set is the batch size, chosen to be 8. For the purpose of the binary segmentation $out_channels = 1$ was set. A legend is then set so that the model is able to perform semantic segmentation using ground truth to identify areas of interest. The images were resized 256x256 and normalised between 0 and 1. Palette parameters were then set to recognise and segment class 0 as *background* and class 1 as *wound*. The number of epoch was initially set to 133 with 5000 iterations: the objective was obtain around 100 epochs for training according to literature founds [59].

$$\text{Number of iterations for each epoch} = \frac{\text{Dimension of dataset}}{\text{Dimension of batch}}$$

$$\text{Number of epochs} = \frac{\text{Number of iterations}}{\text{Number of iterations for each epoch}}$$

Cross entropy loss function was used during the training of the network 4.1. Early stopping was set at 10 epoch patience and did not take place. The last checkpoint saved are used to evaluate results: the palette for visualising results was $background=[0,0,0]$ and $wound=[255,255,255]$. The optimized model after training on thermal images dataset is composed by 124 layers.

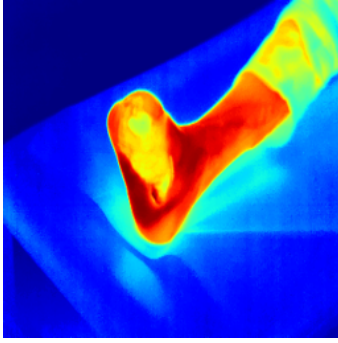


Figure 4.17.



Figure 4.18.



Figure 4.19.

4.3.3 YOLOv8-seg

YOLO (you only look once) is a CNN-type network based on a regression mechanism with a CSPdarknet backbone⁵³ that implements two types of activation functions: on the final layer, the sigmoid function 4.2 is used, while on the intermediate layers, the SiLu function 4.3 is used.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (4.2)$$

$$\text{SiLU}(x) = x \cdot \sigma(x) = x \cdot \frac{1}{1 + e^{-x}} \quad (4.3)$$

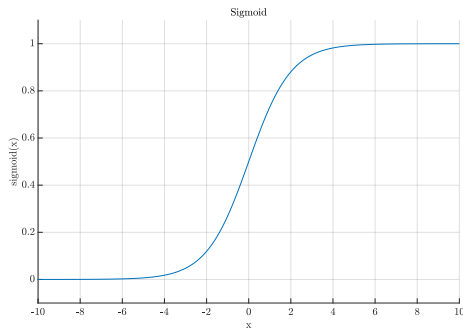


Figure 4.20. Graph showing sigmoid activation function

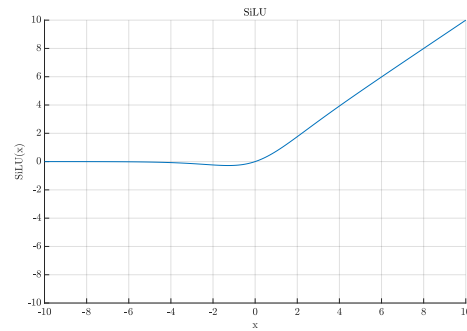


Figure 4.21. Graph showing SiLu activation function

It is a network with a very complex structure 4.22 capable of performing classification, detection and segmentation tasks. Features are extracted via the CSPDarknet53 backbone, also thanks to the C2f module that diversifies the normal YOLO neck structure 4.23. The module precedes the segmentation heads that are able to learn to predict the segmentation based on the annotations provided for the input images. There are 5 detection modules with 1 prediction layer, while the detection heads are similar to those of basic YOLOv8. YOLOv8-Seg model outperforms on various object detection and semantic segmentation benchmarks while maintaining high speed[56]. Since the network is capable of performing several tasks, it implements more than one loss function to define the learning improvement. The different loss functions used are in fact specific to the task whose learning evaluation they perform. In particular, the BCE is used for the classification branch and the Complete IoU together with the distributional focal loss for the branch that deals with segmentation and bounding box prediction [56]. The complete IoU 4.4 is an evolution of the other loss functions based on intersection over union. It takes into account in particular the overlapping areas of the bounding boxes belonging to the ground truth and those predicted, together with the distance between the centres of the bounding boxes [66].

$$CIoU = 1 - IoU + \frac{\rho^2(\mathbf{b}, \mathbf{b}_{gt})}{c^2} + \alpha v \quad (4.4)$$

In the formula [66] (b, b_{gt}) indicates the Euclidean distance between the centres of the two bounding boxes, while c represents the length of the diagonal of the rectangular frame that can contain the two bounding boxes, v and α are balancing parameters [66].

$$DFL(S_i, S_{i+1}) = -((y_{i+1} - y) \log(S_i) + (y - y_i) \log(S_{i+1})) \quad (4.5)$$

In 4.5[67] S_i and S_{i+1} are the model-predicted probabilities for the values y_i and y_{i+1} . y : is the target value we are trying to predict. y_i and y_{i+1} : are the discrete values that bound y . The

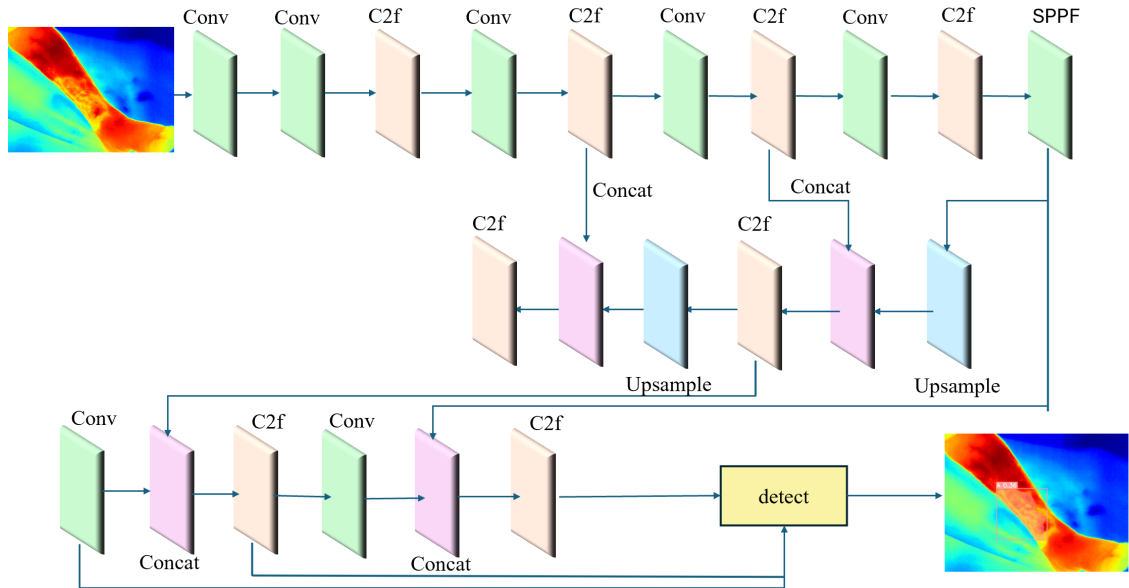


Figure 4.22. Diagram of YOLOv8 detection algorithm

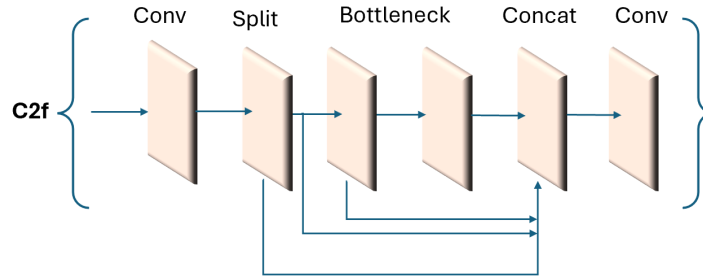


Figure 4.23. Detail of C2f module structure

network was pre-trained on the public COCO(Common object in context) dataset. The configuration file of the pre-trained network has 80 detectable object classes, which are overwritten with the classes of interest: background and wound. The network needs to provide for each training image, thus both training and validation set, a text file annotating the contours of the object to be detected normalised with respect to the size of the image. Together with the normalised contour for each object, the numerical code of the class to which it belongs must be written, which is indicated in the configuration file. It is used with task *segment*, from the available pre-trained models *'yolov8segX.pt'* was chosen (specification: YOLOv8x-seg summary: 401 layers, 71755663 parameters, 71755647 gradients, 344.5 GFLOPs) as a compromise between computational cost and average model accuracy on the COCO (common objects) dataset. As far as data augmentation is concerned, it is implemented automatically by the yolov8.2 model via integration of the *albumentations* library (parameters: albumentations: Blur(p=0.01, blu_limit=(3, 7)), MedianBlur(p=0.01, blur_limit=(3, 7)), ToGray(p=0.01), CLAHE(p=0.01, clip_limit=(1, 4.0), tile_grid_size=(8, 8)). In addition, several parameters are available that can be varied during training to increase the generalisation capabilities of the model such as mosaic=1.0. Mosaic augmentation is a data augmentation technique that combines four different training images into a single composite image in order to improve the robustness and generalisation capabilities of the model, the functionality is deactivated during the last 10 epochs compared to the initially set epochs. We then set flplr to 0.5, horizontal flipping for augmentation. The batch size is set to 8. The optimizer AdamW type is selected automatically by network. For training process, 200 epochs [68][69] were chosen with patience set to 10. The final training is then done with 110 epochs, the value identified by early stopping. The fused model after training on thermal images dataset is composed by 195 layers.

4.4 Classification of wound status

YOLOv8 was chosen for the classification task: *'yolov8segX.pt'* net weights are also used for this training. The classes set are the 4 classes identified by the WBP score which are overwritten to the initial 80 classes. Annotations were provided for each image for the training phase. At each class was associated a number from 0 to 4 in order provide .txt files for the network training.

Class	Code associated
A	—————→ 0
B	—————→ 1
C	—————→ 2
D	—————→ 3

The epochs initially set are 200, value chosen from literature [68][69], decreased to 135 after the early stopping intervention (patience set at 10 epochs without increasing of performance). The fused model after training on thermal images dataset is composed by 206 layers. The network is able also to segment and detect elements, useful since the images are not cropped, but since the main task of interest is classification initial loss function are modified as follow: $box_loss=5$, $cls_loss=3$, $seg_loss=0.5$. The value of class loss is increased from 0.5 to 3 so that the model focuses on the classification task, penalising errors due to the task to a greater extent. For this task WoundViewer algorithm’s classification is set as ground truth.

4.5 Metrics for evaluation of performances

The results obtained were analysed using specific metrics calculated by comparing the ground truth masks with those obtained after training the models. Precision, recall and Dice were chosen. The values for the segmentation task were calculated for each image and then averaged to obtain a value representative of performance. The Python library `classification_report` imported from `sklearn.metrics` was used for the calculation. In particular, the metrics are calculated after obtaining the automatic masks with the different inference processes of the different networks. Pixels with a value of 255 (white) represent the detected object while pixels with a value of 0 (black) represent the background. If the pixels of automatic mask and ground truth have the same value they are true positive if they represent the object to be detected (TP) and true negative (TN) if they represent the background, if the pixels were meant to represent the object but are detected as background they indicate a false negative (FN) and on the contrary a false positive (FP).

Recall (Sensitivity)

$$\text{Recall (Rec)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4.6)$$

Precision (Positive Predictive Value)

$$\text{Precision (Pre)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (4.7)$$

F-Score (Dice)

$$\text{F-Score} = \frac{2 \times \text{Pre} \times \text{Rec}}{\text{Pre} + \text{Rec}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (4.8)$$

These metrics were chosen through literature analysis, as they were found to be more meaningful for assessing the performances of wound segmentation algorithms [70] [71]. Recall and Precision

are essential to understand whether the model is sensitive to over- or under-segmentation. Dice is obtained by combining precision and recall and returns an indicative value of how much the predicted area overlaps with ground truth[72].

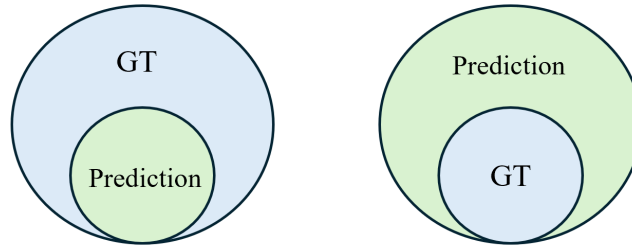


Figure 4.24. Comparison between ground truth(GT) and algorithm segmentation

As shown in the figure 4.24, unless there is a perfect overlap between ground truth and automatic mask, we may find ourselves in one of two situations. With a very high precision and a low recall we know that we will be in an under-segmentation situation, while with a particularly high recall and low precision the prediction will be tending to over-segmentation. These two cases in relation to the clinical case at hand have different values. Over-segmentation would be preferable, as a wound that does not heal and maintains a large size attracts the physician’s attention, whereas if a decrease in size is detected by mistake, some data of fundamental importance for a correct follow-up may be missed.

4.6 Results for segmentation task

All three models were tested on the images of the Test set as they represent 3 alternative pipelines for the intended purpose. The best results are obtained using the PSP-Resnet50 model, neither the U-net nor the YOLO exceeded the 50% Dice on the test set images. There is no significant discrepancy for any model with regard to precision and recall values, indicating that there is no particular tendency for over- or under-segmentation. The coefficient values for the Training set are generally higher than the Validation and Test set values as the model is trained to recognise those images. The YOLO model is the only one that also shows poor results on the Training set.

Table 4.3. U-net segmentation performance metrics

	Training set	Validation set	Test set
Average_Precision	66%	60%	41%
Average_Recall	72%	66%	52%
Average_Dice	63%	57%	39%

Table 4.4. PSP-Resnet50 segmentation performance metrics

	Training set	Validation set	Test set
Average_Precision	83%	80%	75%
Average_Recall	79%	76%	70%
Average_Dice	80%	77%	72%

Table 4.5. Yolov8 segmentation performance metrics

	Training set	Validation set	Test set
Average_Precision	45%	45%	24%
Average_Recall	50%	47%	34%
Average_Dice	44%	42%	21%

4.7 Results for classification task

Confusion matrices were derived for Training, Validation and Test sets. The metrics of Class Accuracy, Balanced Accuracy and Error Rate were calculated from the confusion matrices. Again, it can be seen that better results are obtained for the Training set because the model is trained to recognise those particular images. Speaking of the Validation and Test set, misclassification phenomena only occur for type A images but what impacts performance are the unclassified images.

Class Accuracy

$$\text{Class Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.9)$$

Balanced Accuracy

$$\text{Balanced Accuracy} = \frac{1}{2} \left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right) \quad (4.10)$$

Error Rate

$$\text{Error Rate} = \frac{FP + FN}{TP + TN + FP + FN} \quad (4.11)$$

		True classification			
		A	B	C	D
Classifier result	A	78			
	B		64		
	C			99	
	D				125
Not classified		30	32	27	10

Figure 4.25. Confusion matrix of Training set image classifications

Table 4.6. Fine tuned-Yolov8 classification performances metrics on Training set: metrics for each class

	A	B	C	D
Class accuracy	72%	67%	79%	93%

Table 4.7. Fine tuned-Yolov8 classification performances metrics on Training set: general metrics

Error rate	21%
Balanced accuracy	78%

		True classification			
		A	B	C	D
Classifier result	A	9			
	B	4	13		
	C			13	
	D				11
Not classified		12	9	3	1

Figure 4.26. Confusion matrix of Validation set image classifications

Table 4.8. Fine tuned-Yolov8 classification performances metrics on Validation set: metrics for each class

	A	B	C	D
Class accuracy	36%	59%	81%	92%

Table 4.9. Fine tuned-Yolov8 classification performances metrics on Validation set: general metrics

Error rate	33%
Balanced accuracy	71%

		True classification			
		A	B	C	D
Classifier result	A	11			
	B	5	17		
	C			10	
	D				9
Not classified		9	5	4	1

Figure 4.27. Confusion matrix of Test set image classifications

Table 4.10. Fine tuned-Yolov8 classification performances metrics on Validation set: metrics for each class

	A	B	C	D
Class accuracy	46%	77%	71%	90%

Table 4.11. Fine tuned-Yolov8 classification performances metrics on Validation set: general metrics

Error rate	38%
Balanced accuracy	67%

4.8 Discussion

In the light of the results obtained, useful information can be deduced regarding the analysis of thermal images using deep learning methods, but there is also room for improvement. The greatest difficulty was finding a dataset that lent itself to this analysis: the corresponding RGB images of the thermal images were required in order to create a reliable ground truth from which conclusions could be drawn by comparing performances of the various models. Few images were found for a more in-depth analysis; it is well known that deep learning needs a large amount of data to be able to generalise correctly. An example in the field of wound care is the training of the WoundViewer algorithm, which was performed with around 3,000 images from clinical trials [42]. Another problem was the low resolution of the images, which can further challenge the algorithms to recognise small details. Segmentation has proven to be a difficult task from thermal images, the reason may simply be that thermal images as such highlight temperature differences that in grade A wounds 4.29, thus composed of 100% granulation tissue, are not often evident because they are not characterized by marked vascular reorganization. In that case, the information that the neural network extracts from the thermal image may not be sufficient to highlight a sharp contour. In the pictures 4.28 4.29 in the figure is noticeable how the difference in temperature distribution on the patient’s leg with D wound type is evident 4.28.

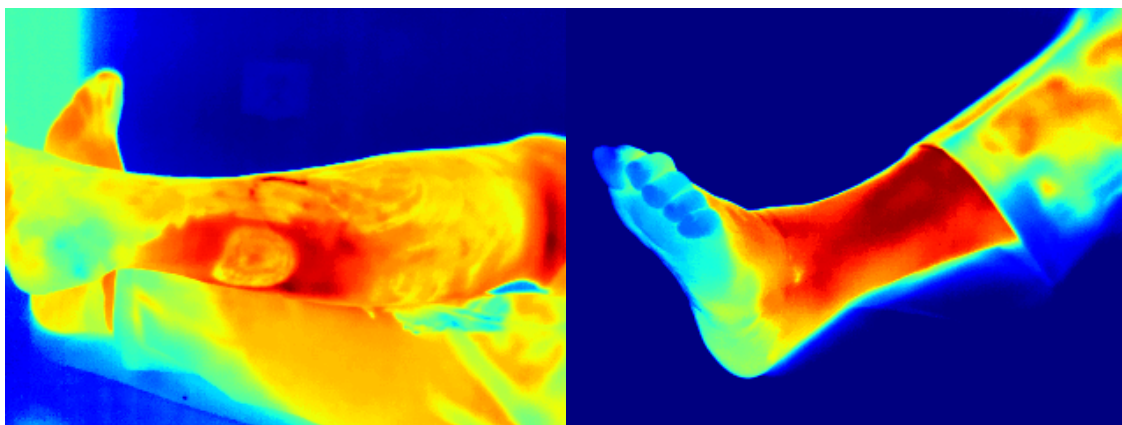


Figure 4.28. Image with a type D wound

Figure 4.29. Image with a type A wound

The YOLO model showed the worst results, probably its more complex structure converges with more difficulty, thus making it uncompetitive with simpler models such as the U-Net and PSP-Resnet50. Different performance is also found between the latter two networks, better for the PSP-Resnet50, attributable to a feature of the structure, the residual module, used by the Resnet50 backbone but not the U-Net. With the use of the residual unit, in fact, the Resnet is able to converge faster while avoiding overfitting. Resnet residual units are designed to solve deep neural network training problems and solve vanishing gradient problems, for this application this difference in architecture may have been crucial and allowed the backbone Resnet type to make a difference. In addition, Resnet has been called the gold standard of wound classification several times in the field of vulnology [53], so using such a backbone for feature extraction, then used by the decoder for segmentation, can be a winning solution. The results show that the Dice on the Test set obtained by the model in question is the only one above 50%, making the model’s ability to recognize wound and background actually useful. Regarding Recall and Precision values, a good balance of the two values is noted, indicating a good estimate of the wound area. It is

important to remember that if a physician support system is to be built, it must show a certain reliability, if not its use becomes superfluous or even misleading, failing the main purpose of the analysis. A very important task in this clinical setting is the correct classification of the degree of wound healing and infection. Incorrect classification can in fact lead to underestimation of a wound, which if left untreated can lead to amputation or death in the worst case. The results of the YOLOv8 model used for detection and classification show balanced accuracy greater than 50 per cent for the Test set indicating a mostly correct classification that can be improved with more extensive training. The network shows that it was able to generalise and recognise all 4 types of injuries. The greatest difficulty is in the identification of type A wounds, evidently the features that the network tries to identify in the thermal image are often not sufficient to make a confident classification. It can be seen that the misclassification is very low and occurs between type A and B images. This is not worrying from the point of view of learning of the algorithm, often in clinical practice these two classes are combined into one. What is of interest is to discern between a wound with more granulation tissue and not infected and a wound that is necrotic and infected. This objective is fulfilled as all type D images in the test set are recognised. The trained algorithm is therefore shown to be useful for the recognition of severe wounds tending to worsen with 90% of accuracy. Precision across all classes is good, but this figure needs to be put into context, as the parameter does not take into account the high number of unclassifieds which are still a problem. This phenomenon can also be attributed to the fact that the input image to be analyzed does not represent only the wound but the whole environment, so some elements may lead the inference procedure not to have a high enough confidence in detecting the area of the image that affects the wound. For this reason YOLOv8 had been chosen, so as to also keep the detection active thus training the network to recognize the wound area and classify only that but not the whole image. The problem could have been solved by choosing to perform cropping on the wound area, but since we wanted to keep a worst case situation and test how the network could behave with this added difficulty the cropping was not done. For this specific application, i.e., classification of chronic wounds using thermal imaging, it has been difficult to find examples in the literature especially for the training parameters of neural network models as it appears to be a still nascent application. It was not possible to find a model that was capable of performing all the tasks of interest, so a combined model can be proposed, with a network capable of segmenting the image well with a network capable of classifying following in cascade. A similar approach is put into practice by the WoundViewer algorithm, where the first network recognizes the area of interest in the image where the wound is located and the second analyzes the area found to find the percentages of granulation or necrotic tissue.

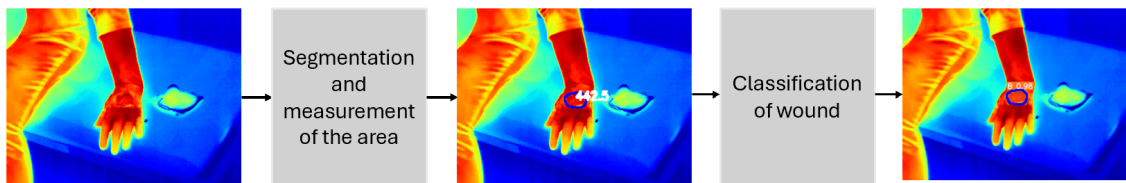
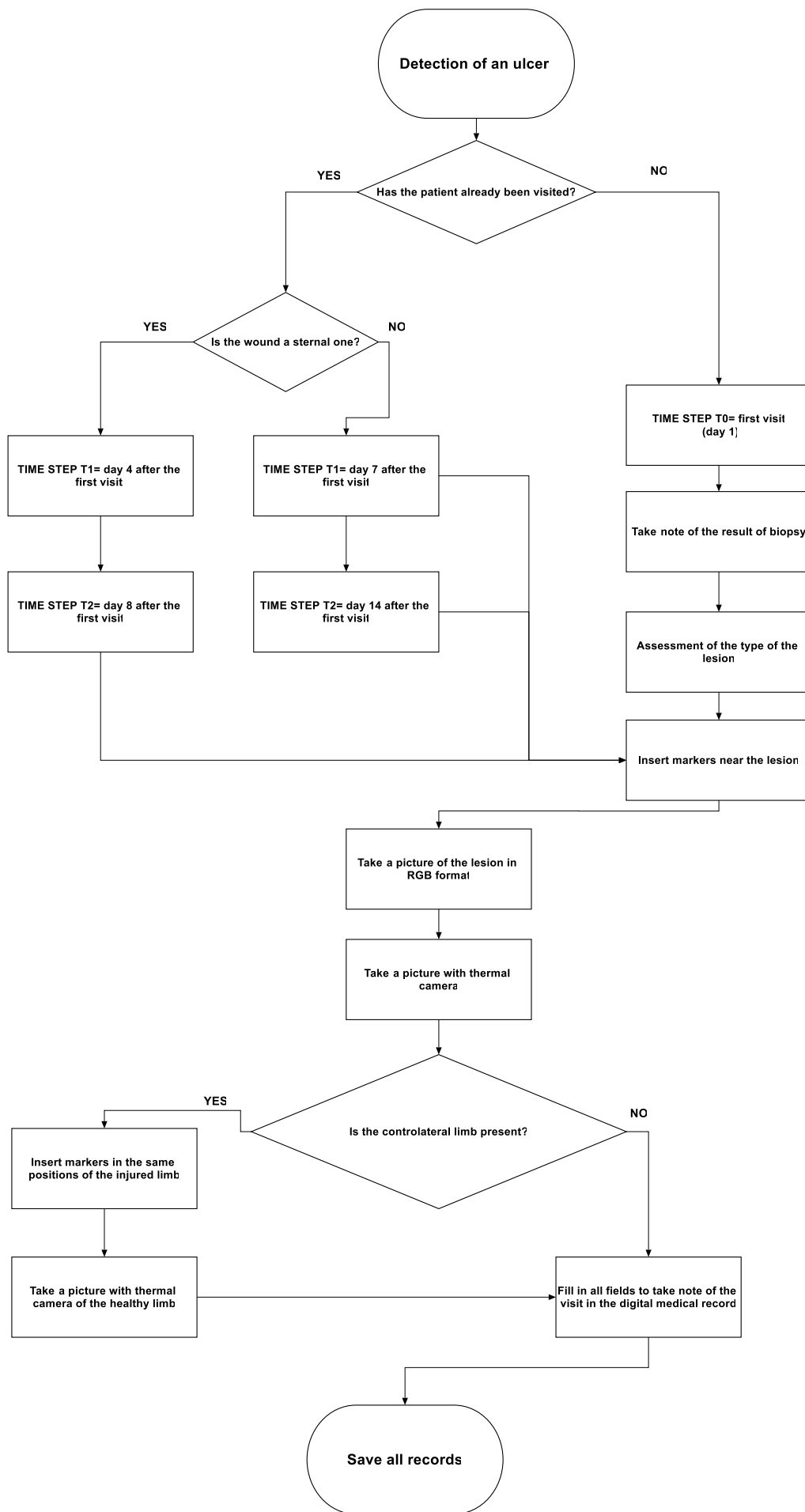


Figure 4.30. Pipeline about the model proposed

As is shown in the **Figure4.30**, the proposed model to obtain a complete assessment of the chronic injury would be the set of the network to segment and measure the area, the measurement of which is expressed in pixels not having the scale factor to carry the measurement in cm, and to follow the neural network model that can classify what has been previously identified. Thermal imaging represents a new frontier with much potential in wound care; however, a lot

of data is needed for analysis with deep learning algorithms. The data needed to develop and implement ad hoc models will be collected at the clinical trial that will take place at Niguarda Hospital. Preliminary results show great promise despite the quality of the dataset. A clinical protocol has been designed so as to collect all the information needed to understand the wound infection status, such as the biopsy result, which will be used as ground truth to classify the wound as infected or not through the images provided. Visits will be made on fixed days so as to have a complete dataset that allows for the development of models that understand the possible evolution of the same wound over time, which was not possible in this preliminary analysis because there were few patients compared to the total number for whom data from multiple visits were available. Key thing will be to acquire both RGB images and thermals so that information can be extracted from both, for example, segmentation and area measurement from RGB images and infection degree measurement from thermals. Even in light of the data provided by the preliminary analysis, it can be seen that the classification task was performed more efficiently on thermal images than segmentation. Thus, the idea of a combined model that can analyze both RGB and thermal images and extract the most interesting features for both types of images is not ruled out.

Thermography has the potential to become the new gold standard for predicting wound infections, aligning well with observations made by wound healing professionals along with new techniques of deep and machine learning. The studies reviewed confirm the reliability of this method. Combining traditional thermal imaging with modern digital processing techniques presents an excellent, non-invasive, cost-effective, and time-efficient solution for diagnosing wound infections. This method's versatility makes it suitable for various types of ulcers and different settings, including remote clinical care, home care, and long-term hospital facilities, underscoring its reliability. Quick diagnosis and prompt medical intervention also reduce hospital costs and enhance patient care efficiency. Establishing a centralized diagnostic method, supported by other biomarkers and parameters, would facilitate more effective patient recovery and healing. Given the promising results from other studies and the high costs and complexity of current medical imaging techniques, thermography appears to be a solid approach. By linking thermal images to a set of static (from single examinations) and dynamic parameters (from multiple medical visits), we can enhance diagnostic precision and specificity of different neural network architectures. A clinical trial, of which the flowchart is shown below, to refine this diagnostic method will improve accuracy, making thermography the preferred technique for diagnosing skin ulcer infections, benefiting both patient health and cost-effectiveness, allowing further investigation of what features of these images can be artificially processed in order to create intelligent and efficient models to support examination and diagnosis.



Appendix A

Clinical trial protocol

Protocollo sperimentazione clinica

“TERMOGRAFIA PER LA DIAGNOSI CLINICA DELLE INFEZIONI NELLE ULCERE CUTANEE ”

Trial spontaneo

Synopsis

Versione 15 novembre 2023

Titolo dello studio: Studio di proof of concept della correlazione tra le immagini termografiche e lo stato di infezione nelle lesioni cutanee croniche ulcerative per determinare la situazione clinico-patologica dell'infiammazione.

Coordinatori dello Studio: Dott. Giovanni Sesana, Chirurgia Vascolare, ASST Grande Ospedale Metropolitano Niguarda (PI clinico)

Prof. Jacopo Secco, Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Torino (PI non clinico)

Razionale: Allo stato attuale l'insorgenza di un'infezione viene osservata tipicamente attraverso le valutazioni soggettive dell'operatore, i parametri fisiologici del paziente (analisi del sangue, urine, temperatura corporea), esami diagnostici strumentali e le analisi microbiologiche. La presenza di un'infezione definisce chiaramente il rallentamento del processo di guarigione della ferita e il suo conseguente peggioramento, con effetti sulla condizione clinica globale del paziente sia nel periodo di ricovero sia nel post-operatorio. I dispositivi attualmente presenti sul mercato sono difficilmente utilizzabili in quanto manca un processo di individuazione automatica dell'infezione della ferita e viene quindi richiesta una partecipazione dell'operatore. In particolare, le tecniche strumentali di CT, MRI, PET sono troppo onerose e complesse per l'analisi dell'ulcera; dall'altro lato le analisi microbiologiche sono lente e costose. La prima diagnosi

dell'infezione deve essere il più possibile rapida e facile, attraverso un protocollo semplice ed efficace d'azione standardizzato per il pronto intervento nella cura della lesione. Lo scopo del presente studio è dunque quello di utilizzare immagini termografiche della ferita che consentano un'automatica identificazione, in maniera precisa, dell'infezione della ferita con un'accuratezza altrettanto adeguata a quella delle tecniche di diagnosi tradizionali. Premesso che la valutazione del letto di ferita è il parametro clinico più importante, utilizzato sia come indicatore dell'evoluzione del trattamento in corso, sia come guida per le decisioni terapeutiche, l'analisi termografica permetterebbe facilmente di valutare anche lo stato infettivo dell'ulcera che spesso rappresenta la prima causa di peggioramento della stessa. La situazione attuale è determinata in modo preponderante dall'esperienza degli operatori, che ovviamente è soggettiva e difficilmente uniformabile. Il nuovo algoritmo di classificazione dello stato infettivo della ferita, basato sulla correlazione tra analisi termografica della ferita e la sua stessa infezione, si propone di superare sia la soggettività insita nel sistema sia le vecchie tecniche di diagnosi attraverso la standardizzazione della valutazione attraverso l'uso combinato di diversi parametri fisiologici del paziente.

Scopo della sperimentazione sarà quello di verificare che l'analisi termografica riesca ad identificare l'effettiva presenza di un'infezione nell'ulcera cutanea di un paziente, acquisendo un'immagine della ferita con opportuna fotocamera termica volta a fornire dati sullo stato termico e infettivo della lesione, saggiando al contempo la percentuale di affidabilità ed il margine di errore, sia sulle singole acquisizioni dati, sia su campioni statisticamente rilevanti.

Disegno dello studio:

Si tratta di uno studio di tipo osservazionale prospettico in quanto verranno prese immagini termografiche di ulcere con il sistema

integrato Wound-Viewer Lite01, Omnidermal Biomedics srl. Il sistema di analisi e classificazione verrà poi valutato in termini di accuratezza e precisione su tutto il campione statistico preso in considerazione. Lo studio avrà caratteristiche di studio monocentrico. Il tipo di studio non è randomizzabile in quanto il confronto avviene verso parametri che sono rilevati strumentalmente dagli operatori sulle stesse lesioni.

Il fine della presente analisi è volta a testare l'accuratezza e la precisione dell'analisi termografica con l'acquisizione delle immagini termografiche di ulcere cutanee, per la diagnosi di una possibile infezione, come metodo alternativo ma soprattutto più efficace rispetto alle altre tecniche diagnostiche in uso. Nello specifico:

- Analisi microbiologiche
- CT, MRI, PET
- Analisi cliniche
- Valutazione clinica

I parametri misurati e integrati con il metodo di classificazione della presenza dell'infezione, sono raccolti con il dispositivo Wound-Viewer e sono i seguenti:

Dati del paziente, questi saranno pseudo-anonimizzati con chiave di cifratura legante il paziente con il suo codice a database. La chiave di cifratura è accessibile solo da parte del PI clinico (Ospedale Niguarda) attraverso username e password privata:

1. Copia del database dei dati paziente e chiave di cifratura ad uso esclusivo del PI clinico Dott. Giovanni Sesana:
 - a. Nome e Cognome
 - b. ID a database (chiave di cifratura)
2. Copia del medesimo database senza nome, cognome e data di nascita, o comunque mancante di qualsiasi riferimento che potrebbe riportare al riconoscimento del paziente. Riferimento al paziente solo tramite ID.

Dati del paziente al momento della visita:

1. Area della ferita in cm^2 ,
2. Profondità della ferita in mm,
3. Volume della ferita in cm^3 ,
4. Dolore percepito dal paziente (su una scala da 1 a 10),
5. Grado di infezione secondo la scala di Cutting e Harding,
6. Classificazione WBP,
7. Classificazione TEXAS (se piede diabetico),
8. Medicazione e trattamento della ferita,
9. Immagine della ferita in RGB,
10. Aspetto del tessuto della ferita (granulazione percentuale rossa, nera, bianca e gialla),
11. Immagine termografica della ferita e della zona peri lesionale immediatamente circostante,
12. Immagine termografica dell'arto controlaterale sano,
13. Esito Biopsia

Dati della ferita:

1. Eziologia,
2. Età dell'ulcera,
3. Stato clinico,
4. Profondità anatomica,
5. Parte del corpo

La combinazione dei seguenti parametri e dell'elaborazione delle rispettive immagini termografiche acquisite sulle ulcere dei pazienti, dovrà permettere di stabilire un criterio di classificazione della lesione nello stato infiammatorio infettivo o non infettivo attraverso algoritmo di machine learning.

Ciascuna immagine verrà pseudo-anonimizzata ai sensi del regolamento GDPR (UE) 2016/679.

I dati raccolti verranno elaborati all'interno di un algoritmo di machine learning con l'intenzione di riuscire a classificare la ferita come infetta o non infetta.

L'algoritmo di machine learning si basa su una rete neurale che riceve in ingresso una serie di feature del paziente: immagine termografica, immagini tridimensionali e bidimensionali, misurazioni della lesione in termini di parametri osservazionali, valori clinici fisiologici, punteggi di valutazione della ferita come il WBP (Wound Bed Preparation Score), composizione tissutale dell'ulcera. L'architettura della rete neurale e il suo algoritmo di classificazione prende ispirazione da metodi di analisi non lineari, sistemi neuromorfici di calcolo, basati sul network delle cellule neurali, random forest o il modello del perceptron.

L'analisi effettuata sui risultati del metodo formulato si baserà sulle decisioni del classificatore nell'indicare la possibilità di infezione o non-infezione su ciascun paziente esprimendo un risultato codificato in maniera binaria. Per far ciò si valuta l'effettiva accuratezza dell'algoritmo di classificazione secondo diversi parametri: accuratezza, recall, F1, score, precisione, curva ROC, mean square error, coefficiente di correlazione, percentuale di falsi negativi e falsi positivi, veri positivi e veri negativi, tramite matrice di confusione. Il valore di tali parametri è da considerarsi come diretta conseguenza del raggiungimento degli endpoint primari e secondari descritti di seguito. A seguito di questa prima valutazione sulla totalità del campione, seguiranno valutazioni dei seguenti parametri per ciascun gruppo eziologico delle ferite. Per il caso di classificatore binario, il criterio di metodo di valutazione sarà basato sul metodo holdout. Per questo i dati raccolti e forniti al classificatore saranno precedentemente divisi in dati di train, test e validation, già targhettati per addestrare l'algoritmo di apprendimento supervisionato.

Questo genere di analisi ha lo scopo di stabilire la sensibilità dell'algoritmo di classificazione e paragonarla con il reale stato infettivo della lesione. La valutazione della bontà del classificatore è basata sulla numerosità del campione scelto.

Popolazione

Verranno arruolati 270 pazienti suddivisi in 6 gruppi in base alle seguenti tipologie di ulcere. La costruzione del campione è stata fatta cercando di arruolare quanti più pazienti possibile compatibilmente con i tempi di valutazione della ferita del medico, la disponibilità di pazienti all'inserimento del trial, la costruzione di un campione statisticamente rilevante, come suggerito dalla letteratura e il soddisfacimento dei criteri di inclusione da parte dei pazienti. Inoltre, il criterio per stabilire la numerosità del campione, affinché fosse statisticamente significativo, si è basato sul reclutamento di dati in maniera da rappresentare equamente ciascun tipo eziologico di lesione. I tipi eziologici scelti sono stati eletti in base alla casistica clinica più comune di lesione, data da indagini preliminari sul luogo della sperimentazione. Per ogni gruppo consideriamo un gruppo di controllo di pazienti non infetti (15 soggetti per ogni eziologia) e un gruppo di lesioni infette già diagnosticate in rapporto 1:2 in maniera evidenziare la sensibilità dell'algoritmo nel diagnosticare l'infezione.

- Ulcere dell'arto inferiore minimo 30 (+ 15 non infetti) lesioni infette
- Lesioni in piede diabetico minimo 30 (+ 15 non infetti) lesioni infette
- Ulcere da pressione minimo 30 (+ 15 non infetti) lesioni infette
- Ulcere acute minimo 30 (+ 15 non infetti) lesioni infette
- Lesioni da patologia autoimmuni minimo 30 (+ 15 non infetti) infette
- Ferite chirurgiche sternali da cardiocirurgia, chirurgia generale e vascolare minimo 30 (+ 15 non infetti) infette.

Numero di visite:

- 3 visite al T0, T1, T2 per tutte le lesioni eccetto le ferite sternali: le visite verranno programmate a distanza di 7 giorni una dall'altra, considerando il T0 al giorno 1, T1 al giorno 7 e T2 al giorno 16.

- 3 visite al T0, T1, T2 per ferite sternali: le visite verranno programmate considerando: T0 al primo giorno post-intervento, T1 al quarto giorno post-intervento, T2 all'ottavo giorno post-intervento.

Centri di studio:

- 1 sito clinico, ASST Grande Ospedale Metropolitano Niguarda, Milano, piazza Ospedale Maggiore 3 – 20162

Strumentazione:

- Biopsia
- Etichetta cartacea da porre nella zona sana limitrofa alla ferita per avere un riferimento nelle immagini.
- Wound Viewer Lite01 by Omnidermal Biomedics srl, integrato con FLIR T620 Thermal Imager, FLIR Systems Boston, MA. Manuale d'uso e certificazione CE sono in allegato nella documentazione.

Criteri di inclusione:

I pazienti devono soddisfare tutti i seguenti criteri per essere ammissibili per l'inclusione nello studio:

1. Paziente di età > 18 anni;
2. Paziente sottoposto a cure standard;
3. Ottenimento del consenso informato;
4. Lesioni > 2 cm²;
5. Lesioni < 100 cm²;

Criteri di esclusione:

I pazienti che hanno incontrato uno dei seguenti criteri (screening visita) non sono ammessi a partecipare nello studio:

1. Mancato ottenimento del consenso informato;
2. Lesioni < 2 cm²;
3. Lesioni > 100 cm²

Criteri di sospensione:

non previsti

EndpointsEndpoint Primario:

- Valutazione della correlazione tra immagini termografiche e stato infiammatorio dell'ulcera tramite coefficiente di correlazione sulla base dei risultati e i parametri di valutazione ottenuti sull'algoritmo di classificazione elaborato.
- Valutazione dell'accuratezza e dell'affidabilità del sistema di classificazione dello stato infettivo dell'ulcera basato sull'analisi termografica. Questo verrà effettuato tramite il parametro di accuratezza ottenuto dai risultati del classificatore, specificatamente dalle percentuali di falsi positivi e negativi e veri positivi e negativi. L'affidabilità verrà valutata tramite i parametri di sensibilità e specificità basati sui risultati del classificatore.

Endpoint Secondari:

- Valutazione della presenza dell'infezione sulla totalità del campione.
- Valutazione della perfusione del sito lesionato.
- Valutazione della presenza dell'infezione su ulcere di diversa etiologia.

Metodi:

Valutazione nella stessa sessione della lesione mediante giudizio clinico con il dispositivo Wound Viewer Lite01 per raccogliere i dati del paziente e lo stato della ferita attraverso parametri già menzionati e immagini fotografiche.

Sono previste 3 visite per ogni ferita: per tutte ad eccezione delle ferite sternali verranno valutate al giorno 1, 7 e 16, mentre le ferite sternali saranno sottoposte a valutazione nel primo, quarto e ottavo giorno post-operatorio.

Tutti i pazienti effettueranno una biopsia per valutare lo stato infetto o meno della ferita. La biopsia consiste nella detersione con soluzione fisiologica della ferita e il prelievo di tessuto profondo che viene poi inviato in laboratorio utilizzando la provetta ESWAB. Il campione viene poi inviato con impegnativa M501UL al laboratorio di analisi che provvede ad elaborare l'esito dell'esame clinico. Successivamente il laboratorio analisi performerà l'esame secondo il protocollo standard allegato nel file 'Protocollo Microbiologia'.

Il risultato permetterà di reclutare pazienti nella giusta coorte di infetti o non infetti fino al raggiungimento del numero minimo di soggetti richiesti dalla definizione della popolazione.

Gli esami e i dati dei pazienti, ricoverati o in ambulatorio, verranno raccolti in centro ospedaliero o ambulatoriale, inviati e archiviati nel

sistema cloud del dispositivo Wound Viewer. Nello specifico l'esito della biopsia verrà trasmesso secondo il protocollo e il sistema sanitario standard al medico richiedente l'esame (PI clinico), il quale registrerà a sua volta l'informazione della biopsia (infetto/non infetto), nell'apposito spazio 'Note' presente nel sistema digitale integrato WV quando si performa la visita del paziente. Il sistema prevede l'utilizzo di CRF digitale già implementata nel dispositivo. Il software implementato all'interno del dispositivo, atto a raccogliere i dati della sperimentazione, è certificato anch'esso come dispositivo medico rientrante nella certificazione del dispositivo stesso, allegata alla documentazione. Si può accedere ai dati del dispositivo esclusivamente tramite credenziali private fornite al PI clinico dello studio. Tali dati vengono automaticamente pseudo-anonimizzati dal sistema (i.e. le chiavi di cifratura rimangono accessibili solo tramite dispositivo e relativo software) prima della condivisione con il PI non clinico.

Di tutte le lesioni verrà eseguita un'acquisizione dei parametri attraverso il dispositivo Wound Viewer Lite01 già integrato di fotocamera termografica FLIR per acquisizione delle immagini termografiche. L'immagine termografica verrà acquisita sull'arto affetto da lesione cutanea e a discrezione dell'operatore sull'arto controlaterale sano. Su entrambi gli arti verrà posizionato un marker non invasivo e anti-irritazione, in una zona sana della cute per poter elaborare le immagini termografiche acquisite con eventuali sovrapposizioni con altre immagini RGB. La posizione di tale marker dovrà risultare nella zona in alto a destra della foto.

Parametri e immagini saranno raccolti ed elaborati nell'algoritmo di classificazione come metodo di identificazione dell'infezione.

L'algoritmo di classificazione viene quindi sviluppato sulla base dei campioni raccolti.

I campioni vengono elaborati con un sistema di intelligenza artificiale integrabile nello stesso dispositivo non invasivo e portatile che raccoglie i dati e le immagini del paziente.

L'algoritmo di machine learning si basa su una rete neurale che riceve in ingresso una serie di feature del paziente: immagine termografica, immagini tridimensionali e bidimensionali, misurazioni della lesione in termini di parametri osservazionali, valori clinici fisiologici, punteggi di valutazione della ferita come il WBP (Wound Bed Preparation Score), composizione tissutale dell'ulcera. L'architettura della rete neurale e il suo algoritmo di classificazione prende ispirazione da metodi di analisi non lineari, sistemi neuromorfici di calcolo, basati sul network delle

cellule neurali, random forest o il modello del perceptron. La dimensione della rete viene gestita in base al numero di dati in input. Il processamento dell'immagine viene sottoposto ad una tecnica di computing basata sul memristor, Memristor Cellular automata, progettata e allenata per questo scopo specifico.

Dai dati in possesso possono essere estratte nuove feature, ancora più significative per l'addestramento della rete, quale il gradiente termico spaziale e temporale della temperatura della lesione. Avere feature pre-elaborate e significative da dare in input alla rete, permette di rappresentare e spiegare meglio i risultati ottenuti.

Analisi statistica:**dimensioni del campione:**

270 ulcere cutanee suddivise in 6 gruppi.

- Ulcere dell'arto inferiore minimo 30 (+ 15 non infetti) lesioni infette
- Lesioni in piede diabetico minimo 30 (+ 15 non infetti) lesioni infette
- Ulcere da pressione minimo 30 (+ 15 non infetti) lesioni infette
- Ulcere acute minimo 30 (+ 15 non infetti) lesioni infette
- Lesioni da patologia autoimmuni minimo 30 (+ 15 non infetti) infette
- Ferite chirurgiche sternali da cardiocirurgia, chirurgia generale e vascolare minimo 30 (+ 15 non infetti) infette.

Tipo di analisi:

Sulla totalità del campione:

1. Calcolo della bontà del classificatore con metodo holdout.
2. Valutazione dei parametri di un classificatore binario: accuratezza, precisione, specificità, sensibilità, recall, F1, score, coefficiente di correlazione, percentuali di falsi negativi e positivi, veri negativi e positivi.
3. L'analisi è valutata tramite matrici di confusione e curve ROC.
4. Il coefficiente di correlazione verrà utilizzato per valutare la correlazione tra l'analisi termografica e il grado di infezione.

Questo processo sarà ripetuto per tutti i gruppi eziologici descritti in precedenza.

Affinché l'analisi dia un esito positivo le seguenti ipotesi devono essere verificate secondo i valori attesi presenti nella seguente tabella

La metodologia statistica è stata sviluppata in collaborazione con il Prof. Mauro Gasparini, Ordinario di Statistica al Politecnico di Torino.

Bibliography

- [1] C. Ligresti Chirurgia Plastica Asti Dr.ssa I. Passaro Residente Ch.Plastica Asti. “Linee guida delle ulcere cutanee croniche: nuovo approccio orientativo nella gestione delle ferite difficili.” In: ().
- [2] Marco Farina and Jacopo Secco. “Live Demonstration: 3D Wound Detection and Tracking System Based on Artificial Intelligence Algorithm”. In: ().
- [3] Chandan K. Sen. “Human Wounds and Its Burden: An Updated Compendium of Estimates”. In: *Advances in Wound Care* 8.2 (2019). PMID: 30809421, pp. 39–48. DOI: [10.1089/wound.2019.0946](https://doi.org/10.1089/wound.2019.0946). eprint: <https://doi.org/10.1089/wound.2019.0946>. URL: <https://doi.org/10.1089/wound.2019.0946>.
- [4] N Graves, CJ Phillips, and K Harding. “A narrative review of the epidemiology and economics of chronic wounds”. In: *British Journal of Dermatology* 187.2 (2022), pp. 141–148.
- [5] Brian C.F. Chan et al. “Lifetime cost of chronic ulcers requiring hospitalization in Ontario, Canada: A population-based study”. In: *Wound Medicine* 20 (2018), pp. 21–34.
- [6] Jacopo Secco. “Imaging and Measurement”. In: *Pearls and Pitfalls in Ulcer Management* (2024).
- [7] Jaymie Panuncialman and Vincent Falanga. “The science of wound bed preparation”. In: *Surgical Clinics of North America* 89.3 (2009), pp. 611–626.
- [8] Xuan Wang et al. “Diabetic foot ulcers: Classification, risk factors and management”. In: *World J Diabetes* 13.12 (2022), pp. 1049–1065. DOI: [10.4239/wjd.v13.i12.1049](https://doi.org/10.4239/wjd.v13.i12.1049).
- [9] Lena Victoria Nordheim, Marianne Tveit Haavind, and Marjolein M Iversen. “Effect of telemedicine follow-up care of leg and foot ulcers: a systematic review”. In: *BMC health services research* 14 (2014), pp. 1–12.
- [10] Shuxin Li et al. “Imaging in Chronic Wound Diagnostics”. In: *Advances in Wound Care* 9.5 (2020), pp. 245–263. DOI: [10.1089/wound.2019.0967](https://doi.org/10.1089/wound.2019.0967). eprint: <https://doi.org/10.1089/wound.2019.0967>. URL: <https://doi.org/10.1089/wound.2019.0967>.
- [11] George Kallstrom. “Are quantitative bacterial wound cultures useful?” In: *Journal of clinical microbiology* 52.8 (2014), pp. 2753–2756.
- [12] Jose L. Ramirez-GarciaLuna et al. “Is my wound infected? A study on the use of hyperspectral imaging to assess wound infection”. In: *Frontiers in Medicine* 10 (2023). ISSN: 2296-858X. DOI: [10.3389/fmed.2023.1165281](https://doi.org/10.3389/fmed.2023.1165281). URL: <https://www.frontiersin.org/articles/10.3389/fmed.2023.1165281>.
- [13] Stefan Andersson-Engels et al. “In vivo fluorescence imaging for tissue diagnostics”. In: *Physics in Medicine & Biology* 42.5 (1997), p. 815.
- [14] Emily Blumenthal and Steven LA Jeffery. “The use of the MolecuLight i: X in managing burns: a pilot study”. In: *Journal of Burn Care & Research* 39.1 (2017), pp. 154–161.

- [15] Monica Monici. “Cell and tissue autofluorescence research and diagnostic applications”. In: *Biotechnology annual review* 11 (2005), pp. 227–256.
- [16] Monique Y. Rennie et al. “Understanding Real-Time Fluorescence Signals from Bacteria and Wound Tissues Observed with the MolecuLight i:X”. In: ().
- [17] Roderick A Thomas et al. “Optimized laser application in dermatology using infrared thermography”. In: *Thermosense XXIV*. Vol. 4710. SPIE. 2002, pp. 424–434.
- [18] BR Mason, AJ Graff, and SP Pegg. “Colour thermography in the diagnosis of the depth of burn injury”. In: *Burns* 7.3 (1981), pp. 197–202.
- [19] James B Mercer, Stig Pors Nielsen, and Gerd Hoffmann. “Improvement of wound healing by water-filtered infrared-A (wIRA) in patients with chronic venous stasis ulcers of the lower legs including evaluation using infrared thermography”. In: *GMS german medical science* 6 (2008).
- [20] Harry H Pennes. “Analysis of tissue and arterial blood temperatures in the resting human forearm”. In: *Journal of applied physiology* 1.2 (1948), pp. 93–122.
- [21] Jose L. Ramirez-GarciaLuna et al. “Is my wound infected?” In: ().
- [22] WCC DWC Bill PT Richlen. “Wound temperature can affect the wound healing process”. In: *Journal of Wound, Ostomy and Continence Nursing* 47.6 (2020), pp. 5204–5220.
- [23] Suzanne Koerner et al. “Use of Thermal Imaging to Identify Deep-Tissue Pressure Injury on Admission Reduces Clinical and Financial Burdens of Hospital-Acquired Pressure Injuries”. In: ().
- [24] Diane K Langemo and James G Spahn. “A reliability study using a long-wave infrared thermography device to identify relative tissue temperature variations of the body surface and underlying tissue”. In: *Advances in skin & Wound care* 30.3 (2017), pp. 109–119.
- [25] Manish Bharara et al. “Wound Inflammatory Index: A "Proof of Concept" Study to Assess Wound Healing Trajectory”. In: *Journal of diabetes science and technology* 4 (July 2010), pp. 773–9. DOI: [10.1177/193229681000400402](https://doi.org/10.1177/193229681000400402).
- [26] Khalad Maliyar, Reneeka Persaud-Jaimangal, and R. Gary Sibbald. “Associations Among Skin Surface pH, Temperature, and Bacterial Burden in Wounds”. In: *Advances in Skin & Wound Care* 33 (2020), pp. 180–185. URL: <https://api.semanticscholar.org/CorpusID:213192918>.
- [27] Carosena Meola. “Infrared thermography recent advances and future trends”. In: (2012).
- [28] Tomasz Sosnowski, Grzegorz Bieszczad, and Henryk Madura. “Image Processing in Thermal Cameras”. In: Jan. 2018, pp. 35–57. ISBN: 978-3-319-64673-2. DOI: [10.1007/978-3-319-64674-9_3](https://doi.org/10.1007/978-3-319-64674-9_3).
- [29] Michał Kręćchwost et al. “Chronic wounds multimodal image database”. In: *Computerized Medical Imaging and Graphics* 88 (2021), p. 101844. ISSN: 0895-6111. DOI: <https://doi.org/10.1016/j.compmedimag.2020.101844>. URL: <https://www.sciencedirect.com/science/article/pii/S0895611120301397>.
- [30] Mirza Cilimkovic. “Neural networks and back propagation algorithm”. In: *Institute of Technology Blanchardstown, Blanchardstown Road North Dublin* 15.1 (2015).
- [31] Teuvo Kohonen. “The self-organizing map”. In: *Proceedings of the IEEE* 78.9 (1990), pp. 1464–1480.
- [32] David J Montana, Lawrence Davis, et al. “Training feedforward neural networks using genetic algorithms.” In: *IJCAI*. Vol. 89. 1989. 1989, pp. 762–767.
- [33] Stephen Grossberg. “Recurrent neural networks”. In: *Scholarpedia* 8.2 (2013), p. 1888.

- [34] Kyunghyun Cho et al. “Learning phrase representations using RNN encoder-decoder for statistical machine translation”. In: *arXiv preprint arXiv:1406.1078* (2014).
- [35] KE ArunKumar et al. “Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends”. In: *Alexandria engineering journal* 61.10 (2022), pp. 7585–7603.
- [36] Dor Bank, Noam Koenigstein, and Raja Giryes. “Autoencoders”. In: *Machine learning for data science handbook: data mining and knowledge discovery handbook* (2023), pp. 353–374.
- [37] Olof Mogren. “C-RNN-GAN: Continuous recurrent neural networks with adversarial training”. In: *arXiv preprint arXiv:1611.09904* (2016).
- [38] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.
- [39] Huikai Wu et al. “Fastfcn: Rethinking dilated convolution in the backbone for semantic segmentation”. In: *arXiv preprint arXiv:1903.11816* (2019).
- [40] Liang-Chieh Chen et al. “Semantic image segmentation with deep convolutional nets and fully connected crfs”. In: *arXiv preprint arXiv:1412.7062* (2014).
- [41] Ayoub Benali Amjoud and Mustapha Amrouch. “Object detection using deep learning, CNNs and vision transformers: a review”. In: *IEEE Access* (2023).
- [42] Gianluca Zoppo et al. “AI technology for remote clinical assessment and monitoring”. In: *Journal of wound care* 29.12 (2020), pp. 692–706.
- [43] Jacopo Secco et al. “Clinically Validated Classification of Chronic Wounds Method with Memristor-Based Cellular Neural Network”. In: (2024).
- [44] Daniel YT Chino et al. “Segmenting skin ulcers and measuring the wound area using deep convolutional networks”. In: *Computer methods and programs in biomedicine* 191 (2020), p. 105376.
- [45] Dominik Spinczyk and Monika Wideł. “Surface area estimation for application of wound care”. In: *Injury* 48.3 (2017), pp. 653–658.
- [46] Hossein Nejati et al. “Fine-grained wound tissue analysis using deep neural network”. In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 1010–1014.
- [47] Jui-Tse Hsu et al. “Chronic wound assessment and infection detection method”. In: *BMC medical informatics and decision making* 19 (2019), pp. 1–20.
- [48] Maitreya Maity et al. “Pixel-based supervised tissue classification of chronic wound images with deep autoencoder”. In: *Advanced Computational and Communication Paradigms: Proceedings of International Conference on ICACCP 2017, Volume 2*. Springer, 2018, pp. 727–735.
- [49] Changhan Wang et al. “A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks”. In: *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2015, pp. 2415–2418.
- [50] Richard Ribón Fletcher et al. “The use of mobile thermal imaging and deep learning for prediction of surgical site infection”. In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021, pp. 5059–5062.

- [51] CJ Siah and C Childs. “A systematic review of the ASEPSIS scoring system used in non-cardiac-related surgery”. In: *Journal of wound care* 21.3 (2012), pp. 124–130.
- [52] Beau J Prey et al. “The use of mobile thermal imaging and machine learning technology for the detection of early surgical site infections”. In: *The American Journal of Surgery* (2023).
- [53] Ruyi Zhang et al. “A survey of wound image analysis using deep learning: Classification, detection, and segmentation”. In: *IEEE Access* 10 (2022), pp. 79502–79515.
- [54] Qinghui Liu, Arnt-Børre Salberg, and Robert Jenssen. “A comparison of deep learning architectures for semantic mapping of very high resolution images”. In: *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*. IEEE. 2018, pp. 6943–6946.
- [55] Glenn Jocher, Ayush Chaurasia, and Jing Qiu. *Ultralytics YOLOv8*. Version 8.0.0. 2023. URL: <https://github.com/ultralytics/ultralytics>.
- [56] Juan Terven, Diana-Margarita Córdova-Esparza, and Julio-Alejandro Romero-González. “A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas”. In: *Machine Learning and Knowledge Extraction* 5.4 (2023), pp. 1680–1716.
- [57] Bader Aldughayfiq et al. “Yolo-based deep learning model for pressure ulcer detection and classification”. In: *Healthcare*. Vol. 11. 9. MDPI. 2023, p. 1222.
- [58] Laith Alzubaidi et al. “Towards a better understanding of transfer learning for medical imaging: a case study”. In: *Applied Sciences* 10.13 (2020), p. 4523.
- [59] Gaetano Scebba et al. “Detect-and-segment: A deep learning approach to automate wound image segmentation”. In: *Informatics in Medicine Unlocked* 29 (2022), p. 100884.
- [60] Taifen Bao et al. “Back Propagation Optimization of Convolutional Neural Network Based on the left and the right hands Identification”. In: *2021 IEEE International Conference on Computer Science, Artificial Intelligence and Electronic Engineering (CSAIEE)*. IEEE. 2021, pp. 182–186.
- [61] Yaoshiang Ho and Samuel Wookey. “The real-world-weight cross-entropy loss function: Modeling the costs of mislabeling”. In: *IEEE access* 8 (2019), pp. 4806–4813.
- [62] Hengshuang Zhao et al. “Pyramid scene parsing network”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 2881–2890.
- [63] Kaiming He et al. “Deep residual learning for image recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.
- [64] MMSegmentation Contributors. *MMSegmentation: OpenMMLab Semantic Segmentation Toolbox and Benchmark*. <https://github.com/open-mmlab/mms Segmentation>. 2020.
- [65] MMPreTrain Contributors. *OpenMMLab’s Pre-training Toolbox and Benchmark*. <https://github.com/open-mmlab/mmpretrain>. 2023.
- [66] Zhaohui Zheng et al. “Enhancing geometric factors in model learning and inference for object detection and instance segmentation”. In: *IEEE transactions on cybernetics* 52.8 (2021), pp. 8574–8586.
- [67] Xiang Li et al. “Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection”. In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 21002–21012.
- [68] KM Veena et al. “FFA-Lens: Lesion detection tool for chronic ocular diseases in Fluorescein angiography images”. In: *SoftwareX* 26 (2024), p. 101646.

- [69] Thi-Loan Pham and Van-Hung Le. “Ovarian Tumors Detection and Classification from Ultrasound Images Based on YOLOv8”. In: *Journal of Advances in Information Technology* 15.2 (2024).
- [70] Kelly H Zou et al. “Statistical validation of image segmentation quality based on a spatial overlap index: scientific reports”. In: *Academic radiology* 11.2 (2004), pp. 178–189.
- [71] Chuanbo Wang et al. “Fully automatic wound segmentation with deep convolutional neural networks”. In: *Scientific reports* 10.1 (2020), p. 21897.
- [72] Intisar Rizwan I Haque and Jeremiah Neubert. “Deep learning approaches to biomedical image segmentation”. In: *Informatics in Medicine Unlocked* 18 (2020), p. 100297.