POLITECNICO DI TORINO

Master's Degree in Biomedical Engineering

FEM-NEURODESIGN

Neuromorphic Approach to the Biomechanical Design of Internal Fixation

Supervisor:

Prof. Jacopo Secco

Co-Supervisor:

Eng. Rosanna Cavazzana

Candidate:

Maria Vittoria Cotzia

Academic Year 2024/2025

Contents

Abstract

Femoral fractures are among the leading causes of disability and mortality worldwide, with a steadily increasing incidence due to aging populations and the growing prevalence of osteoporosis. In Italy alone, over 500,000 fragility fractures occur each year, many involving the proximal femur. Accurate and timely diagnosis is crucial to guide appropriate treatment and reduce complications and healthcare costs. However, traditional radiological assessment still relies heavily on the clinician's subjective judgment, leading to variability and potential diagnostic errors.

This thesis explores the use of artificial intelligence to support fracture diagnosis and treatment planning. A neural network was developed to analyze retrospective femoral X-ray images and automatically detect and classify proximal femoral fractures. In the first phase, a YOLO model was implemented to identify and classify fractures according to the AO classification system, focusing on types A2 (pertrochanteric) and B2 (transcervical). The model achieved a Precision of 80.8% and a Recall of 79.7%, demonstrating strong performance in detection and classification tasks.

In the second phase, key morphometric parameters (such as Pauwels angle, neck–shaft angle, lateral wall thickness, fracture length, and femoral neck diameter) were manually extracted and combined with clinical data (age, osteoporosis, fracture stability). These features were used to train several machine learning models—Random Forest, Support Vector Machine, and k-Nearest Neighbors—optimized through six-fold cross-validation. The SVM achieved the best performance, with a balanced accuracy of 72.8% in predicting the most suitable therapeutic treatment.

Finally, an explainability analysis using the SHAP (SHapley Additive Explanations) framework confirmed that the model's decision-making process aligned with clinical reasoning, as validated by an orthopedic specialist. The promising results of this work have led to the initiation of a retrospective clinical trial in collaboration with the Rizzoli Orthopedic Institute and Ospedale Maggiore (Bologna), laying the foundation for the future integration of this system as a decision-support tool for surgical planning and the standardization of orthopedic treatment protocols.

Chapter 1

Introduction

1.1 Clinical and Social Context of Bone Fractures

The present thesis will focus on the analysis of the phenomenon of fractures, with particular attention to fragility fractures and traumatic injuries. These types of bone injuries represent one of the main causes of disability in the population, with considerable effects on quality of life, motor function, and mortality risk, as well as generating significant social and healthcare costs [1, 2]. Fractures, in many countries, represent an event that is constantly increasing daily, affecting various age groups, mainly young and elderly, due to accidental falls, collisions, or other traumatic events. The causes are diverse and range from stress injuries to imperfect osteogenesis, up to osteoporosis and oncological diseases.
In 2017, fragility fractures in Europe amounted to 2.68 million, with Sweden representing the highest contribution, followed by Italy [3]. In detail, it appears that in Italy in 2017 there were about 560,000 fragility fractures, of which about 20%, or 112,000, involved the femur.[3] From a gender distribution in the over-50 age group, it emerges that women have a fragility fracture rate of 1:3 (34%), while in men this ratio is 1:5 (16%). Considering only femoral fractures, the risk is significantly higher in women: about 15% compared to 6–7% in men. Since these percentages are expected to increase, a total rise in fragility fractures of about 22.4% is predicted, with an estimated total of 690,000 fractures by 2030. The costs associated with femoral fractures will also remain the highest and most clinically relevant. In 2017, the total costs of fragility fractures in Italy were estimated at about 9.45 billion euros, of which 59% were attributable to femoral fractures. Costs are expected to increase by 26.2% by 2030. Finally, the average care burden is estimated at 882 hours of assistance per 1,000 people per year, due to the long-term consequences of fractures [3]. In particular, there has been an increase of almost 35.12% over 30 years, rising from 8.5 million cases in 1990 to over 11 million in 2021. The study under examination also considers other characteristics; in particular, there has been a record growth among the elderly over 75 years old, with an increase of 232.67%, due also to the fact that the elderly are the most vulnerable to falls, osteoporosis, and postural instability. Other age groups have also been affected, particularly the 50–74 age group, which has increased by 115.32% [4]. A further study [5] conducted in 19 countries and regions is based on 20 databases, covering a period from 2005 to 2018,

with a population over 50 years of age hospitalized for femoral fractures. It was found that there were about 4,115,046 total fractures, considering a standardization per 100,000 inhabitants, with the highest incidence in Denmark (315.9) and the lowest in Brazil (95.1). The study found that in many countries the incidence had decreased, but that, computed over time and particularly by 2050, it is expected to at least double. A fundamental characteristic is that men were treated less than women but showed higher mortality compared to the latter. In fact, a detailed analysis shows that the one-year mortality rate is 14.4% in Singapore, which is at the bottom of the list, compared to 28% in the United Kingdom, which is at the top. In fact, men show a much higher mortality rate, caused by underdiagnosis and undertreatment.

1.2 Artificial Intelligence for the Classification of Femoral Fractures

Despite technological progress, the diagnosis and management of fractures continue to rely largely on the clinical experience of the physician, who integrates radiological, tomographic, and anamnestic information to determine the most appropriate treatment. However, with the evolution of artificial intelligence (AI) and predictive models, automated diagnostic support is playing an increasingly important role, helping to improve the speed and accuracy of fracture identification, classification, and evaluation [6]. The use of such technologies allows for automated and timely analysis of medical images, significantly reducing diagnostic times and supporting faster clinical intervention, which is especially useful in emergency settings [7, 8].

Moreover, these systems help optimize the workload of radiologists, allowing them to focus on the most complex cases. AI is capable of detecting fractures that are difficult to identify and assess with the naked eye, ensuring greater diagnostic accuracy and reducing the risk of human error.[8] This support is particularly useful in emergency departments or healthcare facilities with a shortage of specialists, promoting more accessible and uniform analysis. The application of advanced tools for medical image analysis is particularly relevant in the context of femoral fractures. These traumatic events represent a clinical condition of great impact, with potential functional and systemic consequences. The interruption of the continuity of the femoral bone can cause intense pain, loss of the ability to walk, and, in the most severe cases, compromise local vascularization and innervation [9, 10, 11]. In this scenario, the use of automated systems for radiological analysis enables a more timely and accurate diagnosis, promoting personalized and effective therapeutic intervention. In this field, several scientific contributions have been proposed in the literature. For example, Yang et al. (2019)[12] present an approach based on Artificial Neural Networks (ANN) for the detection of fracture lines from X-ray images, concluding that this method allows automatic classification with an accuracy greater than 70%. Another work, by Lim et al. (2004)[13], adopts a different strategy based on the identification of features such as the neck-shaft angle, Gabor texture, and intensity gradient, analyzed using Support Vector Machine (SVM). A third contribution, by Hržić et al. (2019)[14], involves the use of local entropy for the segmentation and classification of pediatric fractures of the ulna and radius, followed by graph theory.

The system achieves accuracy up to 91.16% and a precision of 86.22%.
Finally, Wei et al. (2025)[15], using the YOLOv11 architecture, developed a model capable of classifying and localizing fractures in real time. Training on a large dataset allowed the achievement of a mean Average Precision (mAP) of 96.8%, demonstrating the effectiveness of the method for clinical applications.
The present study aims to develop a neural network for the analysis of retrospective X-ray (RX) images, with the goal of accurately detecting and characterizing femoral fractures, with a specific focus on those involving the proximal femur. The proposed system will be able to automatically extract and interpret fundamental features in the RX images, including patterns not recognizable by the human eye, to perform fracture detection and classification tasks.
The images used will represent fractures classified according to the Arbeitsgemeinschaft für Osteosynthesefragen (AO)[16] system, recognized as the reference standard for traumatic fractures in orthopedics, such as B2 and A2, corresponding to transcervical and pertrochanteric fractures.
A second step of system implementation will follow, during which the morphological characteristics of the bone and the fracture under analysis will be manually extracted. In particular, the quantitative parameters extracted will indicate the length of the fracture, the cervico-diaphyseal angle, the lateral femoral wall thickness (LWT), the femoral neck diameter, and the Pauwels angle. These parameters, associated with information on age group, presence or absence of osteoporosis, fracture stability, and finally the AO fracture class, were used to train three machine learning models, namely Random Forest, K-Nearest Neighbors, and Support Vector Machine, in order to evaluate the optimal model and suggest as output the most appropriate surgical therapy for the specific clinical case. For internal fixation, reference will be made to screws, plates, or intramedullary nails, while in more severe cases, which would lead to loss of mobility with the onset of necrosis due to compromise of the venous compartment of the femur, prosthetic implantation will be used. The orthopedic devices used were selected from the technical catalog of Intrauma S.p.A. company (Bruino). The devices considered were used to simulate realistic scenarios and ensure the biomechanical consistency of the model.
The choices for internal fixation were defined for each clinical case under analysis through direct consultation with a specialist surgeon from Istituto Ortopedico Rizzoli (Bologna), who established the decision-making criteria for selecting the most appropriate therapeutic intervention also based on the morphological characteristics of the fracture. This process made it possible to build a reliable ground truth, essential for training the developed machine learning models. The models were trained to replicate the physician's decision-making criteria, receiving as input the morphological information of the bone and the lesion, together with the clinical data of the patient. The system's output allows automation of the selection of the most suitable fixation device, providing advanced support for treatment planning and contributing to the standardization of therapeutic choices.

1.3 Structure of the Thesis

The thesis is divided into six main chapters, which analyze the various elements necessary to provide a clear and comprehensive overview of the topic addressed. The final chapters describe the work carried out from an operational perspective, the results obtained, and what followed, namely the launch of a clinical trial at Istituto Ortopedico Rizzoli (Bologna) and Ospedale Maggiore (Bologna).
The first chapter provides a general overview of bone, with particular attention to its formation, remodeling, and mechanical properties. The second chapter is dedicated to the structure of the femur, with a specific focus on the proximal portion, analyzing the main types of fractures and the fundamental orthopedic methods for their classification, such as the Pauwels method and the AO system. The following chapters provide an overview of the essential tools for the study, namely: bioimaging, in particular X-rays, internal fixators, and the main therapeutic solutions for this type of fracture. The sixth chapter focuses on the artificial intelligence technologies with specific attention to YOLOv8 and its evaluation metrics. The seventh chapter explores machine learning models, including data preparation, training, hyperparameter optimization, and explainability techniques. Finally, the last chapter will illustrate the procedures adopted and the results obtained, which will be analyzed and critically discussed to evaluate their effectiveness, limitations, and possible applications in the clinical setting. A specific section of the concluding chapter will be dedicated to future approaches.

Chapter 2

Structural Analysis of the Bone

2.1 Introduction: The Skeletal System

The skeletal system is composed of various components in different percentages, each of which performs a specific function, allowing perfect collaboration to support, protect, and ensure the proper functioning of the body. Bone makes up about 18% of body weight and constitutes one of the most complex structures of the human body. There are approximately 206 bone elements in total, which are divided into 5 categories according to their shape [17] [18]:

- **Long bones**: they have an almost cylindrical structure and a longitudinal dimension much greater than the others. They are found in the upper limbs (humerus, radius, and ulna) and lower limbs (femur, tibia, and fibula), but also in the hands and feet. They act as levers since they move in response to muscle contraction.

- **Short bones**: there is no predominant dimension compared to the others (carpal bones, tarsal bones). They provide stability and support while allowing a limited range of movement.

- **Flat bones**: characterized by a thin and flattened shape (bones of the skull, ribs). They serve as points of attachment for muscles and often protect internal organs[17].

- **Irregular bones**: they have a complex shape and cannot be classified in the other groups, such as the vertebrae which support the spinal cord and protect it from compression forces.

- **Sesamoid bones**: they develop within tendons, serving as structural reinforcement. Sesamoid bones protect the tendons, facilitating their passage over compression forces. Their number and position may vary from one individual to another. The patellae represent the only sesamoid bones present in all individuals.

From a chemical point of view, bone consists of an organic matrix (30–35%) and an inorganic component (65–70%) 2.1 [19].

Structure	Composition	Characteristics provided to the bone
Organic component	Collagen	Tensile strength and flexibility
Inorganic component	Hydroxyapatite (HA)	Rigidity and resistance to compression

Table 2.1: Synthetic structure of the bone.

Collagen is organized into triple helices, first forming microfibrils, and then fibers with a diameter of up to 20 µm. These are arranged in parallel lamellae, superimposed and oriented in different directions, creating concentric cylindrical structures called osteons (or Haversian systems). Each osteon has a central Haversian canal, which houses blood vessels and ensures the nutrition of the tissue.[20] In addition to these two components, bone has a cellular component made up of three main types of cells:

- **Osteocytes**: located in lacunae and embedded in the mineralized matrix, they regulate the balance between organic and mineral components.
- **Osteoblasts**: immature cells responsible for the production of collagen and hydroxyapatite, which later transform into osteocytes.
- **Osteoclasts**: cells responsible for the breakdown of the bone matrix, essential in remodeling processes.

The bone, as illustrated in Figure 2.1, is a "living" tissue, as it continuously remodels itself in response to mechanical and metabolic stimuli. This occurs thanks to the cooperation of osteoclasts (resorption of bone tissue) and osteoblasts (formation of new tissue). When this balance is altered, pathologies arise, for example, excessive osteoclastic activity causes osteoporosis, while excessively high osteoblastic activity leads to hypercalcification. Mainly, the combined action of these three elements gives the bone its properties of strength, elasticity, and ability to regenerate.

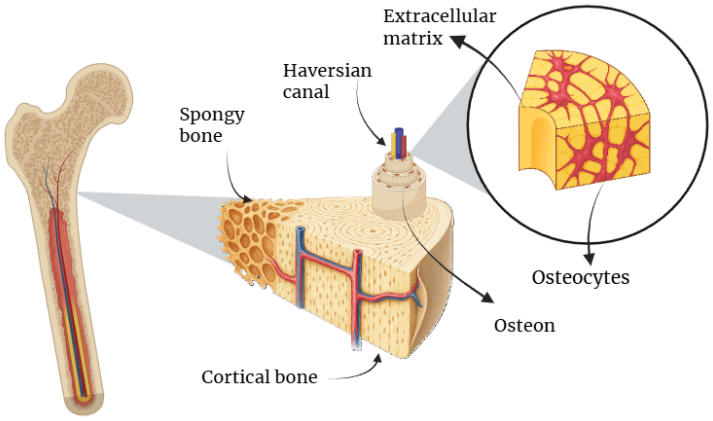

Figure 2.1: Histological structure of human bone: from the medullary cavity to compact bone with osteons. Highlighted are the Haversian canals, osteocytes, and concentric lamellae, fundamental for the nutrition and strength of the bone tissue.

On a macroscopic level, bone consists of two main types of tissue: compact bone (or cortical), in Figure 2.2, and spongy bone (or trabecular), in Figure 2.3.

Compact bone represents the densest and most resistant portion of bone tissue, with a density of about 2 kg/dm^3. It mainly performs support and protection functions and constitutes the central part of the diaphysis of long bones. The fundamental structural unit of this tissue is the osteon. Spongy bone, on the other hand, has a density between 0.15 and 1 kg/dm^3. It is mainly located in the epiphyses of long bones. Bone is called spongy precisely because of the three-dimensional network of trabeculae that gives it a porous appearance. The trabeculae are arranged to follow the lines of force to which the bone is subjected. This trabecular orientation develops during growth and can dynamically change in response to biomechanical variations, such as after a fracture or a change in joint load.

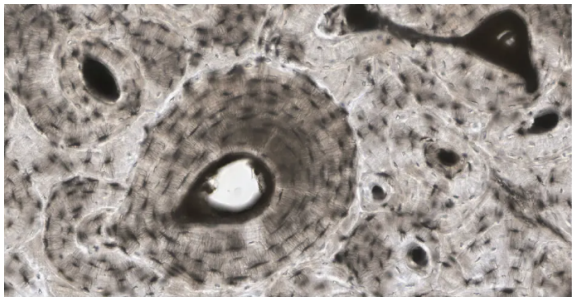

Figure 2.2: Transverse view of compact bone showing the basic structural unit, the osteon. In this micrograph of the osteon, the concentric lamellae and central canals can be clearly seen. LM × 40. (Micrograph provided by the Regents of the University of Michigan Medical School © 2012)

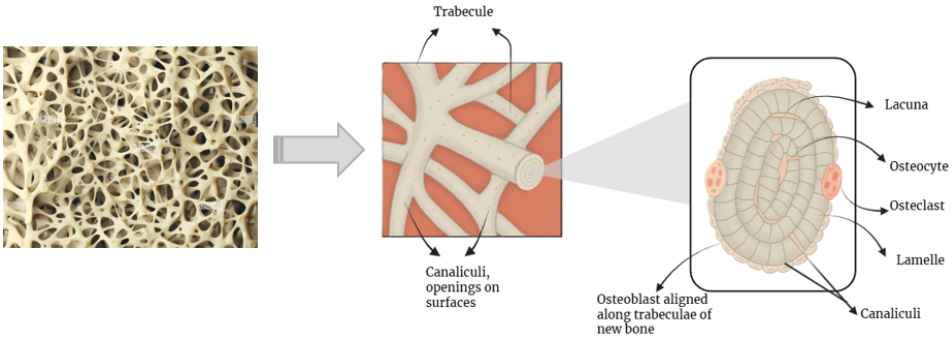

Figure 2.3: Three-dimensional structure of spongy bone tissue: on the left, the reticular appearance visible under the microscope; in the center, an enlargement of a trabecula with canaliculi; on the right, the internal section revealing osteocytes, osteoblasts, osteoclasts, and concentric lamellae.

2.1.1 Bone Remodeling

History of the Birth of Wolff's Theory

The concept of bone remodeling has developed over time thanks to several fundamental events. Jean-Baptiste Bourgery (1832) is considered the first observer of the

architecture of spongy bone. Later, the Swiss anatomist Mayer published in 1867 an article entitled "Die Architektur der Spongiosa," in which he linked the orientation of the trabeculae to the lines of mechanical force. The study presented a drawing of the structure of spongy bone (cancellous bone) observed at the proximal end of the human femur [21].

During the same period, structural engineer C. Culmann noticed the similarity between von Meyer's drawings and the stress trajectories in the curved beams he was designing. Culmann and von Meyer observed that the orientation of the trabeculae in the femur seemed to coincide with the main load lines in a femur subjected to transverse stress.

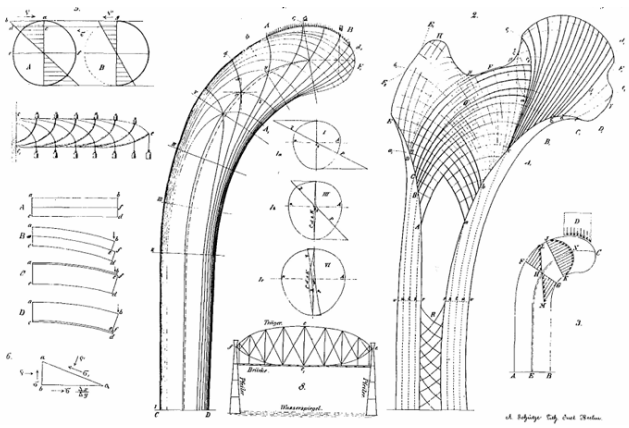

Figure 2.4: On the left is depicted Culmann's crane with the main stress trajectories indicated. On the right is shown von Meyer's sketch of the trabecular architecture in a section of the proximal end of the human femur. Both the femur and Culmann's crane are transversely loaded at their cantilevered ends, as illustrated in the small inset at the bottom right. Taken from Wolff (1872).

Starting from this observation, Julius Wolff in 1872 [22] stated that the directions of the trabeculae (particularly in the proximal femur) were due to the application of bending stresses. This intuition, influenced by the observations of von Meyer and Culmann, led to the formulation of his law:

"The internal architecture of bone adapts to functional loads through structural remodeling."

Wolff, however, never attempted to experimentally demonstrate this statement, which is nevertheless considered a cornerstone in bone biomechanics. Wolff's law is based on two principles:

- **general theory of bone transformation**: in which every variation in functional requirements corresponds to an architectural variation of bone tissue. Such correspondence can be described mathematically.

- **hypothesis of the trajectorial structure of trabecular bone**: where the distribution and direction of the bony trabeculae dynamically change according to the external load history and the average load history to which that volume of bone tissue is subjected.

According to this theory, there are three distinct processes through which a close correlation between mechanical stresses and structural organization of bone tissue manifests itself [20]:

- **Bone modeling**: process of bone formation and growth. The shape of bones adapts optimally to the mechanical functions for which they are intended. This shape is not entirely genetically determined but is also influenced by epigenetic factors related to the environment and functional stresses.
- **Bone healing**: process that occurs following fractures. The morphological and structural reorganization observable, for example, in the case of imperfect fracture healing, constitutes one of the main pieces of evidence supporting Wolff's theory.
- **Bone turnover**: under the influence of mechanical stresses, bone undergoes a constant process of renewal of the extracellular matrix. This renewal is necessary to maintain calcium homeostasis in the blood and the structural integrity of the skeleton. The process takes place through a resorption phase carried out by osteoclasts, followed by a phase of deposition and mineralization of the bone matrix by osteoblasts.

Many studies have attempted to predict bone remodeling in specific sites and under particular conditions. Some support the thesis proposed by Wolff, while others have questioned it.
Among those who confirmed its validity are Roux (1885) [23], with his analysis of the ankylosed knee, and Pauwels [24], who, using photoelastic models, verified the stress trajectories of the femur, confirming Roux's observations.
Subsequently, S.C. Cowin (1986) [21] proposed a mathematical formulation of the law, introducing the fabric tensor as a measure of trabecular architecture.
Among the dissenters, reference is made to Martin (1984) [25], who highlighted a relationship between porosity and specific surface area that does not easily fit into Wolff's model.
Others include Robling & Turner (2009) [26], who analyzed mechanical signaling at the cellular level, focusing on osteocytes and mechanosensors, highlighting a change in bone structure induced by complex cellular responses.

More recent studies, such as that of Razi et al. (2015) [27], have demonstrated dysfunction in mechanically induced bone remodeling with age, suggesting that not only the mechanical factor influences remodeling, but also the genetic one.
We can nonetheless affirm that the skeleton is a *living* structure, subject to numerous changes throughout life. Indeed, the characteristics of bone, in the same individual, vary over time and adapt to the individual's activity. Bone mass goes through different phases during life: there is a growth phase that continues until about the age of 20, during which bone progressively increases in volume and density. This is followed by a consolidation phase between 25 and 30 years of age, culminating in the achievement of the so-called *peak bone mass* [28].
The last phase is that of aging, particularly important and delicate due to demineralization. If the latter exceeds certain limits, osteopenia occurs, which in more severe cases can evolve into osteoporosis.

2.1.2 Mechanical Properties of Bone

Bone represents one of the first biological materials to have been studied in the field of bioengineering, for several reasons. Firstly, its structure and composition show analogies with materials already extensively analyzed in traditional engineering, making it a relatively simple model to interpret. Secondly, the orthopedic field has historically been one of the first sectors where it has been possible to intervene through the use of prostheses, thus promoting the development and application of specific biomaterials for the replacement and support of bone tissue.
A thorough understanding of the mechanical characteristics of bone is essential for multiple reasons. Indeed, it not only allows the assessment of possible variations in structural properties in the presence of pathological conditions, distinguishing between healthy and compromised bone tissue, but it is also essential for the design of substitute materials capable of replicating or supporting the biomechanical functions of natural bone. Finally, the study of mechanical properties plays a key role in analyzing the interaction between bone and other materials, both synthetic and biological, in order to ensure a functional and stable interface in cases of coupling with prostheses or adjacent tissues. As already mentioned, bone tissue is a dynamic structure, subject to continuous remodeling processes throughout an individual's life. However, its mechanical properties do not depend solely on age but are influenced by a multitude of factors. Among these, biological sex plays a significant role, as does the specific anatomical area considered. For instance, the cortical bone of the femur shows different mechanical characteristics compared to those of the ulna or radius. Another determining factor is the history of mechanical stress to which the bone has been subjected. In this sense, a young and physically active subject tends to develop stronger bone tissue adapted to functional loads compared to a peer with a sedentary lifestyle [20].
We can state that bone is an anisotropic material, meaning that its properties vary according to the direction of the load; in fact, the trabeculae of spongy bone do not grow randomly but follow precise trajectories determined by the history of stress [20, 19].
When characterizing bone, test specimens must be prepared on which mechanical tests are carried out. In fact, all mechanical properties reported in the literature should be interpreted carefully because they are data obtained from tests on dead bone [19]. To characterize bone, there are mainly two methods: strain gauges (which change resistance depending on whether they are stretched or shortened, thus changing the potential difference read) that can be applied to bone segments, and ultrasound. The latter are preferred as they can be used multiple times, whereas the former often lead to specimen fracture.

Elastic Modulus or Young's Modulus

The Young's modulus E, or modulus of elasticity, is a quantity that expresses the tendency of materials to elongate or shorten as a result of an applied load. Mathematically, it is expressed as shown in Eq. (2.1).

$$\sigma = E \cdot \varepsilon \tag{2.1}$$

Where σ is the stress and ε is the strain. The higher the modulus, the stiffer the material.

Type of bone	**Young's Modulus** E
Cortical	17 GPa to 20 GPa
Trabecular	0.1 GPa to 2 GPa (variable)

Table 2.2: Indicative values of the Young's modulus for the main types of bone tissue.

As shown by the E values in Tab. 2.2, cortical bone exhibits high stiffness and good tensile and compressive strength, while trabecular bone is lighter and more porous. Its elastic modulus is therefore much lower and strongly depends on density and load direction. This is due to its trabecular structure, which makes it less rigid but more capable of absorbing energy.

Transverse Elasticity Modulus or Poisson's Ratio

The Poisson's ratio, expressed in Eq. (2.2), represents the ratio between lateral strain and longitudinal strain in a material under load.

$$\nu = -\frac{\varepsilon_{\text{transverse}}}{\varepsilon_{\text{longitudinal}}} \tag{2.2}$$

Type of bone	**Poisson's coefficient** ν
Cortical bone	≈ 0.30
Trabecular bone	0.20 to 0.30 (variable)

Table 2.3: Indicative values of the Poisson's coefficient for the main types of bone tissue.

From the values shown in Tab. 2.3, cortical bone behaves transversely isotropically, while trabecular bone is less resistant and more deformable, with a structure that absorbs energy but does not withstand concentrated loads.

Stiffness

Stiffness is defined as the measure of a body's resistance to deformation under load, as expressed in Eq. (2.3). This characteristic depends on both geometry and material.

$$k = \frac{F}{\delta} \tag{2.3}$$

Where F is the applied force and δ the resulting deformation.
Typically, stiffness is measured through axial compression tests, ultrasound, or FEM modeling that simulates physiological or pathological conditions.
It has been observed that values vary with age, sex, and loading history.

Stress-Strain Relationship

The mechanical behavior of bone under load is described by the stress-strain curve, which represents the material's response to an external force [31]. Stress, defined as force per unit area, can be compressive, tensile, or shear. Strain occurs when a force modifies the length of an object, leading to its shortening or elongation. This curve is fundamental for understanding the bone's response to physiological, traumatic, or surgical loads and for designing compatible orthopedic devices.

$$\sigma = \frac{F}{A} \qquad \text{and} \qquad \varepsilon = \frac{\Delta L}{L_0} \tag{2.4}$$

Where:

- F is the applied force [N]
- A is the cross-sectional area [m^2]
- σ is the resulting stress [Pa]
- ΔL is the change in length [m]
- L_0 is the initial length [m]
- ε is the strain (dimensionless)

The curve presents several regions:

- **Elastic region**: The initial phase of the curve is linear and follows Hooke's law, which states that if deformation is proportional to the applied load, the material tends to return to its native shape. Bone returns to its original shape once the load is removed. The slope of this region represents the *Young's modulus* (E).
- **Yield point**: Marks the beginning of plastic deformation.
- **Plastic region**: A phase in which the material undergoes irreversible deformation. Beyond the yield point, the relationship is no longer linear. In bone, it can still bear load, but with structural damage compromising functionality.
- **Fracture point (ultimate stress)**
  The final phase of the graph occurs after reaching maximum tensile strength. It represents the maximum stress the bone tissue can withstand, beyond this point, a fracture occurs.

Cortical bone, being stiffer, withstands high stresses but with low strain ($\approx 2\%$); trabecular bone, being more ductile, resists up to 7% before fracture.

Fatigue Strength

Like other materials, bone exhibits fatigue behavior under cyclic loads repeated over time. This means that it has a trend called the "S-N curve", where S is stress and N the number of cycles. In fact, S can appear in two ways:

- Low stress → gives high durability
- High stress → leads to premature fracture

Seireg & Kempke (1969) [32] conducted a study on the right tibia of mice and discovered that the maximum stress value that bone can withstand for a given number of cycles without breaking is about 42% of the static breaking stress, a value between 15–20 Newtons (N), resulting in Eq. 2.5.

$$\sigma_{\text{fatigue}} \approx 0{,}42 \cdot \sigma_{\text{fracture}} \tag{2.5}$$

In general, bone shows good fatigue and compression resistance but moderate tensile strength, while ensuring high flexibility. Specifically, it has been observed that regarding fatigue resistance, bone can withstand a greater number of cycles before fracture when the applied load decreases. Fatigue behavior is also influenced by the internal bone architecture, particularly the presence of Haversian canals, which act as crack arresters, preventing crack propagation and thus improving fatigue resistance.

Anisotropy of Bone

Bone tissue, both cortical and trabecular, exhibits anisotropic behavior, meaning that its mechanical properties vary according to the loading direction.
In cortical bone, anisotropy is related to the orientation of osteons: studies such as that of Abdel-Wahab (2010) [33] show that the elastic modulus and mechanical strength vary significantly between longitudinal and transverse directions. Reilly and Burstein (1975) [34] also confirm that cortical bone is stronger along the longitudinal axis, with elastic modulus values around 17 GPa, and demonstrate that bone is stronger in compression and more deformable in tension, with a well-defined directional mechanical response. Finally, the study by Meyers et al. (2008) [36] further investigates the role of microstructure, showing how the arrangement of collagen fibrils and hydroxyapatite crystals influences stiffness and energy absorption capacity. These findings explain why fractures occur differently depending on the loading direction and the bone structure involved. Compact bone is stronger along the diaphysis, while trabecular bone, being less dense and stiff, absorbs energy better in compression but is more vulnerable in tension.

Viscoelastic Behavior of Bone

Being a composite material, bone exhibits viscoelastic characteristics [38].

This means that its response depends not only on the applied load but also on the duration over which this force acts. Therefore, bone deforms based on both the duration and the rate of load application, in addition to the load intensity [20].

Typically, when a viscoelastic material sample is subjected to various mechanical tests, the following results are obtained [38, 39]:

- Creep test: constant stress, and deformation increases over time.
- Stress relaxation: deformation is kept constant, and stress progressively decreases over time.

- Constant strain rate: the rate of deformation is constant, and the material's effective stiffness depends on it.
- Cyclic loading: the hysteresis phenomenon occurs, representing a dissipation of mechanical energy within the material.

Lakes, R. S. (1998) [40], proposed a modeling of viscoelastic behavior through equivalent mechanical systems, consisting of springs (for the elastic component) and dashpots (for the viscous component).

Influence of Load Application Rate

Since bone is viscoelastic, both fracture load and elastic modulus depend on the rate of application. The absorbed energy is proportional to the area under the load curve in the stress-strain diagram. It is observed that as the load rate increases, the absorbed energy also increases. One of the first studies to analyze the behavior under dynamic loads, i.e., at high speeds, was conducted by J.H. McElhaney [41]. The study showed that the tibia can absorb up to 45% more energy, at speeds comparable to those of a trauma, than during a slow load application.

Bone Behavior with Age

Regarding the correlation between bone and age, two main studies were considered. The first, by Burnstein (1976) [34], stated that the reduction in fracture strain constitutes the greatest change that occurs in bone over time. The experiment involved bone samples taken from subjects aged between 21 and 86 years, subsequently subjected to tension tests. It was found that fracture strains decreased by 5% per decade in the femur and 7% in the tibia. This resulted in a decrease in absorbed energy, respectively 32% and 42% before tensile fracture between the 3^{rd} and 9^{th} decades of life. The decrease in tensile strength over time was observed in both femoral and tibial bones, with a decrease of about 2% and 1% per decade, respectively [19].

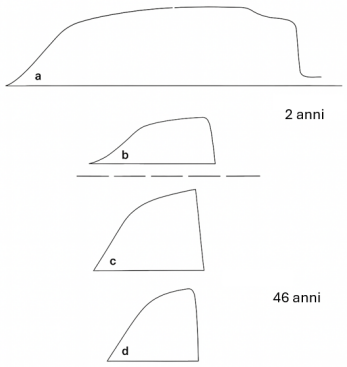

Figure 2.5: Load-displacement curves obtained with bone specimens loaded in bending; curves a) and b) show greater and lesser deformation obtained with bone specimens taken from a 2-year-old individual; curves c) and d) are comparable and were obtained with bone specimens taken from a 46-year-old individual.

Regarding the second experiment, by Lindahl *et al.* [42], they analyzed how age affects femur and humerus bones, from ages 15/20 up to 89 years. This experiment supports the previous one, and it could be stated that the skeletal structure would reduce so much as to be unable to support itself around 130–140 years of age, assuming no biological interventions occur. In both studies, no substantial differences in mechanical characteristics between males and females were found [19].

2.1.3 Osteoporosis

The term osteoporosis refers to a pathological condition characterized by reduced bone strength, leading to an increased risk of fractures [43]. It mainly manifests as a loss of bone mass in both cortical and trabecular bone. This is evident due to cortical thinning, where the number of trabeculae begins to decrease in size, leading to increased porosity [44].

Osteoporosis is divided into primary and secondary forms, the former being the most common and classified as type I and type II [45].

- **Type I osteoporosis**: typical of postmenopausal women, characterized by loss of trabecular bone. Possible causes include age, heredity, diet, alcohol consumption, and smoking.

- **Type II osteoporosis**: also called senile, as it mainly affects individuals over seventy of both sexes. It affects both cortical and trabecular bone and may be caused by reduced activity, low intake of calcium, magnesium, and vitamins.

Secondary osteoporosis occurs as a consequence of another pathology that leads to the onset of the osteoporotic condition.

This confirms that there is an intrinsic link between fracture and osteoporosis, due to the progressive thinning of bone and the loss of connections between the trabeculae. The loss of trabeculae leads to a reduction in the global mechanical strength of bone under stress [46].

At the femoral neck level, this is much more evident; for example, the fibers that describe Ward's triangle (the intersection of three trabecular bundles), the arcuate and cephalic fibers undergo fragmentation, creating a zone more susceptible to fracture [46].

Chapter 3

Femoral Anatomy and Related Fractures

3.1 Femur

After outlining the general framework, the analysis will now proceed with the anatomical structure of main interest in this paper, namely the femur. The femur, shown in Figure 3.1, is the longest, heaviest, and strongest bone in the human body. It is located in the lower limb and connects the hip to the knee. It is present in both lower limbs with a similar shape, although not perfectly symmetrical, and is classified as a long bone, composed of a diaphysis and two epiphyses [47]. The femur is a fundamental bone for the transmission of body weight and for gait stability during walking. It is stabilized by capsular and ligamentous structures that limit its movements and protect its integrity. These structures are part of the hip joint (coxofemoral joint), which is a ball-and-socket (synovial enarthrosis) joint, covered with cartilage to reduce friction. The round ligament (ligamentum teres femoris) connects the fovea capitis to the acetabulum but plays a secondary role compared to the capsular ligament [48, 49, 50]. Its morphological characteristics reflect functional adaptations to upright posture and walking [51]. At the proximal end, the pyramidal neck connects the spherical head to the apex and the cylindrical diaphysis. There are two well-defined bony prominences: the greater trochanter and the lesser trochanter, which serve as insertion points for the muscles involved in hip and knee movements. Other important anatomical elements include the adductor tubercle, the insertion point for the posterior part of the adductor magnus muscle, and the linea aspera. The hip joint is a ball-and-socket joint formed by the femoral head that fits into the acetabulum of the pelvis. The head is oriented medially, superiorly, and slightly anteriorly. The diaphysis shows a moderate anterior curvature. At the distal end, it widens into a conical shape up to a cuboidal base formed by the medial and lateral condyles. These condyles articulate with the tibia, forming the knee joint [47].

The synovial joints involved are the hip and knee, both of which are lined with cartilage to reduce friction and improve range of motion. The axes of the femur are measured by considering bony landmarks, such as the prominences [54, 55, 47].

The main axes are:

- **Mechanical axis**: ideally passes through the center of the ankle, knee, and

hip. It is inclined at approximately 3 degrees from the vertical.

- **Diaphyseal (or anatomical) axis**:this is the longitudinal line that passes through the center of the femoral diaphysis, following its natural curvature. It is inclined at approximately 6 degrees to the mechanical axis and 9 degrees to the vertical.

The main angles are:

- **Anteversion angle**: orientation of the head relative to the diaphysis, approximately 15°. It influences hip rotation: an alteration can cause pathological rotations of the femur.
- **inclination of the femoral neck**:the angle of inclination of the femoral neck relative to the diaphyseal axis is considered physiological if it is between 125° and 130°. It is typically around 128 degrees in the average adult, but tends to decrease with age [52, 53, 47]. Outside these values, two conditions can be distinguished:
  - Femoral valgus: increased neck-diaphysis angle (>135°), the neck is more vertical and the head moves away from the diaphysis.
  - Femoral varus: reduced neck-diaphysis angle (<120°), the neck is more horizontal and the head moves closer to the diaphysis.

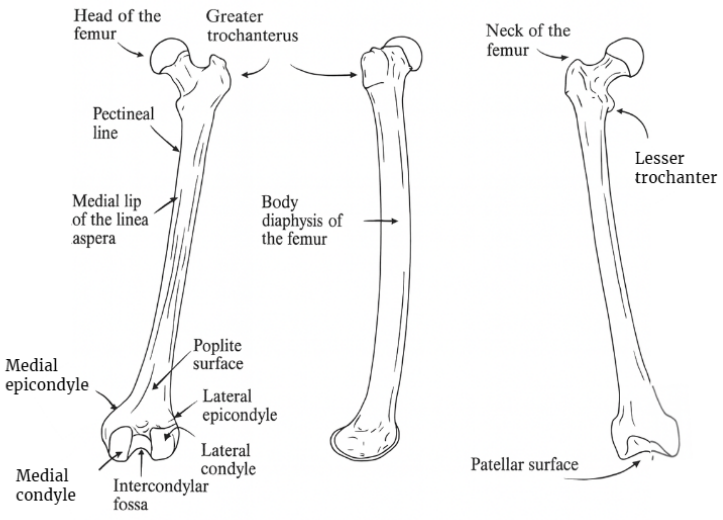

Figure 3.1: Anatomical views of the human femur: anterior, medial, and posterior. The illustration highlights key bony landmarks such as the femoral head, neck, trochanters, linea aspera, condyles, and epicondyles.

Anatomy of the Proximal Femur

The head of the femur, as shown in Figure3.2, articulates with the lunate surface of the acetabulum of the pelvis, forming a ball-and-socket synovial joint known as the

acetabular-femoral joint, more commonly called the hip joint.
The surface of the femoral head is smooth and covered with hyaline cartilage, except for a small medial area called the fovea capitis, a bony depression that serves as the insertion point for the ligament of the femoral head (ligamentum teres femoris). The Latin term fovea means "dimple" or "small cavity," accurately describing this anatomical structure. This ligament is a loose, non-tensioned structure that does not contribute significantly to joint stability but plays an important role in the vascularization of the femoral head, especially in young people. The neck of the femur is located distally from the head and represents its anatomical continuation. It connects the head to the body of the femur and is covered by the fibrous joint capsule of the hip, together with the supporting ligaments. At the base of the neck are two osseous prominences: the greater trochanter, laterally, and the lesser trochanter, medially. The latter is a conical eminence of variable size, which projects from the medial side of the proximal end of the femur and constitutes the distal insertion point of the iliopsoas muscle, which is fundamental in hip flexion.
The greater trochanter is a palpable bony prominence, approximately quadrangular in shape, located in the superolateral portion of the proximal end of the femur. It projects superiorly from the junction between the neck and the femoral diaphysis. This structure is an important site of muscle insertion, housing the tendons of: Gluteus maximus, Gluteus medius, Gluteus minimus, Piriformis, Obturator internus, Gemellus superior, and Gemellus inferior On the medial surface of the greater trochanter there is a deep depression known as the trochanteric fossa, into which the tendon of the obturator externus muscle inserts.
The third trochanter is a rounded bone prominence, variably present, located near the gluteal tuberosity on the proximal end of the femur. When present, it serves as an additional site of insertion for the ascending tendon of the gluteus maximus muscle.
The intertrochanteric line is a bony ridge located on the anterior surface of the femur, at the junction between the neck and the diaphysis. It extends from the apex of the greater trochanter to a small bony prominence located inferior to the lesser trochanter. This line represents an insertion point for the hip joint capsule, the ileofemoral ligament, and the vastus medialis and vastus lateralis muscles. The intertrochanteric crest, on the other hand, is a posterior osseous elevation that marks the transition between the neck and the body of the femur. It extends superolaterally from the greater trochanter to the lesser trochanter and is an important anatomical landmark for the posterior surface of the proximal femur.
On the posterior and medial side, the intertrochanteric crest houses the insertion of the hip joint capsule, contributing to the stability of the joint. Approximately halfway along the crest is a bone protuberance known as the quadrate tubercle (tuberculum quadratum), which is the insertion point of the quadrate muscle of the femur (quadratus femoris), involved in lateral rotation of the hip and posterior stabilization of the hip joint [56].

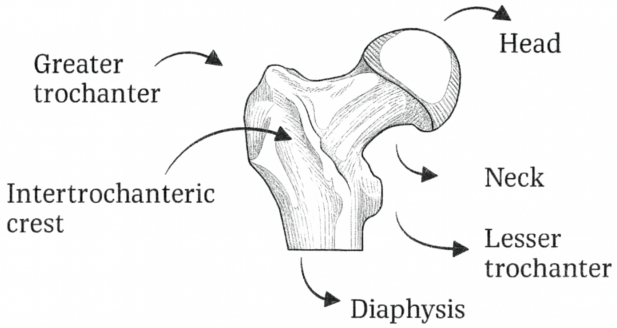

Figure 3.2: Anatomical posterior views of the human femur. The illustration highlights the landmarks such as the femoral head, neck, and trochanters.

3.2 Femur Fractures

When considering the proximal femur, it is essential to also include the arterial compartment in this region presented in Figure 3.3. In fact, the arteries that run through the proximal femur are responsible for supplying blood to most of the neck and head of the femur, and must be carefully evaluated in the event of a fracture of the proximal femur.

Injury to these vessels can seriously compromise the vascularization of the femoral head, with potentially serious clinical consequences such as avascular necrosis. For this reason, knowledge of the vascular anatomy of the femur is essential in the diagnosis, surgical planning, and prognosis of proximal fractures. Blood supply is divided into three distinct components, as described by Crock [55, 61]:

- Extracapsular arterial ring, originating anteriorly from the lateral circumflex femoral artery and posteriorly from the medial circumflex femoral artery.
- Ascending intracapsular cervical branches of the extracapsular ring, known as retinacular arteries.
- The artery of the round ligament (ligamentum teres).

The retinacular arteries form a subsynovial ring along the surface of the femoral neck [62]. The medial circumflex femoral artery provides the greatest blood flow to the femoral head, particularly the upper and lateral areas, which are typically subjected to the most stress during weight bearing [63, 62]. The lateral circumflex femoral artery supplies the anteroinferior portion of the femoral head via the inferior metaphyseal artery [63, 64].

The round ligament artery and the lateral and medial epiphyseal branches contribute less to blood flow.

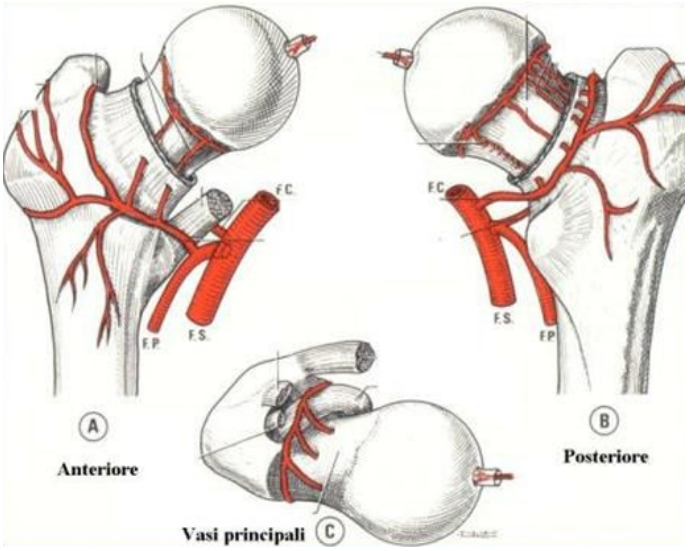

Figure 3.3: Vascularization of the femoral head: (A) anterior view, (B) posterior view, (C) summary of the main vessels.

In particular, intracapsular fractures expose the blood supply to a very high risk, with the possibility of delayed healing and necrosis [60, 61]. The neck is a very delicate area, as it is the most common point of penetration for the various lateral epiphyses. In fact, if this area is involved, the risk is very high but tends to decrease as one moves distally away from the femoral neck [65].
Mainly fractures, illustrated in Figure 3.4, involving the head, subcapital, and transcervical areas must be treated by opting for preservation, restoration, or prosthetic replacement because they are the ones that most compromise blood flow. Basicervical and intertrochanteric fractures, on the other hand, are treated with early immobilization or functional unloading during healing because they are mainly focused on stabilization with implants as they present a minimal risk of blood flow interruption. [61][66]

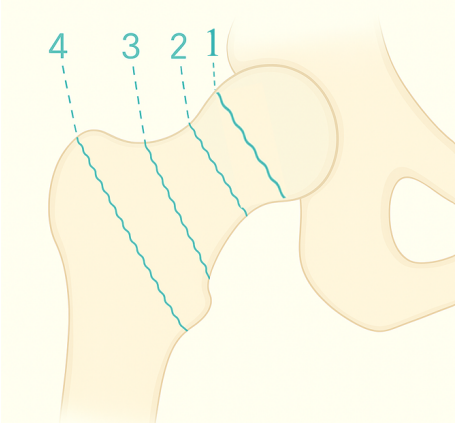

Figure 3.4: Visualization of proximal femur fractures. The fractures are called 1, 2, 3, and 4, respectively: subcapital, transcervical, basicervical, and transtrochanteric.

3.2.1 Main Causes of Femur Fractures

There are essentially two main causes of femur fractures: falls and/or osteoporosis. With regard to falls, there are two types of cases:

- high-energy trauma, such as a road accident or a fall from a height, typically seen in young people
- low-energy trauma, which is more common in the elderly and is caused by a simple fall.

A low-energy trauma fracture is known as a fragility fracture. It occurs in the presence of skeletal fragility due to qualitative and quantitative alterations in bone tissue, as in osteoporosis [57]. Femur fractures in the elderly are often the result of the combined action of a fall and osteoporosis, and are therefore classified as fragility fractures or low force fractures, i.e., fractures caused by minor trauma [58]. Low-energy trauma fractures are closely related to osteoporosis. The associated risk factors are linked to lifestyle, genetics, and clinical context, and are classified as modifiable and non-modifiable.

Modifiable risk factors include: smoking, alcohol, sedentary lifestyle, low body weight, poor diet, vitamin D and calcium deficiency, eating disorders (*e.g.* anorexia, bulimia), malabsorption, use of medications (*e.g.* glucocorticoids, antidepressants, anticonvulsants, antiandrogen therapy), and falls.

Non-modifiable risk factors include: advanced age, female gender, Caucasian ethnicity, previous fractures, family history of fractures, and menopause.

Although non-modifiable factors cannot be corrected, knowledge of them is essential for implementing effective preventive strategies. Like many age-related conditions, fragility fractures result from multiple causes and factors [59].

3.3 Classification of Femoral Fractures

3.3.1 AO Classification System

The classification takes its name from the foundation that created it, namely the Association for the Study of Internal Fixation (AO Foundation). It was founded in Switzerland in 1958 by 13 surgeons and orthopedists, with the aim of developing techniques and instruments for the treatment of bone fractures.

The AO system is now an international standard for fracture classification, providing a structured and shared basis for diagnosis and treatment planning.

The AO classification consists of a 4-digit alphanumeric code. This unique code describes the exact location and severity of the fracture. It is mainly used in the context of fractures of long tubular bones, such as the femur, humerus, tibia, and radius, but it is also applicable to other bones in the body [67]. The AO classification method is based on four fundamental elements: identification of the bone segment, location of the fracture, severity scale, and an additional parameter to further refine the classification. Each skeletal segment is represented by a number that uniquely identifies it, as reported in Figure 3.5.

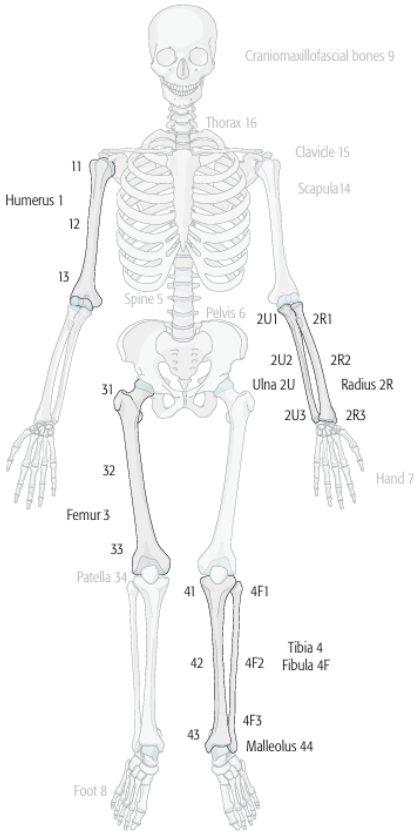

Figure 3.5: Assignment of the AO classification to the various body segments. The image provides a useful anatomical reference for locating fractures and applying AO classifications.

After determining the affected bone segment, proceed with locating the fracture. This applies to long bones, according to the following numerical code (Figure 3.6):

1. Proximal (proximal epiphysis, the part closest to the center of the body)
2. Intermediate (diaphysis)
3. Distal (distal epiphysis, the part furthest from the center of the body)

Figure 3.6: The figure shows three cases of femur fractures. The first case shows a type 31 fracture, involving the proximal femur. The second case present a type 32 fracture, involving the intermediate femur. Finally, the third case shows a type 33 fracture as it involves the distal femur.

Once the segment and specific portion containing the fracture have been identified, the fracture type is graded using a severity scale from A to C, where A indicates the least severe pattern, B an intermediate level of severity, and C the most severe form. An additional three-point parameter is added for further precision. For clarity, the subdivisions by location of the injury and severity are shown in the following paragraph.

Proximal segment of the femur (Code 31)

31A. **Trochanteric region (extra-articular fractures):**

- 31A1: Simple pertrochanteric fracture.
- 31A2: Multifragmentary pertrochanteric fracture with incompetent lateral wall (thickness ≤ 20.5 mm).
- 31A3: Intertrochanteric fracture with reverse obliquity.

31B. **Femoral neck (extra-articular fractures):**

- 31B1: Subcapital fracture.
- 31B2: Transcervical fracture.
- 31B3: Basicervical fracture.

31C. **Femoral head (articular fractures):**

- 31C1: Splinter fracture.
- 31C2: Depression fracture.

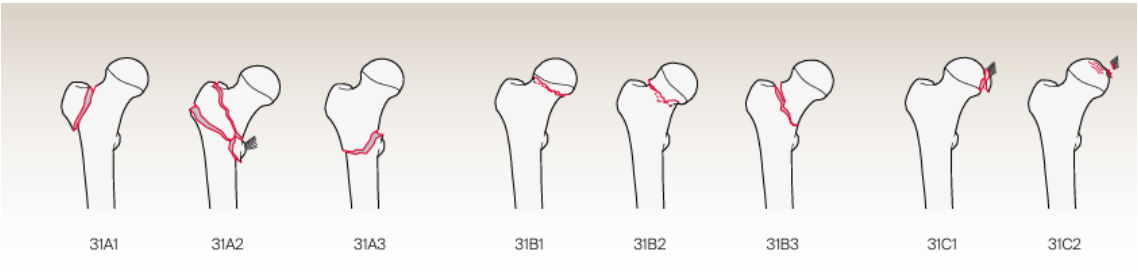

Figure 3.7: Classification A0 of proximal femur fracture, all cases with code 31 are highlighted.

Diaphyseal segment of the femur (Code 32)

32A. **Simple fractures:**

- 32A1: Spiral fracture.
- 32A2: Oblique fracture (angle $\geq 30°$).
- 32A3: Transverse fracture (angle $< 30°$).

32B. **Wedge fractures:**

32B2: Fracture with intact wedge.

32B3: Fracture with fragmented wedge.

32C. **Multifragmentary fractures:**

32C2: Segmental fracture with intact segment.

32C3: Segmental fracture with fragmented segment.

The AO classification for diaphyseal fractures of the femur (code 32) includes additional qualifications to describe the location and morphology of the fracture more precisely.

For fractures 32A and 32B, the qualifications are:

a: Proximal third of the diaphysis.

b: Middle third of the diaphysis.

c: Distal third of the diaphysis.

These letters indicate the longitudinal position of the fracture along the femoral diaphysis.

For 32C fractures, the qualifications are:

d: Fracture involving the junction between the diaphysis and the proximal metaphysis.

e: Fracture confined to the diaphysis alone, without extension to the metaphysis.

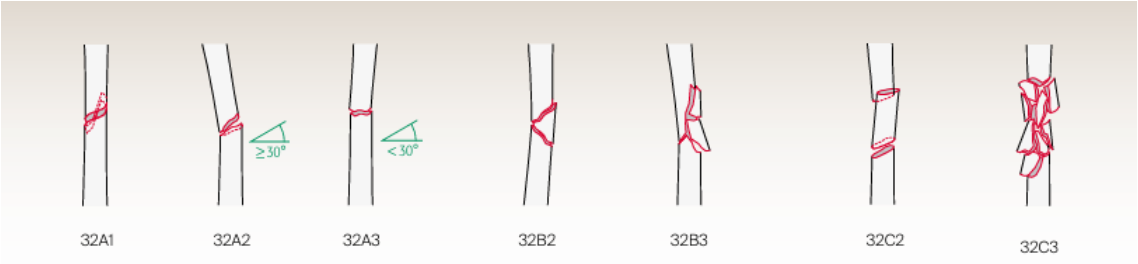

Figure 3.8: Classification A0 of proximal femur fracture, all cases with code 32 are highlighted.

Distal segment of the femur (Code 33)

33A. **Extra-articular fractures**:

33A1: Avulsion fracture.

33A2: Simple fracture.

33A3: Wedge or comminuted fracture.

33B. **Partially articular fractures**:

33B1: Fracture of the lateral condyle in the sagittal plane.

33B2: Fracture of the medial condyle in the sagittal plane.

33B3: Fracture in the frontal/coronal plane.

33C. **Completely articular fractures**:

33C1: Simple articular fracture with simple metaphysis.

33C2: Simple articular fracture with wedge or multifragmentary metaphysis.

33C3: Multifragmentary articular fracture with simple, wedge or multifragmentary metaphysis.

Figure 3.9: Classification A0 of proximal femur fracture, all cases with code 33 are highlighted.

3.3.2 Pauwels Classification

Pauwels' classification was first published in 1935 [68, 69] in a German publication.

We can say that Pauwels' angle is defined as the angle formed between the fracture line of the femoral neck and a horizontal line drawn parallel to the pelvic plane (usually at the level of the iliac crests).

Pauwels highlighted the fundamental role of the angle of the femoral neck fracture, emphasizing that the more vertical the fracture line, the greater the risk of non-union or delayed healing. Therefore, according to Pauwels, a high-angle fracture in the transcervical area creates a mechanical problem that often cannot be resolved through screw fixation.

In his work, Pauwels defined his classification based on the forces acting on the fracture. The values that Pauwels indicated as maximum for each degree are 30, 50, and 70 degrees. Therefore, neck fractures are divided into three classes [70], reported in Figure 3.10:

- **Type I**: compression forces are present; in other words, internal fixation is not necessary. This corresponds to a fracture with an angle of $<30°$ to the horizontal.

- **Type II**: there is a shear force that can negatively affect bone healing. Internal fixation is therefore recommended to eliminate this shear force. It has an angle between 30° and 50°.

- **Type III**: Shear force dominates and is associated with a varus force, known as the inclination moment. This causes displacement of the fracture and collapse into varus. Typically, a corrective osteotomy in valgus is used. It has an angle >50° to the horizontal.

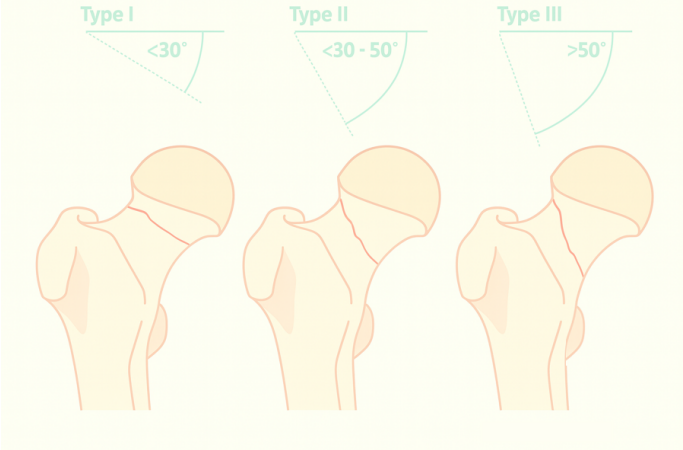

Figure 3.10: The Pauwel's classification of femoral neck fractures.

3.4 Fracture Stability

An important parameter to consider is the stability of the fracture; this, together with other characteristics, allows for the identification of the most appropriate surgical solution. The fracture can primarily be classified as either stable or unstable, as shown in Figure 3.11 [71].

- **Stable**: is characterized by having two completely intact surfaces; typically, stability is based on reduction and preservation of the medial cortex. The main types are A1 and A2.1 according to the AO classification.

- **Unstable**: is characterized by comminution of the postero-medial cortex, reverse patterns, and those with subtrochanteric extension. These include types A2.2, A2.3, and A3 of the AO classification.

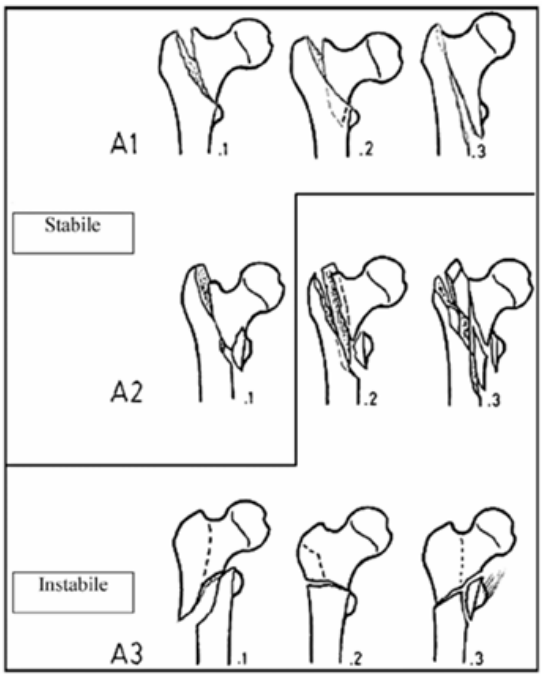

Figure 3.11: AO classification of lateral fractures of the upper end of the femur distinguishes between two types of fracture: stable and unstable. The figure was taken from Cavenago, C., & Prestianni, C. (2010)

Chapter 4

Bioimages

The images discussed in the following chapter are bioimages, this term mainly refers to figures that represent the anatomy or physiognomy of internal parts of the body. The images can be divided into categories based on three different classification criteria, according to the acquisition technique and the intended use, as shown in Tab. 4.1.

Criterion	Categories
Function	• **Morphological** techniques: provide information on anatomical structure (e.g., CT, anatomical MRI, ultrasound). • **Functional** techniques: show tissue activity or metabolism (e.g., PET, fMRI, SPECT).
Clinical use	• Analysis of anatomical structures • Diagnostics • Pathology analysis • Monitoring or follow-up
Characteristics	• Use of **ionizing radiation** (e.g., CT, PET) or **non-ionizing** (e.g., MRI, ultrasound) • **Invasiveness** of the technique • Whether or not **contrast agent administration** is required

Table 4.1: Classification of medical imaging techniques.

Images are mainly generated by exploiting the interaction of a certain type of

energy with the human body through the physical mechanisms listed below:

- **Absorption**: this physical phenomenon is based on the fact that the form of energy is absorbed differently depending on the density and composition of the tissue under analysis. This principle is typical of X-ray and CT images.
- **Reflection**: a high-frequency wave generated by a transducer is used, after encountering a structure (bones, organs, fluids), it is partially reflected and returns to the receiver or probe. Typical images that exploit this physical phenomenon are ultrasound scans, commonly known as sonograms.
- **Diffusion**: a radiopharmaceutical is administered to the patient, which emits radioactive particles in a band that diffuse in all directions. By capturing these outgoing particles, it is possible to reconstruct the bioimage. This is used in nuclear medicine, for example in SPECT and PET techniques.

4.1 X-ray Images

X-rays are a form of energy with a very short wavelength, ranging approximately between 10 nm and 10 nm. In fact, they are invisible to the human eye, but they can interact effectively with human tissue. They belong to the category of ionizing radiation, where radiation refers to the phenomenon of energy transfer, and ionizing means that, once it hits a given neutral medium, it ionizes it, i.e., it creates ions. Consequently, if a measurement is taken, a certain amount of charge can be detected within the tissue. In particular, X-rays can penetrate certain materials such as water and muscles, while being blocked by bones. It is precisely because of the different ways in which tissues absorb rays that images are constructed, and it is this characteristic that is exploited. Typically, the human body is able to repair the damage caused by these ionization phenomena. However, if these phenomena exceed the body's repair capabilities, they can become a health problem, even giving rise to long-term issues. In specific cases, i.e., when exposed to high doses of ionizing radiation or low doses for very long periods of time, ionization can affect the cell nucleus, where the chromosomes are located. At this point, the cell can begin to replicate in an altered manner, leading to the formation of tumors [72].

Despite the potential risks, ionizing radiation is widely used in medicine according to the principle of minimum necessary dose (As Low As Reasonably Achievable, ALARA).

X-ray Generation

The physical principle underlying the generation of X-rays in modern radiological devices has remained almost unchanged from Roentgen's original model. The source used is the so-called X-ray tube, a cylindrical vacuum glass structure that houses two fundamental components: one responsible for generating charges (cathode) and the other serving as the target (anode), with a potential difference applied between these two parts [73]. Typically, the radiogenic tube in Figure 4.1 is contained within a metal casing (external) internally lined with a lead sheath to shield the stray radiation

emitted by the tube during X-ray production. The inner part of the casing contains a liquid called mineral oil, which serves to cool the glass envelope and electrically insulate the anode and cathode. The envelope itself ensures a high vacuum precisely to prevent residual gases from altering the velocity of the electrons and becoming sources of secondary electrons.
The operation is as follows: The cathode, consisting of a tungsten filament, is heated by a current generated by a generator until it reaches a temperature that allows the emission of electrons. In this way, the electrons can be released and exit the metal to propagate. Their direction is random, so all directions in space are equally probable; therefore, the filament is said to be isotropic. Given the isotropic nature of the filament, a support is needed to focus the electrons onto the anode. To do this, a focusing cup is used, which serves to absorb the electrons that do not go in the direction of the target.
By increasing the intensity of the current flowing through the cathode, the number of emitted electrons can be increased and, consequently, the flow of X-ray photons produced as a result of collisions in the anode [74]. It has already been mentioned that there is a potential difference between the source and the target, on the order of tens or hundreds of kV; this serves to give the correct direction to the outgoing electrons. Indeed, the potential difference is continuous because if it were alternating, the electrons, being negative charges, would change direction with each cycle.
By increasing the potential difference between the cathode and the anode, the kinetic energy of the individual electron will increase and, consequently, higher-energy and more penetrating X-ray photons will be produced [74].
The anode, also made of tungsten or other high atomic number metals, serves as the target for the electrons. In more advanced models, the anode is mounted on a rotating axis (rotating anode tube), so as to distribute the heat generated by the impact of the electrons over a larger surface, avoiding localized damage.
The target against which the electrons strike is called the focal spot, which allows the size of the X-ray beam to be modulated. If we consider an anode with a very large focal spot, the geometric dimensions of the beam are larger.

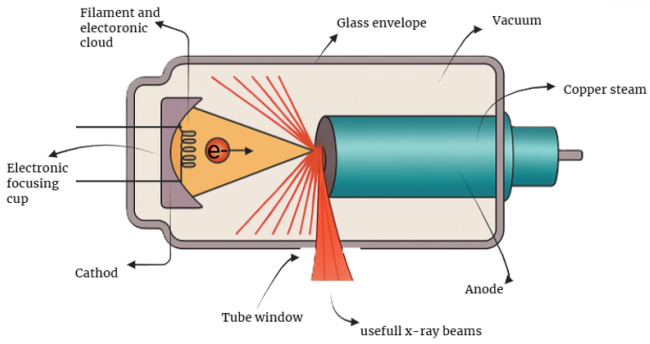

Figure 4.1: Graphic representation of a cross-section of an X-ray tube. On the left is the cathode with a heated filament and focusing cup; on the right is the inclined anode. The X-ray beam, generated by the interaction between electrons and the anode material, emerges in a collimated form through the tube window.

When the electron beam strikes the target, two types of interactions occur, one

of which is dominant over the other.

- **Braking radiation** : in this case, the electron impacts with a certain speed but is deflected, changing direction and losing part of its energy. The energy variation is transformed into an X-ray photon. This phenomenon is also known as braking.
- **Characteristic radiation**: The incident electron collides with another electron in the atom of the anionic material, which is deflected and forms a gap. From here, an electron from the more external orbitals fills the void left by the electron, performing what is called a quantum leap between orbitals. This leap allows energy to be released in the form of an X-ray photon.

Material Interaction

There are mainly three types of interaction mechanisms between X-rays and matter:

- **Scattering:** it is not used for medical imaging purposes, as it has low energy (about 10 keV) and a random emission direction.
- **Photoelectric effect:** occurs when an electron strikes the innermost electron of the material. This phenomenon allows for very strong ionization, through which all the energy is converted and no secondary emission occurs.
- **Compton effect:** occurs when the photon strikes an electron in the outer orbital and displaces it. In this case, there is secondary emission since the emitted photon has lower energy than the incident one. This mechanism is the main cause of image noise, as the random radiation settles on the image like a veil.

There are two other mechanisms that occur at very high energies: pair production (1.02 MeV) and the photodisintegration effect (10 MeV). Pair production occurs when the photon interacts with the force field of the nucleus and disappears, simultaneously generating two particles: an electron and a positron. The electron produced causes ionization, while the positron undergoes annihilation [75], resulting in the production of two gamma rays of 0.511 MeV emitted in diametrically opposite directions [76]. As for photodisintegration, the incident photon is absorbed by the nucleus, which, once excited, expels a nucleon or another nuclear fragment. We can state that among all these interaction mechanisms, the ones useful for imaging purposes are only the Compton effect—although less useful for contrast—and the photoelectric effect, as it is the mechanism that creates contrast between soft tissues and dense structures, and is therefore essential for obtaining sharp images. To understand the formation of the X-ray image, it is necessary to introduce a fundamental concept: the Lambert-Beer law (Eq. (4.1)). This principle explains how many photons have been absorbed by the material compared to those emitted. In fact, once X-rays enter the material, they propagate along a certain path . However, the material has its own linear attenuation coefficient x, which provides information on how much that material absorbs.
From this, it can be easily inferred that the photons emerging from the medium will certainly be fewer than those emitted, precisely because of this coefficient. It must

be considered that the coefficient μ does not depend exclusively on the material: in fact, the coefficients vary according to the different energies of the incident photon. Therefore, the material absorbs differently depending on the incoming energy. For the law to be applicable, we must assume that all photons have the same energy. This law allows us to distinguish different tissues irradiated by the same beam, as they have different μ values and therefore absorb differently from one another. This information is captured by the sensors, which translate the number of photons into grayscale levels, creating what becomes the image. The Lambert-Beer law is expressed as follows:

$$N(x) = N_0 \cdot e^{-\mu x} \tag{4.1}$$

Where N is the number of emerging photons and N_0 the number of incident photons. As previously mentioned, the value of μ depends on the incident energy of the photons: the more energetic the photon, the less the material absorbs. The typical trend between the value of μ and the energy E is almost a decreasing exponential. The unit of measurement of the coefficient is μ = [cm-1].

Radiological Device

Considering now the actual device, shown in Figure 4.2, it mainly consists of an X-ray tube, which is typically attached to the ceiling. It is also equipped with handles so that the structure can be moved to correctly irradiate the body. Since X-rays are not visible to the human eye, the device incorporates a light bulb that emits a beam of light with the same geometry as the beam emitted by the instrument.
Finally, there is a bed, also called a patient support, on which the digital sensors are located.

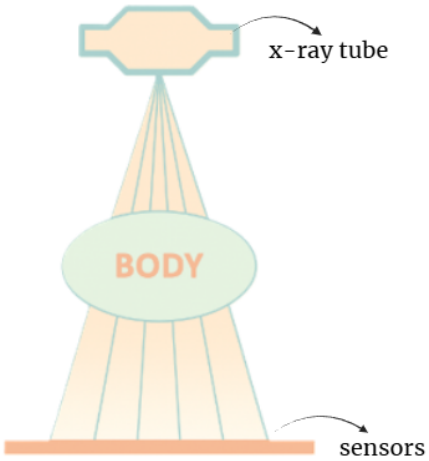

Figure 4.2: Simplified diagram of radiographic acquisition: the X-ray tube is visible at the top, from which a beam of X-rays represented by diverging lines originates. The beam passes through the patient's body and continues downward, where it is detected by the acquisition plane.

The sensors used today are digital sensors, but there are also analog sensors still in use, albeit very rarely, called brightness amplifiers [72].

There are two main types of digital sensors: Computed Radiography (CR) systems and Digital Radiography (DR) systems. Both provide a digital image, but they have very different operating principles.

- **C**R: it is a type of indirect conversion, as panels made of photostimulable phosphors are used. In fact, when the photon enters the phosphor, visible photons are generated, which, however, remain confined within the thickness of the panel. The generated image is latent and, due to its nature, unusable, therefore, this panel is placed on another reading device that uses a red/infrared laser (630–680 nm) to read the image and return the digital image. To reuse the panel, the phosphors are treated with an appropriate voltage so that the latent image is erased and the device can be ready for the next use. For normal use, a lifespan of around 10 years is usually provided, although this excludes mechanical issues [72, 78].

- **D**R: allows digital data output without the need for intermediate procedures (they work in real-time). Like CR sensors, DRs are also panels, but they are composed differently. These feature a silicon semiconductor technology, where particular impurities of sodium iodide are present inside, capable of interacting well with X-rays. Once the ray hits the structure, electric charges are generated and accumulated within capacitors to generate the image. The greater the charge created at a point, the greater the number of photons that have hit it. This capacity is provided by metallizations placed above and below the silicon panel; in particular, the upper part must be thin so as not to attenuate the X-rays, while the lower part is composed of many metallizations [72, 78].

Chapter 5

Internal Fixators

Internal fixators are implantable devices used to promote bone healing following a fracture. Unlike external fixators, these devices are placed directly in contact with the bone and remain in place throughout the entire healing process. They comprise a broad family, within which plates, intramedullary nails, screws, and Kirschner wires can be found.
In the present study, plates, nails, screws, and prostheses were considered. The latter play a major role in all cases where it is not possible to intervene on the fracture in any way, or when large-caliber arteries have been damaged. Prostheses, for example, are highly recommended in elderly patients, in whom it is necessary to accelerate postoperative recovery times, in addition to the fact that in many cases they would not be able to undergo a revision surgery. It should be noted that the solutions analyzed are taken from the catalog of the company Intrauma.

5.1 Plates

These devices, defined by the ASTM F382–24 standard as elements equipped with at least two holes or slots and a two-dimensional cross-section (width and thickness), are designed to ensure fracture healing through the alignment and stabilization of separated bone segments. Plates are fixed to the bone using cortical, locking, or polyaxial screws, or through cerclage wires. The geometry of the plate may vary depending on the anatomical site: proximal plates for the neck and femoral trochanters, diaphyseal plates for the central portion of the femur, and distal plates for the femoral condyles. Plates may feature uniform width or tapered sections, with greater thickness at the center and less at the edges, to optimize stress distribution. Threaded holes and preassembled bushings allow modular screw insertion, with the possibility of angular orientation in the case of polyaxial screws, which are particularly useful in multifragmentary fractures. The ASTM F382–24 standard also establishes test methods for evaluating the mechanical properties of plates, including bending stiffness, bending strength, and cyclic fatigue, providing a technical reference for the design, labeling, and clinical management of these devices.

As mentioned, plates are classified according to the anatomical region to which they refer. Naturally, they have different shapes, as they must follow the structure of the bone to best perform their function. This thesis will mainly discuss the plates

used for the proximal femur.

PFF Rail Proximal Femur Plate

The plates for the proximal femur reported in Figure 5.1 are made of Ti6Al4V, in accordance with ISO 5832-3. Ti-6Al-4V is a titanium alloy composed mainly of titanium, to which approximately 6% aluminum is added to improve mechanical strength and 4% vanadium to increase toughness and fatigue resistance. This alloy is regulated by the ASTM F136 standard, which specifies its properties for implantable use [80]. The sizes available for this type are short (169 mm), medium (219 mm), long (269 mm), and extra long (319 mm).

Figure 5.1: Medium proximal femoral plate, front, side, and rear views.

5.2 Endomedullary Nail

Intramedullary nails, on the other hand, comply with ASTM F1264. These devices are implanted in the medullary canal and then fixed in place using proximal and distal screws to ensure proper stability. Their cylindrical or slightly conical shape is ideal for adapting to the profile of the medullary canal. Mechanically, the nail allows stress to be transferred from the bone to the metal while promoting the bone remodeling process.

ELOS - Trochanteric Nail

The nail and the long nail both comply with the same ASTM standard, ISO 5832-3, and are therefore made from the same Ti6Al4V material. Essentially, the long nail is an extension of the first and, for this reason, they will be presented together. The trochanteric nail is available in various lengths: 165, 180, and 240 mm for the normal version, and 300, 340, 360, 380, and 400 mm for the long version. These different lengths are the result of anatomical structures that vary from individual to individual, while the choice between a short and long nail is closely related to the stability of the fracture (an aspect that will be discussed in more detail later). It should also be borne in mind that the nail is not only composed of the vertical structure, but also includes the cephalic screw. Indeed, the angle formed between them,that is, the longitudinal axis of the nail and that of the cephalic screw, is of particular importance. It can be 122° or 127°, in order to respect the natural anatomical conformation as much as possible.

Figure 5.2: ELOS - Trochanteric nail 240 mm diameter 10 mm - 127° with cephalic screw.

5.3 Prostheses

Prostheses are implantable devices designed to replace an anatomical segment. Like other devices, they must comply with specific international standards, including ISO 7206, ISO 5832, and ASTM F75, in order to fulfill their purpose, which is to restore the functionality of the replaced limb, relieve pain, and improve the patient's quality of life. Generally, a prosthesis consists of a femoral stem, which is inserted into the medullary canal and made of Ti6Al4V alloy to ensure strength and biocompatibility. The ball that replaces the femoral head can be made of ceramic, metal, or polyethylene. If the hip joint also needs to be replaced, the prosthesis includes an acetabular cup, which fits into the acetabulum and can also be made of polyethylene, metal, or ceramic. The system is completed by the insert, which acts as the interface between the femoral head and the cup, contributing to the stability and fluidity of joint movement.

Stem P - Femoral Stem

Forged in biocompatible Ti6Al4V titanium alloy in accordance with ISO 5832/3, it is coated with a layer of CaP that promotes osseointegration (Figure 5.3). It features a straight profile with a rectangular cross-section, designed to offer good stability in the proximal portion. The entire surface of the stem is coated with calcium phosphate, which allows and promotes the osseointegration process. The distal part of the stem is tapered to improve its adaptability to the most common bone morphologies. Vertical grooves are present in this area to ensure effective rotational stability. In the proximal portion, on the other hand, the stem has horizontal striations: these serve to prevent the device from sinking and contribute to achieving secondary stability.

Figure 5.3: Stemp P femoral stem, intrauma catalog

5.4 Screws

Orthopedic screws are devices designed to ensure mechanical stabilization in the femoral neck and trochanteric region. Unlike more complex solutions such as intramedullary nails or plates, screws can be used as stand-alone solutions, especially in non-displaced fractures or in patients with good bone quality. They are generally made of Ti6Al4V or 316L stainless steel, in compliance with ISO 5832-1 and ISO 5832-3 standards. They have partial or total threads depending on the desired compression. The structure is cylindrical with a conical profile and has diameters ranging from 6 mm to 8 mm. The length depends on the bone and the location of the fracture and is decided directly by the doctor during the operation.

Chapter 6

Automatic Detection and Classification of Fractures

6.1 Artificial Intelligence, Machine Learning, Deep Learning

To understand the study, it is necessary to clarify four fundamental terms that are central to this study. Artificial intelligence (AI), machine learning (ML), and deep learning (DL) are often confused or even interchanged, but although they are different, they are not entirely unrelated. AI can be considered a large group in which ML and DL are subsets, structured as shown in Fig. 6.1. In particular, DL is a subset of ML, which in turn falls within the broader scope of artificial intelligence.

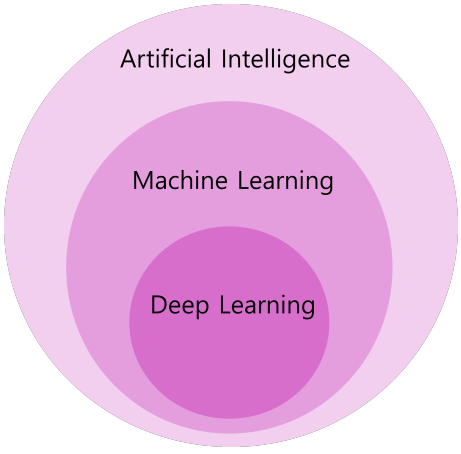

Figure 6.1: Hierarchical relationship between artificial intelligence, machine learning, and deep learning.

As shown in Table 6.1, the main difference between the two techniques lies in the learning mode. Machine learning is a branch of artificial intelligence [79] which uses mathematical models and methods that allow computers to learn patterns from data and make predictions or decisions without explicit rules, mimicking human learning mechanisms. This method is called " non-deep" and relies more heavily on human assistance. The typical flow of building a machine learning model includes: data

Method	Description	Manual Feature Extraction
ML	Models that learn from data, often requiring manually designed features (e.g., SVM, Random Forest, regression).	YES
DL	Deep neural networks that learn representations and classification *end-to-end* (e.g., CNN, Transformer).	NO

Table 6.1: Comparison between ML and DL. The table highlights the fundamental differences between the two learning techniques.

collection, cleaning/feature engineering, model selection, training and validation, evaluation, and finally, deployment. Finally, ML is divided into three categories based on the method and purpose of training:

- **Supervised:** learns from labeled examples to predict values or classes.
- **Unsupervised:** uses unlabeled data to learn latent structures such as clusters or principal components.
- **Reinforcement:** an agent learns to minimize a reward through trial and error.

DL, contrarily, is a subset of ML that uses deep neural networks to learn data representations automatically. These neural networks consist of layers of nodes, connected to each other via weights and activation functions. Typically, each node receives an input and, after processing, transmits the output to the next node. During training, the weights that interconnect the nodes of the network are updated according to a backpropagation method on large data sets, calculating loss functions and applying an essential flow from the output to the input that updates the weights by minimizing errors. Given the higher level of complexity of deep learning models, this technique allows for the recognition of more complex structures and relationships, such as images or text, and automatically determines the information that allows for differentiation between classes or categories of data. We can therefore consider deep learning as "scalable machine learning." [81].

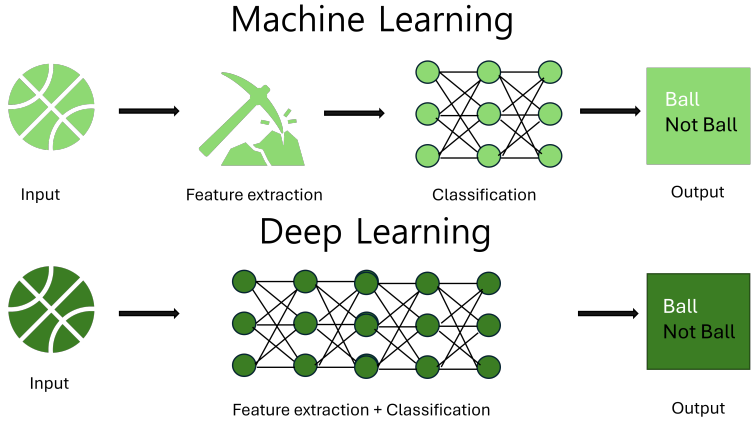

Figure 6.2: Comparison between machine learning (ML) and deep learning (DL) approaches. In ML, feature extraction is performed manually before the task (in this example, classification), while in DL, both phases are integrated within the neural network.

Figure 6.2 illustrates the fundamental difference between an ML and DL approach to image classification. While ML requires a preliminary feature extraction phase in which image features are manually extracted to train the model, DL uses a neural network that integrates the two steps of feature extraction and classification.

Assisted diagnosis models are becoming increasingly popular, as they improve the interpretation of medical images compared to traditional early diagnosis. Indeed, conventionally, fracture diagnosis is performed manually by experts through the analysis of X-ray images, a process that can be time-consuming and sometimes prone to error. In the medical field, these models are used in numerous contexts, and one of the areas in which they are expanding most effectively is the automatic detection of fractures, helping to reduce the workload of doctors and allowing them to focus on more complex cases. There are several studies in the literature that use these technologies to classify bones as fractured or healthy. Yadav *et al.* [90] proposed a model able to automatically classify bones fractures into healthy and fracture classes, starting from X-ray images from a public dataset. The model performed better than previous approaches, achieving an accuracy of 92.44%. Wint Wah Myint *et al.* [91] presented a new method to automatically identify and classify transverse, oblique, and comminuted tibial fractures, using X-ray images. After a pre-processing phase and the use of an algorithm for detecting angular points, the authors adopted classification models such as Decision Tree, KNN, and SVM, obtaining successfully results in terms of reliability in detection and classification. A further study [92] highlights the integration between ML and DL, giving rise to a hybrid model for the automatic detection of healthy and fractured bones. The proposed model combines a convolutional neural network (CNN) with a Canny edge algorithm, which highlights the edges in the images, thus emphasizing the fractures. The model achieved significant results with an accuracy of 99.12%.

6.2 DL Methodology For Automatic Classification

In this thesis, the You Only Look Once model, version 8 (YOLOv8) was used for the automatic classification of proximal femur fractures. YOLOv8 was released by Ultralytics on January 10, 2023, and offers very accurate and fast performance that is suitable for real-time object detection. Performance is improved by the architecture, which includes components such as the backbone and neck. In addition, it can be used for different tasks, including segmentation, object detection, and classification, making it a versatile tool. The choice of YOLOv8 was motivated not only by its technical performance, but also by evidence reported in the literature. For example, Ju *et al.* applied YOLOv8 to the detection of pediatric wrist fractures, achieving an average precision of 0.638 [94]. Verma *et al.* [95] used an earlier version of the model (YOLOv4) to detect humerus fractures on public datasets. Although the system produced promising results, some limitations emerged, including the difficulty in accurately locating fractures in radiographic images. Finally, another study [96] used YOLOv8 in combination with other deep learning algorithms, such as EfficientDet and DETR3, to develop a hybrid model for the automatic detection of hand fractures. This approach demonstrated how the integration of different architectures can improve detection capabilities in complex clinical settings, contributing to more timely and accurate diagnoses.

Convolutional Neural Network

Convolutional neural networks (CNNs) are some of the most widely used architectures for processing two-dimensional images and signals. These neural networks are based on the principle of convolution, a mathematical operation that allows local information to be analyzed, identifying recurring patterns and structures in the data.

CNNs are mainly based on three repeating computational layers[85]:

1. **Convolutional layer**: in this layer, kernels, or weight matrices, are used to scan the image and detect all the main features, such as edges, corners, and textures, which allow a feature map to be created. This step can be performed on subsequent layers, as CNNs are hierarchical structures. Specifically, kernels are typically 7x7, 5x5, or 3x3 matrices that are scanned over the image and produce output values for each position corresponding to the scalar product between the matrix and the portion of the image (Eq. (6.1)). They can produce low-level features, corresponding to lines or curves, or high-level features, which allow more complex objects to be identified. However, it should be noted that the convolution operation results in a loss of information due to a loss of image size. To avoid this, the zero padding technique is used. This technique allows the image size to be increased by adding a layer of values, usually 0, to the input image matrix. For a general grayscale image x and a kernel k, the two-dimensional discrete convolution is defined as follows:

$$y(i,j) = (x * k)(i,j) = \sum_m \sum_n x(i+m,\ j+n)\, k(m,n) \tag{6.1}$$

where, $x(i,j)$ is the image value at coordinates (i,j), $k(m,n)$ is the kernel

weight at coordinates (m, n), and $y(i, j)$ is the output value in the feature map at coordinates (i, j).

2. **Pooling Layer**: this allows the size of the maps to be reduced by eliminating all information deemed superfluous, thus reducing complexity and retaining only the most useful information. This operation is performed by dividing the input image into sub-regions and then taking the highest pixel value for each one.

3. **Activation function**: this layer introduces nonlinearity into the neurons, allowing the network to learn complex patterns such as shapes, textures, or objects. Specifically, YOLOv8 architecture implements the SiLU function (Eq. (6.2)) in its intermediate layers, and the Sigmoid function (Eq. (6.3)) in the final layer.

$$SiLU(x) = x \cdot \frac{1}{1 + e^{-x}} \tag{6.2}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{6.3}$$

A final fully connected layer completes the CNN architecture. In this layer, all neurons are connected one to the other, allowing all extracted features to be combined into a final decision, which is the classification of the object.

In the present study, the CNN architecture employed is the open-source YOLOv8, which was trained by adjusting configuration parameters that must be tuned according to the specific training conditions, such as dataset size, data characteristics, and the final task. The final parameters used for fractures classification with X-ray images are shown in Table 6.2.

Category	**Parameter**	**Description**	**Set value**
Architecture	model	YOLOv8 version	`yolov8n.pt`
Task	task	Problem type	`detect`
Epochs	epochs	Number of training epochs	`100`
Batch size	batch	Number of images per iteration	`16 (default)`
Learning rate	lr0	Learning rate	`0.01 (default)`
Optimizer	optimizer	Optimization method	`SGD (default)`
Momentum	momentum	Acceleration factor in weight update	`0.937 (default)`
Weight decay	weight_decay	Penalty to reduce overfitting	`0.0005 (default)`
Image size	imgsz	Image input resolution	`640 px (default)`
Early stopping	patience	Epochs without improvement to stop training	`50 (default)`

Table 6.2: Configuration parameters for YOLOv8 training

In order to evaluate the optimal number of epochs in the training process, the trend of the loss function calculated on the validation set was analyzed. As shown in Figure6.3, the box_loss, cls_loss, and dfl_loss values tend to decrease in the first iterations. Furthermore, around the 70-80 epoch, the curves stabilize, showing only small variations. Here, we can see the presence of an elbow, which suggests the end

of significant improvements in the ability to generalize on the validation data. This allows us to conclude that the choice of 100 epochs is appropriate and allows the model to converge without overfitting.

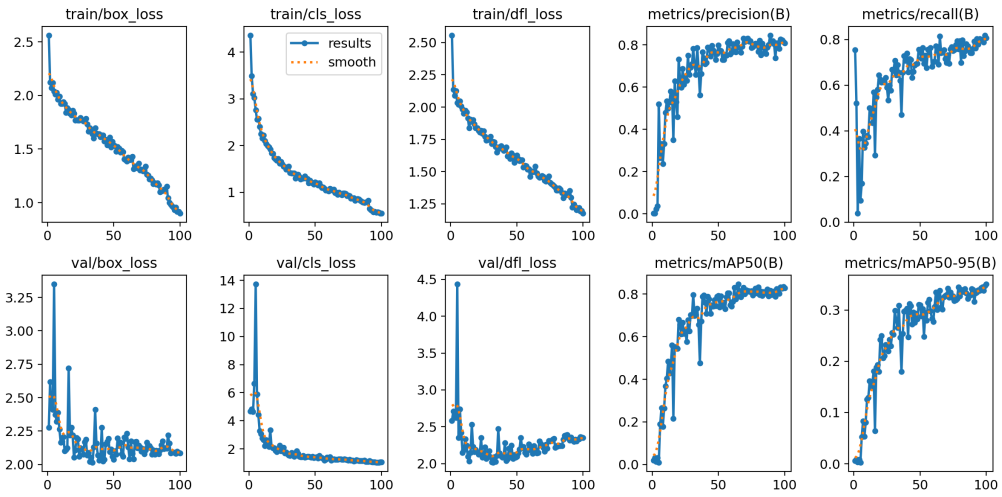

Figure 6.3: The figure shows the evolution of the main loss functions (box_loss, cls_loss, dfl_loss) and evaluation metrics (precision, recall, mAP50, mAP50-95) calculated on both the training set (first row) and the validation set (second row) during the 100 training epochs.

6.2.1 Data Preparation

For the purposes of this thesis, X-ray images from public datasets representing proximal femur fractures were used. The images were labeled with A2 and B2 types of fractures according to the standard AO orthopedic scale. All analyses presented below were conducted in consultation with a specialist surgeon, Dario Haddad, from the Istituto Ortopedico Rizzoli (Bologna, Italy). The correct matching between each fracture and the corresponding label was verified for greater certainty by the physician, who confirmed almost all of the images, discarding a total of three. Specifically, out of a total of 717 images, in addition to the three already excluded, another four images were eliminated in which the physician was unable to identify the fracture due to the presence of a single projection, probably, with additional projections not in our possession, it would have been easier to identify it. A total of 710 images were obtained and subsequently partitioned into training (237), validation (135), and test (338) sets. Within the dataset, there are images that exhibit extreme contrast: some are markedly underexposed (high contrast), while others are overexposed (low contrast). These artifacts are likely due to different acquisition conditions or device settings, such as illumination, exposure time, gain, or post-processing parameters. Such images were retained within the dataset in order to ascertain the deep learning model's ability in recognizing fractures even in suboptimal conditions and to simulate an operating condition in a realistic clinical setting. Indeed, images with both contrasts were retained in the test set to evaluate the latter's robustness and generalization ability.

The YOLOv8 model trained for the automatic detection and classification of the fracture using the dataset described above. To this end, bounding boxes were manually drawn over the images to identify the fracture position, as shown in Figure 6.4. The bounding box coordinates were extracted and, by associating them with the corresponding fracture class, the labels required to train the network for detection and classification tasks were generated. Indeed, Yolov8 model expects ground-truth annotations as `txt` files in the following format. The first value is the class ID (here, 0 for A2 and 1 for B2). The next four values are: the bounding box center coordinates, each normalized by the image width and height, respectively, followed the bounding box width and height, each normalized by image width and height.

Figure 6.4: coordinates of the label for class image A2: 0 0.675607 0.477103 0.399798 0.411076

Once the labels for all the images were obtained, the dataset was created, dividing it into a construction set and a test set, with a split of 70% and 30%, respectively. The first group was further divided into a training set and a validation set, again with different percentages, 60% and 40% respectively. The training set is used to train the network, while the validation serves to fine tune the network parameters over further different samples and avoid overfitting on a limited set of samples. Finally, the test set was used to evaluate the model's generalization to unseen samples and to validate its performance.

6.2.2 Evaluation metrics

In the context of classification, the most commonly used quantitative metrics for calculating model performance are:

- **Precision**: indicates the percentage of correct positive predictions among all

those classified as positive.

$$\text{Precision} = \frac{TP}{TP + FP}$$

It is useful when it is important to reduce false positives.

- **Recall**: represents the model's ability to identify all real positive samples.

$$\text{Recall} = \frac{TP}{TP + FN}$$

Important in cases where it is preferable to reduce false negatives.

- **Specificity**: measures the model's ability to correctly identify negative samples.

$$\text{Specificity} = \frac{TN}{TN + FP}$$

It complements recall and is useful for evaluating the model's ability to avoid false alarms.

- **F1-score**: combines precision and recall into a single metric by calculating their harmonic mean.

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

It provides a balance between precision and recall, which is useful in the presence of unbalanced datasets.

- **Accuracy**: measures the proportion of correct predictions out of the total number of samples.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

However, it is not very representative in the case of unbalanced datasets.

- **Balanced Accuracy**: measures the average between the true positive rate (Recall) and the true negative rate (Specificity).

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

Particularly suitable when classes are highly unbalanced, as it considers performance on both classes equally.

- **Confusion Matrix**: One way to qualitatively evaluate the performance of a classifier is the confusion matrix. In the case of a binary classifier, the confusion matrix is represented by a 2×2 table; however, in this case we have 5 classes, so a 5x5 table will be used. The rows represent the actual classes (ground truth), while the columns represent the classes predicted by the model. This

structure allows us to evaluate how many elements have been classified correctly, represented by the main diagonal, and which classes are confused with each other, represented by the values outside the diagonal [93]. So the values outside the diagonal indicate where the model is going wrong, and analyzing them allows us to understand which classes are being confused with each other. For example, if a column has many values outside the diagonal, then the predicted class is too generic, which means it has low precision. On the other hand, if a row has many values outside the diagonal, then the real class is often confused, so it has low recall.

In addition to the main metrics, there are also others:

- **Mean Average Precision**: referred to as mAP@50, it calculates how closely the predicted bounding boxes correspond to the actual ones. They are evaluated for each class and then averaged:

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^{N} \text{AP}_i$$

- **Box(P)**: indicates how well the drawn boxes match the actual ones.

- **Box Loss**: a consequence of the parameter calculated above, it indicates the error in object localization.

$$L_{\text{box}} = 1 - \text{IoU}(B_p, B_t)$$

- **Class Loss**: measures the error in classification for each detected object.

$$L_{\text{cls}} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

Chapter 7

Automatic Surgical Treatment Classification

The second part of this study focuses on classifying the optimal surgical treatiment of the fracture, given information such as the prevoiusly extracted AO class and other morfological characteristics of the bone, derived from the very same X-ray images. The present analysis can be divided into three main phases: manual feature extraction, model training and validation, and explainability analysis. For this purpose, three different ML models were implemented with a 5-class classification task for the type of surgical intervenion, as shown in Table 7.1.

Class	**Synthetic method**
1	Nail
2	Plate
3	Long nail
4	Prosthesis
5	Screws

Table 7.1: Association between class and means of synthesis.

7.1 Machine Learning

Machine learning (ML) is a technique that uses mathematical models to enable computers to learn from experience provided by data. Specifically: *"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."* [82].

Where T is the task and indicates all the operations or duties that must be completed. Among these, the most common are:

- **Classification:** supervised learning technique in which a function

$$f : \mathcal{X} \to \mathcal{Y}$$

is learned that assigns each input $x \in \mathcal{X}$ a discrete label $y \in \mathcal{Y}$ (class). Classification can be binary (2 classes) or multi-class (more than 2 classes).

- **Regression:** supervised learning task in which a function

$$f : \mathcal{X} \to \mathbb{R}$$

is learned that associates each input $x \in \mathcal{X}$ with a continuous value $y \in \mathbb{R}$. The goal in this case is to estimate a quantitative relationship between the input variables and the output.

- **Clustering:** unsupervised method that aims to group data into sets (clusters) of elements that are similar to each other and different from other groups. No known labels are used, but latent structures in the data are sought, typically based on similarity or distance measures. It can be used to detect hidden patterns, reduce the dimensionality of the data, or pre-process data for subsequent tasks.

P represents the value on which we base our assessment of whether the system is accurate, that is, whether what it returns as output is actually correct and to what extent. This is calculated based on a dataset that is different from the one used during the training phase, in order to effectively evaluate its ability to generalize and apply the learned experience to a new set of data [83].

The letter E, instead, identifies experience. Based on the type of experience to which the model is subjected, the training approach can be considered supervised (presence of labels) or unsupervised (absence of labels).

In the present study, it was considered appropriate to test and compare three different supervised learning models in order to evaluate both predictive performance and robustness with respect to the available data. The selected models are Random Forest, Support Vector Machine, and K-Nearest Neighbors. This allowed for the selection of the model that showed the best performance on the test set, on which the explainability analysis will then be carried out that is, the ability of a model to make its decisions understandable and interpretable, showing which data features most influence the predictions.

7.1.1 Random Forest

Random Forest (RF) is a supervised learning method based on decision trees. As the name suggests, it consists of a large set of trees that are combined to generate a single prediction. This method belongs to the ensemble family, which allows for more accurate and robust predictions by combining the predictions of multiple models. [84]. The model is based on a combination of two fundamental techniques: bagging (bootstrap aggregating) and random feature selection. Bagging involves randomly dividing the data, thus generating high variance between trees, as each is trained on a different subset of the dataset. However, by aggregating the results of many models, a significant reduction in variance is achieved, improving the overall stability and accuracy of the system.

The term "Random Forest" derives from random decision forests, a concept first introduced by Tin Kam Ho of Bell Labs in 1995. RF can be used for both classification and regression tasks. The outputs of the individual trees are aggregated using: the majority vote in the case of classification, the arithmetic mean in the case of regression. We can say that the construction of the Random Forest (RF) is divided into two main phases:

- Tree creation: each tree in the set is constructed from q data samples extracted from an initial training set using sampling with replacement, known as bootstrap sampling. This means that each data point is selected randomly from within the set, and a given sample may be chosen multiple times, even for different trees.
- Random feature selection:
  - Instead of using all M available features, a random subset of m variables is considered in each node. This ensures that the most predictive features do not dominate the others, improving the diversity of the tree.
  - Among the m variables selected, the one that optimizes data separation is selected.
  - The node is divided into two based on the condition found, and the process continues recursively on each branch.

Parameters

- k: is the number of trees. A high value allows for better performance and reduces overfitting, but also involves greater computational cost.
- q: is the number of samples selected to build each tree. This value must be less than n, where n is the total number of samples available.
- m: is the number of features that each tree considers. This value must be less than M, where M represents the total number of features available in the dataset.

To evaluate the optimal k value, the out-of-bags (OOBs) error is typically calculated (Eq. (7.1)). This measure provides an internal estimate of the generalization error and is calculated by evaluating each sample only on the trees that did not see it during the training phase, as well as the OOBs [85].

$$\text{OOB Error} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}\big(\hat{y}_i^{(\text{OOB})} \neq y_i\big) \tag{7.1}$$

Where:

- N is the total number of samples;
- y_i is the actual label of sample i;

- $\hat{y}_i^{(\text{OOB})}$ is the average prediction (or the most voted class) obtained by considering only the trees for which sample i was not included in the bootstrap;
- $\mathbb{I}(\cdot)$ is the indicator function (equal to 1 if the condition is true, 0 otherwise).

The value of k can therefore be chosen as the number of trees beyond which the OOB error no longer decreases.

7.1.2 Support Vector Machine

The Support Vector Machine (SVM) is a supervised classification model based on finding a hyperplane that separates classes with the maximum margin. It is one of the most widely used classifiers, particularly in the biomedical field, thanks to its ability to handle complex data and perform well even in the presence of non-linearly separable features.

The model was originally developed as a binary classifier, but was instead designed to distinguish between two classes. However, when classes are not linearly separable in the space of the original variables, a space transformation is used. In this case, the variables are projected into a higher-dimensional space, where they are more likely to become linearly separable via a line (in 2D), a plane (in 3D), or a hyperplane (in higher spaces).

Linear problems

If the elements are already linearly separable in the original plane and the problem is binary, linear SVM can be applied directly.

In this case, the algorithm constructs the line (in 2D) that maximizes the margin between the two classes.

The margin is defined as the distance between the separating hyperplane and the closest elements of the two classes. These elements are called support vectors and represent the most critical points to classify, as they are located close to the decision boundary.

Given a supervised training set as input, the model implements the function described in Eq. (7.3), constructing a line, plane, or hyperplane that maximizes the margins.

In a T-dimensional space, the equation of the hyperplane is:

$$f(x) = \mathbf{w}^T x + b \tag{7.2}$$

where:

- $\mathbf{w}$ is the weight vector, which determines the direction of the hyperplane;
- b is the bias term, which represents the shift of the hyperplane;

The classification is based on the **sign** of the function $f(x)$:

$$\text{Class}(x) = \begin{cases} +1 & \text{se } f(x) > 0 \\ -1 & \text{se } f(x) < 0 \end{cases}$$

All points above the hyperplane take on a positive value, while those below take on a negative value.

As already mentioned, in two-dimensional space, we can visualize the central hyperplane and two parallel lines that pass through the support vectors of each class. During training, the algorithm attempts to maximize the distance between these two parallel lines. Since optimization problems in mathematics are often formulated as minimization problems, the objective is rewritten in inverse form: finding the hyperplane that maximizes the margins is equivalent to minimizing the norm of the weight vector.

Nonlinear problems

In the real world, it is rare for classes to be linearly separable in the space of the original variables. In such cases, a simple straight line (or hyperplane) is not sufficient to correctly distinguish between classes. There are two main strategies for dealing with this situation:

- **Soft Margin Classification**: consists of maintaining the original space, but allowing for a certain number of misclassification errors. In this case, there is a very important parameter, called C, known as the penalty term. It controls how much the algorithm tolerates misclassification errors. High values of C heavily penalize errors, so the model tries to classify each point correctly. For low values of C, on the other hand, the model accepts that some points are misclassified, favoring wider margins and greater generalization capacity.
- **Non-Linear SVM**: consists of remapping the points in a new implicit feature space, ensuring that the elements belonging to the different classes are linearly separable. This makes it possible to separate the classes with a hyperplane that maximizes the margins.

This thesis will focus on non-linear SVM. This model uses the kernel trick to maximize the margin in an implicit feature space using a kernel function. The primary problem with soft margin and map $\phi(\cdot)$ is expressed by Eq. (7.3).

$$\min_{w,b,\xi} \ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i \quad \text{s.t.} \quad y_i\big(w^\top\phi(x_i)+b\big) \geq 1-\xi_i,\ \xi_i \geq 0. \tag{7.3}$$

Only scalar products appear in the dual form. $K(x_i, x_j) = \langle\phi(x_i), \phi(x_j)\rangle$:

$$f(x) = \operatorname{sign}\left(\sum_{i=1}^{N}\alpha_i y_i K(x_i, x) + b\right). \tag{7.4}$$

The basic idea is to map the points into a higher-dimensional space where the points are linearly separable. There are different types of kernels, which are mainly divided into linear and nonlinear kernels, as detailed in Table 7.2.

To build an SVM classifier in the presence of non-linearly separable data, a kernel function is applied that allows implicitly working in a new feature space of higher dimensions. In this space, during the training phase, the hyperplane and the

Kernel	Formula	Description
Linear	$K(x_i, x_j) = x_i^\top x_j$	Used for linearly separable problems, calculates the scalar product in the original space.
Polynomial	$K(x_i, x_j) = (x_i^\top x_j + c)^d$	Models curved decision boundaries; d controls the degree (with c bias term).
Radial Basis Function	$K(x_i, x_j) = \exp(-\gamma\|x_i - x_j\|^2)$	Measures similarity based on Euclidean distance; suitable for non-linearly separable problems.
Sigmoid	$K(x_i, x_j) = \tanh(\alpha\, x_i^\top x_j + c)$	Similar to an artificial neuron; α and c control the shape of the function.

Table 7.2: Types of kernels used in Support Vector Machines (SVM).

maximum margin are estimated with soft margin, balancing the margin width and classification errors. During the prediction phase, each new point is evaluated using the chosen kernel, and the decision is obtained from the sign of the function in Eq. (7.4), determining whether the point falls above or below the hyperplane in features space.

To handle a multi-class problem, the task is broken down into multiple binary problems, the results of which are then combined according to two different approaches:

- **One vs All**, or **1-vs-Rest**: k binary classifiers are constructed, where k is the number of classes. For each classifier, one class is considered positive and the others negative. In this case, each classifier provides an output, and the classifier with the highest output will be the one that assigns the class.

- **One vs One**, or **1-vs-1**: all possible classifiers are constructed considering all possible pairs of classes, ignoring the rest of the classes. In this case, when classifying a new element, each binary classifier is used, the results are aggregated with majority voting, and the class that obtained the most votes is assigned.

7.1.3 K-Nearest Neighbors

K-Nearest Neighbors (KNN) is a supervised learning model based on similarity or "instance-based learning." It is used for both classification and regression tasks. The model is based on the assumption that data can be classified by considering the labels of the closest neighbors. In a classification task, the model assigns a label to an instance x based on the labels of the k closest neighbors according to a distance metric $d(\cdot, \cdot)$. Once the distances have been calculated, the set $\mathcal{N}_k(x)$ of the k closest neighbors is selected and the prediction is calculated as expressed in Eq. (7.5) in the case of classification and as in Eq. (7.6) in the case of regression.

For classification:

$$\hat{y}(x) = \arg\max_{c} \sum_{i \in \mathcal{N}_k(x)} w_i \, \mathbb{I}[y_i = c], \tag{7.5}$$

For regression:

$$\hat{y}(x) = \frac{\sum_{i \in \mathcal{N}_k(x)} w_i \, y_i}{\sum_{i \in \mathcal{N}_k(x)} w_i} \tag{7.6}$$

The parameter w_i represents the weighting and determines how much each neighbor counts in the decision. The two most common types of weighting are expressed in Eq. (7.7).

$$w_i = \begin{cases} 1 & \text{Uniform weighing} \\ \frac{1}{d(x,x_i)+\varepsilon} & \text{Reverse weighing at a distance} \end{cases} \tag{7.7}$$

with $\varepsilon > 0$ small for numerical stability. By applying uniform weighting, all k neighbors count equally, while inverse distance weighting (Shepard) allows for greater consideration of very close neighbors by increasing the parameter p.

Once the parameter w_i has been defined, a majority evaluation is then performed among the labels of the neighborhood considered $\mathcal{N}_k(x)$. The most frequent class among the k neighbors is the class that will be assigned by the model to instance x in the case of classification. If the task is regression, the arithmetic mean of the values of the k neighbors is calculated.

The selection of the k parameter is therefore crucial in the construction of a KNN model, as it controls the bias-variance trade-off. Typically, small values of k, for example 1 or 3, can cause the model to overfit the training set instances, while values that are too large (e.g., greater than 20) cause the model to generalize better but can introduce bias. Values between 7 and 15 are mainly considered in order to have a more robust model that is able to generalize during prediction [86].

Various distance metrics are used to determine the neighborhood $\mathcal{N}_k(x)$. The distance metrics most commonly used in classification tasks are listed below.

Euclidean Distance

The Euclidean distance between two vectors x and y in an n-dimensional space is defined in Eq. (7.8) :

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{7.8}$$

It is the most common metric, based on Pythagoras' theorem, and calculates the straight-line distance between two points in Euclidean space [87, 84].

Manhattan Distance (or City Block)

The Manhattan distance measures the sum of the absolute differences between coordinates, as shown in Eq. (7.9) below:

$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i| \tag{7.9}$$

This parameter is useful when the given points are structured in a grid-like pattern; in fact, it is also called taxi geometry or block distance [85].

Cosine Distance

The cosine distance, in Eq. (7.10), evaluates the directional similarity between two vectors:

$$d(x, y) = 1 - \frac{x \cdot y}{\|x\| \cdot \|y\|} \tag{7.10}$$

where $x \cdot y$ is the scalar product and $\|x\|$, $\|y\|$ are the Euclidean norms of the vectors. We can say that when the distance between two points increases, the similarity decreases. Conversely, if the distance decreases, the similarity increases. Therefore, we can say that if two points are very similar, then I will obtain a cosine distance equal to 1; if they are dissimilar, then the cosine distance is equal to -1 [88].

Mahalanobis distance

The Mahalanobis distance, represented in Eq(7.11), takes into account the correlation between variables:

$$d(x, y) = \sqrt{(x - y)^T S^{-1} (x - y)} \tag{7.11}$$

where S is the covariance matrix of the features and S^{-1} is its inverse. The latter serves to "normalize" the directions based on variance and correlation. This distance, therefore, taking into account the correlation between variables, is ideal for datasets where the features are not independent. Furthermore, it does not consider the unit of measurement of the features as the distance is automatically corrected. However, it is very sensitive to outliers that can distort the covariance [89].

7.1.4 Feature Extraction

The X-ray image dataset used in this phase of the study is the same as that employed in the first part and previously described. In this second phase, to train the ML models, an initial step of manual feature extraction was carried out to identify the most relevant information within the images and convert it into numerical features interpretable by each model. Additional information that was not available in the original dataset was obtained in consultation with the orthopedic surgeon from the Istituto Ortopedico Rizzoli, together with the ground-truth labels of the surgical treatment class for each fracture.

The features were extracted by identifying key reference points as landmarks on the bone, as shown in Figure 7.1, and computing the corresponding distances and angles, using `Matlab` toolbox. Below is a list of the key features and how they were calculated:

- **S1, S2, S3**: reference points for defining the axes of the femur. The femoral neck axis is defined as the straight line passing through point S1 and the midpoint of the segment connecting C1 and C2. The femoral axis is instead represented

by the straight line passing through points S2 and S3. The intersection of these two axes forms an angle that allows us to determine whether the femur is in a varus or valgus position, a condition that can cause instability.

- **C1 and C2**: identify the ends of the femoral neck. They are essential for identifying the midpoint (CM), through which the femoral neck axis must pass.
- **F1 and F2**: represent the ends of the fracture line. Taking them into account allows the length of the fracture to be calculated.
- **LC1**: lateral reference point, located precisely 3 cm below the vastus ridge. It is used to draw the 135° reference line that will intersect the fracture line.
- **D1 and D2**: ends of the medullary canal, allowing the diameter of the bone canal to be calculated.

Starting from the anatomical points saved as coordinates (X,Y), the following parameters can be calculated:

- **Pauwels angle** : angle formed by the intersection between the horizontal axis and the fracture line, computed as in Eq. (7.12).

$$\text{Pauwels} = |\theta| = \arctan\left(\frac{x_2 - x_1}{y_2 - y_1}\right) - \text{refH} \tag{7.12}$$

  The value refH is the horizontal line closest to the fracture.

- **Lateral Femoral Wall Thickness (LWT)**: corresponds to the thickness of the lateral wall of the femur and is calculated using point LC1. The LWT value is given by the distance between LC1 and an intersection point, as reported in Eq. (7.13). The latter is given by the intersection of two lines, the first being a line inclined at 135 degrees to the diaphyseal axis passing through point LC1 and the line connecting the ends of the fracture line.

$$\text{LWT} = \|\text{LC1} - P_{\text{intersezione}}\| \tag{7.13}$$

- **Neck–Shaft Angle (NSA)**: angle that allows us to assess deformities such as varus or valgus. It is calculated as the angle subtended between the femoral neck axis uN (given by points S1 and CM) and the diaphysis axis uS (given by S2 and S3), as shown in Eq. (7.14).

$$\text{NSA} = 180° - \cos^{-1}\left(\frac{\vec{u}_N \cdot \vec{u}_S}{\|\vec{u}_N\| \, \|\vec{u}_S\|}\right) \tag{7.14}$$

- **Fracture length**: direct distance between the two ends of the fracture line F1 and F2 (Eq. (7.15)).

$$\text{FractureLength} = \|F_2 - F_1\| \tag{7.15}$$

- **Medullary canal width**: measures the internal diameter of the femur and it is computed as Eq. (7.16).

$$\text{Diameter} = \|D_2 - D_1\| \tag{7.16}$$

- **Scale factor**: reference used to normalize data, its value is given in pixels as shown in Eq. (7.17).

$$\text{ScaleFactor} = \|S_1 - CM\| \tag{7.17}$$

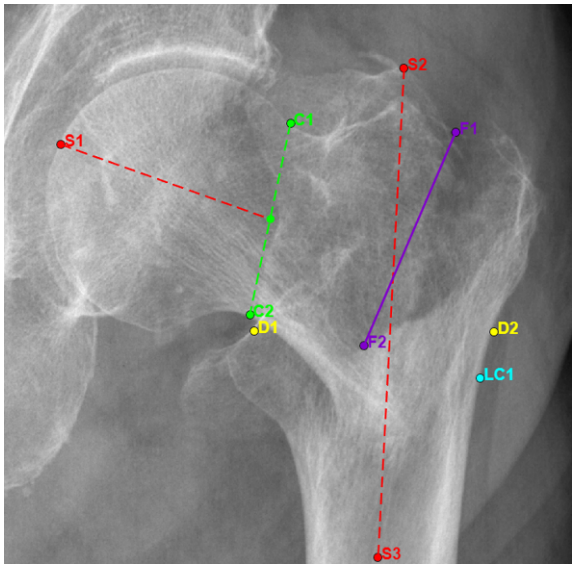

Figure 7.1: The figure shows the fundamental anatomical points of the proximal femur.

The features extracted from each image were then normalized using a scale factor computed directly from the image itself. This procedure standardizes the measured dimensions across images, ensuring consistency between cases and making the features independent of the acquisition distance.

As mentioned above, some clinical parameters provided by the physician were then integrated into the dataset according to the following coding:

- **Age**: value 0 for young subjects (between 18 and 65 years old), value 1 for elderly subjects (65 and over), and value 0.5 for intermediate cases, for instance, patients who, despite being elderly (e.g., 80 years old), remain physically fit thanks to physical activity, or young subjects (in this case, between 50 and 60 years old) with a sedentary lifestyle. Although these values are not objectively measurable, they reflect the physician's judgment based on experience, in the absence of precise personal data.

- **Osteoporosis**: value 0 if the condition is not present, value 1 if it is present.

- **AO**: value 0 for type A2 fractures, value 1 for type B2 fractures.

- **Stability**: value 0 for fractures considered stable, value 1 for unstable fractures.

- **Labels**: 5 means of synthesis were used, assigning each a value from 1 to 5.

These characteristics emerge directly from the image, although not always in such a simple way. The doctor confirmed that the above values can be verified. For example, the patient's age can be deduced from the thickness of the cortex in the diaphyseal region of the femur, a parameter that tends to decrease with age. The second parameter, relating to the presence of osteoporosis, was assessed by observing bone density, in other words, variations in contrast, appearing lighter or darker in different areas of the femur. In fact, most of the samples analyzed showed mild osteoporosis; some patients showed signs of arthritis, while others were in good bone condition.
The mechanical stability of a fracture is the ability of the bone and/or its fragments to maintain their original position without collapsing when a load is applied. Therefore, a fracture is considered stable if it allows correct load transmission along the axis of the femur; conversely, it is defined as unstable if it tends to give way, causing varus deformation of the bone.
To assess stability, two main anatomical landmarks are considered: the lesser trochanter and the lateral wall of the femur, which involves the greater trochanter. Actually, the physician observes whether the fracture line extends beyond the lesser trochanter. The more the fracture line extends distally from this landmark, the greater the degree of instability. As for the lateral wall, it must be intact and able to accommodate a fixation device, typically a screw, in order to ensure stability. If the lateral wall is excessively thin, it will not be able to hold the fixation device.
The physician provided further clinical input regarding the choice of fixation devices. Specifically, for class A2 fractures, the main orthopedic solutions recommended by the physician were nails, plates, and long nails, while for class B2 fractures, the options included prostheses and screws. Although this division represents a general classification, there are cases in which an A2 fracture may require prosthetic treatment, for example, following a more detailed analysis of the stability of the fracture or the quality of the bone. Similarly, some typical class A2 solutions may occur in fractures classified as B2.
The doctor emphasized that the choice of prosthesis is often tailored to the patient's characteristics. Generally speaking, for a young person, a plate is preferable to a prosthesis, while for an elderly patient, a prosthesis is recommended, as it tends to promote rapid mobilization. Furthermore, an older person may not be able to undergo a second surgical procedure and therefore it is more appropriate to proceed directly with the prosthetic implant.

7.1.5 Training and Hyperparameters Optimization

For the purposes of this second phase of the study, three different ML models were implemented: Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). For each model, training and hyperparameters optimization were performed through a grid-search using k-fold cross-validation, in order to compensate for the limited size of the available dataset. The aim was to identify, for each model, the optimal combination of hyperparameters that maximizes balanced accuracy on the construction set.

Due to the limited size of the dataset, dividing it into fixed training, validation, and testing sets would result in a loss of useful data. Furthermore, since the classes are unbalanced, this could lead to divisions that do not allow for a correct representation of all classes. To overcome these problems, k-fold cross-validation is used. The dataset is divided into k subsets, known as folds. At each iteration, one of the folds is used as a validation set and the remaining $k-1$ are kept for model training. This method ensures that each fold maintains the proportion between classes, improving robustness. Performance metrics are calculated for each fold, and the results obtained in the k experiments are averaged.
Once the optimal parameters have been identified for each ML model, the models were instantiated using these optimal settings and trained on the construction set. Finally, each model were tested using the test set to obtain a final performance estimate.
The optimal value of k, given the size of the construction set, was tuned by evaluating the balanced accuracy obtained through cross-validation and selecting the value that maximized it, as shown in Table7.3. The selected k value is equal to 6.

Final Best Model K = 4				
RF	Trees=50	Leaf=9	Mtry=5	Acc=72.67%
SVM	Kernel=rbf	C=10	Scale=10	Acc=73.10%
kNN	K=13	Dist=cosine	Weight=equal	Acc=73.13%

Final Best Model K = 5				
RF	Trees=50	Leaf=3	Mtry=3	Acc=73.10%
SVM	Kernel=linear	C=1	Scale=NaN	Acc=73.43%
kNN	K=5	Dist=euclidean	Weight=equal	Acc=72.60%

Final Best Model K = 6				
RF	Trees=50	Leaf=9	Mtry=5	Acc=73.07%
SVM	Kernel=linear	C=1	Scale=NaN	Acc=74.23%
kNN	K=5	Dist=cosine	Weight=equal	Acc=72.77%

Final Best Model K = 7				
RF	Trees=200	Leaf=5	Mtry=3	Acc=73.17%
SVM	Kernel=linear	C=1	Scale=NaN	Acc=73.90%
kNN	K=7	Dist=cityblock	Weight=equal	Acc=72.76%

Table 7.3: Result for various k folds of parameter optimization

The hyperparameters considered vary depending on the model. For RF, the following were tested:

- **Number of trees in the forest:** it represents the number of trees generated, each of which produces an independent prediction. The values tested were 50, 100, 150, and 200.

- **Minimum leaf size:** indicates the minimum number of samples that a node must contain to be considered a leaf, that is, a terminal node of the tree. This parameter allows you to control the complexity of the model. The values tested were 3, 5, 7, and 9.

- **Number of predictors at each split**: after each split, the model randomly selects a subset of features. This parameter introduces variability between trees, preventing them all from using the same variables. The values tested correspond to the total number of features and half the total number of features.

For multi-class SVM using the One-vs-One method.

- **Kernel Function**: type of kernel used to separate classes, mainly defines the transformation function that maps data in space. A linear kernel (creating a linear hyperplane in the original data), a radial basis function kernel (using a Gaussian function), and a polynomial kernel (using a linear function, typically considered for curvilinear or parabolic patterns) were considered.

- **C penalty factor**: penalty considered for classification errors. It acts as an adjustment parameter, compromising the margin width and the number of classification errors. The values tested were 0.1, 1, 10, 100.

- **Kernel Scale**: used only for RBF and serves to control the width of the Gaussian function bell curve. The testes values are 0.1, 0.3, 1, 3, 10.

- **Polynomial Order**: defines the order if the polynomial kernel is used. The testes values are 2, 3, and 4.

And finally, for KNN:

- **Number of neighbors (K)**: indicates how many neighbors are taken into consideration to assign the class to the point under examination. The values tested are 5, 7, 9, 11, 13, and 15.

- **Distance metric**: defines how the distance between points in the feature space should be calculated. This decision affects how neighbors are identified. The distances tested are Euclidean, cityblock, cosine, and Mahalanobis.

- **Neighbor weight**: defines how neighbors should contribute. Two types were tested: equal, where all neighbors contribute equally, and inverse, where points at a shorter distance have more influence. In the code, we have only kept the first case.

Model	Optimal hyperparameters	Accuracy (%)
RF	trees = 50, leaf = 9, mtry = 5	73.07
SVM	Kernel = linear, C = 1	74.23
KNN	K = 13, Distance = cosine, Weight = equal	72.77

Table 7.4: Best models hyperparameters configurations, considering the balanced accuracy metric after k-fold cross-validation.

7.1.6 Explainability

Understanding how and why the model makes certain decisions is important. This process falls within the scope of explainability, or rather, the ability to interpret and justify the model's behavior.
In this study, the SHapley Additive exPlanations (SHAP) method was adopted, which draws on the concept of cooperative games. In this context, the contribution that each player makes to the overall result is measured. In our case, the players correspond to the features of the model, while the result is represented by the predictions.

Mathematically, for a prediction $f(x)$, the SHAP value ϕ_i associated with feature i is computed as reported in Eq. (7.18).

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|! \cdot (|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)] \tag{7.18}$$

Where:

- F is the complete set of features,
- S is a subset of features that does not include i,
- $f(S)$ is the model's prediction considering only the features in S,
- the combinatorial coefficient weights the contribution of each possible coalition.

Conceptually, an initial prediction is made by the model using all features, then the extent to which the prediction changes when one feature or groups of features are removed is measured. In fact, the Shapley value represents the extent to which a given feature influences the prediction towards a given class.

7.1.7 Evaluation metrics

For the evaluation of the performance of the models implemented in this second phase of the study, the following metrics were computed: precision, recall, specificity, F1-score, accuracy, balanced accuracy, and the confusion matrix, already discussed in Section 6.2.2.

Chapter 8

Results and Discussion

This section will present the results obtained during the two phases of the study. The data will be presented and explained using the evaluation metrics described in the previous chapters.

8.1 DL For Fracture Classification

The results for this first part will be presented both qualitatively, through images, and quantitatively, through evaluation metrics. Qualitatively, it can be seen in Figure8.1 how the network highlights the region of interest through the bounding box in red. This visual analysis allows us to verify the correctness of the fracture location, observing that it is able to generalize even in images with variable contrasts and different anatomies.

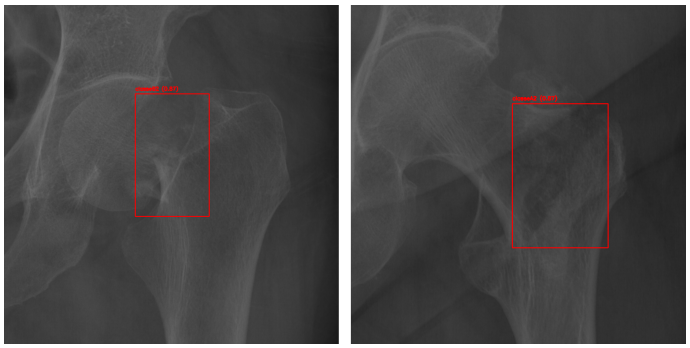

Figure 8.1: The images are annotated with colored bounding boxes that identify the region of interest and the predicted class, along with the confidence score that quantifies how strongly the model associates the region with that specific class.

Quantitatively, the results obtained are available in Table8.1 and graphically represented in the confusion matrix in Figure8.2.

Class	Images	Istances	Box(P)	Recall	mAP@50	mAP@50–95
All	232	232	0.808	0.797	0.828	0.35
Classe 1	116	116	0.794	0.793	0.829	0.316
Classe 2	116	116	0.822	0.802	0.827	0.385

Table 8.1: Evaluation metrics in femur fracture classification with YOLOv8 model. The first row of results refers to the overall average performance, while the following rows report the values obtained for each class.

The results indicate that the model is able to correctly localize and classify almost all instances belonging to both classes. Focusing on the values, we notice that the precision exceeds 80%, meaning that only 20% of predictions are incorrect. Furthermore, the recall value of 0.797 shows that the model detects almost 80% of the objects present in the image. These values appear balanced, confirming that the model does not tend to be biased either toward detection loss or an excess of false positives. The mAP@50 value of 0.828 can be considered satisfactory. Analyzing the individual classes, we observe that class B2 achieves slightly better performance, although the two are quite similar. This suggests that the model can effectively distinguish B2 from A2.
On the other hand, it tends to confuse class A2 with the background, as shown in the confusion matrix 8.2. Specifically, we observe:

- 103 correct predictions for A2, with 30 FN classified as background.
- 99 correct predictions for B2, with 13 FN classified as background and 2 wrongly classified as A2.

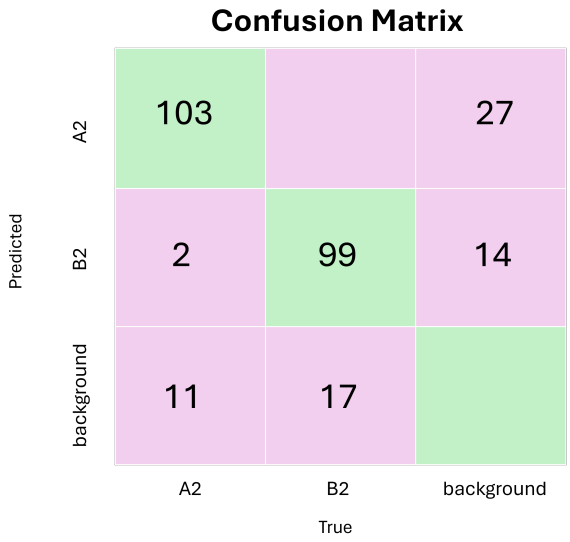

Figure 8.2: Confusion matrix obtained during model training.

Overall, the performance is good, and the model shows generalization capability. A possible improvement could involve training optimization, perhaps by increasing

the number of images through data augmentation to diversify them and reduce background confusion.

8.2 ML Models Surgical Treatment Classification

In this section, we present the results obtained from training the three ML models used for the classification of the optimal surgical treatment for each specific fracture case.

In Figure 8.3, the confusion matrices obtained on the construction set are reported for RF, SVM, and KNN respectively.

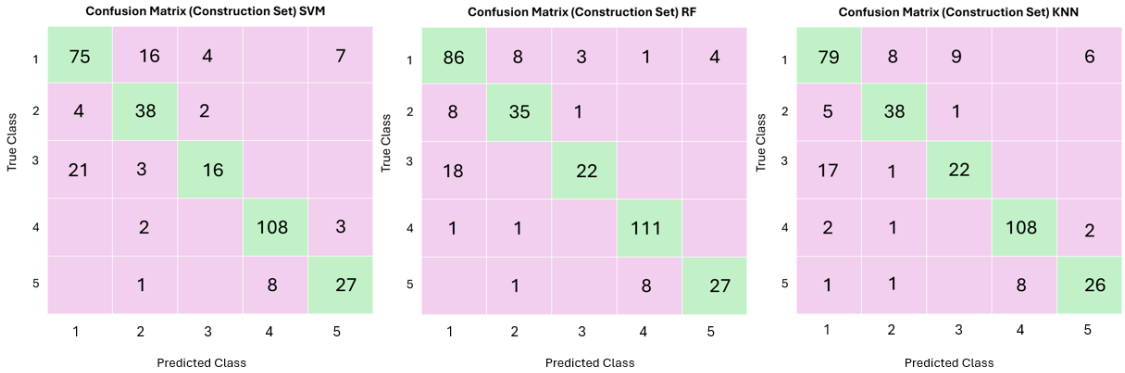

Figure 8.3: Confusion matrices obtained during model training. From left to right: RF, SVM, KNN.

For all three models, high values are observed along the main diagonal, which implies correct classification during training. However, some classes show non-zero values outside of it. In detail, we observe that:

- RF: the model was able to capture the fundamental patterns, but some classes, such as class 3, were confused with class 1, and class 1 was often mistaken for class 2.
- SVM: in this case as well, the most indistinct classes are class 1 and 2, and class 3 and 1. Compared to RF, some classes show improvement in the diagonal values.
- KNN: the results are still good, but there is a slight decrease in precision for some classes, with more cross-classification errors present.

In fact, we can generalize by stating that among all models, the SVM performs best, as it is more selective compared to KNN, which shows a more generic behavior, and to RF, which, although achieving good values, is less precise. It should be noted that this does not indicate superiority of the model itself, but rather that each has correctly learned the main characteristics of the data. Moreover, the choice of SVM is reinforced by the values in Table7.4 in the previous chapter and by the graphical representation of the test set confusion matrix shown in Figure 8.4. For this reason,

only the test results for this model will be presented.
Considering both the values in Table 8.2 and the SVM's confusion matrix 8.4, we can explain the results:

Metric	Value (%)
Balanced Accuracy	72.82
Macro Precision	67.05
Macro Recall	72.82
Macro F1-score	68.70

Table 8.2: SVM model testing. The table reports quantitative results obtained on the test set.

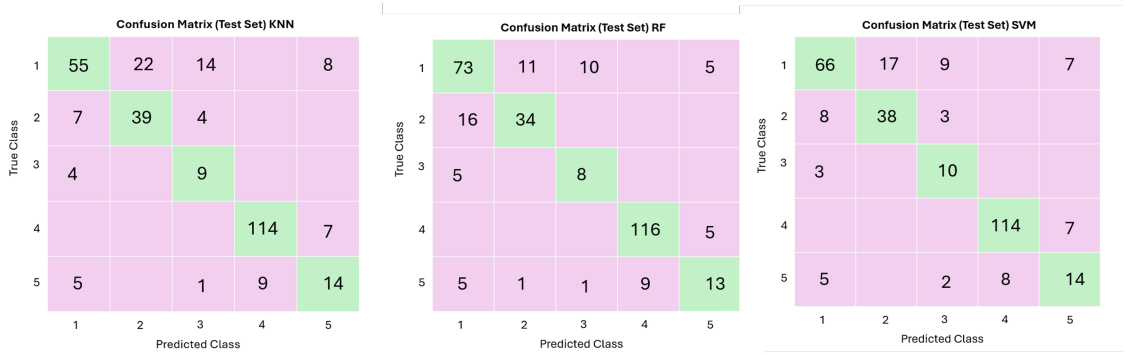

Figure 8.4: KNN,RF,SVM model testing. The confusion matrix obtained on the test set is reported.

Taking the average model values as reference:

- **Balanced Accuracy (72.82%)**: shows that, despite class imbalance, the model does not reach particularly high accuracy levels. This suggests that, while well-represented classes like class 4 perform well, others such as 3 and 5 are more problematic.

- **Macro Precision (67.05%)**: indicates that the model is overall precise but shows wide margins for improvement. It correctly classifies a good percentage of cases but, due to non-uniform class distribution, struggles to define clear decision boundaries.

- **Macro Recall (72.82%)**: indicates that most of the distinctive features among classes are effectively recognized by the model.

- **Macro F1-score (68.70%)**: measures the balance between precision and recall, showing that the model makes non-negligible errors in less represented classes.

Analyzing the individual classes in detail:
Class 4 is the best identified, with a precision of 99%, a recall of 94.2%, and an F1-score of approximately 97.0%. These results are consistent with the fact that class 4 is one of the most populated classes in the available dataset, together with class 1.
Class 3 shows a high recall of approximately 76.9%, but a very low precision close to 33.3%, indicating that it is frequently misclassified in other classes. This is probably due to the fact that class 3 is the least represented class in the dataset.
Class 5 exhibits comparable precision and recall, at around 50% and 48% respectively, indicating intermediate performance. It is frequently misclassified as classes 1 and 3 and is the second least represented class in the dataset.
Classes 1 and 2 achieved good performance, with recall values of 66.7% and 77.6%, and precision values of 80.5% and 66.7%, respectively. Overall, this indicates generally reliable classification, although some degree of confusion between the two classes is still present.
After discussing the results reported above with the medical expert, we were able to justify the misclassifications observed.
It was confirmed that the use of the plate or nail is often at the surgeon's discretion, who decides after careful evaluation whether to use one or the other depending on fracture stability and patient age. Essentially, their interchangeability was justified depending on the case, stating that the nail is generally more stable. For this reason, in cases of uncertainty, the nail is prioritized. Obviously, for younger patients, plates are used more frequently, since inserting a nail into the medullary canal causes greater damage than a plate. Therefore, in younger patients, the plate is preferred. This consideration can be regarded as a valid explanation for the value 8 represented in the confusion matrix position $M_{2,1}$.
As for values 3 and 9, $M_{3,1}$ and $M_{1,3}$ respectively, the physician provided an even stronger justification. He confirmed that, since these refer to the same fixation device, their interchangeability is even more acceptable. The long nail is preferred in all cases where there is doubt about subtrochanteric involvement, since, by entering deeper into the medullary canal, it allows for longer-lasting fracture stability. He also confirmed that the postoperative healing time is practically the same. The physician further supported this statement by noting that, in current practice, there is a progressive tendency to prefer the use of the long nail, and such statements are increasingly discussed in specialist conferences.
Now, regarding screws, we attempt to explain values 8 and 5 in positions$M_{5,4}$ e $M_{5,1}$. The doctor agreed with both cases, as placing screws requires ensuring that the fracture is stable, which must be evaluated based on the fracture line: the more horizontal it is, the more stable it is, the more vertical, the more unstable. In other words, a higher success rate is expected using a prosthesis than with screws for the same fracture. Thus, value 8 is justified, as the failure rate is higher when using screws in uncertain cases. Using screws results in a much longer postoperative healing time compared to a prosthesis.
Regarding value 5, the doctor did not disagree, as the nail also includes a cephalic screw; therefore, in cases of uncertainty between the two, it is safer to opt for the solution that ensures a faster recovery.

One result that raised some concern again involves class 1: the model incorrectly classified 7 features into class 5. This discrepancy raised doubts from the physician, who questioned the clinical coherence of such an assignment, considering it poorly justified. A possible explanation for this result is likely the limited number of images available for that class. In such a case, the lack of representativeness within the dataset may have prevented the model from clearly learning its features, leading to significant errors.
Overall, the values support the idea that the SVM can learn discriminative features from the data but needs improvements to fully capture their complexity.
To ensure the reliability of the model's results, the values were further justified through explainability analysis, shown in Figures 8.5, 8.68.7, 8.8, 8.9. This approach allows us to evaluate which features influenced the predictions, enabling verification of whether the automatic decision-making process actually reflects the clinician's reasoning.

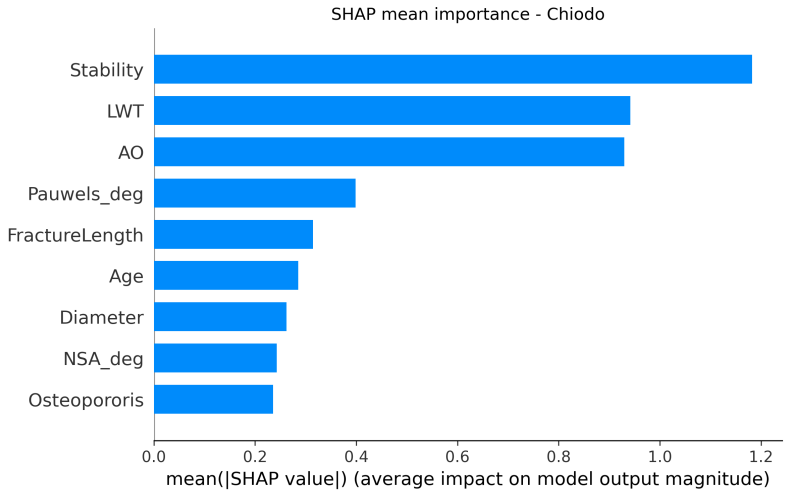

Figure 8.5: SHAP values:nail class

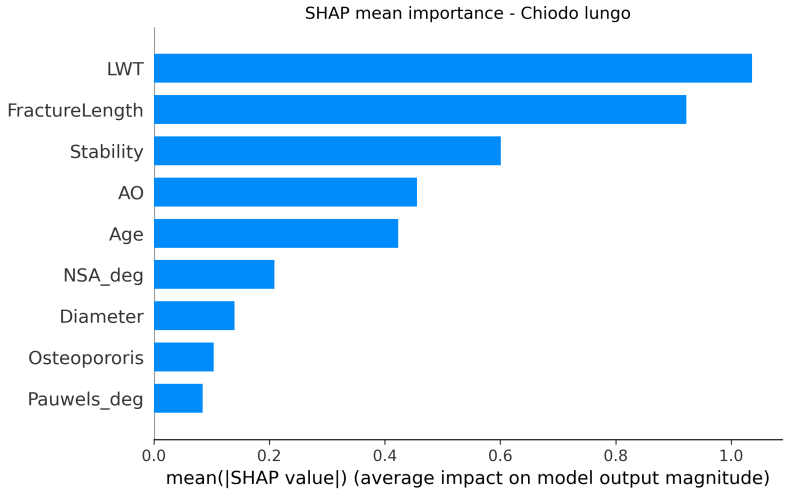

Figure 8.6: SHAP values: long nail class

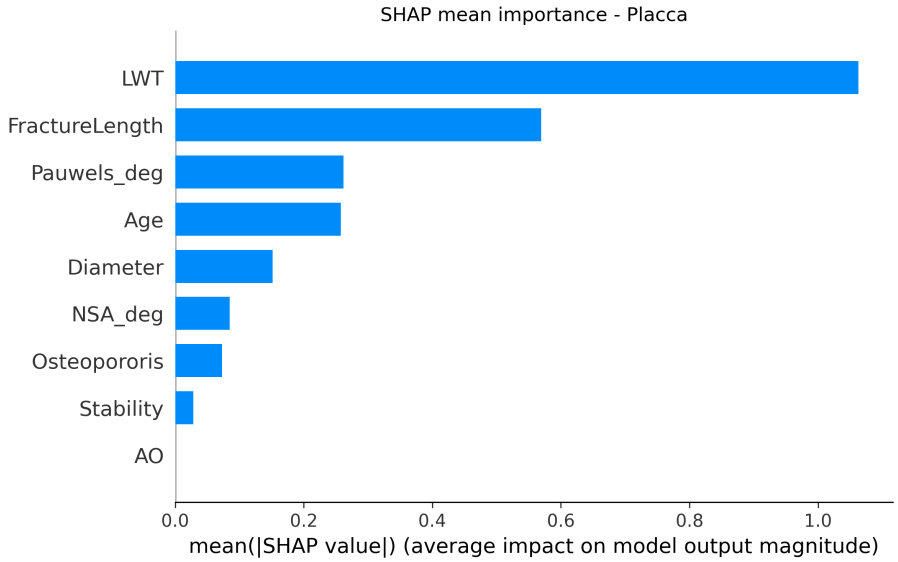

Figure 8.7: SHAP values: plate class

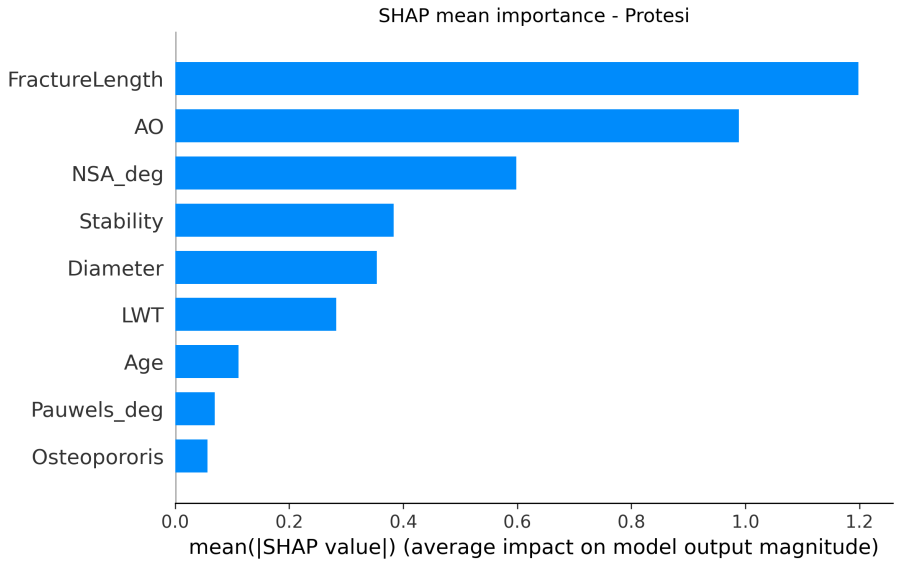

Figure 8.8: SHAP values: prosthesis class

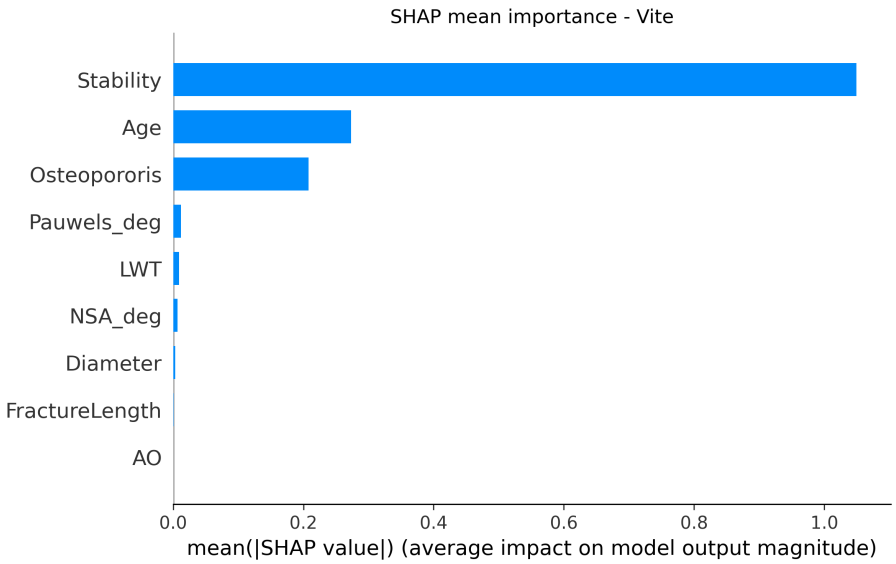

Figure 8.9: SHAP values: screw class

To assess the coherence produced by the model, a ranking of the features was built together with the physician, based on their degree of importance. This phase was carried out before showing the SHAP analysis to the physician, in order to avoid any form of conditioning in the clinical judgment. In summary, the general consideration expressed by the physician was positive; he expressed a favorable and satisfactory opinion for most of the analyzed cases. Below, the clinical observations he formulated are reported in detail.

The first one to be analyzed is the nail 8.5: the first three features are correctly positioned and coherent. However, he raised a criticism regarding the length of the fracture, considering that this parameter should have appeared earlier in the ranking, specifically before the Pauwel's angle, which he believes has a less decisive impact in this context. The features located below were considered negligible or of little decision-making relevance. It is interesting to note that, precisely in this class, the model showed a superior ranking ability compared to the initial classification proposed by the physician.

For the long nail class8.6, things are somewhat different: in this case, the physician had considered the AO feature on par with length, placing both in the first position. Subsequently, he disagreed with the LWT, stating that it would be more appropriate in the fourth position, giving greater importance to stability by placing it second.

Regarding the plate8.7, the physician expressed a specific criticism about the position of the stability feature, considering it too low in the hierarchy proposed by the model. According to his evaluation, it should occupy at least the second place, if not the first, among the most relevant features. This observation is motivated by the fact that plate and nail represent interchangeable therapeutic solutions in many clinical cases, and that the only discriminating factor between the two could be the patient's age.

For the prosthesis class8.8, the judgment was overall very satisfactory. Particularly consistent were the AO and fracture length features. The NSA angle was also recognized as relevant, although it was noted that stability should be placed in the

third position in the feature ranking.
With reference to the last class8.9, the physician confirmed the correctness of the position attributed to stability and age. However, he observed that, instead of osteoporosis, he would have given greater weight to the Pauwel's angle, considering it more crucial in the therapeutic choice for this class.

8.3 Conclusion

The work aimed to develop and analyze classification models, such as SVM, KNN, and RF, in order to support the selection of the most appropriate fixation method in the treatment of proximal femur fractures, highlighting the coherence of the decisions made by the model. The analysis identified SVM as the best-performing model, with a Balanced Accuracy of 72.82% and a Macro F1-score of 68.70%. Although these values are not particularly high, they show the model's ability to learn discriminative relationships between the classes. Specifically, the good performance obtained for the "Plate" (4) class should be noted; however, it is also necessary to consider the classes with more weaknesses, such as Long Nail (3) and Screw (5), probably due to the scarcity of available data. A fundamental contribution was given by the Doctor, who made it possible to validate and clinically contextualize the main discrepancies in the confusion matrix. In fact, many misclassifications turned out to be plausible choices, due to the interchangeability of therapeutic solutions.
The explainability analysis using SHAP further reinforced this coherence, showing that the model tends to base its decisions on features that are actually relevant from the physician's perspective. Nevertheless, although in some cases the importance order does not perfectly match the physician's opinion, most evaluations were aligned and interpretable, leading to the conclusion that the model seeks to mirror clinical reasoning.
Globally, the results obtained are satisfactory and confirm that machine learning models can be integrated as a support tool in the medical field. On the other hand, some aspects remain to be improved, such as the expansion of the dataset with additional images and the integration of objective data such as chronological age and the presence of possible comorbidities, avoiding reliance solely on the physician's experience.

8.4 Future Approaches

8.4.1 Start of Clinical Trial

Given the promising results obtained from the present study, a clinical trial has been initiated with the objective of addressing the main issue of the study. The lack of data will therefore be resolved through the collaboration between the Politecnico di Torino and two hospital facilities in Bologna: Istituto Ortopedico Rizzoli and Ospedale Maggiore. This will make it possible to obtain X-ray and CT images that can better represent the less numerous classes, and at the same time allow the fracture to be observed from multiple projections, thus achieving a more precise localization. For

the drafting of the protocol, the assistance of specialist doctors was necessary, who made themselves available both for the classification and for the identification of the most favorable treatment for each fracture. The protocol begins with an introduction to the main problem that motivated the initiation, first of the thesis and then of the current project. This section is called the Rationale, within which the pathology under investigation is extensively described, highlighting its clinical relevance through the support of epidemiological data drawn from previous publications. Through the analysis of the latter, the current limitations in the management of femoral fractures have been highlighted, explaining how the study can contribute to filling some of these gaps. The protocol continues with the second section called study design. The trial is defined as a retrospective observational investigation, based on the analysis of radiological images (CT and X-ray) of the femur that are already available in the facilities. At this point, a detailed description of all the phases of the process was made, such as data selection, up to the processing of the expected results. As already mentioned, the data will be acquired from facilities that already have them within their clinical archives, making it necessary to include an anonymization phase in compliance with the General Data Protection Regulation GDPR (EU 2016/679), in order to protect patients' privacy. These images will constitute the basis for the training and validation of a machine learning algorithm, developed with the aim of classifying femoral fractures and supporting therapeutic planning. The dataset will be divided into three distinct sets: training set, validation set and test set, and the model will be trained to recognize the distinctive characteristics of the fractures. In addition to femur classification, we will implement a system that allows the network to autonomously extract morphological and structural parameters such as bone characteristics (such as discontinuities, density, orientation of fragments) and anatomical dimensions (length, diaphyseal diameter, joint angles, anteversion), using the images obtained through femoral CT scans.

To make the dataset as uniform as possible, the population has been specifically described. In this section, the characteristics that patients must possess to be eligible have been specified in detail. It was then necessary to distinguish between subjects to whom the Vancouver classification applies and those for whom the AO classification is adopted. For the AO classification, the protocol stipulated that the images be free of medical devices such as screws, nails, plates or fixators, therefore representing only the femur. This choice was motivated by the desire to focus exclusively on the morphology of the native bone, without interference due to the presence of metallic or implantable elements. Conversely, for the Vancouver classification, it was required that the images clearly include both the fracture and the entire hip prosthesis, in order to allow an accurate evaluation of the periprosthetic fracture. Regardless of the classification system, the images must be divided according to biological sex (male or female) and age group. Specifically, the protocol distinguishes between young subjects, aged between 18 and 60 years, and adults, aged 60 years or older. This distinction is essential precisely because of the change in femoral morphology with age. In fact, in elderly patients, due to degenerative diseases, hip prosthesis implantation is more frequently required. In young patients, on the other hand, such invasive interventions tend to be avoided, preferring approaches aimed at preserving bone tissue. Furthermore, the lifestyle between the two cases is very

different; typically, physical activity in younger subjects, being more intense, would lead to greater mechanical stress on the prosthesis and increase the risk of premature wear, leading to complex revisions in the case of implantation.
In order to obtain a balanced dataset among the classes, for the AO classification, the collection of a minimum of 200 images related to male patients and 200 related to female patients aged between 18 and 60 years has been planned. Furthermore, for each of the four identified classes (males 18–60, males ≥ 60, females 18–60, females ≥ 60), a minimum of 400 images has been required, equally divided between X-ray and CT (200 for each modality). As for the Vancouver classification, reserved for patients aged over 60 years, the protocol provides for the collection of at least 200 X-ray images and 200 CT images for each sex.
If the images obtained are not sufficient to reach these numbers, it has been planned to integrate them with public databases such as Musculoskeletal Radiographs (MURA), Open-i, and DeepLesion. Continuing with the description, the instrumentation has been listed, represented not only by the images but also by the networks and the patient's medical record. The latter is essential in order to identify the appropriate fixation device for each individual. In fact, the main objective of the study is precisely to build a predictive model capable of recognizing not only the fracture but also of suggesting the subsequent solution through the choice of the fixation device most suitable for the patient's needs, taking into account both the morphology of the fracture and the patient's clinical condition. The latter includes fundamental parameters such as age, bone quality, the presence of comorbidities, and other pre-existing conditions, which can significantly influence therapeutic choices.
The image processing will be preceded by a pre-processing phase to normalize the data and eliminate noise. Subsequently, the images will be analyzed using convolutional neural networks (CNN), and afterwards, machine learning systems dedicated to the selection of the most suitable orthopedic plate will be used, depending on the fracture and patient characteristics. In fact, for this purpose, the Intrauma S.p.A. company has made available a plate catalog that collects the main available solutions. The entire process has been designed with the purpose of providing decision support to the clinician, offering a personalized solution.

Appendix 1

Clinical Trial Protocol

Protocollo sperimentazione clinica

"Approccio neuromorfico al design biomeccanico delle placche femorali"

Acronimo

"FEM-NEURODESIGN"

Trial Spontaneo, Osservazionale Retrospettivo No profit
Synopsis

Versione 8 Agosto 2025

Titolo dello studio:	Analisi TC e RX del femore per la classificazione delle fratture e la selezione automatizzata del design della placca tramite IA
Promotore dello Studio:	Politecnico di Torino, Dipartimento di ingegneria biomedica, Corso Duca degli Abruzzi, 24, 10129 Torino TO
Coordinatori dello Studio:	Prof. Jacopo Secco, Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Torino (PI non clinico) Dott. Dante Dallari, Direttore della SC Chirurgia Ortopedica Ricostruttiva Tecniche Innovative-BTM, IRCCS Istituto Ortopedico Rizzoli, Bologna
FIRMA	Firma del Promotore Prof. Jacopo Secco _____ Firma del Co-PI Dr. Dante Dallari

Razionale:

Nonostante i progressi tecnologici, la diagnosi e la gestione delle fratture continuano a basarsi in larga misura sull'esperienza clinica del medico, il quale integra le informazioni radiologiche, tomografiche e anamnestiche per determinare il trattamento più appropriato. Tuttavia, con l'evoluzione dell'intelligenza artificiale (IA) e dei modelli predittivi, il supporto automatizzato alla diagnosi sta assumendo un ruolo sempre più rilevante, contribuendo a migliorare la rapidità e l'accuratezza nell'identificazione, classificazione e valutazione delle fratture.

L'impiego di reti neurali consente un'analisi rapida e automatica fondamentale soprattutto nei contesti di emergenza quali pronto soccorso o in strutture sanitarie con carenza di specialisti. Inoltre, questi sistemi permettono non solo di alleggerire il lavoro dei radiologi ma anche di ridurre gli errori legati al fattore umano, facilitando l'individuazione di fratture di difficile rilevazione e lasciando che gli specialisti si concentrino su casi più complessi. Questi strumenti sono di enorme aiuto, come nel caso delle fratture femorali, incluse anche quelle associate alla presenza di impianto protesico. Tali eventi traumatici comportano potenziali conseguenze sia funzionali che sistemiche. L'interruzione della continuità dell'osso femorale può comportare dolore intenso, perdita della capacità di deambulazione e, nei casi più gravi, compromettere la vascolarizzazione e l'innervazione locale.

Il presente studio si propone di sviluppare e addestrare una rete neurale con l'obiettivo di analizzare immagini retrospettive di Tomografia Computerizzata (TC) e Raggi X (RX), al fine di individuare e valutare con precisione le fratture del femore. L'analisi condotta prevede la

suddivisione delle fratture femorali in due principali categorie, basata sull'età del paziente e sul contesto clinico di insorgenza.

La prima categoria, che comprende pazienti tra i 18 e i 60 anni in assenza di impianti protesici, riguarda i traumi ad alta energia.

La seconda categoria riguarda, invece, pazienti di età superiore ai 60 anni dotati di impianto protesico, fratture che coinvolgono il femore in corrispondenza o in prossimità dell'impianto.

Questa distinzione è cruciale per classificare correttamente e ottimizzare la gestione terapeutica. Infatti, le fratture femorali native (ovvero quelle fratture che non sono alterate da interventi chirurgici, protesi, o trattamenti invasivi) e quelle associate a impianti presentano differenze dal punto di vista biomeccanico e chirurgico. In particolare, la necessità di adottare delle scale ortopediche specifiche è data dalla complessità della frattura, della stabilità dell'impianto, della qualità dell'osso residuo e della tipologia della protesi addoprata.

Il sistema proposto sarà in grado di estrarre e interpretare parametri fondamentali, tra cui le dimensioni ossee e le discontinuità della struttura corticale.

Servendosi di tali dati, la rete neurale procederà alla classificazione delle fratture secondo il sistema Arbeitsgemeinschaft für Osteosynthesefragen (AO), riconosciuto come standard di riferimento per le fratture in ortopedia.Tuttavia, in presenza di una protesi, si rende indispensabile l'impiego della classificazione di Vancouver.

L'impiego di algoritmi IA garantisce una diagnosi più standardizzata e uniforme. A differenza di altre soluzioni attualmente disponibili, il modello sviluppato in questo studio non si limiterà all'analisi delle sole immagini TC e RX, ma integrerà anche dati clinici rilevanti, come la storia anamnestica del paziente, comportandosi come strumento accurato e

oggettivo a supporto del medico nei processi di diagnostica e definizione del piano terapeutico.

Disegno dello studio:

Lo studio in oggetto adotta un disegno osservazionale retrospettivo, basandosi sull'analisi di immagini di TC e RX del femore acquisite in precedenza per finalità cliniche. Per garantire il rispetto delle normative etiche e la tutela della privacy, tutte le immagini utilizzate saranno preventivamente anonimizzate, eliminando qualsiasi informazione che possa ricondurre all'identità dei pazienti ai sensi del regolamento GDPR - UE 2016/679. Lo studio consentirà di sfruttare dati preesistenti per individuare correlazioni, riconoscere pattern diagnostici e testare nuovi algoritmi di analisi automatizzata.

Le immagini TC e RX considerate saranno classificate in modo da rispecchiare al meglio la frattura sotto esame attraverso un algoritmo di machine learning.

Per la classificazione del primo gruppo indicato come giovane, si utilizzerà il sistema AO, uno standard internazionale che permette di descrivere le fratture in modo univoco e universale, utile per la comunicazione tra specialisti, la pianificazione del trattamento e la valutazione della prognosi.

Il metodo di classificazione AO si basa su quattro elementi fondamentali: l'identificazione del segmento osseo, la localizzazione della frattura, la scala di gravità e un parametro aggiuntivo per affinare ulteriormente la classificazione. Ogni segmento scheletrico è rappresentato da un numero che lo identifica in modo univoco; nel caso di studio descritto, il femore ha il numero identificativo 3.

Dopo aver determinato il segmento osseo interessato, si procede con la localizzazione della frattura, secondo il seguente codice numerico:

1. Prossimale
2. Intermedia
3. Distale

Successivamente, il tipo, la morfologia e la gravità della frattura vengono classificati secondo un codice alfanumerico come segue:

Segmento prossimale del femore (Codice 31)

31A – Regione trocanterica (fratture extra-articolari):

- **31A1:** Frattura pertrocanterica semplice.
- **31A2:** Frattura pertrocanterica multiframmentaria con parete laterale compromessa (spessore ≤ 20,5 mm).
- **31A3:** Frattura intertrocanterica con obliquità inversa.

31B – Collo femorale (fratture extra-articolari):

- **31B1:** Frattura subcapitale.
- **31B2:** Frattura transcervicale.
- **31B3:** Frattura basicervicale

31C – Testa femorale (fratture articolari):

- **31C1:** Frattura a spacco.
- **31C2:** Frattura da depressione.

Segmento diafisario del femore (Codice 32)

32A – Fratture semplici:

- **32A1:** Frattura spirale.
- **32A2:** Frattura obliqua (angolo ≥ 30°).
- **32A3:** Frattura trasversale (angolo < 30°).

32B – Fratture a cuneo:

- **32B2:** Frattura con cuneo intatto.
- **32B3:** Frattura con cuneo frammentario.

32C – Fratture multiframmentarie:

- **32C2:** Frattura segmentaria con segmento intatto.
- **32C3:** Frattura segmentaria con segmento frammentario.

La classificazione AO per le fratture diafisarie del femore (codice 32) include qualificazioni aggiuntive per descrivere con maggiore precisione la localizzazione e la morfologia della frattura.

Per le fratture 32A e 32B, le qualificazioni sono:

- **a**: Terzo prossimale della diafisi.
- **b**: Terzo medio della diafisi.
- **c**: Terzo distale della diafisi.

Queste lettere indicano la posizione longitudinale della frattura lungo la diafisi femorale.

Per le fratture 32C, le qualificazioni sono:

- **i**: Frattura che coinvolge la giunzione tra la diafisi e la metafisi prossimale.
- **j**: Frattura confinata alla diafisi pura, senza estensione alle regioni metafisarie.

Segmento distale del femore (Codice 33)

33A – Fratture extra-articolari:

- **33A1:** Frattura da avulsione.
- **33A2:** Frattura semplice.
- **33A3:** Frattura a cuneo o multiframmentaria

33B – Fratture parzialmente articolari:

- **33B1:** Frattura del condilo laterale in piano sagittale.
- **33B2:** Frattura del condilo mediale in piano sagittale.
- **33B3:** Frattura in piano frontale/coronale.

33C – Fratture completamente articolari:

- **33C1:** Frattura articolare semplice con metafisi semplice.
- **33C2:** Frattura articolare semplice con metafisi a cuneo o multiframmentaria.
- **33C3:** Frattura articolare multiframmentaria con metafisi semplice, a cuneo o multiframmentaria.

Per il secondo gruppo di pazienti, ovvero coloro che presentano una frattura femorale in presenza di un impianto protesico d'anca, si adotta il sistema di classificazione Vancouver, il quale verrà descritto nel dettaglio. Tale classificazione pone particolare attenzione alla stabilità dello stelo protesico e alla quantità e qualità del tessuto osseo residuo

(bone stock), elementi fondamentali per la pianificazione del trattamento chirurgico e per la prognosi.

A. Fratture del trocantere:
   - A_G : che coinvolgono il grande trocantere
   - A_L : che coinvolgono il piccolo trocantere.

B. Fratture che coinvolgono il femore a livello dello stelo protesico:
   - B1 Protesi stabile
   - B2 Protesi mobilizzata, buon bone stock
   - B3 Protesi mobilizzata, bone stock carente.

C. Fratture distali alla protesi.

D. Fratture interprotesiche (fra protesi d'anca e protesi di ginocchio)

L'algoritmo di machine learning elaborerà i parametri estratti dalle immagini con l'intento di assegnare automaticamente a ciascuna frattura la sua classificazione.

Il processo di sviluppo e validazione del classificatore si articola in due fasi (addestramento e testing) andando a suddividere il dataset di immagini in tre set differenti: Training Set, Validation Set e Test Set.

Durante la fase di addestramento, il modello analizza il training set e il validation set, contenenti le immagini TC e RX del femore lesionato associate ciascuna alla corretta classificazione secondo il sistema AO oppure secondo Vancouver, fornita dal clinico di riferimento. Attraverso un processo iterativo, la rete neurale apprende a riconoscere le caratteristiche distintive delle fratture, ottimizzando progressivamente i propri parametri. Confrontando le previsioni con le classificazioni corrette, l'algoritmo riduce gli errori mediante tecniche di ottimizzazione, affinando così la sua capacità di identificare e classificare accuratamente le fratture.

Inoltre, al fine di creare un training set bilanciato, sarà necessario avere un numero di immagini omogeneo per ciascuna classe della scala di A0 e Vancouver da riconoscere. Questo contribuirà a ridurre il sovradattamento (overfitting), ovvero la capacità del modello di adattarsi troppo ai dati di addestramento, ottenendo l'incapacità di riconoscere invece nuove immagini.Per poter dimostrare le capacità effettive del modello si utilizza un ulteriore set di immagini, ovvero il test set. In questa fase, successiva all'addestramento, il modello osserverà delle immagini mai viste prima e dimostrerà la sua capacità di classificare nove fratture. Per capire quanto il classificatore sia efficace, vengono utilizzate diverse metriche di valutazione:

L'accuratezza misura quante fratture sono state classificate correttamente rispetto al totale delle immagini analizzate.
La precisione indica la percentuale di previsioni positive corrette rispetto a tutte quelle che il modello ha classificato come positive.
Il recall (o sensibilità) Rappresenta la capacità del modello di individuare tutti i campioni realmente positivi. È particolarmente importante nei casi in cui si vuole ridurre il numero di falsi negativi.
L'F1-score bilancia precisione e recall, fornendo un quadro più completo delle prestazioni del modello.
La matrice di confusione è uno strumento utile per valutare le prestazioni di un classificatore. Mette in luce quali tipi di frattura vengono spesso confusi tra loro.
Infine, per garantire che il modello sia realmente utile in ambito clinico, si effettuano test su dati provenienti da fonti esterne, simulando scenari reali. Se il modello mantiene buone prestazioni anche su questi nuovi dati, significa che ha una buona capacità di generalizzazione e può essere utilizzato in contesti pratici per supportare la diagnosi medica delle fratture femorali.

L'analisi basata su reti neurali si concentrerà in secondo luogo sull'estrazione di un'ampia gamma di parametri, suddivisi in due categorie principali: caratteristiche strutturali dell'osso e dimensioni anatomiche del femore.

1. Caratteristiche strutturali dell'osso:
    a. Discontinuità ossee: Visualizzazione di interruzioni nella continuità dell'osso, indicative di una possibile frattura.
    b. Linee di frattura: rilevamento di linee sottili o crepe che attraversano la struttura ossea.
    c. Deformazioni ossee: osservazione di modificazioni della forma dell'osso che possono suggerire la presenza di una frattura.
    d. Densità ossea: valutazione delle variazioni nella densità, utili a individuare zone di frattura o altre patologie.
    e. Posizione e orientamento: analisi della disposizione spaziale e dell'orientamento dei frammenti ossei per stabilire il tipo e la gravità della lesione.
    f. Tessuti circostanti: esame dei tessuti molli adiacenti per rilevare eventuali danni associati alla frattura.
2. Parametri dimensionali del femore
    a. Lunghezza del femore: è la distanza tra la testa femorale e i condili distali, misurata nella vista sagittale. Serve per avere una valutazione accurata della morfologia dell'osso.
    b. Diametro diafisario: indica la larghezza trasversale della diafisi femorale. Serve per poter scegliere placche e viti adeguate, prendendo in esame sia la sezione più ampia che quella più stretta.
    c. Curvatura anteriore del femore: corrisponde al raggio di curvatura della superficie anteriore della diafisi.

Si ottiene a partire da punti di riferimento utilizzati con un algoritmo di fitting. Questo permette di verificare la compatibilità con placche.

d. Angolo cervico-diafisario (CD angle): normalmente compreso tra 120° e 135°, è fondamentale per la scelta delle placche prossimali.Si calcola considerando i due assi fondamnetali del femore ovvero l'asse del collo femorale e quello della diafisi presi frontalmente.
e. Angolo condilo-diafisario: di solito tra 80° e 85°, è rilevante per la selezione delle placche distali. L'analisi si basa sul tracciamento dell'asse diafisario e di una linea lungo la superficie articolare del femore distale.
f. Antiversione femorale: misura la rotazione del collo femorale rispetto all'asse della diafisi. È un parametro cruciale per evitare disallineamenti dopo l'intervento. Si determina tracciando una linea tangente al profilo posteriore dei condili femorali (riferimento per l'asse distale) e una seconda linea che collega il centro della testa femorale con il centro della base del collo. L'angolo tra queste due linee rappresenta il valore dell'antiversione.
g. Contorni della frattura: La rete è capace di visualizzare ed estrarre i contorni e le linee della frattura, che influiscono sulla differenziazione tra fratture lineari da quelle complesse
h. Dislocazione dei frammenti: La rete esamina la quantità di spostamento o angolazione dei frammenti ossei, che è cruciale per determinare la gravità della frattura.
i. Comportamento angolare: La rete esamina la rotazione e l'inclinazione dei frammenti ossei.

Una volta ottenuta la classificazione della frattura ed estratti i parametri strutturali dell'osso, è possibile determinare il tipo di placca più idoneo per il trattamento. Attualmente, questa scelta viene effettuata dal medico specialista, il quale identifica la classificazione della lesione e seleziona la placca più appropriata. Tuttavia, questo processo verrà integrato all'interno della rete neurale, che una volta allenata con i dati estratti sarà in grado di riprodurre i criteri decisionali del medico.

Popolazione

Lo studio prevede l'analisi di immagini retrospettive di Tomografia Computerizzata (TC) e Radiografie (RX) del femore, con l'obiettivo di raccogliere un volume di dati quanto più esteso e rappresentativo possibile. Il dataset così ottenuto sarà utilizzato per l'addestramento di reti neurali, al fine di migliorarne progressivamente l'accuratezza nelle attività di rilevamento e classificazione delle fratture.

Le immagini saranno acquisite presso l'Unità Operativa di Radiologia dell'Ospedale Maggiore di Bologna e IRCCS Istituto Ortopedico Rizzoli di Bologna. Esse dovranno rispettare le stesse caratteristiche previste dal protocollo di studio, essere opportunamente anonimizzate e conformarsi ai requisiti che verranno dettagliati successivamente.

Le immagini che verranno classificate secondo il sistema AO dovranno raffigurare esclusivamente femori fratturati privi di mezzi di sintesi, come viti, chiodi, placche o fissatori, in modo da garantire un dataset omogeneo e focalizzato sulla struttura dell' osso . Per il sistema Vancouver, invece, le immagini dovranno includere chiaramente sia la frattura che l'intera protesi, per un'adeguata valutazione della frattura periprotesica.

Indipendentemente dalla classificazione di riferimento, tutte le immagini dovranno essere preventivamente anonimizzate e suddivise in base a due criteri: sesso biologico (maschile/femminile) e fascia d'età, distinguendo tra soggetti giovani (18 < età < 60 anni) e adulti (età ≥ 60

anni). Questa distinzione si rende necessaria in quanto l'impianto di una protesi d'anca risponde a indicazioni cliniche differenti in funzione dell'età del paziente.

Negli anziani, infatti, si osserva una maggiore incidenza di patologie articolari come coxartrosi, osteonecrosi e fratture del femore prossimale. In questa fascia di età, il deterioramento della cartilagine articolare e la conseguente compromissione funzionale rendono spesso inefficaci i trattamenti conservativi, orientando l'approccio verso soluzioni definitive come la protesi d'anca. Un altro fattore determinante è la ridotta aspettativa di vita e la minore richiesta funzionale: negli anziani l'obiettivo principale è il miglioramento della qualità della vita, e la durata media della protesi (15–20 anni) è solitamente sufficiente. Al contrario, nei pazienti giovani, si preferiscono tecniche conservative che ritardino l'impianto protesico, come osteotomie, infiltrazioni o terapie rigenerative, al fine di preservare il tessuto osseo e articolare. Inoltre, i giovani, essendo più attivi, esercitano un maggiore stress meccanico sulla protesi, aumentando il rischio di usura precoce e di interventi di revisione complessi.

Sarà altresì fondamentale garantire una distribuzione equilibrata delle immagini tra le diverse classi, in modo da ottenere un dataset adeguatamente rappresentativo per ciascuna categoria e idoneo all'addestramento della rete neurale. Per la classificazione AO è prevista la raccolta di un minimo di 400 immagini (almeno 200 immagini per RX e 200 immagini per TC)per ciascuna delle quattro classi di popolazione individuate: maschi di età compresa tra 18 e 60 anni, maschi di età superiore ai 60 anni, femmine di età compresa tra 18 e 60 anni e femmine di età superiore ai 60 anni.

Il totale previsto per ogni centro è quindi di 1.600 immagini.

Per la classificazione secondo il sistema Vancouver, relativa a pazienti di età superiore ai 60 anni, è richiesta la raccolta di almeno 200 immagini

RX e 200 immagini TC per ciascun sesso. Ciascuna delle due categorie – maschi e femmine – prevede quindi 400 immagini complessive (200 RX + 200 TC), per un totale di 800 immagini per centro.. Qualora non fosse possibile raggiungere il numero minimo richiesto, si provvederà a integrare il dataset mediante l'utilizzo di immagini provenienti da database pubblici, al fine di garantire una rappresentatività adeguata delle diverse tipologie di frattura.

Uno dei principali database pubblici per le radiografie muscoloscheletriche è il MURA (Musculoskeletal Radiographs), che contiene fratture ossee in diverse regioni del corpo. Sebbene il dataset non sia focalizzato esclusivamente sul femore, alcune immagini potrebbero riguardare fratture del femore prossimale o altre anomalie ossee correlate. [1],[4]

Altri Database pubblici utili per la ricerca nel campo della rilevazione automatica di fratture ossee, tra cui Open-i, DeepLesion.

Open-i è un motore di ricerca open access che fornisce accesso a una vasta raccolta di immagini mediche, tra cui radiografie (RX), TC, e risonanze magnetiche (RM). Il database contiene immagini relative a patologie muscoloscheletriche, neurologiche e cardiovascolari, incluse fratture ossee (compreso il femore) e altre lesioni. [2]

DeepLesion è un dataset che include migliaia di immagini di lesioni ossee, tra cui fratture del femore, colonna vertebrale, bacino e altre aree corporee, raccolte da scansioni TC. [3]

Oltre a queste informazioni, sarà essenziale disporre dell'anamnesi patologica remota del paziente, al fine di valutare eventuali condizioni pregresse che possano influenzare lo stato dell'osso. Infine, si richiederà una descrizione anatomica dettagliata del femore analizzato, per garantire l'accuratezza dei dati raccolti.

Centri arruolanti:

IRCCS Istituto Ortopedico Rizzoli, Bologna, via Giulio Cesare Pupilli 1 – 40136

Ospedale Maggiore Carlo Alberto Pizzardi, Bologna (BO), Largo Bartolo Nigrisoli 2 - 40133

Strumentazione:

- Dataset di immagini di RX: un database di immagini anonimizzate di fratture femorali etichettate secondo la classificazione AO o Vancouver.
- Dataset di immagini TC: un database di immagini anonimizzate di fratture femorali etichettate secondo la classificazione AO o Vancouver.
- Database pubblici di immagini mediche: risorse utilizzate solo in caso di insufficienza di dati raccolti per garantire un training adeguato del modello. I database che verranno utilizzati sono i seguenti: MURA (Musculoskeletal Radiographs), Open-i (Open Access Biomedical Image Search Engine), DeepLesion
- Reti neurali convoluzionali (CNN) per il riconoscimento dei pattern ossei e la classificazione delle fratture.
- Modelli di machine learning.
- Catalogo di placche femorali, impiegato per la selezione di soluzioni più adatte considerando le caratteristiche della frattura e i parametri anatomici del paziente.
- Quadro clinico del paziente: informazioni essenziali, tra cui età, condizioni ossee preesistenti e comorbidità,
- Metodi di elaborazione di immagini per il pre-processing del dataset.

Criteri di inclusione:

I pazienti e le relative immagini devono soddisfare tutti i seguenti criteri di inclusione nello studio:

1. Anonimizzazione completa per il rispetto delle normative etiche e della privacy.
2. La frattura femorale deve essere stata confermata classificabile secondo il sistema AO o Vancouver.
3. Disponibilità di immagini TC e RX del femore con qualità diagnostica adeguata.

4. Disponibilità dei dati clinici del paziante che potrebbero essere rilevanti (storia clinica, trattamento ricevuto).
5. Acquisizione delle immagini secondo protocolli standardizzati, al fine di garantire uniformità nei dati.
6. Risoluzione sufficiente per un'analisi accurata delle caratteristiche ossee e della frattura.
7. Pazienti maschi e femmine di età superiore a 18 anni, visitati/con accesso e immagini radiologiche presenti in cartella presso i centri IOR e AUSL dal 2014 al 2024
8. Formato DICOM (Digital Imaging and Communications in Medicine) per assicurare la compatibilità con i software di analisi automatizzata. Formati alternativi come PNG o JPEG possono essere comunque ottimali.
9. Consenso informato per l'utilizzo dei dati a fini di ricerca, nel rispetto delle normative etiche e sulla privacy.

Criteri di esclusione:

I pazienti e le immagini che presentano almeno uno dei seguenti criteri di esclusione non saranno ammessi allo studio:

1. Immagini TC e RX di qualità insufficiente o dataset incompleti, che potrebbero compromettere l'accuratezza dell'analisi.
2. Presenza di malformazioni congenite o patologie ossee preesistenti, quali metastasi o osteoporosi avanzata, che potrebbero alterare i parametri strutturali dell'osso.
3. Fratture che non sono state classificate secondo la scala di interesse al caso in studio ovvero AO e Vancouver.
4. Assenza del consenso informato per l'utilizzo dei dati a fini di ricerca, nel rispetto delle normative etiche e sulla privacy.

Criteri di sospensione: Non previsti

Endpoints:

Endpoint Primario:

- Valutazione del legame tra le caratteristiche estratte dalle immagini TC e RX e la classificazione della frattura secondo i sistemi AO e Vancouver.

- Valutazione della correlazione tra la classificazione AO o Vancouver e la scelta della placca più appropriata. Sarà analizzata la relazione tra la tipologia di frattura e la selezione della placca effettuata dall'algoritmo, confrontandola con la valutazione clinica del medico specialista.

- Valutazione degli angoli fondamentali del femore: analisi delle misurazioni angolari rilevanti, come l'angolo cervico-diafisario e l'angolo di inclinazione femorale, per determinare eventuali variazioni morfologiche associate alle diverse classi di fratture.

Endpoint Secondari:

- Valutazione delle prestazioni su sottogruppi specifici: come pazienti con condizioni ossee preesistenti, per individuare eventuali limiti o necessità di adattamenti del modello.

Metodi:

Nella fase iniziale del progetto, verrà effettuata una valutazione dell'integrazione di database pubblici con l'obiettivo di aumentare il numero di immagini TC o RX nelle classi di fratture che presentano dati insufficienti. L'integrazione di dataset esterni sarà presa in considerazione solo se i dati disponibili non risultano adeguati a garantire un'analisi affidabile e rappresentativa. I criteri per l'integrazione dei database pubblici includono la varietà delle classi, per cui i dataset esterni devono coprire un'ampia gamma di fratture, con particolare attenzione alle categorie meno rappresentate nel dataset originale, e la completezza dei metadati, in quanto i metadati associati alle immagini devono essere completi e compatibili con quelli necessari per l'addestramento corretto del modello. Solo se i dati raccolti soddisfano questi criteri, l'integrazione avverrà con l'obiettivo di migliorare la rappresentatività delle classi di fratture. In caso contrario, qualora i dati non siano adeguati o sufficienti, non verrà effettuata alcuna integrazione, e si procederà utilizzando esclusivamente il dataset originale per l'analisi.

Le immagini, prima di essere elaborate dalla rete neurale, necessitano di un processo di pre-elaborazione fondamentale per standardizzarle e ottimizzare le prestazioni del modello. Questo passaggio riduce il rumore e le incoerenze, adattando le immagini ai requisiti del modello, al fine di garantire risultati più precisi e affidabili nell'analisi e nella classificazione delle fratture.

Il processo di pre-processing ha inizio con la rimozione del rumore, essenziale per migliorare la qualità dell'immagine e ottimizzare l'analisi successiva.

Esistono diversi tipi di rumore che affliggono le immagini TC e RX. Nella letteratura è stato evidenziato che le immagini RX sono particolarmente influenzate dal rumore “salt and pepper”, che si manifesta come punti casuali chiari e scuri distribuiti in tutta l'immagine. Questo tipo di rumore è generalmente causato da malfunzionamenti durante la fase di acquisizione o trasmissione delle immagini. La presenza di questo disturbo può compromettere la qualità visiva dell'immagine, rendendo necessaria l'applicazione di tecniche di riduzione del rumore. Una delle soluzioni più efficaci per mitigare il rumore “salt and pepper” senza compromettere i dettagli dell'immagine consiste nell'utilizzo di un filtro gaussiano. Inoltre, qualora un pixel presenti una deviazione significativa rispetto ai suoi vicini, il suo valore può essere sostituito con il valore mediano dei pixel circostanti, contribuendo così a mantenere l'integrità dell'immagine pur riducendo il rumore.

Il rumore “salt and pepper”può apparire nelle immagini TC, ma non è il più comune. I tipi di rumore più frequenti sono il rumore di Poisson, causato dalle fluttuazioni statistiche nei fotoni rilevati dai sensori, e il rumore gaussiano, dovuto a fluttuazioni elettroniche nei sensori.

Per ridurre il rumore di Poisson, si utilizzano tecniche di denoising come la trasformata wavelet o i denoising autoencoders, che mantengono i dettagli dell'immagine. Il rumore gaussiano, invece, può essere eliminato tramite il filtro gaussiano o il filtro mediano, che preservano le informazioni strutturali evitando di compromettere l'immagine.

Oltre all'eliminazione del rumore, diverse tecniche di pre-elaborazione possono migliorare la qualità delle immagini. In letteratura, l'Equalizzazione dell'Istogramma Adattativa (AHE) è utilizzata per aumentare il contrasto localmente, migliorando la visibilità dei dettagli senza alterare uniformemente l'intera immagine. Questo approccio è particolarmente utile nelle immagini mediche, dove il contrasto può variare a causa di fattori come l'illuminazione e la densità dei tessuti.

L'applicazione combinata di queste tecniche di pre-elaborazione consente di ottenere immagini più chiare e utili per l'analisi diagnostica delle immagini RX e TC.

Successivamente, viene applicata la normalizzazione, che trasforma i valori di intensità dei pixel in un intervallo standardizzato, assicurando uniformità tra le diverse immagini. Un altro passaggio essenziale è il ridimensionamento, che consente di adattare le immagini alle dimensioni richieste dalla rete neurale convoluzionale (CNN), garantendo così una corretta elaborazione.

Infine, viene individuata la Regione di Interesse (ROI - Region of Interest), ovvero l'area specifica dell'immagine contenente la frattura. Questo permette di escludere le parti non rilevanti e di ottimizzare l'analisi, concentrando il modello solo sulle informazioni utili per la classificazione della frattura.

Una volta completata la fase di pre-processing, il dataset viene suddiviso in tre insiemi distinti: il training set, utilizzato per addestrare la rete neurale e consentirle di apprendere le caratteristiche delle fratture; il validation set, impiegato per ottimizzare i parametri del modello e prevenire il rischio di overfitting; e infine il test set, che permette di valutare in modo oggettivo le prestazioni del modello su dati mai presentati alla rete, verificandone l'accuratezza e la capacità di generalizzazione.

Una volta effettuato il pre-processing e suddiviso il dataset nei rispettivi set, le immagini vengono poi fornite come input a reti neurali per l'identificazione delle fratture. In letteratura, diversi tipi di reti hanno dimostrato di essere particolarmente efficaci in questo compito, tra cui U-Net, YOLO e Mask R-CNN.

La U-Net, progettata per la segmentazione delle immagini mediche, è in grado di identificare e segmentare con precisione le fratture grazie alla sua architettura a forma di U.[5]

YOLO (You Only Look Once) è utilizzata per il rilevamento veloce delle fratture, individuando in tempo reale i confini delle fratture nelle immagini.[6]

Mask R-CNN estende Faster R-CNN, aggiungendo la capacità di segmentare le fratture e produrre una maschera dettagliata, permettendo una localizzazione precisa e una visualizzazione della forma della frattura.[7],[8], [9]

Queste reti offrono potenti strumenti per l'analisi automatica delle fratture, migliorando la diagnosi e l'efficienza nel trattamento.

Una volta individuata la frattura, il passo successivo consiste nella sua classificazione secondo le scale precedentemente descritte, ovvero AO e Vancouver.
L'approccio adottato prevede l'impiego di una rete neurale unica con task multipli, in cui le predizioni vengono concatenate per ottenere un'uscita strutturata. Questo metodo, noto come Multi-Task Learning (MTL), consente alla rete di analizzare simultaneamente diverse caratteristiche di un'immagine, ottimizzando l'efficienza computazionale e migliorando la capacità di generalizzazione attraverso la condivisione delle feature tra i diversi compiti di classificazione.
A tal fine, la rete neurale fornisce in uscita i parametri distinti, fondamentali per la classificazione della frattura. Per la classificazione con scala AO il primo parametro riguarda la struttura ossea coinvolta, che nel caso specifico è il femore. Il secondo parametro riguarda la regione del femore, classificata in tre possibili valori: 1, 2, 3, corrispondenti rispettivamente alle diverse sezioni del femore (prossimale, distale e diafisario). Il terzo parametro identifica il tipo di frattura, suddiviso in tre categorie: A, B e C, dove il tipo A rappresenta fratture semplici, il tipo B fratture a cuneo e il tipo C fratture segmentali. Infine, il quarto parametro riguarda la gravità della frattura, valutata in base al livello di compromissione strutturale.
La classificazione Vancouver si basa su tre criteri fondamentali: la localizzazione della frattura rispetto allo stelo protesico, la stabilità dell'impianto e la qualità del bone stock residuo. Il parametro di output è uno dei tre nelle principali categorie. Le categorie sono: fratture di tipo A, interessano il trocantere e si distinguono ulteriormente in AG, quando coinvolgono il grande trocantere, e AL, se colpiscono il piccolo trocantere. Fratture di tipo B le quali coinvolgono la zona dello stelo protesico e si suddividono in tre sottotipi: B1, in cui la protesi risulta stabile; B2, caratterizzata da una mobilizzazione dello stelo protesico ma con un bone stock ancora adeguato; e B3, dove la mobilizzazione dello stelo si accompagna a una significativa compromissione della qualità e quantità dell'osso femorale residuo. Infine, fratture di tipo C che si verificano distalmente all'impianto protesico, ovvero al di sotto dell'estremità inferiore dello stelo.

Il "ground truth" della classificazione viene definita da specialisti del settore medico, i quali etichettano manualmente le immagini, assegnando a ciascuna le informazioni corrette relative alla struttura ossea, alla localizzazione anatomica, alla tipologia e alla gravità della frattura. Tali dati di riferimento vengono impiegati per l'addestramento

della rete neurale, consentendo al modello di apprendere con precisione le caratteristiche distintive di ciascuna classe. Il confronto tra le predizioni del sistema e il ground truth permette di affinare progressivamente le prestazioni del modello, garantendo un'elevata affidabilità nella diagnosi automatizzata.

Una volta determinata la classificazione della frattura secondo le scale, è necessario definire la correlazione tra la tipologia di frattura e il tipo di placca più appropriata per il trattamento. Il modello di associazione tra frattura e placca deve essere validato confrontando le predizioni del sistema con le scelte effettuate dai medici specialisti, analogamente al ground truth utilizzata per la classificazione delle fratture.

Il medico specialista che fornirà le informazioni sulle fratture per la definizione del ground truth sarà lo stesso che indicherà il tipo di placca più adatto per ciascun caso, garantendo così coerenza nei dati di riferimento.

Per l'addestramento di un sistema di associazione, è possibile utilizzare modelli di apprendimento supervisionato, tra cui Support Vector Machines (SVM), Gradient Boosting e Random Forest, che consentono di elaborare le caratteristiche estratte dalle immagini e determinare la placca ottimale per ciascun caso clinico. Tali modelli saranno addestrati utilizzando la classificazione AO o Vancouver ottenuta nello step precedente e i parametri del femore estratti.

Ruolo del Promotore:

Il Promotore assume un ruolo centrale nel coordinamento e nella supervisione di tutte le fasi legate all'elaborazione delle immagini e alla gestione dei dati raccolti. In particolare, è responsabile di:

- Garantire la qualità e l'integrità dei dati raccolti, coordinando la raccolta, l'archiviazione e la verifica dei dataset, compresi quelli integrati da database pubblici, secondo criteri di completezza e rappresentatività
- Gestisce e supervisiona l'intera pipeline di elaborazione, dalla pre-elaborazione delle immagini (denoising, normalizzazione, ridimensionamento, selezione della ROI) attraverso l'addestramento e la validazione, fino alla fase di testing, garantendo coerenza, qualità e riproducibilità dei risultati.
- Coordinare con i medici ortopedici al fine di realizzare il "ground truth", ovvero le etichette delle immagini.
- È responsabile del rispetto delle normative relative alla privacy, sicurezza dei dati e etica nell'uso delle immagini mediche e delle informazioni personali durante tutto il ciclo di vita del progetto.
- È responsabile della scrittura e implementazione del codice sorgente su cui si basa l'addestramento della rete neurale, inclusi gli script e i modelli specifici utilizzati per il rilevamento e la classificazione delle fratture. Tutte le attività di training della rete neurale si basano su questi codici, che sono progettati e sviluppati direttamente dal Centro Promotore.

In sintesi, il Centro Promotore non solo supervisiona il progetto, ma è l'ingegnere responsabile della progettazione, implementazione e gestione tecnica del codice e della pipeline su cui si basa l'algoritmo di rete neurale, garantendo che il sistema operi secondo le specifiche cliniche e tecniche richieste.

Analisi statistica:

dimensioni del campione:

Per la classificazione AO è prevista la raccolta di un minimo di 200 immagini RX e 200 immagini TC per ciascuna delle seguenti quattro categorie:

- maschi di età superiore ai 60 anni
- maschi di età compresa tra 18 e 60 anni
- femmine di età superiore ai 60 anni
- femmine di età compresa tra 18 e 60 anni

Ogni categoria richiede quindi 400 immagini (200 RX + 200 TC), per un totale di 1.600 immagini per centro. Considerando i due centri coinvolti, il numero complessivo di immagini da acquisire sarà pari a 3.200.

Per la classificazione Vancouver è prevista la raccolta di un minimo di 200 immagini RX e 200 immagini TC per ciascun sesso, riferite a pazienti di età superiore ai 60 anni. Le categorie sono:

- maschi di età superiore ai 60 anni
- femmine di età superiore ai 60 anni

Ogni categoria richiede quindi 400 immagini (200 RX + 200 TC), per un totale di 800 immagini per centro. Per entrambi i centri complessivamente, il totale sarà di 1.600 immagini.

Totale (AO + Vancouver):

- per ogni centro: 2.400 immagini
- per entrambi i centri: 4.800 immagini

È fortemente consigliato che le immagini TC e RX provengano dallo stesso paziente e si riferiscano alla stessa lesione. Nel caso in cui non sia possibile rispettare tale criterio, saranno comunque accettate tutte le immagini TC disponibili, purché rientranti nell'arco temporale considerato dallo studio e riferite a fratture del femore.

Qualora il numero di immagini disponibili non fosse sufficiente, si procederà con l'integrazione del dataset tramite immagini provenienti da database pubblici (DeepLesion, Open-i, MURA).

Tipo di analisi:

Lo studio si concentrerà sull'analisi e la classificazione automatica delle fratture del femore tramite l'uso di immagini RX e TC, con l'obiettivo di sviluppare un sistema basato sull'intelligenza artificiale in modo da rendere più agevole la diagnosi e di conseguenza il trattamento della frattura.

L’analisi prevede l’utilizzo di reti neurali convoluzionali (CNN) per l’identificazione e la classificazione delle fratture in scala AO e Vancouver e modelli di machine learning per successiva selezione automatizzata del design della placca. Le fasi del progetto possono essere suddivise in:

- Pre-processing delle immagini: dove le immagini sono sottoposte a denoising, miglioramento del contrasto, normalizzazione e ridimensionamento.

- Classificazione: Le immagini vengono poi processate tramite reti neurali convoluzionali (CNN) come U-Net, YOLO, o Mask R-CNN, che sono in grado di identificare le fratture all'interno delle immagini e classificare la loro tipologia in base alla scala AO e Vancouver. La rete fornisce in uscita parametri relativi alla struttura ossea (femore), alla regione, al tipo di frattura e alla gravità della frattura, stabilità dell’impianto e qualità/quantità dell’osso residuo. Alcuni dei parametri relativi all'estrazione delle caratteristiche dalle immagini tramite reti neurali convoluzionali (CNN), che sono stati trovati in letteratura, includono: simmetria, contorni della frattura, dislocazione dei frammenti, tratti lineari e angoli.

- Estrazione dei parametri morfologici e strutturali del femore attraverso algoritmi.

- Assegnazione della placca: una volta identificati la classificazione e i parametri dell’osso, il sistema deve associare il tipo di frattura al trattamento più appropriato, in particolare il tipo di placca. Per fare ciò, vengono utilizzati algoritmi di apprendimento supervisionato, come Support Vector Machines (SVM), Gradient Boosting, e Random Forest.

L'obiettivo del progetto è sviluppare uno strumento che possa aiutare i medici nella diagnosi delle fratture, fornendo rapidità, precisione, e un sistema valido nella selezione del mezzo di sintesi attraverso la classificazione automatica delle fratture.

Durata prevista dello studio

Si ritiene che la durata complessiva dello studio possa essere di circa otto mesi, considerando che sarà necessario acquisire le immagini e procedere con la loro successiva analisi. Tale tempistica permetterà di garantire una raccolta accurata dei dati e una valutazione approfondita di ciascuna immagine, al fine di ottenere risultati significativi e affidabili.

PRIVACY

Tutti i pazienti inclusi nello studio saranno identificati con un codice numerico, in modo che i dati sensibili saranno resi pseudoanonimi e utilizzati nel rispetto della normativa in vigore in materia di privacy. I dati verranno conservati dallo sperimentatore per il tempo necessario alla produzione scientifica. Al fine di garantire la riservatezza dei dati delle sperimentazioni cliniche come disposto dalla normativa nazionale ed europea applicabile, i dati saranno accessibili solo al promotore dello studio e i suoi designati, per le procedure di monitoraggio/auditing, allo sperimentatore e i collaboratori, e al Comitato etico del centro in cui viene condotta la ricerca e alle autorità sanitarie preposte. Lo sperimentatore e l'Istituto consentiranno l'accesso ai dati e alla documentazione di origine per il monitoraggio, l'audit, la revisione del Comitato etico e le ispezioni dell'Autorità sanitaria, ma conservando riservatezza dei dati personali in accordo alla normativa vigente

Bibliografia:

[1] https://stanfordmlgroup.github.io/competitions/mura/
[2] https://openi.nlm.nih.gov/faq?download=true
[3] https://paperswithcode.com/dataset/bone-fracture-multi-region-x-ray-data

[4] Choudhary, R. R., Choudhary, S., & Meena, G. (2021). Abnormality Detection in Musculoskeletal Radiographs. *IOP Conference Series: Materials Science and Engineering, 1020*(1), 012009. https://doi.org/10.1088/1757-899X/1020/1/012009

[5] Cai, D., Zhou, Y., He, W., Yuan, J., Liu, C., Li, R., Wang, Y., & Xia, J. (2024). Automatic segmentation of knee CT images of tibial plateau fractures based on three-dimensional U-Net: Assisting junior physicians with Schatzker classification. *Journal of Medical Imaging and Health Informatics*, 14(2), 123-130. https://doi.org/10.1016/j.ejrad.2024.111605

[6] W. Wei, Y. Huang, J. Zheng, Y. Rao, Y. Wei, X. Tan, and H. OuYang, "YOLOv11-based multi-task learning for enhanced bone fracture detection and classification in X-ray images," Version of Record 24 January 2025. https://doi.org/10.1016/j.jrras.2025.101309

[7] Aydın, A., & Özcan, C. (2024). Upper and lower extremity bone segmentation with Mask R-CNN. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, 13*(1), 358-365. https://doi.org/10.17798/bitlisfen.1413650

[8] S. Chai, P. Gao, M. Li, and L. Jin, "Mask R-CNN for Predicting Rib Fractures on CT Images with Interpretability and ChatGPT-Based Structured Outcomes," Radiology Department, Huadong Hospital, Fudan University, Shanghai, PR China. https://doi.org/10.1016/j.eswa.2025.127047

[9] Paik, S., Park, J., Hong, J. Y., & Han, S. W. (Anno). *Deep learning application of vertebral compression fracture detection using Mask R-CNN*. DOI: 10.1038/s41598-024-67017-6

[10] M. S. Ahammed, A. Ahamed, M. H. Huda, M. R. Mridha, M. J. Chaudhary, and M. F. N. Nakib, "Bone Fracture Classification in X-ray Images: A Deep Learning Approach Leveraging Transfer Learning

[11] Mamun, S., Al Amin, M., Ali, A. J., & Li, J. (2024). Bone Fracture Detection from X-ray Images using a Convolutional Neural Network (CNN). *Preprints.org*. DOI:10.20944/preprints202410.1320.v1

Bibliography

[1] Ratti, C. (2023). Le fratture da fragilità. Azienda Sanitaria Universitaria Giuliano Isontina (ASUGI)

[2] Consiglio Superiore di Sanità. (2021). Patologia ortopedica nelle condizioni di fragilità. Ministero della Salute.

[3] International Osteoporosis Foundation (IOF) (2018). Ossa spezzate, vite spezzate: un piano d'azione per superare l'emergenza delle fratture da fragilità in Italia. Nyon (CH): IOF.

[4] Fu F., Liu B., Pu H., Wang Y., Zhang P., Wei S., Gu H., Zhang Q., Ye H. (2025). Global trends in the incidence and primary causes of femoral fractures, excluding femoral neck fractures: a global epidemiological study. *Risk Management and Healthcare Policy*, 18, 117–129. https://doi.org/10.2147/RMHP.S498918

[5] Sing, C.-W., Lin, T.-C., Bartholomew, S., Bell, J.S., Bennett, C., et al. (2023). Global Epidemiology of Hip Fractures: Secular Trends in Incidence Rate, Post-Fracture Treatment, and All-Cause Mortality. Journal of Bone and Mineral Research, 38(8), 1064–1075. https://doi.org/10.1002/jbmr.4821

[6] Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sánchez, C.I. Asurvey on deep learning in medical image analysis. Med. Image Anal. 2017, 42, 60–88.

[7] Boginskis, V.; Zadoroznijs, S.; Cernavska, I.; Beikmane, D.; Sauka, J. Artificial intelligence effectivity in fracture detection. Med. Perspect. 2023, 3, 68–78.

[8] Abdellatif, N., El-Rawy, A., Abdellatif, A. et al. Assessment of artificial intelligence-aided X-ray in diagnosis of bone fractures in emergency setting. Egypt J Radiol Nucl Med 56, 160 (2025). https://doi.org/10.1186/s43055-025-01580-4

[9] Società Italiana di Ortopedia e Traumatologia (SIOT) & Società Italiana di Geriatria Ospedale e Territorio (SIGOT). (2021). Frattura del femore prossimale nell'anziano: Linee guida cliniche. SIGOT

[10] Katsoulis M, Benetou V, Karapetyan T et al. Excess mortality after hip fracture in elderly persons from Europe and the USA: the CHANCES project. J Intern Med. 2017 Mar;281(3):300-310.

[11] Guzon-Illescas O, Perez Fernandez E2, Crespí Villarias N et al. Mortality after osteoporotic hip fracture: incidence, trends, and associated factors. J Orthop Surg Res. 2019 Jul 4;14(1):203.

[12] Yang, A. Y., Cheng, L., & Shimaponda-Nawa, M. (2019). Long-Bone Fracture Detection using Artificial Neural Networks based on Line Features of X-ray Images. In 2019 Conference on Electrical and Information Engineering.

[13] Lim, S. E., Xing, Y., Chen, Y., Leow, W. K., Howe, T. S., & Png, M. A. (2004). Detection of femur and radius fractures in X-ray images. In Medical Image Computing and Computer-Assisted Intervention. Springer

[14] Hržić, F., Štajduhar, I., Lerga, J., Tschauner, S., & Sorantin, E. (2019). Local-Entropy Based Approach for X-Ray Image Segmentation and Fracture Detection. Entropy, 21(4), 338.

[15] Wei, W., Huang, Y., Zheng, J., Rao, Y., Wei, Y., Tan, X., & OuYang, H. (2025). YOLOv11-based multi-task learning for enhanced bone fracture detection and classification in X-ray images. Journal of Radiation Research and Applied Sciences, 18(1), 101309.

[16] Buckley, R., Moran, C., & Apivatthakakul, T. (2020). Principi AO per il trattamento delle fratture (Opera in due volumi). Edra.

[17] Betts, J. G., Desaix, P., Johnson, E., Johnson, J. E., Korol, O., Kruse, D., ... & Wise, J. A. (2023). *Anatomy and Physiology 2e* (Section 6.2: Bone Classification). OpenStax, Rice University. https://openstax.org/books/anatomy-and-physiology-2e/pages/6-2-bone-classification

[18] Anatomia Umana – Basi Anatomiche per le Neuroscienze tratto dal Prometheus Vol. II di M. Schünke, E. Schulte e U. SchumacherI, I ed. 2021, EDISES.

[19] Verne', E. (2023). Materiali per la bioingegneria [Lezione struttura e proprietà dell'osso]. Politecnico di Torino, Torino.

[20] Bignardi, C. (2024). Bioingegneria Meccanica [Lezione Osso e anatomia]. Politecnico di Torino, Torino.

[21] Cowin, S. C. (1986). Wolff's Law of Trabecular Architecture at Remodeling Equilibrium. Journal of Biomechanical Engineering, 108(1), 83–88. https://doi.org/10.1115/1.3138584

[22] Wolff, J. (1870). Über die innere Architektur der Knochen und ihre Bedeutung für die Frage vom Knochenwachstum. Archiv für pathologische Anatomie und Physiologie und für klinische Medizin (Virchow's Archiv), Vol. 50, pp. 389–453.https://doi.org/10.1007/BF01944490

[23] Roux, W. (1885). Gesetz der Transformation der Knochen. Leipzig: Wilhelm Engelmann.

[24] Pauwels, F. (1976). Biomechanics of the Locomotor Apparatus: Contributions on the Functional Anatomy of the Locomotor Apparatus. Berlin: Springer-Verlag.

[25] Martin, R. B. (1984). Porosity and specific surface of bone. Critical Reviews in Biomedical Engineering, 10(3), 179–222. PMID: 6368124.

[26] Robling, A. G., & Turner, C. H. (2009). Mechanical signaling for bone modeling and remodeling. Critical Reviews in Eukaryotic Gene Expression, 19(4), 319–338.

[27] Razi, H., Birkhold, A. I., Weinkamer, R., Duda, G. N., Willie, B. M., & Checa, S. (2015). Aging Leads to a Dysregulation in Mechanically Driven Bone Formation and Resorption. Journal of Bone and Mineral Research, 30(10), 1864–1873.

[28] Società Italiana dell'Osteoporosi, del Metabolismo Minerale e delle Malattie dello Scheletro (SIOMMMS). (n.d.). *Le diverse età dell'osso*. https://www.siommms.it/area-pazienti/le-diverse-eta-dellosso/

[29] Hoffler et al. (2000) – Age, Gender, and Bone Lamellae Elastic Moduli

[30] Reilly, D. T., & Burstein, A. H. (1975). The elastic and ultimate properties of compact bone tissue. Journal of Biomechanics, 8(6), 393–405. https://doi.org/10.1016/0021-9290(75)90075-5

[31] https://kdmfab.com/it/curva-sforzo-deformazione-2/

[32] A. Seireg, W. Kempke, Behavior of in vivo bone under cyclic loading, Journal of Biomechanics, Volume 2, Issue 4, 1969, Pages 455-461, ISSN 0021-9290, https://doi.org/10.1016/0021-9290(69)90020-7.

[33] Abdel-Wahab, A. A., Alam, K., & Silberschmidt, V. V. (2011). Analysis of anisotropic viscoelastoplastic properties of cortical bone tissues. Journal of the Mechanical Behavior of Biomedical Materials, 4(5), 807–820.

[34] Burstein, A. H., Reilly, D. T., & Martens, M. (1976). Aging of bone tissue: mechanical properties. Journal of Bone and Joint Surgery – American Volume, 58(1), 82–86. https://pubmed.ncbi.nlm.nih.gov/1249116/

[35] Turner et al. (1999) Turner, C. H., Rho, J. Y., Takano, Y., Tsui, T. Y., & Pharr, G. M. (1999). The elastic properties of trabecular and cortical bone tissues are similar: results from two microscopic measurement techniques. Journal of Biomechanics, 32(4), 437–441. https://doi.org/10.1016/S0021-9290(98)00177-8

[36] Meyers et al. (2008) Meyers, M. A., Chen, P. Y., Lin, A. Y. M., & Seki, Y. (2008). Biological materials: Structure and mechanical properties. Progress in Materials Science, 53(1), 1–206. https://doi.org/10.1016/j.pmatsci.2007.05.002

[37] Katz, J. L., & Meunier, A. (1987). The elastic anisotropy of bone. Journal of Biomechanics, 20(11–12), 1063–1070. https://doi.org/10.1016/0021-9290(87)90024-8

[38] Currey, J. D. (2002). Bones: Structure and mechanics. Princeton University Press

[39] Bertozzo, S. (2023). Analisi del comportamento viscoelastico del tessuto osseo (Tesi di laurea triennale, Università degli Studi di Padova).

[40] Lakes, R. S. (1998). Viscoelastic solids. CRC Press.

[41] McElhaney, J.H. (1966). Dynamic Response of Bone and Muscle Tissue. Journal of Applied Physiology, 21, 1231–1236.

[42] Lindahl, O., & Lindgren, Å. G. H. (1968). Cortical Bone in Man: III. Variation of Compressive Strength with Age and Sex. Acta Orthopaedica Scandinavica, 39(1–3), 129–135. DOI: 10.3109/17453676808989447

[43] C. Verdoia, C. Corradini, and L. Pietrogrande, "Osteoporosi e fratture da fragilità," in Ortopedia, 2013

[44] MSD Manuale Professionale. (2025). Osteoporosi

[45] Curate et al. (2009) Curate, F., Piombino-Mascali, D., Tavares, A., & Cunha, E. (2009). Assottigliamento corticale del femore e fratture da fragilità ossea: uno studio della Collezione Scheletrica Identificata di Coimbra (Portogallo)

[46] Tarantino et al. (2014) Tarantino, U., Feola, M., Piccirilli, E., Marziali, D., & Rao, C. (2014). Fratture da osteoporosi: lo specchio di una fragilità sottostante. Focus, 40, 289–295

[47] Chang, A., Breeland, G., Black, A. C., & Hubbard, J. B. (2018). Anatomy, Bony Pelvis and Lower Limb: Femur.

[48] Wagner, F. V., Negrão, J. R., Campos, J., Ward, S. R., Haghighi, P., Trudell, D. J., & Resnick, D. (2012). Capsular ligaments of the hip: Anatomic, histologic, and positional study in cadaveric specimens with MR arthrography. Radiology, 263(1), 189–198.

[49] Fernandez, M. A., Griffin, X. L., & Costa, M. L. (2015). Management of hip fracture. British Medical Bulletin, 115(1), 165–172.

[50] van Arkel, R. J., Amis, A. A., & Jeffers, J. R. T. (2015). The envelope of passive motion allowed by the capsular ligaments of the hip. Journal of Biomechanics, 48(14), 3803–3809.

[51] Aiello, L., & Dean, C. (1990). An Introduction to Human Evolutionary Anatomy. Academic Press.

[52] Boese CK, Dargel J, Oppermann J, Eysel P, Scheyerer MJ, Bredow J, Lechler P. The femoral neck-shaft angle on plain radiographs: a systematic review. Skeletal Radiol. 2016 Jan;45(1):19-28

[53] Reynolds A. The fractured femur. Radiol Technol. 2013 Jan-Feb;84(3):273-91; quiz p.292-4

[54] Li M, Cole PA. Anatomical considerations in adult femoral neck fractures: how anatomy influences the treatment issues? Injury. 2015 Mar;46(3):453-8

[55] Carballido-Gamio J, Nicolella DP. Computational anatomy in the study of bone structure. Curr Osteoporos Rep. 2013 Sep;11(3):237-45.

[56] Crumbie, L., & Osika, A. (2023). Femur bone anatomy: Proximal, distal and shaft. Kenhub. Retrieved from https://www.kenhub.com/en/library/anatomy/femur

[57] Società Italiana di Endocrinologia. (2024). Focus sulle fratture da fragilità.

[58] Scottish Intercollegiate Guidelines Network (SIGN). (2002). Prevention and management of hip fracture in older people: A national clinical guideline.

[59] World Health Organization. (2024, September 25). Fragility fractures. https://www.who.int/news-room/fact-sheets/detail/fragility-fractures

[60] Crock HV. An atlas of the arterial supply of the head and neck of the femur in man. Clin Orthop Relat Res 1980;(152):17–27

[61] Sheehan, S. E., Shyu, J. Y., Weaver, M. J., Sodickson, A. D., & Khurana, B. (2015). Proximal femoral fractures: What the orthopedic surgeon wants to know. RadioGraphics, 35(5)

[62] Chung SM. The arterial supply of the developing proximal end of the human femur. J Bone Joint Surg Am 1976;58(7): 961–970

[63] Trueta J, Harrison MH. The normal vascular anatomy of the femoral head in adult man. J Bone Joint Surg Br 1953;35- B(3):442–461.

[64] Ehlinger M, Moser T, Adam P, et al. Early prediction of femoral head avascular necrosis following neck fracture. Orthop Traumatol Surg Res 2011;97(1):79–88

[65] Damany DS, Parker MJ, Chojnowski A. Complications after intracapsular hip fractures in young adults: a metaanalysis of 18 published studies involving 564 fractures. Injury 2005;36(1)

[66] Claffey TJ. Avascular necrosis of the femoral head: an anatomical study. J Bone Joint Surg Br 1960;42-B:802–809

[67] Meinberg, E. G., Agel, J., Roberts, C. S., Karam, M. D., & Kellam, J. F. (2018). Fracture and dislocation classification compendium – 2018. Journal of Orthopaedic Trauma, 32(Suppl 1), S1–S170. https://doi.org/10.1097/BOT.0000000000001063

[68] Pauwels F. Der Schenkelhalsbruch, ein mechanisches Problem. Grundlagen des Heilungsvorganges. Prognose und kausale Therapie. Stuttgart, Beilageheft zur Zeitschrift für Orthopädische Chirurgie, Ferdinand Enke, 1935

[69] Bray T. Femoral neck fracture fixation. Clin Orthop 1997;339:20–31

[70] Bartoníček J. Pauwels' classification of femoral neck fractures: Correct interpretation of the original. J Orthop Trauma 2001;15(5):358–360. doi:10.1097/00005131-200106000-00009

[71] Cavenago, C., & Prestianni, C. (2010). Criteri decisionali per la concessione del carico precoce nelle fratture laterali dell'estremo superiore del femore. Giornale Italiano di Ortopedia e Traumatologia, 36, 25–35. Recuperato da

[72] Agostini, M. (2020). Bioimmagini – Lezioni del corso A.A. 2020/2021. Politecnico di Torino, Corso di Laurea in Ingegneria Biomedica.

[73] Gonzalez, R. C., & Woods, R. E. (2018). Digital image processing

[74] ETSRM – Ente Toscano di Sicurezza Radiologica in Medicina. (s.d.). Tubo a raggi X

[75] Bushberg, J.T., Seibert, J.A., Leidholdt, E.M., Boone, J.M. (2011). The Essential Physics of Medical Imaging (3rd ed.). Lippincott Williams & Wilkins.

[76] PhysicsOpenLab. (2016, 9 febbraio). Annichilazione del positrone.

[77] Catrambone, V. (2024). Principi di Bioingegneria – Lezione Bioimmagini. Università di Pisa, Corso di Laurea in Ingegneria Biomedica

[78] RGMedicali. (2019, 19 marzo). Radiologia: radiologia analogica e digitale, quali differenze. RGMedicali Diagnostica per Immagini.

[79] Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge University Press.

[80] https://it.lyntop.com/blog/ti-6al-4v-eli/

[81] Fridman, L. (2025). Slide del corso "Introduction to Deep Reinforcement Learning". Lex Fridman

[82] Mitchell, T. M. (1997). Machine Learning. McGraw-Hill.

[83] Lollato, F. (2023). Deep learning per la classificazione di immagini: un approfondimento sulle reti neurali profonde

[84] La Scienza dei Dati. (s.d.). Il metodo Random Forest: la foresta magica della statistica.

[85] IBM. (2025). Che cos'è la foresta casuale? IBM Think.

[86] Statistics Easily. (s.d.). Che cos'è l'algoritmo KNN? Una guida completa

[87] ServiceNow. (2025). Che cos'è l'algoritmo K-Nearest Neighbors (KNN)?

[88] Ichi.pro.. (s.d.). Capitolo 1: K-Nearest Neighbors – Algoritmo di apprendimento automatico supervisionato

[89] FasterCapital. (2025). K-Nearest Neighbors (KNN) e Mahalanobis Distance: una combinazione per un'accurata classificazione.

[90] D. P. Yadav and S. Rathor, "Bone Fracture Detection and Classification using Deep Learning Approach," in Proc. 2020 Int. Conf. on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC), Mathura, India, Feb. 2020, pp. 1–5. doi: 10.1109/PARC49193.2020.236611.

[91] W. W. Myint, K. S. Tun, H. M. Tun, "Analysis on Leg Bone Fracture Detection and Classification Using X-ray Images," Machine Learning Research, vol. 3, no. 3, pp. 49–59, 2018. doi: 10.11648/j.mlr.20180303.11

[92] D. P. Yadav, A. Sharma, S. Athithan, A. Bhola, B. Sharma, and I. B. Dhaou, "Hybrid SFNet Model for Bone Fracture Detection and Classification Using ML/DL," Sensors, vol. 22, no. 15, p. 5823, Aug. 2022. doi: 10.3390/s22155823.

[93] Rajesh, S. (s.d.). Confusion Matrix for Multiclass Classification. Medium

[94] R.-Y. Ju and W. Cai, "Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm," Scientific Reports, vol. 13, p. 20077, Nov. 2023. doi: 10.1038/s41598-023-47460-7.

[95] A. Verma, V. Kumar, R. K. Yadav, and S. Sharmila, "Humerus Bone Fracture Detection Utilizing YOLOv4 Algorithm: A Deep Learning Approach," in Proc. 2024 2nd Int. Conf. on Disruptive Technologies (ICDT), Ghaziabad, India, 2024, pp. 1–6. doi: 10.1109/ICDT61202.2024.10489429.

[96] Midigudla, R.S., Dichpally, T., Vallabhaneni, U. et al. A comparative analysis of deep learning models for waste segregation: YOLOv8, EfficientDet, and Detectron 2. Multimed Tools Appl 84, 35941–35964 (2025). https://doi.org/10.1007/s11042-025-20647-y