جامعة أبو بكر بلقايد- تلمسان

University Abou Bakr Belkaid of Tlemcen

Faculty of Technology

Department of biomedical engineering

Final year project report

MASTER in BIOMEDICAL ENGINEERING

Speciality: Medical Imaging

Presented By: OUAHAB Soundous Selsabile

Estimating Age and Gender Using Artificial Intelligence in Digital Forensics

---

Defended on the 29th of June, 2025.

Examination Board jury

Mr. HADJ SLIMANE Zine Eddine	Professor	President
Mr. BEHADADA Omar	MCA	Examinator
Mrs. BENCHAIB Yasmine	MCA	Supervisor
Mr. BEREKSI REGUIG Fethi	Professor	Co-supervisor

Academic year: 2024-2025

Acknowledgements

I begin by offering my sincere and heartfelt thanks to God. This journey, marked by both perseverance and hardship, would not have been possible without His enduring presence. Through every challenge and moment of doubt, I found strength in faith and comfort in knowing I was never alone. The completion of this thesis is more than an academic achievement, it is also a reflection of the grace and guidance that carried me through each step of the way.

To my supervisors, Mr. Bereksi Reguieg and Mrs. Benchaib, thank you for believing in my proposed theme from the very beginning. Your expert guidance, thoughtful feedback, and continued support were instrumental at every stage of this research. I am particularly grateful for your responsiveness, your encouragement, and the clarity you helped me bring to what once felt like a distant vision.

I extend sincere thanks to the University of New Mexico for granting access to the New Mexico Decedent Image Database. This thesis would not have been possible without it. Your resources have made a significant contribution to this piece of work.

To Mr. Hadj Slimane and Mr. Behadada, thank you for agreeing to judge my work. I am grateful for your time and consideration.

A quiet thank you to the few teachers who truly loved their craft, who taught to ignite understanding, not just deliver content. You may never know the impact you had, but your influence lives in these pages.

And finally, to my parents and sister, you were my lifeline. When I lost faith in myself, you believed enough for all of us. When I wanted to quit, you anchored me. This degree is not mine alone. It is a reflection of your sacrifices, your endless encouragement, and your unwavering love. I hope this achievement honors everything you have given me.

Dedication

In a world where education remains a privilege denied to far too many, this work is a humble testament to those who never had the chance to complete their own. To the students of Congo, Sudan, Palestine, Yemen, and every corner of the world where war, oppression, and injustice have silenced voices and severed dreams: this is for you. For the theses left unwritten, the degrees left unearned, and the boundless potential extinguished too soon, may your struggles never be forgotten. May your courage in the face of adversity inspire those of us who are fortunate enough to learn freely, to do so with purpose, gratitude, and a commitment to justice. This work honors your memory, your resilience, and the futures that were taken too soon.

To my parents and sister, this is for you. Your sacrifices laid the foundation on which this entire journey was built. Your patience gave me the space to grow, your encouragement lifted me in moments of doubt, and your faith in me became the strength behind every challenge I overcame. You nurtured not only my ambitions but also my values, and for that, I am eternally grateful. This degree, while bearing my name, belongs just as much to you. It is the product of your love, your dedication, and the many unseen acts of support that made each page possible. I carry your legacy into every endeavor, and this achievement stands as a shared milestone of everything we've endured and accomplished together.

Abstract

Accurate identification of human remains is critical in forensic science, yet traditional morphological methods are subjective and limited. Post-mortem computed tomography (PMCT) offers non-invasive, high-resolution skeletal analysis, but requires advanced tools for data interpretation. This study explores machine learning (ML) to objectively analyze PMCT scans for estimating age and sex, enhancing reliability in biological profiling. We evaluate the diagnostic values of skeletal features, compare anatomical regions (skull, torso, limbs), and validate ML models for forensic applications. By integrating computational methods with PMCT, the study aims to establish standardized, data-driven techniques, advancing accuracy and reproducibility in forensic anthropology for criminal, disaster, and archaeological investigations.

Keywords

Forensic science, Post-mortem computed tomography, Machine learning, Age estimation, Sex estimation, XGBoost.

Résumé

L’identification précise des restes humains est essentielle en science médico-légale, mais les méthodes morphologiques traditionnelles demeurent subjectives et limitées. La tomodensitométrie post-mortem (PMCT) permet une analyse squelettique non invasive et à haute résolution, mais elle requiert des outils avancés pour interpréter efficacement les données. Cette étude explore l’utilisation de l’apprentissage automatique (machine learning, ML) pour analyser objectivement les scans PMCT en vue d’estimer l’âge et le sexe, renforçant ainsi la fiabilité des profils biologiques. Nous évaluons la valeur diagnostique des caractéristiques squelettiques, comparons différentes régions anatomiques (crâne, tronc, membres) et validons des modèles ML dans un contexte médico-légal. En intégrant des approches computationnelles à la PMCT, cette recherche vise à établir des techniques standardisées et fondées sur les données, améliorant la précision et la reproductibilité en anthropologie médico-légale, dans des contextes criminels, de catastrophes ou archéologiques.

Mots-clés

Science médico-légale, Tomodensitométrie post-mortem, Apprentissage automatique, Estimation de l’âge, Estimation du sexe, XGBoost.

الملخص

تُعدّ عملية تحديد هوية الرفات البشرية بدقة من الأمور الأساسية في علم الطب الشرعي، غير أن الطرق الشكلية التقليدية تبقى محدودة وذات طابع شخصي. توفّر تقنية التصوير المقطعي المحوسب بعد الوفاة تحليلاً هيكليًا غير تدخلي وعالي الدقة، لكنها تتطلب أدوات متقدمة لتفسير البيانات. تهدف هذه الدراسة إلى استخدام تقنيات التعلم الآلي لتحليل صور التصوير المقطعي بعد الوفاة بشكل موضوعي بغرض تقدير العمر والجنس، مما يعزز موثوقية تحديد السمات البيولوجية. تم تحليل الخصائص الهيكلية للعظام وتحديد مدى إسهامها في عملية التقدير، مع مقارنة المناطق التشريحية المختلفة مثل الجمجمة والجذع والأطراف، والتحقق من دقة نماذج التعلم الآلي في سياقات الطب الشرعي. من خلال دمج الأساليب الحاسوبية مع تقنية التصوير المقطعي بعد الوفاة، تهدف هذه الدراسة إلى إرساء تقنيات معيارية تعتمد على البيانات، مما يُسهم في تحسين الدقة وقابلية التكرار في علم الأنثروبولوجيا الشرعية، سواء في التحقيقات الجنائية أو الكوارث أو الدراسات الأثرية .

الكلمات المفتاحية

الطب الشرعي، التصوير المقطعي المحوسب بعد الوفاة، التعلم الآلي، تقدير العمر، تقدير XGBoost. الجنس،

Table of Contents

List of Figures

List of Tables

List of Acronyms

PMCT: Post-Mortem Computed Tomography

CT: Computed Tomography

MRI: Magnetic Resonance Imaging

QCT: Quantitative Computed Tomography

DEXA: Dual-Energy X-ray Absorptiometry

pQCT: Peripheral Quantitative Computed Tomography

MDCT: Multidetector-Row Computed Tomography

2D-DWT: Two-Dimensional Discrete Wavelet Transform

PCA: Principal Component Analysis

PLS: Partial Least Squares Regression

NMDID: New Mexico Decedent Image Database

OMI: Office of Medical Investigation

DICOM: Digital Imaging and Communications in Medicine

kVp: Kilovolt Peak

mAs: Milliampere-Seconds

FOV: Field of View

HU: Hounsfield Unit

GLCM: Gray-Level Co-occurrence Matrix

NIfTI: Neuroimaging Informatics Technology Initiative

ML: Machine Learning

AI: Artificial Intelligence

XGBoost: Extreme Gradient Boosting

MAE: Mean Absolute Error

MSE: Mean Squared Error

Introduction

The identification of human remains is a fundamental yet complex task in forensic science, essential in contexts ranging from criminal investigations to mass disasters and archaeological excavations. Accurately determining biological characteristics such as age, sex, and ancestry from skeletal remains is crucial for establishing identity. Traditionally, forensic anthropologists have relied on morphological analysis, employing visual examination and manual measurements to estimate these parameters. While valuable, these methods are inherently subjective, leading to variability in interpretation and reduced reliability, particularly when dealing with incomplete or compromised remains.

Advancements in technology are transforming the landscape of forensic anthropology. High-resolution medical imaging techniques, particularly post-mortem computed tomography (PMCT), offer non-invasive, detailed examinations of skeletal structures. Unlike traditional methods that may involve destructive sampling, PMCT preserves the integrity of remains while providing comprehensive three-dimensional data. However, the potential of this imaging technology is not fully realized without advanced analytical tools to interpret the extensive anatomical information it generates.

This is where modern computational methods, particularly through the use of machine learning, come into play. These algorithms excel at recognizing patterns within large datasets, a crucial capability for extracting meaningful biological information from post-mortem computed tomography (PMCT) scans. By integrating these advanced technologies, we can develop objective, precise, and reproducible techniques for estimating age and sex from skeletal remains. Such an approach enhances the reliability and validity of forensic assessments, ultimately leading to improved outcomes.

The current research addresses significant gaps in forensic methodologies. First, while various skeletal features are known to indicate age and sex, their relative diagnostic values remain inadequately explored. Second, a systematic evaluation of different anatomical regions such as the skull, torso, and lower extremities for biological profiling is necessary using modern analytical techniques. Third, the application of machine learning to PMCT data for forensic purposes requires rigorous validation to ensure reliability in real-world scenarios. By addressing these challenges, this study aims to establish best practices for biological profile estimation in the era of digital forensic anthropology.

The first chapter frames the research within the scientific and practical frameworks of forensic anthropology, highlighting its significance and current limitations in biological profile estimation. It discusses the critical challenge of differentiating between chronological and biological age, influenced by various environmental and genetic factors. This chapter also explores contemporary methodological approaches, particularly recent studies that combine PMCT imaging with machine learning for age and sex estimation, culminating in an overview of PMCT as a transformative technology in forensic analysis.

The second chapter delves into the anatomical foundations of biological profile estimation, systematically reviewing skeletal markers indicative of sexual dimorphism and age-related changes. It contrasts the diagnostic value of cranial features with postcranial characteristics, particularly pelvic morphology. The section on age estimation outlines the developmental timeline of skeletal changes, from growth plate fusion in young adults to degenerative alterations in older individuals, setting the stage for understanding how machine learning can quantify these morphological patterns in PMCT data.

The third chapter presents the study's methodological framework, detailing the PMCT dataset from the New Mexico Decedent Image Database and the image processing system that converts the dataset into analyzable three-dimensional models. Emphasis is placed on the feature extraction process, which bridges anatomical observations with computational analysis. Each skeletal region like the skull, torso, and lower extremities is examined to identify and quantify relevant biological markers, followed by a discussion of the machine learning approach utilizing the XGBoost algorithm for both classification (sex estimation) and regression (age estimation) tasks.

The final chapter presents and interprets the research findings, demonstrating the varied reliability of different skeletal regions for biological profile estimation. Notably, cranial features are highlighted for their high accuracy in both sex determination and age estimation. The discussion contextualizes these findings within established anatomical principles and existing literature, analyzing how estimation accuracy varies by sex and across age groups. The chapter concludes with practical implications for forensic casework and outlines potential directions for future research.

CHAPTER I

General Overview

1 General Overview

This chapter provides an overview of current methodologies in forensic science, specifically focusing on post-mortem computed tomography (PMCT) and machine learning for estimating biological profiles. It starts with an introduction to the importance of forensic science in criminal investigations. The chapter then discusses the challenges of estimating age and gender from skeletal remains, particularly in cases of decomposition, and highlights the discrepancies between biological age and chronological age, which can complicate forensic assessments. It reviews relevant studies that have used imaging techniques and machine learning to improve accuracy in these estimates. Finally, it addresses the limitations of existing methods and the difficulties in integrating these technologies into routine forensic practice.

1.1 Forensic science

Forensic science is essential in crime investigations, applying scientific principles to the analysis of evidence [1]. This field includes various disciplines such as forensic biology, toxicology, latent fingerprint analysis, and digital forensics [2]. Each area contributes to reconstructing events and identifying individuals involved in criminal cases, mass disasters, or archaeological findings. A primary objective of forensic science is the development of biological profiles, which encompass key characteristics like age, gender, and ancestry.

Recent advancements in technology have significantly improved forensic science, particularly through imaging techniques. Non-invasive methods such as computed tomography (CT) and magnetic resonance imaging (MRI) have transformed the examination of human remains. Among these, post-mortem computed tomography (PMCT) has emerged as a valuable tool, enabling detailed scans of deceased individuals. PMCT allows visualization of internal structures, including bones and soft tissues, without altering the physical state of the remains. This is crucial for preserving the integrity of the body in legal investigations or culturally sensitive contexts.

Unlike traditional autopsies, which involve physical dissection, PMCT offers a digital approach to analyzing skeletal and soft tissue morphology. This method is effective for identifying injuries and fractures that inform the biological profile. For example, PMCT can reveal age-related changes in bone density and pelvic features indicative of gender. Furthermore, integrating PMCT with advanced computational techniques, such as machine learning (ML), promises to enhance the accuracy of biological profile estimations by uncovering patterns that traditional methods may overlook.

Despite the potential benefits of PMCT, its routine integration into forensic practice remains limited. Current methods often rely on subjective expert assessments, which can introduce errors, particularly with decomposed or severely damaged remains. There is an urgent need for more objective, data-

driven approaches to enhance traditional methods and provide greater reliability in forensic investigations.

This Master's Thesis focuses on combining PMCT with machine learning models to address these challenges. By generating 3D models from images of human remains and using these models to train algorithms for predicting age and gender, this research aims to establish a robust framework for biological profile estimation. The proposed approach seeks to improve the accuracy and efficiency of forensic investigations, representing a significant step toward incorporating modern technologies into the field of legal medicine.

1.2 Motivation

Estimating age and gender at the time of death is a fundamental aspect of forensic science, forming the backbone of a biological profile. Such a profile is essential in numerous scenarios, including the identification of victims in mass disasters, solving criminal cases, and uncovering the identities of unknown remains. However, traditional methods for biological profiling often face significant limitations, particularly when dealing with bodies that have undergone decomposition, trauma, or environmental alterations, such as burning or mummification. These factors can lead to subjective assessments with a high margin of error, reflecting a pressing need for more objective and precise techniques.

Emerging technologies, particularly post-mortem computed tomography (PMCT), offer a promising solution to these challenges. PMCT provides non-invasive imaging of internal structures, allowing forensic scientists to visualize anatomical features that might otherwise remain hidden. For instance, it can reveal skeletal characteristics, such as bone density and morphological changes, useful for age estimation, as well as pelvic structures that assist in determining gender. Additionally, PMCT preserves the integrity of the remains, which is crucial in legal and ethical contexts where maintaining the physical state of the body is paramount.

Despite its advantages, PMCT alone does not fully address the limitations of traditional methods. The interpretation of PMCT data often relies on expert analysis, which can still introduce subjectivity and variability. To enhance the precision of biological profile estimation, integrating machine learning (ML) techniques can be beneficial. By utilizing large datasets derived from PMCT scans, ML models can identify patterns and features that may not be readily observable by human experts. This integration has the potential to significantly reduce human error and improve the accuracy of age and gender predictions, even in cases involving severely degraded remains.

The motivation for our research comes from the need to bring new technologies, like PMCT and machine learning, into forensic science to improve how we estimate age and gender. By combining these imaging techniques with advanced computational tools, this work aims to develop methods that

are more accurate and reliable than traditional approaches. The integration of these techniques not only addresses the weaknesses of existing methods but also helps establish more consistent practices in forensic investigations. Ultimately, the goal is to enhance the ability of forensic scientists to provide clear and objective information in critical situations.

1.3 Problem formulation

Forensic experts estimate age and gender from skeletal remains by understanding the nature and timing of skeletal changes throughout a person's life. The process involves distinguishing between chronological age, which refers to the total time a person has lived, and biological age, also known as physiological age. Biological age reflects the physiological changes that occur over time and can be assessed through various biomarkers and developmental stages observable in bones.

Establishing a connection between an individual's biological and chronological age is crucial for accurately determining the age of unidentified remains. However, as illustrated in Figure 1:1, discrepancies often arise between these two measures. Biological age may not always align with chronological age due to a phenomenon known as the trajectory effect [3]. This effect tends to increase with age and is influenced by a range of factors, including genetics, nutrition, environmental conditions, and levels of physical activity.

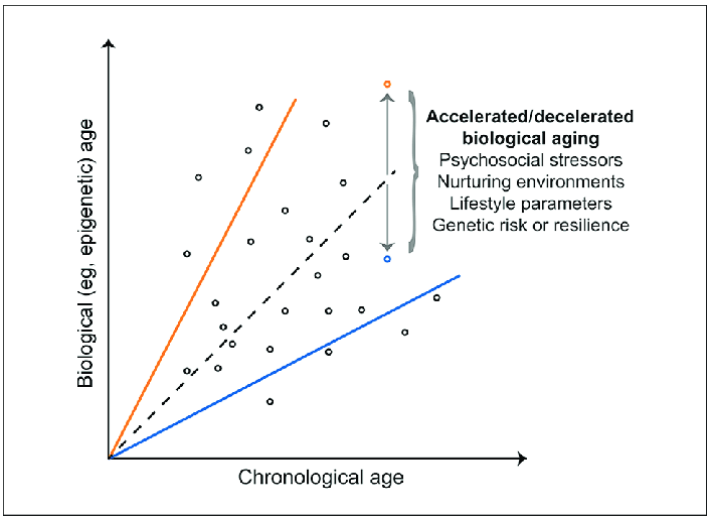

Figure 1:1 Relationship between chronological age and biological age [4].

Understanding the discrepancies between chronological and biological age is crucial in forensic science, as these differences can complicate age assessments and lead to misinterpretations in investigations. Estimating biological profiles typically relies on expert assessments and available documentation. When information is lacking, forensic scientists can infer important details by examining physiological markers and the developmental stages of bones; which raises critical questions about which body regions are most effective for determining age and gender. Furthermore, how can integrating post-mortem computed tomography (PMCT) and artificial intelligence (AI)

improve the accuracy of these assessments? Addressing this question is essential for refining forensic methodologies, as it could enhance the distinction between chronological and biological age, ultimately aiding in the identification of remains.

1.4 Aim

The main goal of this Master's Thesis is to explore the integration of emerging technologies in forensic science, focusing specifically on the combination of machine learning models with post-mortem computed tomography (PMCT). We aim to improve the accuracy of estimating biological profiles, which include factors like age and gender, from skeletal remains.

To achieve this, useful data from PMCT scans is extracted and used to generate three-dimensional models of the skeletal structures. These models will then be integrated into machine learning algorithms to enhance the precision of age and gender estimations.

A key aspect of our research is to develop methods that can reliably establish the identity of individuals, regardless of the circumstances surrounding their death. Ultimately, this work aims to contribute to the fields of digital forensics and legal medicine by providing more robust tools for identifying human remains, thereby aiding investigations and supporting legal processes.

1.5 Related work

In the area of age and gender estimation, researchers have used various medical imaging techniques in their studies. Early efforts with machine learning primarily focused on X-ray and MRI images, especially of the hand, pelvis, and knee. X-ray imaging of the hand is often seen as the best method for estimating age and gender in living patients. This is because it reduces exposure to nearby organs while offering a clear picture of bone development.

On the other hand, imaging the pelvis and knee can be more complicated. These images can be harder to interpret, particularly for those without specialized training, which could lead to mistakes in assessing age and gender. Additionally, pelvic and knee imaging may expose surrounding tissues to ionizing radiation, raising concerns about safety. While X-ray imaging of the hand is effective and safer, there is a need for more research to improve how we use pelvic and knee imaging for age and gender estimation. It could include better imaging techniques or machine learning tools that make these images easier to understand while minimizing radiation exposure.

Recent advancements have shifted attention towards specialized applications of computed tomography (CT), particularly Quantitative Computed Tomography (QCT) and Post-Mortem Computed Tomography (PMCT). QCT is a diagnostic method that measures bone mineral density (BMD) using a phantom scanning technique. It involves placing a calcium hydroxyapatite phantom beneath the patient's bone, which allows for a correlation between Hounsfield units (HU) and measurable

quantities of bone mass. Unlike dual-energy X-ray absorptiometry (DEXA), which provides a two-dimensional assessment of bone density, QCT offers a three-dimensional analysis, giving a more detailed view of bone structure and density. Researchers [6] have found that QCT yields slightly better results than DEXA, making it a valuable tool for age estimation, even in cases of putrefaction, with moderate correlations reported (R^2 = 0.33-0.42).

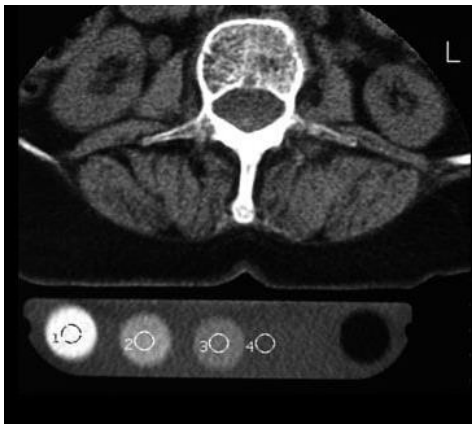

Figure 1:2 CT of the lumbar spine measuring the density of the bone with a calibration phantom underneath the patient [7].

The integration of machine learning has revolutionized age and gender estimation techniques in medical imaging. One notable method is peripheral quantitative computed tomography (pQCT), which provides high-resolution, three-dimensional images of bone and soft tissue. The technique is particularly useful for assessing bone density and structure at peripheral sites, such as the forearm or lower leg. A recent study [8] highlighted the effectiveness of deep learning algorithms in this context. For instance, a deep learning model utilizing a LeNet5 architecture achieved an accuracy of 86.4% in sex estimation from a single pQCT slice of the fourth lumbar vertebra (L4). The study analyzed 117 vertebrae (58 males and 59 females) and serves as a proof-of-concept, demonstrating that deep learning can effectively leverage individual skeletal traits to provide accurate sex estimation.

Furthermore, dental analysis has become a useful method for estimating age, especially by looking at the size of the pulp cavity in teeth. Research [9] using images from multidetector-row computed tomography (MDCT) of mandibular first premolars has shown that the volume of the pulp cavity has a significant relationship with age, with a correlation coefficient of 0.76. In this study, 136 mandibular first premolars from individuals of known age were analyzed, suggesting that MDCT is a practical tool for gathering information needed for age prediction in forensic cases. MDCT is a type of imaging that provides detailed cross-sectional images of the teeth and surrounding structures using multiple rows of detectors, allowing for faster and more accurate scans.

Recent studies underscored the effectiveness of multifactorial approaches to age estimation. A novel deep learning technique [10] utilizing 3D PMCT scans of both the mandible and femur achieved a mean absolute error (MAE) of 5.15 years and a concordance correlation coefficient (CCC) of 0.80

using a sample of 814 cadavers (619 males and 195 females). The method demonstrated improved accuracy over using either skeletal element alone, indicating that the integration of multiple anatomical indicators enhances the reliability of age assessments. Pham evaluated mandibles and femurs separately, achieving MAEs of 7.07 ± 0.51 years and 5.74 ± 0.04 years, respectively, and combined data yielded an MAE of 5.41 ± 0.26 years.

Another study [11] introduced a method for estimating age using post-mortem computed tomography (PMCT) images of bones, focusing on areas such as the vertebral body, ischial tuberosity, iliac crest, and femur. The study analyzed PMCT images from a total of 807 decedents, comprising 617 males and 190 females. They implemented a two-dimensional discrete wavelet transform (2D-DWT) to capture high-frequency features from the images and employed machine learning techniques, particularly support vector regression, for analysis. Additionally, the study incorporated principal component analysis (PCA) and partial least squares regression (PLS) to facilitate dimensionality reduction. Validation through a 10-fold cross-validation method revealed a mean absolute error (MAE) of 7.92 years for male vertebral bodies. Notably, the MAEs were higher for males in the extreme age groups of 20–30 years and 80–90 years compared to those in the mid-range groups of 40–70 years, while in females, elevated MAEs occurred only in the youngest age group (20–30 years).

In addition to age estimation, several studies have explored sex estimation through innovative methods. For example, one of these [12] studies investigated the potential for sex estimation via the analysis of finger lengths using postmortem CT data, involving 205 cases of known sex. Similarly, another study [13] examined CT-derived femoral images from a contemporary Black South African population, analyzing 280 femurs (140 males and 140 females) and achieving classification accuracies ranging from 79.3% to 91.4% using discriminant function and logistic regression analyses. These findings reinforce the potential of CT-derived measurements in forensic applications, providing reliable and reproducible data that surpass traditional methods reliant on older osteometric collections.

Collectively, these studies illustrate the versatility of imaging techniques, particularly CT, in forensic age and gender estimation. The integration of advanced methods and machine learning not only enhances the accuracy of these estimations but also paves the way for future advancements in forensic anthropology, offering non-invasive and reliable approaches for postmortem identification.

1.6 Limitations

The use of postmortem computed tomography (PMCT) in the estimation of biological profiles, particularly age and gender, remains limited within the existing literature. This constraint is primarily due to the complexities involved in obtaining relevant data from deceased individuals. Ethical considerations, including the necessity for ethical approval and consent, pose significant challenges in conducting research in this sensitive area. Additionally, there are currently no standardized databases

or established methodologies for employing deep learning (DL) or machine learning (ML) techniques to estimate age and gender based on PMCT data.

Moreover, recent studies often focus on a single anatomical region, which can limit the comprehensiveness of findings. For instance, focusing exclusively on a single part of the body may limit our overall understanding of the biological profile. Researchers have underscored the importance of large sample sizes, as a more extensive dataset, enhances the reliability of ML models. A larger pool of data allows for more robust training of algorithms, ultimately leading to more accurate and trustworthy results.

Another significant limitation arises from the condition of the skeletal remains. Advanced stages of decomposition, mummification, or burns can severely impede the visibility and integrity of bone structures. These factors not only complicate the assessment of age and gender but also introduce variability that can affect the accuracy of any estimations made. Consequently, the interplay of these limitations underscores the need for further research and the development of standardized practices in the application of PMCT for biological profile estimation.

Conclusion

In conclusion, this chapter has highlighted the integration of post-mortem computed tomography and machine learning in estimating biological profiles, emphasizing the challenges related to biological and chronological age discrepancies. The next chapter will build on this foundation by providing a detailed background on relevant medical concepts. It will also cover the fundamentals of machine learning and explore performance evaluation methods, along with the importance of feature selection in enhancing predictive accuracy.

Bibliography

[1] NIST, “Forensic Science,” *NIST*, Aug. 20, 2013. Available: https://www.nist.gov/forensic-science

[2] “Forensic Sciences Command,” *isp.illinois.gov*. Available: https://isp.illinois.gov/Forensics/Specialties

[3] noble.dana, “Understanding the difference between biological age and chronological age - Mayo Clinic Press,” *Mayo Clinic Press*, Jul. 25, 2024. Available: https://mcpress.mayoclinic.org/healthy-aging/understanding-the-difference-between-biological-age-and-chronological-age/

[4] Zannas, Anthony. (2019). Epigenetics as a key link between psychosocial stress and aging: concepts, evidence, mechanisms. Dialogues in clinical neuroscience. 21. 389-396. 10.31887/DCNS.2019.21.4/azannas.

[5] “Age Determination - an Overview | ScienceDirect Topics,” *www.sciencedirect.com*. Available: https://www.sciencedirect.com/topics/nursing-and-health-professions/age-determination

[6] M. Barszcz and K. J. Woźniak, "A review of methods of age estimation based on postmortem computed tomography," *Forensic Sciences Research*, vol. 2024, no. 1, pp. 1-21, 2024. DOI: 10.1093/fsr/owae036.

[7] “Quantitative Computed Tomography (QCT),” *UCSF Radiology*, May 02, 2017. Available: https://radiology.ucsf.edu/patient-care/services/QCT

[8] P. Oura, N. Korpinen, A. L. Machnicki, and J.-A. Junno, "Deep learning in sex estimation from a peripheral quantitative computed tomography scan of the fourth lumbar vertebra—a proof-of-concept study," *Forensic Science, Medicine and Pathology*, vol. 19, pp. 534–540, 2023. DOI: 10.1007/s12024-023-00586-6.

[9] A. Sakuma, H. Saitoh, Y. Suzuki, Y. Makino, G. Inokuchi, M. Hayakawa, D. Yajima, and H. Iwase, "Age estimation based on pulp cavity to tooth volume ratio using postmortem computed tomography images," *Journal of Forensic Sciences*, vol. 58, no. 6, pp. 1531-1535, 2013.

[10] C. V. Pham, S. J. Lee, S. Y. Kim, S. Lee, S. H. Kim, and H. S. Kim, "Age estimation based on 3D post-mortem computed tomography images of mandible and femur using convolutional neural networks," *PLoS ONE*, vol. 16, no. 5, e0251388, 2021. DOI: 10.1371/journal.pone.0251388.

[11] K. Imaizumi, S. Usui, K. Taniguchi, Y. Ogawa, T. Nagata, K. Kaga, H. Hayakawa, and S. Shiotani, "Development of an age estimation method for bones based on machine learning using post-mortem computed tomography images of bones," *Forensic Imaging*, vol. 26, 2021, Art. no. 200477, doi: 10.1016/j.fri.2021.200477.

[12] T. Ikeda, K. Miyamoto, N. Tani, S. Oritani, T. Michiue, F. Morioka, and T. Ishikawa, "Forensic evaluation of sex estimation via measurements of adult index and ring finger lengths using postmortem computed tomography," *Egyptian Journal of Forensic Sciences*, vol. 8, no. 43, pp. 1-9, 2018. DOI: 10.1186/s41935-018-0075-5.

[13] Ujaddughe, O.M., Haberfeld, J., Bidmos, M.A. *et al.* Evaluation of standards for sex estimation using measurements obtained from reconstructed computed tomography images of the femur of contemporary Black South Africans. *Int J Legal Med* (2025). https://doi.org/10.1007/s00414-025-03430-4

CHAPTER II

Background

2 Background

The chapter provides an overview of the key concepts related to age and gender estimation from skeletal remains. As forensic science advances, accurately estimating an individual's age and gender has become increasingly important. The process relies on understanding the biological changes that bones undergo with age and recognizing the differences between male and female skeletal structures.

We will discuss the physiological changes in bone composition and structure over a person's lifetime, emphasizing how these changes vary between genders. Additionally, we will examine specific anatomical features critical for determining biological sex, focusing on both cranial and post-cranial differences. Furthermore, we will introduce the background of machine learning techniques that can enhance the accuracy of age and gender estimations. This background will set the stage for the methods used in age and gender estimation.

2.1 Medical Background

As individuals age, their bones undergo significant changes in appearance and structure, which can differ markedly between males and females. Understanding these changes is crucial for age and gender estimation in both forensic and medical contexts. Bones are dynamic organs composed of various tissues that provide structural integrity to the body, serve as attachment points for muscles, facilitate movement, and protect vital organs. They work in conjunction with muscles and joints, enabling movement through contractions. Tendons connect muscles to bones, while ligaments connect bones to one another, providing stability. Additionally, cartilage cushions joints and reduces friction between bones during movement, as illustrated in Figure 2:1.

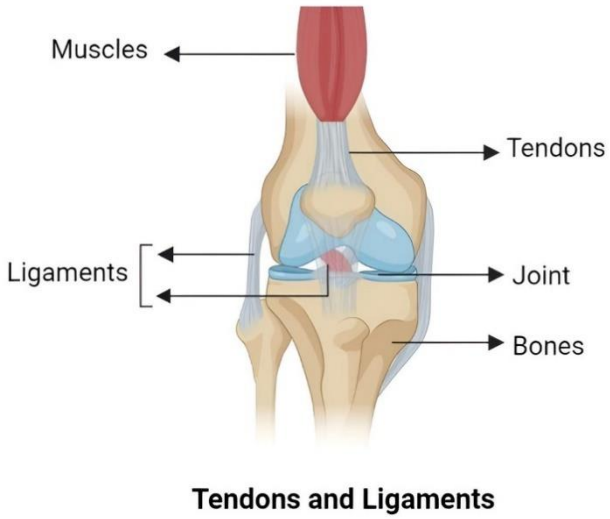

Figure 2:1 The structure of bones and connective tissues [14].

Bone tissue consists of specialized cells known as osteocytes, which play crucial roles in maintaining bone health. Within this framework, there are two main types of cells: osteoclasts and osteoblasts. Osteoclasts are responsible for dissolving and absorbing bone tissue, a process that is essential for bone remodeling and repair. In contrast, osteoblasts are involved in building new bone tissue, facilitating growth and the healing of fractures. The interplay between these two cell types is vital for maintaining the balance of bone density and structural integrity.

Bones are composed of two primary types of tissue: spongy bone and compact bone. Spongy bone, found within the interior of bones, is lighter and less dense, which helps reduce overall body weight while providing strength. Compact bone, on the other hand, forms the outer layer and is dense and robust, offering protection and support. Additionally, bone marrow is present within the cavities of bones; red marrow, located in the epiphysis, is responsible for producing red blood cells, while yellow marrow, which is primarily fatty, is found in the shaft of the bone. The structure and functions of these tissues are further illustrated in Figure 2:2.

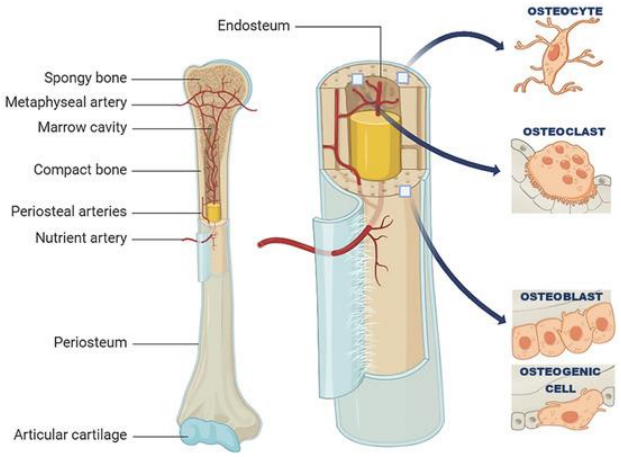

Figure 2:2 Bone tissue: structure and function [15].

At birth, humans possess approximately 300 bones; however, as they grow, some of these bones fuse together, resulting in a total of 206 bones in adulthood [16]. This process, known as ossification, involves the gradual hardening of cartilage into bone, which is vital for developing a strong and functional skeletal system. The structure of bones is essential for their function, with specific regions playing key roles in growth and development.

The cap, or epiphysis, is located at both ends of long bones and is covered by cartilage that facilitates joint movement. The growth plate, situated between the epiphysis and the diaphysis (the shaft of the bone), is where most of the bone's growth occurs. This growth plate produces new cells that push older cells inward, leading to the elongation of the bone. As a person reaches maturity, these growth plates eventually close and are replaced by the epiphyseal line, marking the conclusion of bone growth and

the transition to a fully developed skeletal structure. The components of this process are depicted in Figure 2:3 below.

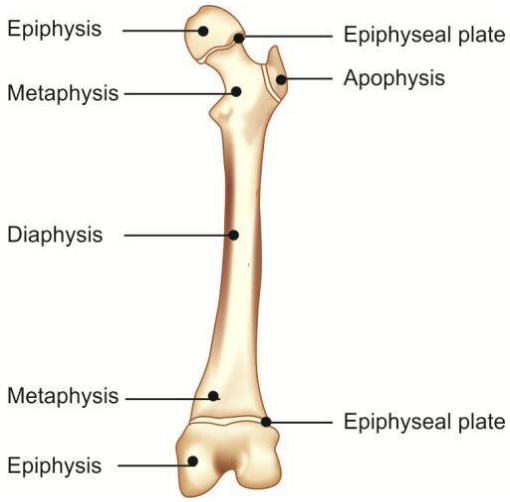

Figure 2:3 Anatomy of long bones [17].

After the age of 30, bones begin to lose minerals, leading to a higher risk of various conditions that can significantly impact bone health. One of the most concerning of these is osteoporosis, a condition in which bones become fragile and more susceptible to fractures. As bone density decreases, individuals may also experience other issues, such as osteoarthritis, a degenerative joint disease characterized by pain and stiffness in the joints. Additionally, bone spurs, which are small protrusions that can develop on bones, may emerge, further complicating mobility and comfort. Another potential concern is scoliosis, a condition that results in spinal misalignment, which can lead to additional pain and functional limitations. Collectively, these age-related changes highlight the importance of maintaining bone health throughout life.

2.1.1 Sex Determination

Determining biological sex is often more straightforward than estimating age or ancestry, as humans are typically classified as either male or female. While genetic anomalies and hormone therapies can introduce some ambiguity, such cases are relatively rare and generally do not obscure the primary determination. This distinct classification is primarily due to the clear physical differences that develop between the sexes during puberty [18]. Sexual dimorphism refers to the observable variations in size and shape between males and females, which become more pronounced after puberty due to hormonal influences. These differences manifest in various anatomical structures, particularly in the skull and pelvis, providing valuable insights for biological sex determination.

2.1.1.1 Cranial Differences

Determining biological sex through cranial features involves analyzing distinct differences observable from both frontal and lateral views of the skull. Key structures of the cranium include the frontal bone, orbits, zygomatic processes, mandible, and brow ridge, all of which provide valuable insights into biological sex. Figure 2:4 illustrates these components, highlighting the frontal bone, parietal bone, sphenoid, ethmoid, lacrimal, nasal bones, zygomatic bone, zygomatic process, maxilla, alveolar margin, mandible, mastoid process, and occipital bone.

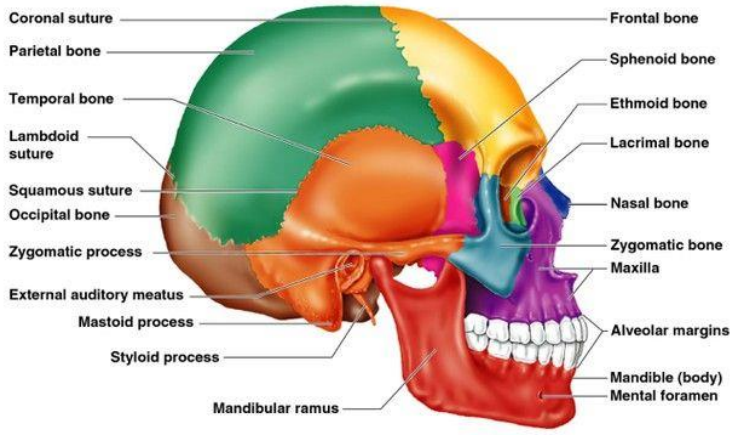

Figure 2:4 The skull anatomy [19].

From the frontal view, male skulls typically exhibit a lower and more sloping frontal bone, while female skulls tend to have a higher and more rounded shape. The orbits are generally square in males, contributing to a more robust facial structure, whereas females display more rounded orbits, which enhance the softness of their features. Additionally, the zygomatic process extends further toward the ear in males, indicating a broader facial width, while in females, it stops short. The mandible in males appears more squared, contrasting with the V-shaped mandible commonly found in females. Moreover, the brow ridge is notably more pronounced in males, imparting a rugged appearance, while females usually have a thinner brow ridge that results in a softer facial contour.

From the side view, male skulls are characterized by a rougher and more robust surface texture and a larger, pointier mastoid process. Males often possess a squarer chin compared to the more rounded chin of females, and their foreheads are typically more sloping. In contrast, female foreheads are rounder and globe-shaped, further differentiating the two sexes. Figure 2:5 depicts the comparison of the side and frontal views of male and female skulls, effectively highlighting these distinctive characteristics.

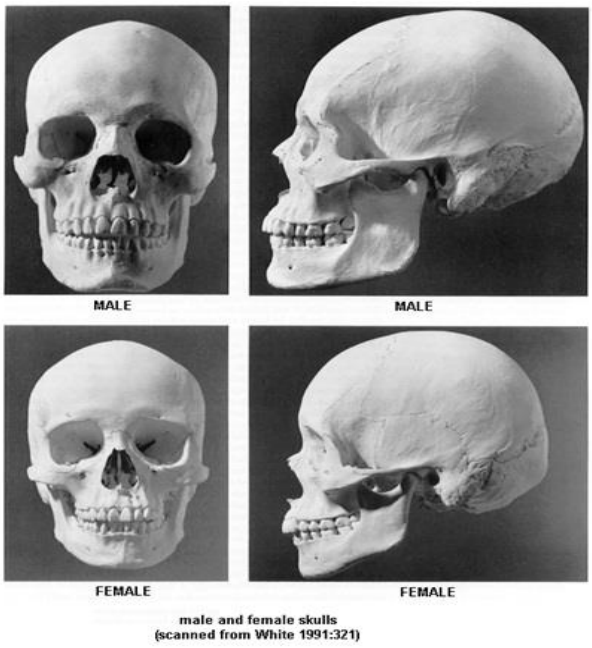

Figure 2:5 Sexual dimorphism in skull anatomy [20].

2.1.1.2 Post Cranial Differences

In post-cranial assessments, the pelvis is the most definitive anatomical feature for determining biological sex. Accurate sex determination requires evaluating several key areas, including the ilium (the upper portion of the pelvis), sacrum, coccyx, pubic bone, and ischium, all of which are illustrated in Figure 2:6. These anatomical differences not only aid in sex determination but also underscore the biological adaptations related to reproduction. Additionally, males typically have longer and more robust long bones, such as those in the arms and legs, reflecting greater overall size and muscularity. The dimensions and shapes of these long bones, along with pelvic characteristics, provide further context for understanding skeletal morphology. While the femur can be measured to estimate a person's height, it is important to note that it is not the most reliable indicator for determining biological sex, as variations in size and shape can occur across individuals.

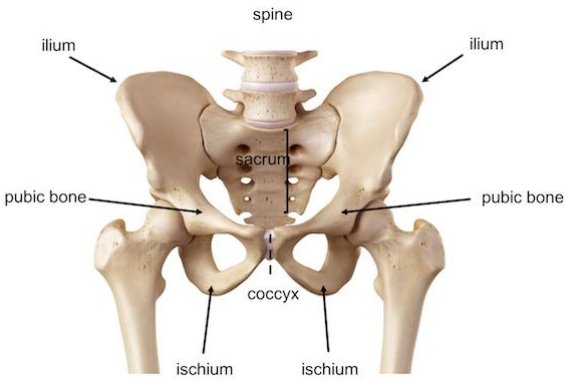

Figure 2:6 Anatomy of the pelvis [21].

The pelvis is the most definitive anatomical feature used to distinguish between male and female skeletons, exhibiting distinct sexual dimorphic traits [22]. The male pelvic cavity is typically narrow and heart-shaped, while the female pelvic cavity is broader and more accommodating, which is essential for reproductive functions. Generally, females possess larger and broader pelvises that facilitate childbirth. The sacrum, located at the base of the spine, also differs between sexes as it is less curved in females, contributing to a wider pelvic cavity, whereas the male sacrum tends to be narrower and more curved. It is important to note that the accuracy of sex estimation based on pelvic characteristics improves significantly after puberty, as hormonal changes during this stage greatly influence sexual dimorphism.

Key features of the female pelvis include the sciatic notch, located at the back of the pelvis. In females, this notch is wider and has a more open angle, facilitating a broader birth canal, while males have a narrower sciatic notch that aligns with the overall constricted structure of their pelvis. Additionally, the pelvic outlet shows significant differences: in females, it is more open and less obstructed by the tailbone, aiding in childbirth, while in males, the pelvic outlet is more constricted, with the tailbone tucking underneath. Another critical feature is the pubic arch, formed where the two pubic bones meet at the front of the pelvis. In males, the pubic arch is typically V-shaped due to the narrower configuration of their pelvis, whereas females exhibit a U-shaped pubic arch, indicating greater openness. The subpubic angle, formed below the pubic symphysis, further differentiates the sexes; in adult males, this angle is narrow, measuring less than 90 degrees, while in females, it is wider, usually exceeding 90 degrees. Figure 2:7 below illustrates the full male and female pelvises, highlighting these distinctions, including the differences in the sciatic notch and pubic arch.

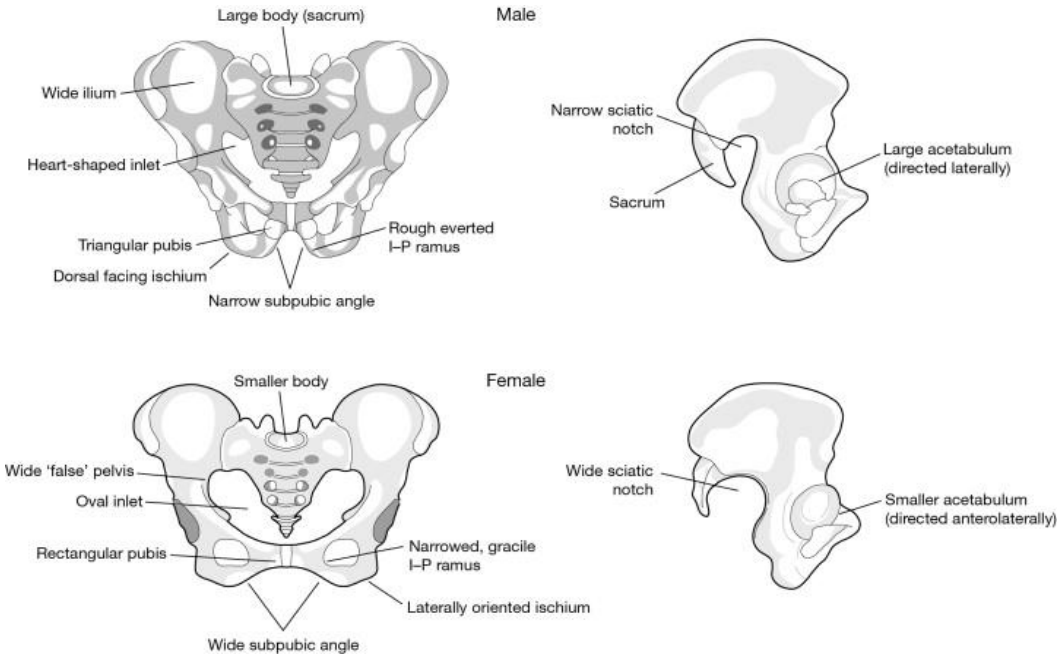

Figure 2:7 Sexual dimorphism in the pelvis anatomy [23].

2.1.2 Age Estimation

After determining a person's sex, the next step in forensic analysis is estimating age. The process becomes increasingly complex, particularly as individuals grow older. Accurate age estimation is crucial in various fields, including anthropology, archaeology, and forensic science, as it helps provide context to skeletal remains and informs legal and medical investigations.

2.1.2.1 Cranial Differences

Cranial features offer valuable insights for age estimation, especially through the examination of cranial sutures. In infants, gaps between the skull bones allow for growth and development, reflecting the dynamic nature of the skull during early life. As individuals age, these sutures gradually fuse, providing a timeline for age estimation. For instance, the coronal suture typically closes around age 50, while the sagittal suture, located on the top of the head, usually closes by about age 32. The lambdoidal suture begins to close around age 21, accelerates at age 26, and is generally closed by age 30. The squamosal suture, situated at the side of the skull, starts to close after age 60. At approximately 20 years of age, sutures remain open, indicating that the various parts of the skull are still growing together. As individuals age into their 70s and 80s, sutures tend to completely obliterate, particularly the sagittal and lambdoid sutures, due to the process of fusion [24]. Figure 2:8 illustrates the upper portion of the skull, highlighting the coronal, sagittal, and lambdoid sutures. Additionally, Figure 2:4 provides a side view of the squamosal suture, further emphasizing the significance of these anatomical features in age estimation. Understanding the timeline of suture closure not only aids in estimating age but also enhances our comprehension of the biological processes that occur throughout a person's life.

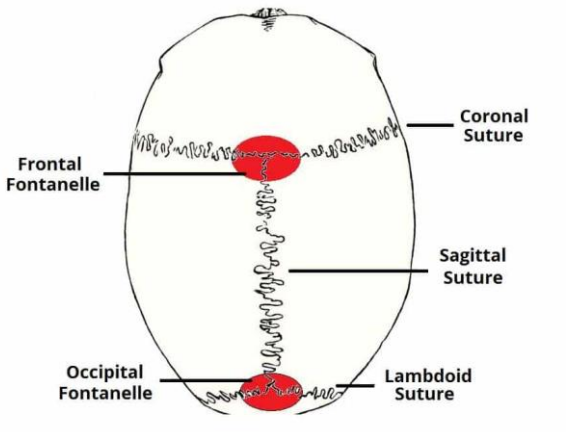

Figure 2:8 Upper skull anatomy: sutures and their connection [25].

2.1.2.2 Post Cranial Differences

The assessment of post-cranial features is essential for estimating age, as various anatomical structures provide valuable insights into an individual's developmental timeline. In examining the shoulder region, the clavicle and sternum typically complete their fusion between the ages of 22 and 30 [26]. This fusion signifies the final stages of skeletal maturation, where the bones solidify into a stable structure. Figure 2:9 below illustrates anterior views of the sternum and thorax, which showcases their structural features.

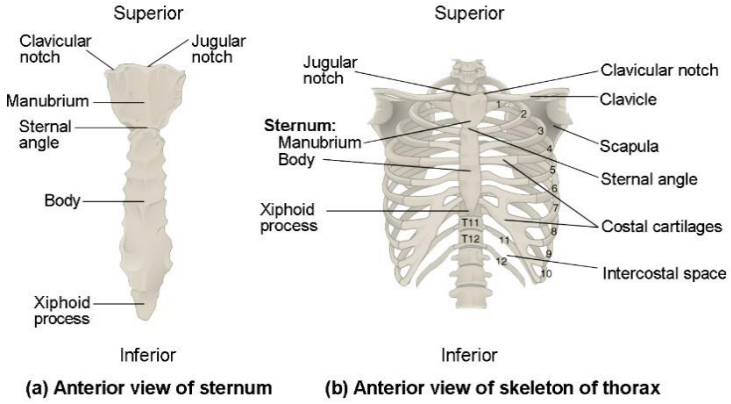

Figure 2:9 Anterior views of the sternum and thorax [27].

In the pelvic region, the fusion of the pubic bone and ischium occurs at approximately ages 7 to 8. This early fusion is crucial for understanding skeletal development. By the ages of 20 to 25, the ilium, ischium, and pubic bones undergo complete ossification, while the sacrum becomes fully united, marking a significant point in skeletal maturity [28]. The pelvic structure not only plays a role in locomotion but also reflects adaptations related to reproduction, making its assessment vital for both age estimation and biological sex determination.

The femur, recognized as the longest bone in the human body, is particularly important in age assessment. Notably, around age 4, the greater trochanter begins to appear, serving as a key attachment site for muscles that aid in movement. By ages 13 to 14, the lesser trochanter becomes visible, indicating further development of the bone. The head of the femur fuses to the shaft between the ages of 16 and 18, which is a critical milestone in the maturation of the skeletal system. Figure 2:10 below depicts the development of the trochanters throughout the first year of life and beyond, illustrating their progression from infancy through adolescence.

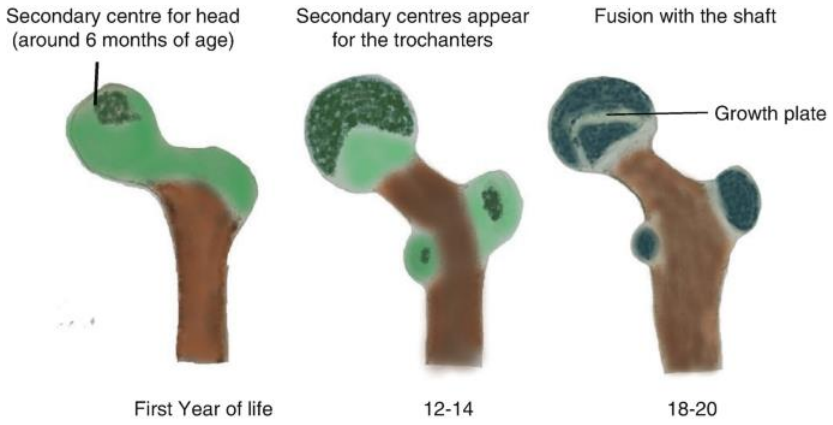

Figure 2:10 Trochanter development and femoral shaft fusion across key growth stages [29].

In terms of epiphyseal fusion, it is important to note that at birth, bones are not fully fused; they consist of distinct sections. The femur, for example, has a diaphysis at the bottom and an epiphysis at the top, with growth plates located between them. These growth plates close at specific intervals during late adolescence and early adulthood, providing key markers for age estimation. Additionally, the deterioration of joint surfaces and bony erosion can offer further insights into an individual's age as they mature. Figure 2:11 depicts the stages of epiphyseal fusion in the knee joint, showcasing X-ray images that illustrate the fusion process of the distal femur and the proximal tibia and fibula at various development stages. Image A shows stage I, where the epiphysis is not fused at the distal end of the femur and the proximal ends of the fibula and tibia. Image B depicts stage II, with the epiphysis fully ossified and the epiphyseal scar visible at the distal end of the femur and the proximal end of the tibia. Image C presents stage III, where the epiphysis scar is fully ossified and the epiphyseal scar is no longer visible at the distal end of the femur and the proximal end of the tibia.

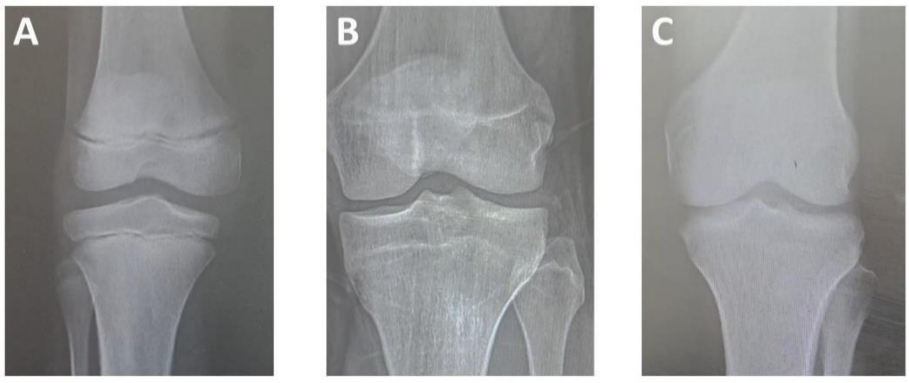

Figure 2:11 Epiphyseal union stages in the knee joint [30].

Sex determination and age estimation from skeletal remains involve the analysis of extensive datasets containing cranial and post-cranial measurements. These biological indicators exhibit intricate variations that arise from developmental processes and characteristics unique to specific populations. Traditional methods of manual analysis often face challenges in addressing these complex,

multidimensional relationships. Therefore, the application of machine learning techniques becomes crucial. Such methods can effectively find small patterns in the data while also handling problems, such as missing values and measurement errors, that are often found in biological datasets.

2.2 Machine Learning

Working with large and complex datasets is a significant challenge, as it requires humans to find meaningful connections and patterns. Machine learning methods are crucial in this context, as they help identify relationships and trends within the data.

Machine learning (ML) is a branch of artificial intelligence (AI) and computer science that focuses on developing algorithms and statistical models. These tools enable computers to learn patterns from data [31]. With this capability, computers can classify, predict, or make decisions based on extensive datasets.

There are two main types of machine learning methods: supervised learning and unsupervised learning. In supervised learning, algorithms are trained on labeled datasets, where each example has a known outcome. This training allows the algorithm to make predictions. In contrast, unsupervised learning works with unlabeled datasets, requiring algorithms to discover hidden patterns and relationships on their own. However, in some cases, basic machine learning models tend to underperform with the limited dataset. To overcome this issue, Ensemble learning can be used. Ensemble learning is a ML method that combines two or more learners in order to produce better results and predictions. It has been adopted in cases where limited datasets cause issues [32].

2.2.1 Extreme Gradient Boosting

Extreme Gradient Boosting (XGBoost) is a powerful tree-based ensemble machine learning algorithm developed by Tianqi Chen in 2016 [33]. It belongs to the gradient boosting family, where models are constructed sequentially to minimize a loss function through additive optimization. XGBoost has demonstrated exceptional performance across various predictive modeling tasks, including both regression and classification problems.

In regression tasks, XGBoost is particularly effective for predicting numerical values based on input samples. In this thesis, age estimation is a key application of regression, where the model predicts an individual's age based on various features derived from the dataset. This approach allows for nuanced insights into age-related trends and variations within the data. In contrast, classification tasks can be binary (male or female) or multi-class [34]. For our purposes, binary classification is utilized for gender estimation. XGBoost is especially advantageous in biomedical data analysis due to its ability to model complex, non-linear relationships and effectively manage high-dimensional structured datasets.

The algorithm operates through a sequential procedure that builds trees iteratively, focusing on refining predictions at each step. Figure 2:12 below illustrates the iterative tree-building procedure in the XGBoost algorithm.

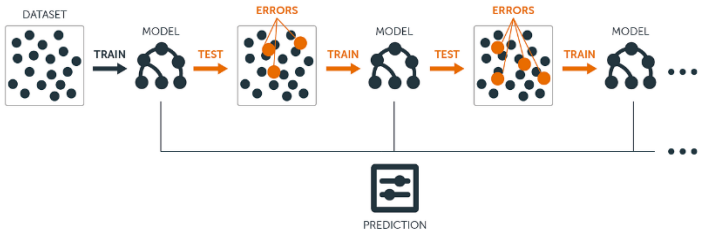

Figure 2:12 The iterative tree-building procedure in the XGBoost algorithm [35].

Initially, XGBoost constructs the first model using the input features, which provides a baseline prediction for the dataset. After this initial model is created, the algorithm calculates the error for each observation, identifying the discrepancies between the predicted and actual values. Subsequent models are then developed specifically to predict these residuals (errors), effectively learning from the mistakes of the previous models. The iterative process culminates in an ensemble of trees that collectively enhance the model's predictive capability.

The approach allows XGBoost to strike a balance between bias and variance, addressing a common limitation in traditional gradient boosting methods, which often overfit the training data. XGBoost includes regularization terms that enhance model generalization.

Each tree is built under the assumption that the features predict the residuals. If the model's predictions closely match the true values, it may indicate overfitting. To mitigate this, XGBoost employs a learning rate that ranges from 0 to 1. This parameter scales the contributions of new trees, ensuring that adjustments are made gradually. By taking smaller steps, the model achieves improved predictions on the test dataset, resulting in lower variance. XGBoost constructs an ensemble of decision trees, where the prediction for a given input is the sum of outputs from all individual trees. Figure 2:13 illustrates the key steps in the XGBoost algorithm showcasing the binary trees and how their collective outputs contribute to the final prediction.

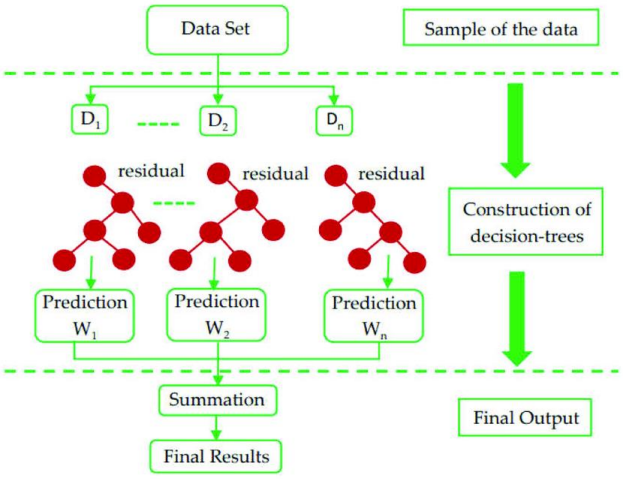

Figure 2:13 Key steps in the XGBoost algorithm [36].

In the following sections, we will also present equations that further elucidate the mathematical underpinnings of this process [37].

2.2.1.1 Mathematical Overview

As described above, XGBoost constructs an ensemble of decision trees, where the prediction for a given input is the sum of outputs from all individual trees. This process is set mathematically by the following equations, Where

$$\hat{y}_i = \sum_{k-1}^{K} f_k(x_i), \quad f_k \in \mathcal{F} \qquad \textit{Equation 1}$$

Here, $\hat{y}_i$ is the predicted output, f_k denotes the k-regression tree, and $\mathcal{F}$ is the set of all possible trees. Each tree is added to minimize the overall objective function $\mathcal{L}^{(t)}$:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \qquad \textit{Equation 2}$$

Where l is the loss function measuring the difference between the true label y_i and the predicted value $\hat{y}_i^{(t-1)}$. And $\Omega(f_t)$ is the regularization term that penalizes the complexity of the model to avoid overfitting.

The regularization term typically takes the form:

$$\Omega(\mathrm{f}) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \qquad \textit{Equation 3}$$

where:

- T is the number of leaves in the tree,

- w_j are the weights of the leaf nodes,
- γ and λ are the regularization parameters.

2.2.2 Performance Evaluation Metrics

Employing a variety of performance evaluation metrics is crucial for a comprehensive assessment of model effectiveness. Each metric offers unique insights into different aspects of model performance. For instance, while accuracy provides a general overview of predictive success, Mean Absolute Error (MAE) specifically highlights the magnitude of prediction errors.

2.2.2.1 Accuracy

Accuracy is one of the most commonly used metrics to evaluate the performance of a machine learning model, particularly when the classes are balanced. It provides an overall measure of how often the model correctly predicts the outcome compared to the total number of predictions.

$$\text{Accuracy} = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \qquad \textit{Equation 4}$$

Where the number of correct predictions refers to the number of times the model's predicted label matches the actual label, which includes true positives and true negatives. The total number of predictions however refers to the total count of all predictions made by the model, which is the sum of true positives, true negatives, false positives and false negatives.

2.2.2.2 Mean Absolute Error

Mean Absolute Error (MAE) is a metric used to evaluate the performance of a model by measuring the magnitude of errors between the predicted and actual voxel intensities in a dataset [38]. It is calculated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i| \qquad \textit{Equation 5}$$

where n is the total number of voxels, y_i represents the actual voxel for the i:th observation, and x_i represents the predicted voxel for the i:th observation. The absolute difference between each predicted and actual voxel is computed, and then the mean of these absolute differences across the entire dataset is calculated to obtain the MAE.

MAE is a useful evaluation metric for several reasons. Firstly, it retains the same unit as the original voxel intensities, making it intuitive and easy to interpret. Additionally, MAE is less sensitive to outliers compared to other metrics like Mean Squared Error (MSE), which can be disproportionately affected by large errors.

Conclusion

In conclusion, the chapter outlined the fundamental biological and anatomical principles essential for age and gender estimation. By examining the dynamics of bone development and the distinguishing characteristics of male and female skeletons, we establish a background for the subsequent discussion on methodologies. The following sections will detail the methods employed in these estimations, particularly the application of machine learning techniques to improve accuracy and effectiveness in forensic analysis.

Bibliography

[14] H. Yousefi, "Comprehensive Understanding of Musculoskeletal Anatomy and Physiology for Holistic Practitioners," *Medium*, Feb. 14, 2024. Available: https://medium.com/@hamedyousefi/comprehensive-understanding-of-musculoskeletal-anatomy-and-physiology-for-holistic-practitioners-f12b15f51b10

[15] K. M. Percival, V. Paul, and G. A. Husseini, "Recent Advancements in Bone Tissue Engineering: Integrating Smart Scaffold Technologies and Bio-Responsive Systems for Enhanced Regeneration," *Int. J. Mol. Sci.*, vol. 25, no. 11, p. 6012, 2024. [Online]. Available: https://doi.org/10.3390/ijms25116012

[16] J. Roland, "How Many Bones Are Babies Born With and Why Do They Have More Than Adults?," *Healthline*, Jun. 26, 2019. Available: https://www.healthline.com/health/how-many-bones-does-a-baby-have

[17] "JaypeeDigital | Anatomy of Bone and Fracture Healing," *Jaypeedigital.com*, 2015. Available: https://www.jaypeedigital.com/book/9789351968085/chapter/ch2.

[18] German A, Hochberg Z. Sexual Dimorphism of Size Ontogeny and Life History. Front Pediatr. 2020 Jul 24;8:387. doi: 10.3389/fped.2020.00387. PMID: 32793524; PMCID: PMC7394004.

[19] "Skull Right Lateral View Labeled - Biology Forums Gallery | Skull anatomy, Anatomy bones, Human skull anatomy," *Pinterest*, 2024. Available: https://www.pinterest.com/pin/75787206205379156/

[20] *Blogspot.com*, 2025. Available: https://lemissingtooth.blogspot.com/2012/10/male-skull-female-skull-2.html.

[21] "Pelvis Hip Anatomy," *anatomy.lexmedicus.com.au*. Available: https://anatomy.lexmedicus.com.au/collection/pelvis-hip

[22] Chaudhary, Alok Kumar1; Jain, Sanjeev Kumar2,. Osteoscopic Assessment of Sexual Dimorphism in Hip Bone. Acta Medica International 1(1):p 28-31, Jan–Jun 2014.

[23] T. D. White, M. T. Black, and P. A. Folkens, “Pelvis,” *Human Osteology*, pp. 219–240, 2012, doi: https://doi.org/10.1016/b978-0-12-374134-9.50011-8

[24] Ruengdit, Sittiporn & Case, D. & Mahakkanukrauh, Pasuk. (2019). Cranial Suture Closure as an Age Indicator: A review. Forensic Science International. 307. 110111. 10.1016/j.forsciint.2019.110111

[25] Sparke, “Bones of the skull - structure - fractures - teachmeanatomy,” *Teachmeanatomy.info*, 2014. Available: https://teachmeanatomy.info/head/osteology/skull/

[26] Zaki, Eman A1; Reda, Alaa M2; Lashin, Heba I1; Kabbash, Abdel-moty MK1. Age Estimation Based on Medial Clavicular Epiphyseal Union in A Sample of Egyptian Population Using Multi-Detector Computed Tomography Scan. Tanta Medical Journal 52(1):p 53-61, January-March 2024. | DOI: 10.4103/tmj.tmj_35_23

[27] OpenStaxCollege, “The Thoracic Cage,” *pressbooks-dev.oer.hawaii.edu*, Mar. 2013, Available: https://pressbooks-dev.oer.hawaii.edu/anatomyandphysiology/chapter/the-thoracic-cage/

[28] T. M. Hoagland, “Hip Joint Anatomy: Overview, Gross Anatomy,” *Medscape.com*, Feb. 19, 2025. Available: https://emedicine.medscape.com/article/1898964-overview?utm_source=chatgpt.com&form=fpf.

[29] C. Kumari and K. Parmar, “Femoral Head: Anatomical Considerations,” pp. 1–11, Jan. 2023, doi: https://doi.org/10.1007/978-981-99-1346-6_1

[30] J. A. M. Alzyoud, E. Rababah, M. H. O. Almuhaisen, and A. I. Al-Qtaitat, "Bone Age Determination of Epiphyseal Fusion at Knee Joint and Its Correlation with Chronological Age," *Medicina*, vol. 60, no. 5, pp. 779, 2024. doi: 10.3390/medicina60050779.

[31] IBM, “What Is Machine learning?,” *Ibm.com*, Sep. 22, 2021. Available: https://www.ibm.com/think/topics/machine-learning

[32] IBM, “Ensemble learning,” *Ibm.com*, Mar. 18, 2024. Available: https://www.ibm.com/think/topics/ensemble-learning

[33] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 785-794.

[34] J. Brownlee, “A Gentle Introduction to XGBoost Loss Functions - MachineLearningMastery.com,” *MachineLearningMastery.com*, Mar. 21, 2021. Available: https://machinelearningmastery.com/xgboost-loss-functions/.

[35] Shilpa kamishetty, “XGBoost - Shilpa kamishetty - Medium,” *Medium*, Apr. 21, 2019. Available: https://medium.com/@shilpakamishetty/xgboost-350530593cb5.

[36] K. Khan, W. Ahmad, M. Amin, A. Ahmad, S. Nazar, and Anas Abdulalim Alabdullah, “Compressive Strength Estimation of Steel-Fiber-Reinforced Concrete and Raw Material Interactions Using Advanced Algorithms,” *Polymers*, vol. 14, no. 15, pp. 3065–3065, Jul. 2022, doi: https://doi.org/10.3390/polym14153065

[37] D. Leventis, “XGBoost Mathematics Explained,” *Medium*, Jan. 02, 2022. Available: https://dimleve.medium.com/xgboost-mathematics-explained-58262530904a

[38] P. Schneider, “Mean Absolute Error - an overview | ScienceDirect Topics,” www.sciencedirect.com, 2022. Available: https://www.sciencedirect.com/topics/engineering/mean-absolute-error

CHAPTER III

Method

3 Method

The chapter tackles the approach and the methods employed in this project. Initially, it explores relative details about the dataset and the method on how it was collected. Subsequently, it explores various ways used in this study to process and handle data accordingly. The project workflow, encompassing its various components, is illustrated in Figure 3:1.

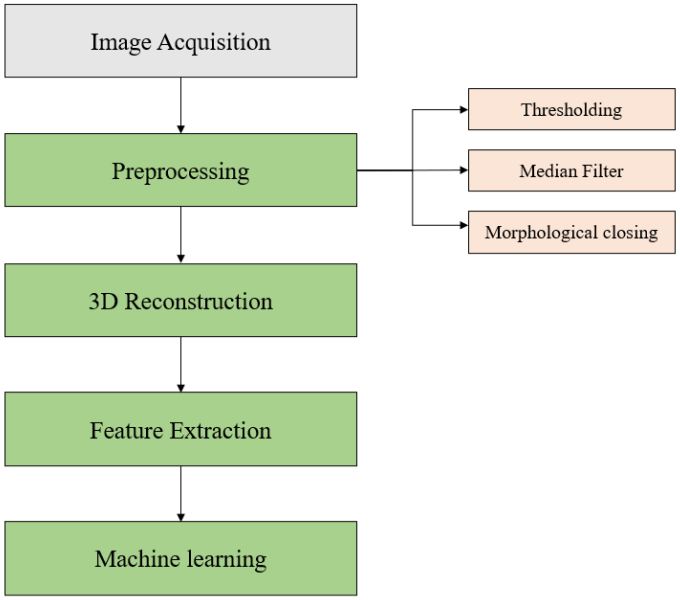

Figure 3:1 An overview of the project workflow, where different colors represent various components: grey indicates steps not conducted in this project, green denotes intermediate steps, and pink highlights corresponding parts of those steps.

3.1 Framework

All models and analysis in this project were conducted using 3D Slicer (version 5.6.2) [39] for preprocessing steps and 3D reconstruction. Additionally, Python (version 3.12.7) [40] was utilized on the Spyder platform for feature extraction and machine learning processes. The computational workflow was executed on a workstation with these specifications described in the table down below:

Table 3:1 Hardware setup for computational processing.

Component	**Details**
Processor	Intel® Core™ i3-8145U
Memory	8.00 GB DDR4 RAM
Operating System	Windows 10 Pro (64-bit, Version 22H2, Build 19045.5737)

3.2 Dataset

The study uses a dataset from the New Mexico Decedent Image Database (NMDID), which offers a unique collection of postmortem computed tomography (PMCT) scans derived from individuals of various ages and genders. The PMCT scans were conducted between 2010 and 2017 during postmortem examinations at the Office of Medical Investigation (OMI) in Albuquerque, New Mexico [41].

The NMDID contains high-resolution PMCT scans along with associated metadata, providing a comprehensive overview of diverse cases, each with distinct causes of death. For this study, we implemented specific inclusion criteria when selecting scans. Only individuals aged 18 to 80 years with verified identities were included. Cases involving significant artifacts, compromised bone structures, or extensive bone damage were excluded to maintain data integrity. We aimed for a balanced representation of age and gender, which is essential for developing an accurate estimation model. Ultimately, our analysis included 202 PMCT scans, comprising 111 male subjects and 91 female subjects. Figure 3:2 depicts the sex distribution between the individuals in the dataset.

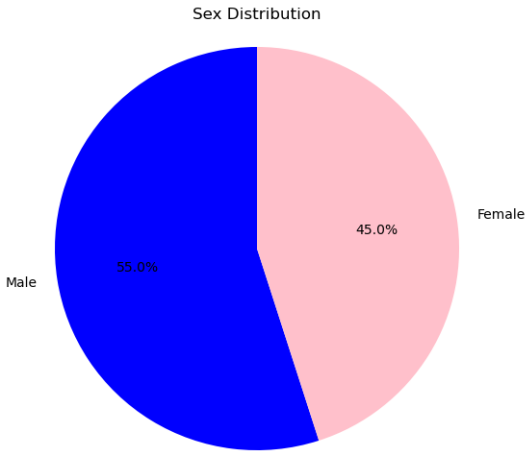

Figure 3:2 Sex distribution in the dataset.

In addition to gender representation, we focused on a wide range of ages within the 18 to 80-year spectrum, ensuring that our dataset reflects the age distribution relevant to forensic studies. The following graph represents the age distribution by brackets, illustrating the specific age groups we aimed for to ensure balanced representation and proper results.

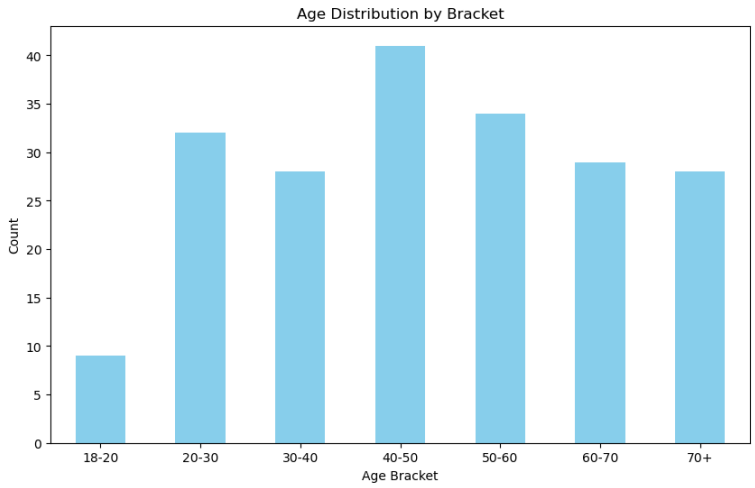

Figure 3:3 Age distribution across age groups

In addition to balancing age and gender representation, we also looked closely at the causes of death among the analyzed cases. The causes varied significantly, with substance intoxication being the most common. This category includes deaths related to drugs, alcohol, and poisons. Following substance intoxication, gunshot wounds and multiple injuries were also frequently observed. Table 3:2 depicts the primary causes of death within our dataset.

Table 3:2 Primary causes of death and corresponding case numbers.

The primary cause of death	Number of cases
Substance intoxication (drug, poison, alcohol, etc.)	51
Gunshot wound	23
Multiple injuries (fractures, lacerations to internal organs)	23
Cardiac arrhythmia	19
Head and neck injuries (cervical, cranio-, cerebral)	8
Hypertension (hypertensive cardiovascular disease)	7
Narcotic abuse	7
Hanging	7
Pneumonia (bronchitis)	7
Other causes of death	50
Total	202

The "Other causes of death" category includes various conditions such as chronic alcoholism, seizure disorders, thermal injuries, and congenital defects like Down syndrome. We also noted cases of Alzheimer's disease and cerebrovascular incidents, including strokes and anoxia. This wide range of causes adds depth to our dataset and highlights the complexity of mortality, offering valuable insights

for ongoing forensic research. Understanding these causes helps improve forensic practices and enhances the accuracy of age and gender estimation models.

3.2.1 CT-Scan protocol

The images used in this study are CT-scan images formatted as DICOM, captured following a specific protocol to ensure high quality and consistency. The imaging protocol used in this study was designed to systematically capture high-quality CT scans of three key anatomical regions: the head-neck, torso, and lower extremities. The protocol aimed to balance image clarity with efficient scanning by using consistent settings where possible while making necessary adjustments for each body area, as detailed in Table 3:3.

Table 3:3 CT Scan parameters for different anatomical regions.

	Head-neck	Torso	Lower extremities
kVp	120	120	120
mAs	300	300	200
Scan length	600-800 mm	600-800 mm	800-1000 mm
Scan FOV	500-699 mm	350 -699mm	500 mm
Pitch	0.567	0.817	0.942
Collimation	16 × 0.75	16 × 0.75	16 × 0.75
Rotation time	1.0 s	1.0 s	1.0 s
Matrix	512 × 512	512 × 512	512 × 512

All scans utilized a tube voltage of 120 kVp (kilovolt peak) for uniform tissue penetration. The kVp represents the maximum voltage applied across the x-ray tube, directly influencing both the energy of the produced x-rays and their ability to penetrate different tissue densities. A detector configuration of 16 × 0.75 mm was employed throughout, allowing the simultaneous acquisition of 16 thin slices per rotation, each just 0.75 mm thick. This configuration provides excellent spatial resolution while minimizing partial volume effects. Figure 3:4 illustrates the geometric relationship between the x-ray source, collimation system, field of view (FOV), and detector array in CT scanning. The collimator precisely shapes the x-ray beam to match the selected FOV, while the multi-row detector captures the attenuated radiation to reconstruct the thin slices used in this study. The rotation time remained constant at 1.0 second per rotation, and all images were reconstructed in a standardized 512 × 512 matrix to ensure consistent resolution across the dataset.

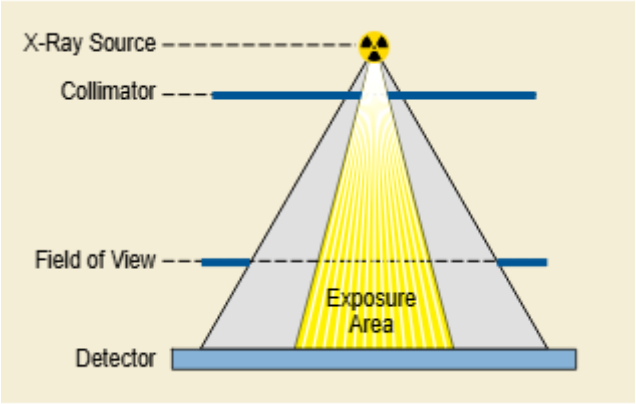

Figure 3:4 Geometric relationships in CT scanning [42].

For head-neck imaging, the protocol applied 300 mAs (milliampere-seconds), a measure of the total x-ray exposure. This relatively higher setting ensured sufficient signal-to-noise ratio for visualizing the complex bony structures of the skull and cervical spine. The scan length ranged from 600 to 800 mm to fully capture the anatomical region, with a field of view (FOV) between 500 and 699 mm. The FOV, which determines the diameter of the imaged area, was selected to include all relevant structures while minimizing unnecessary radiation to surrounding tissues. A pitch of 0.567 was used, providing moderate overlap between consecutive slices to enhance image quality for detailed anatomical assessment.

Torso scans maintained the 300 mAs setting to adequately penetrate the thorax and abdominal structures, but employed a narrower FOV range of 350 to 699 mm. This adjustment helped focus on internal organs while reducing scatter from peripheral tissues. The pitch increased to 0.817 for torso imaging, striking an optimal balance between image detail and scanning efficiency given the larger anatomical area requiring coverage.

For the lower extremities, the protocol incorporated several modifications to accommodate the region's simpler anatomy and greater length. The mAs was reduced to 200, as the predominantly bony structures required less radiation exposure to achieve diagnostic image quality across the extended 800 to 1000 mm scan length. A fixed 500 mm FOV provided consistent coverage of the limbs, while the pitch was increased to 0.942 to allow faster table movement during the longer scan without compromising the diagnostic value of the images.

Through these carefully selected and regionally optimized parameters, the protocol successfully produced clear, high-resolution images while maintaining efficient scan times and minimizing radiation exposure. The systematic approach to parameter selection ensured reliable, comparable data acquisition across all subjects - a critical foundation for the subsequent age and sex estimation analyses. The consistent use of thin-slice acquisitions (0.75 mm) proved particularly valuable for

detailed morphological assessments and 3D reconstructions, while the standardized reconstruction matrix guaranteed uniform image quality throughout the study.

3.3 Preprocessing

After the acquisition of the images, it is essential to preprocess them correctly to ensure accurate 3D reconstruction. The preprocessing procedure consists of three major steps: thresholding, filtering, and morphological operations. All preprocessing steps were performed using 3D Slicer.

3.3.1 Thresholding

Thresholding is a fundamental technique used in image segmentation to differentiate between various structures within medical images. This process involves selecting a specific intensity value, known as the threshold, to classify each pixel. Pixels with intensities above this threshold are categorized as foreground, representing the structures of interest, while those below are classified as background [43].

To facilitate thresholding, the Hounsfield Unit (HU) is considered, which quantifies the radiodensity of different tissues in computed tomography (CT) imaging. The Hounsfield scale for various tissues and materials is summarized in Table 3:4 Hounsfield scale values for various tissues/materials [44].below. A threshold value of 300 HU was applied to ensure the inclusion of both trabecular bone (300-800 HU) and cortical bone (>1000 HU), while effectively excluding softer tissues and background materials. This value successfully segmented bone structures from surrounding tissues and other materials, providing a clear structure for further analysis.

Table 3:4 Hounsfield scale values for various tissues/materials [44].

Tissue/Material	Hounsfield Scale
Air	-1000
Fat	-100
Water	0
Soft Tissue	30 to 45
Acute Blood	60 to 95
Trabecular Bone	300 to 800
Cortical Bone	>1000

This global thresholding approach converted grayscale images into binary formats, clearly delineating the bone features for analysis. The figure below illustrates the thresholding process: the first image represents the grayscale image before thresholding, while the second image displays the binary format obtained after applying the threshold. This visual comparison highlights the effectiveness of the thresholding technique in isolating the bone structures of interest.

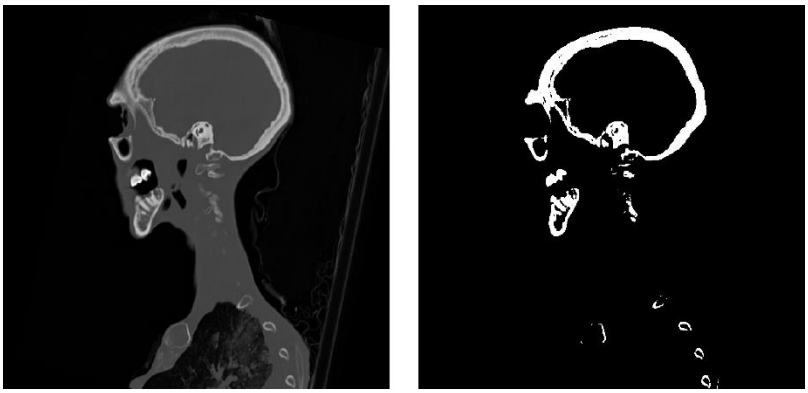

Figure 3:5 Thresholding process: grayscale to binary conversion.

3.3.2 Median filter

The median filter is an effective technique for reducing noise, particularly the 'salt-and-pepper' noise that often affects medical images. Sometimes, the application of a threshold value can result in small artifacts appearing in the segmented images. To address this, the median filter works by replacing each voxel's value with the median of its 3x3x1 neighborhood. This kernel size is widely adopted as a standard in filtering because it provides a balanced trade-off between noise suppression and edge preservation. A 3x3x1 neighborhood (effectively a 3x3 kernel applied slice-by-slice in volumetric data) is large enough to eliminate isolated noise pixels without excessively blurring fine structures or altering nearby anatomical details. Larger kernels (for example 5x5x1) risk oversmoothing or distorting edges, particularly in high-resolution medical images where subtle boundaries must remain intact. Thus, the 3x3x1 configuration suppresses outlier intensities while preserving important structural edges.

While individual 2D axial slices may show subtle improvements, the cumulative effect of median filtering becomes significantly apparent in the 3D reconstruction. This filtering technique helps to eliminate scattered artifacts and residual soft tissue fragments that may remain after the bone segmentation process. The accompanying figure illustrates this enhancement, comparing 3D torso reconstructions before and after median filtering. The filtered result reveals substantially cleaner bone structures with reduced noise contamination, underscoring the filter's importance in improving image quality.

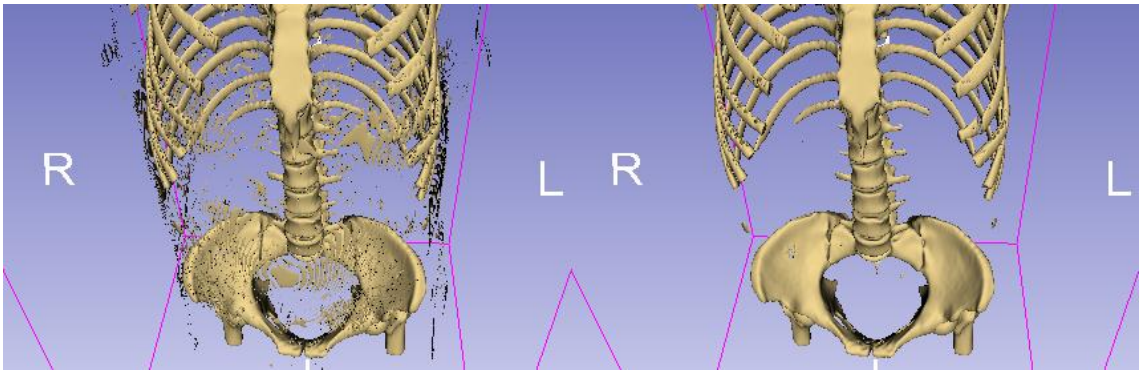

Figure 3:6 Median filtering in 3D torso reconstructions.

3.3.3 Morphological closing

Morphological operations are essential for processing object shapes in images, with closing being particularly useful for improving 3D segmentation by addressing small gaps left by thresholding. In this study, we employed a standard 3×3×1 structuring element. This compact size effectively fills minor discontinuities without distorting structures, unlike larger kernels which may artificially modify structures, while a smaller one might not fully eliminate discontinuities. Thus, the 3×3×1 structuring element provides an optimal balance between gap filling and shape preservation.

Morphological closing consists of two key steps: dilation followed by erosion. Dilation expands the boundaries of objects, filling in small holes and connecting disjointed regions. Erosion then refines the result by trimming back the outer edges, ensuring that the objects retain their original shape without unnecessary extra pixels [45]. Mathematically, closing is expressed as:

$$Closing(A, B) = (A \oplus B) \ominus B \quad \textit{Equation 6}$$

Where A represents the input image, B is the structuring element, $\oplus$ denotes dilatation and $\ominus$ denotes erosion.

By applying morphological closing, the segmentation of 3D images becomes more robust to thresholding artifacts, ensuring continuous structures and eliminating small gaps. This improvement is particularly critical for tasks like age and gender estimation, where accurate shape representation directly impacts predictive models. To better understand the effect, Figure 3:7 compares the rib cage's upper side view before and after morphological closing, demonstrating how the operation enhances structural continuity.

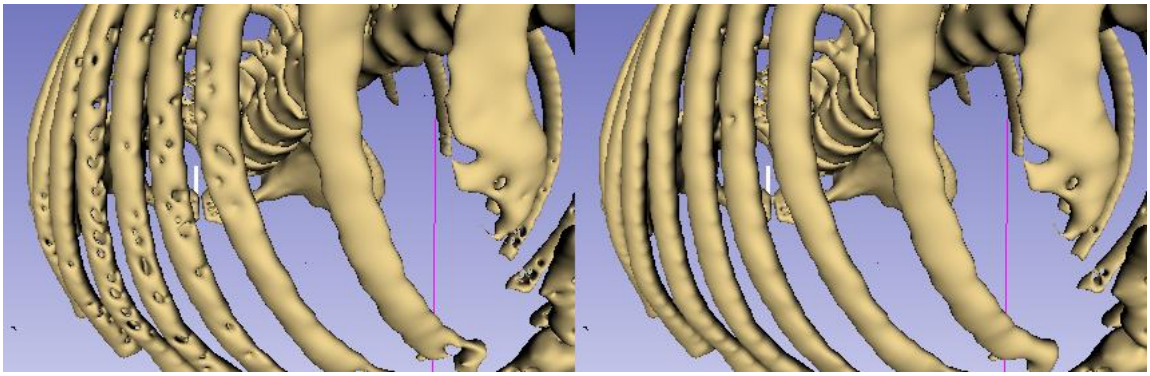

Figure 3:7 Effect of morphological closing.

3.4 3D reconstruction

Following the preprocessing steps, the two-dimensional slices for each anatomical region were systematically stacked to create three-dimensional volumes. The resulting volumetric datasets were tailored to the specific anatomical areas of interest: the head and neck region was reconstructed at a resolution of 512 × 512 × 120 voxels, the torso at 512 × 512 × 270 voxels, and the lower extremities at 512 × 512 × 320 voxels. A uniform voxel spacing of 3 × 3 mm was maintained throughout the reconstruction process to ensure anatomical accuracy.

The reconstructed volumes were saved in the NIfTI format, a widely accepted standard in medical imaging that facilitates high-dimensional data storage and processing [46]. A binary segmentation scheme was implemented within these files, where voxels corresponding to bone tissue were assigned a value of 1, and all non-bone regions were set to 0. This approach allowed for straightforward differentiation between bone and background during subsequent analysis and visualization.

For visualization and analysis of the reconstructed volumes, the open-source software 3D Slicer was utilized. Its built-in volume rendering tools enabled immediate inspection of the three-dimensional datasets.

3.5 Feature Extraction

After the 3D reconstruction of the segmented images, the next step is feature extraction. The feature extraction process utilized pre-segmented 3D postmortem computed tomography (PMCT) reconstructions that included only bone structures. Using the NiBabel library in Python, the 3D NIfTI files were imported and prepared for analysis.

Features were systematically extracted from three key anatomical regions: the skull, torso, and lower extremities. Each region was analyzed independently to capture distinct morphological characteristics that are relevant to age and sex estimation.

3.5.1 Cranial Features extraction

The cranial feature extraction process used image processing techniques to analyze age- and sex-related characteristics from 3D skull reconstructions. For age estimation, the focus was on cranial suture fusion patterns, which are important indicators of skeletal maturation.

The process began with edge enhancement using a 3x3x3 3D Sobel filter, which was specifically selected since it captures subtle gradient changes in all three anatomical planes simultaneously. Unlike 2D filters that might miss depth-wise variations in suture morphology, this three-dimensional approach highlighted the boundaries of sutures by calculating gradients in axial, sagittal, and coronal planes. The method proved particularly valuable for identifying partial fusion patterns, where sutures showed incomplete edges, as it could detect faint edge continuations across slices that might be overlooked in separate 2D analyses. The filtered images were then analyzed in two ways: first, by using Gray-Level Co-occurrence Matrices (GLCM) to evaluate texture changes related to fusion, and second, by measuring the gaps between sutures through connected-component labeling.

A simple classification system was used to categorize the fusion states of sutures:

- **Score 0** for unfused sutures, which had continuous high-intensity edges,
- **Score 1** for partially fused sutures, which showed intermittent edge patterns with areas of bony bridging,
- **Score 2** for fully fused sutures, characterized by complete bony union with no detectable edges.

The classification combined edge measurements from the Sobel filter with GLCM texture features like contrast, entropy, and homogeneity to ensure reliability.

For sex estimation, the analysis centered on identifying features that indicate sexual dimorphism. Male skulls typically have larger dimensions, so total cranial volume was calculated by summing voxel values. This overall measurement was supplemented with regional analyses to capture specific sex differences.

For example, the angle of the frontal bone was measured to reflect the steeper slope found in male skulls, while the shape of the eye sockets was analyzed using elliptical fitting to differentiate between the rounder female orbits and the squarer male orbits. The zygomatic arch was assessed through both linear measurements and surface curvature to identify the more pronounced projection and flatter curvature typical of male skulls. Additionally, the mandible was examined for its angle, which tends to be sharper in males compared to females, further supporting sex determination.

3.5.2 Torso Features extraction

After analyzing the skull, the system processed torso images to extract features for age and sex estimation. For age estimation, the focus was on developmental markers in the femur and clavicle.

The neck-shaft angle was calculated, and trochanter development was assessed by measuring protrusions from the femoral axis. Fusion states of the epiphyseal plates were evaluated using texture analysis (GLCM) in the metaphyseal regions, particularly looking at closure patterns between ages 16 and 25. For the clavicle, Sobel edge detection identified fusion status at the sternal end, which is important for age estimation between 22 and 30 years.

For sex estimation, the analysis concentrated on differences in pelvic morphology. Male skeletons usually have greater bone mass and density, so total pelvic volume was calculated by summing voxel values. This overall measurement was supported by regional analyses to capture specific differences.

For instance, the width and angle of the sciatic notch were measured using 3D edge detection, as females typically have a wider notch. The subpubic angle was assessed using landmark detection on axial slices, with female pelvises generally exceeding 90 degrees. The shape of the pelvic inlet was classified using moment invariants and aspect ratio calculations, distinguishing the broader circular female inlet from the narrower heart-shaped male inlet.

Additional metrics included the depth of the greater sciatic notch and the curvature of the pelvic brim, both calculated through surface normal analysis of the bone. These measurements highlighted the differences between male and female pelvises, with males showing larger sizes and thicker bones.

3.5.3 Lower extremities extraction

After analyzing cranial and torso features, the system processed lower extremity CT scans to extract features for age and sex estimation. This methodology focused on measurable differences in bone structure and development patterns.

For age estimation, the system examined growth plate fusion patterns in the femur and tibia, which occur according to predictable timelines during skeletal development. It assessed epiphyseal fusion status using texture analysis and edge detection algorithms applied to the metaphyseal regions. Bone lengths were measured from segmented reconstructions, accounting for natural curvature through alignment along the anatomical axis. Cortical thickness measurements also provided insights into age, as bone density and structure change with age. These metrics were compared against known developmental timelines to estimate age ranges.

For sex estimation, the analysis utilized recognized differences in lower extremity bone morphology. Male specimens typically have more robust bones, so total femoral and tibial volumes were calculated

by summing voxel values as key metrics of sexual dimorphism. This overall measurement was supported by regional analyses to capture specific sex differences. The analysis included measurements of femoral head diameter, cortical bone thickness, and bone lengths, all of which consistently show sexual dimorphism. Male specimens generally exhibited larger bone diameters, thicker cortices, longer bones, and more pronounced muscle attachment sites, indicating greater strength and robustness in male lower extremities.

3.6 Machine learning

The system integrated features extracted from the skull, torso, and lower extremities for classification purposes, with each region analyzed separately. For gender classification, we utilized XGBClassifier, while XGBRegressor was employed for age estimation. Both models are part of the XGBoost framework.

3.6.1 Skull analysis

The machine learning approach analyzed cranial features extracted from 3D skull reconstructions to estimate age and sex. Two specialized XGBoost models were developed: an XGBRegressor for continuous age prediction and an XGBClassifier for binary sex determination. Both models were carefully configured to address the challenges of our dataset comprising 202 individuals.

For age prediction, we used a learning rate of 0.01 to gradually capture the relationship between cranial suture fusion patterns (unfused, partially fused, fully fused) and biological age. The model's complexity was controlled by limiting tree depth to 4, preventing overfitting while maintaining interpretability. We incorporated L1 regularization (reg_alpha=0.1) to automatically select the most informative features by shrinking less important ones, such as minor texture variations, toward zero. Simultaneously, L2 regularization (reg_lambda=1.0) smoothed predictions by discouraging over-reliance on any single morphological trait.

The sex classification model employed a slightly higher learning rate (0.05) and deeper trees (max_depth=5) to better capture subtle differences in sexually dimorphic traits like frontal bone angle and orbital shape. The same L1/L2 regularization strengths (reg_alpha=0.1, reg_lambda=1.0) helped manage feature selection and prevent overfitting, particularly important given the imbalanced datset distribution.

Performance was evaluated through rigorous 5-fold cross-validation. The dataset was divided into five equal subsets (40-41 skulls each), with each subset serving as the test set once while the remaining four trained the model. Standard KFold partitioning sufficed for age prediction, while StratifiedKFold preserved the sex ratio in each fold for classification tasks.

Age prediction accuracy was measured using Mean Absolute Error (MAE), which provided intuitive interpretation in years. Sex classification performance was assessed through overall accuracy, supplemented by gender-specific metrics to ensure balanced results. This approach combined robust machine learning techniques with domain-specific considerations for anthropological applications.

3.6.2 Torso analysis

The extracted pelvic and femoral features were implemented in XGBoost models, using XGBRegressor for age estimation and XGBClassifier for sex determination, specifically optimized for torso characteristics. For age estimation, we configured three critical parameters: a conservative learning rate (0.01) to gradually capture maturation patterns, paired with L1 regularization (α=0.3) to emphasize the most predictive features like sciatic notch dimensions, and L2 regularization (λ=1.5) to smooth predictions across age groups. A tree depth of 5 enabled modeling complex feature interactions while maintaining interpretability.

The sex classification model employed distinct parameterization: a higher learning rate (0.1) accelerated convergence for dimorphism detection, balanced by equal L1/L2 regularization (α=1.0, λ=1.0) to prevent over-reliance on any single pelvic metric. Shallower trees (depth=4) optimally distinguished sex-specific morphology without overfitting. Both models used early stopping to dynamically determine boosting rounds.

We employed repeated stratified 5-fold cross-validation on our dataset of 202 individuals. The data was split into 5 subsets, ensuring each fold maintained class proportions. In each iteration, one fold served as the validation set while the other four (approximately 162 individuals) were used for training. This process was repeated five times, allowing each fold to be validated once. This approach confirmed that our parameter choices effectively balanced predictive power and generalization.

3.6.3 Lower extremities analysis

The extracted femur and tibia features were implemented in XGBoost models, using XGBRegressor for age estimation and XGBClassifier for sex determination. For age prediction, we employed a deeper tree structure (max_depth=6) to capture complex skeletal maturation patterns across different bone metrics, combined with a conservative learning rate (0.01) for stable optimization. The model incorporated strong L1 regularization (α=1.0) to emphasize key predictors like femur-tibia length ratios and cortical thickness, while enhanced L2 regularization (λ=1.5) smoothed predictions across the age spectrum. Gamma regularization (0.3) provided additional protection against overfitting to local bone morphology variations.

For sex classification, we implemented shallower trees (max_depth=4) to maintain generalizability, with balanced L1/L2 regularization (both 1.0) preventing dominance by any single extremity feature. The model used gamma=0.2 to control node splitting and a reduced learning rate (0.01) for precise optimization of dimorphic traits. These parameters worked synergistically with our engineered features - including senior-specific bone ratios and limb proportion metrics - to achieve robust classification.

The validation framework implemented 5-fold KFold for age regression and 10-fold StratifiedKFold for gender classification, ensuring robust evaluation across diverse age groups with customized binning. Similar to the previous analyses, mean absolute error (MAE) was used for age prediction, while accuracy was employed for gender classification.

Conclusion

The chapter outlines the methods and steps in the workflow used for this project, emphasizing the preprocessing steps, 3D reconstruction, feature extraction, and the machine learning process for age and gender estimation across different anatomical regions. In the next chapter, we will present the results obtained from these analyses, providing insights into the models' performance and discussing their implications for forensic applications. We will also explore how the findings contribute to the understanding of anatomical characteristics in relation to age and gender estimation.

Bibliography

[39] Fedorov A., Beichel R., Kalpathy-Cramer J., Finet J., Fillion-Robin J-C., Pujol S., Bauer C., Jennings D., Fennessy F., Sonka M., Buatti J., Aylward S.R., Miller J.V., Pieper S., Kikinis R. 3D Slicer as an Image Computing Platform for the Quantitative Imaging Network. Magnetic Resonance Imaging. 2012 Nov;30(9):1323-41. PMID: 22770690.

[40] P. Raybaut, "Spyder-documentation," 2009. [Online]. Available: pythonhosted.org.

[41] Edgar, HJH; Daneshvari Berry, S; Moes, E; Adolphi, NL; Bridges, P; Nolte, KB (2020). New Mexico Decedent Image Database. Office of the Medical Investigator, University of New Mexico. doi.org/10.25827/5s8c-n515.

[42] L. Innolitics, “Exposed Area Attribute – DICOM Standard Browser,” *Innolitics.com*, 2017. Available: https://dicom.innolitics.com/ciods/digital-x-ray-image/x-ray-acquisition-dose/00400303.

[43] N. Buhl, “Image Thresholding in Image Processing,” *encord.com*, Sep. 12, 2023. Available: https://encord.com/blog/image-thresholding-image-processing/

[44] K. Greenway, “Hounsfield unit | Radiology Reference Article | Radiopaedia.org,” *Radiopaedia*, Jan. 22, 2024. Available: https://radiopaedia.org/articles/hounsfield-unit

[45] A. Sachdev, “Closing (Morphological Operation) — Image Processing,” *Medium*, Mar. 17, 2024. Available: https://medium.com/@anshul16/closing-morphological-operation-image-processing-59a0ef6210e3

[46] “NIfTI: — Neuroimaging Informatics Technology Initiative,” *nifti.nimh.nih.gov*. Available: https://nifti.nimh.nih.gov/

CHAPTER IV

Results and Discussion

4 Results and Discussion

The chapter presents the results and discussion of our research on sex determination and age estimation using skeletal remains. We analyzed the accuracy of various skeletal regions, including the skull, torso, and lower extremities, to evaluate their effectiveness as indicators of biological sex and age. By examining these anatomical features, we aim to provide insights that can enhance forensic applications. Our results highlight the significance of anatomical variations and their implications for accurate sex classification and age estimation.

4.1 Sex Determination

The first aspect we analyzed was the accuracy of sex determination across three different skeletal regions: the skull, torso, and lower extremities. The table below illustrates the results of this analysis, highlighting the varying levels of accuracy in sex classification for each region.

Table 4:1 Overall accuracy of sex determination by skeletal region.

Region	Overall Accuracy
Skull	**90.46%**
Torso	86.13 %
Lower extremities	77.91%

The skull demonstrates superior performance at 90.46% accuracy, reflecting its numerous highly dimorphic features. The distinct anatomical differences between male and female skull, which include a variation in frontal bone slope, orbital shape and the mandible structure. These features provide multiple reliable indicators for sex classification. The torso, particularly the pelvis, follows closely at 86.13%, consistent with its reproductive adaptation that create clear anatomical differences in pelvic inlet shape, sciatic notch width and subpubic angle between sexes. Lower extremities on the other hand, show markedly reduced accuracy (77.91%), as limb bones rely more on generalized size and robustness rather than the distinct shape characteristics seen in cranial and pelvic structures.

When compared to recent advances in skeletal sex estimation, our cranial approach demonstrates clear advantages. Oura et al. [8] developed a deep learning model using LeNet5 architecture to analyze pQCT scans of the fourth lumbar vertebra, achieving 86.4% accuracy. Focusing on a single vertebral element lacks the needed markers available through comprehensive cranial analysis. Our method not only achieves higher accuracy but also uses multiple features for greater reliability.

The limitations of single-metric approaches become particularly apparent when examining Ikeda et al.'s [12] work on finger length ratios. Their study of 205 individuals using postmortem CT data found that a 6.0 mm difference in ring-to-index finger length could distinguish sex with only 64% sensitivity and specificity. This moderate performance reflects the variability in digit ratios. Unlike our multi-trait cranial method, which accounts for natural biological variation through multiple markers, finger ratio analysis depends entirely on a single measurement with substantial overlap between sexes. Furthermore, such metric-based approaches often require population-specific calibration, while cranial morphology maintains more consistent sexual dimorphism across different ethnic groups.

4.1.1 Sex-Specific Classification Performance

To gain a deeper understanding of the sex determination accuracy, we further analyzed the sex-specific classification performance across the three skeletal regions. The table below presents the accuracy rates for male and female identification, highlighting the disparities between sexes in each anatomical region.

Table 4:2 Sex-specific accuracy of determination by skeletal region.

Region	Male only accuracy	Female only accuracy
Skull	**91.50%**	**89.23%**
Torso	87.62%	84.31%
Lower extremities	84.40%	70.00%

All regions showed higher accuracy for male identification, most likely due to the fact that the sample data included more male samples than females. However, this effect varied substantially by anatomical region. The disparity was most pronounced in lower extremities, where female accuracy dropped to 70% compared to 84% for males. This likely reflects both the sample imbalance and the fundamental biological reality that male limb bones tend to follow more predictable robust patterns, while female morphology shows greater natural variability, particularly in regions like the limbs where dimorphic traits are less pronounced. Figure 4:1 depicts the sex accuracy across regions.

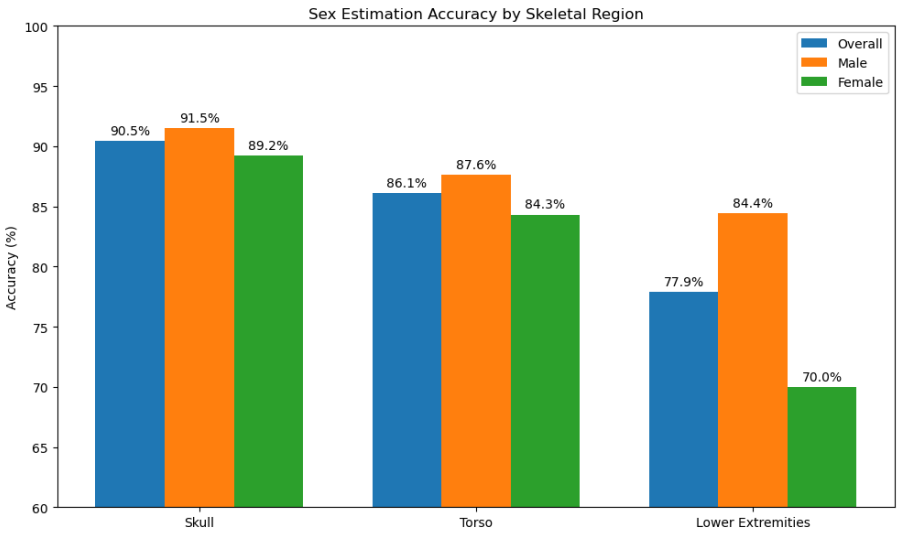

Figure 4:1 Sex accuracy across skeletal regions: overall, male-specific, and female-specific rates.

The skull maintained a balanced performance between sexes (94% for males and 89% for females), demonstrating how its multiple dimorphic features provide robust classification power, regardless of biological sex. Pelvic analysis showed a moderate accuracy gap favoring male identification, which may reflect both the sample characteristics and the broader natural variation in female pelvic anatomy that can occur due to childbirth history.

The results highlight that while all skeletal regions show some degree of sexual dimorphism, their reliability for sex determination varies based on both anatomical factors and sample characteristics. It showcase the strength and consistency of sexual dimorphism expressed in each region's anatomical structure. This has important implications in forensic applications, suggesting that experts should prioritize cranial and pelvic indicators while using limb bone data more cautiously. Additionally, the results highlight the need for balanced training samples that adequately represent both sexes equally. Population variation in skeletal sexual dimorphism further complicates this issue as genetic, environmental, and biomechanical factors can influence the expression of sex-specific traits differently across populations [47].

When compared to current methods in skeletal sex estimation, Ujaddughe et al.'s [13] focused analysis of the femur provides valuable insights that confirm our understanding of lower extremity dimorphism. Their CT-based study achieved classification accuracies of 82.5-91.4% for the femur, with males consistently showing higher accuracy rates (85-95%) compared to females (75-85%). These results align remarkably well with our broader findings for lower extremities, where we observed an 84.4% male versus 70% female accuracy rate, while demonstrating that the femur itself performs somewhat better than other limb bones.

The consistency between these studies is remarkable, especially considering their different methodologies. While our analysis examined lower extremities as a general category, Ujaddughe's approach using population-specific measurements confirms that the femur maintains stronger sexual dimorphism than other limb elements. Their finding shows that proximal femoral measurements (84.3% accuracy) outperformed distal ones (79.3%) helps explain why the femur shows better performance than other lower extremity bones in our data. This proximal-distal difference likely reflects the femur's unique biomechanical relationship to the pelvis and its more pronounced muscle attachment sites.

Both studies reveal the same persistent challenge in skeletal sex estimation: the male accuracy advantage remains significant regardless of methodology or specific bone examined. Ujaddughe's results, despite using a balanced sample (140 male and 140 female), still showed a 10-20% accuracy gap favoring males. This strongly supports our conclusion that this disparity reflects fundamental biological differences rather than methodological limitations, as male postcranial elements simply follow more robust and predictable morphological patterns, while female bones exhibit greater natural variability.

In conclusion, while cranial and pelvic features remain the standard for sex estimation, their reliability must be interpreted within the context of sample composition and population variability.

4.2 Age Estimation

After establishing sex determination, the next step is to estimate age. The table below presents the overall mean absolute error (MAE) in years for age estimation across different skeletal regions, highlighting the accuracy of age assessments in the skull, torso, and lower extremities.

Table 4:3 Overall Mean Absolute Error (MAE) by skeletal region.

Region	Overall MAE
Skull	**1.56 years**
Torso	2.58 years
Lower extremities	2.74 years

The skull demonstrated the highest precision in age estimation, with a mean absolute error (MAE) of just 1.56 years. This accuracy reflects the reliability of cranial suture fusion patterns as age markers. For instance, if an individual’s actual age is 35 years, our system would typically predict an age within

1.5 years younger or older. The skull achieves this level of precision through observable changes that occur at specific life stages.

In contrast, the torso showed a slightly reduced performance with an MAE of 2.58 years. This decrease can be attributed to the anatomical characteristics of post-cranial bones, such as the clavicle and pelvis, which undergo fewer distinct age-related changes compared to the skull. However, when available, these markers still provide reliable indicators for age estimation.

Analysis of the lower extremities exhibited the lowest accuracy among the regions, with an MAE of 2.74 years. While the skull provides multiple precise indicators and the torso offers reliable markers like clavicular fusion and pelvic changes, the long bones of the legs present greater challenges for age estimation in adults. This limitation stems from the fact that long bones complete their primary growth and development by late adolescence, resulting in fewer distinct markers for age beyond the ossification of growth plates.

Our research demonstrates significant advancements in skeletal age estimation accuracy compared to current methods. The skull emerged as the most reliable region, achieving an MAE of 1.56 years, an improvement over Pham et al.'s [10] mandible-focused approach (MAE: 7.07 years) and Imaizumi et al.'s [11] vertebral body results (MAE: 7.92 years). The superior performance of cranial analysis is due to the well-defined timeline of suture fusion patterns, which provide consistent age markers.

For post-cranial structures, our methods also outperformed existing techniques. The torso analysis yielded an MAE of 2.58 years by incorporating clavicular and pelvic markers, surpassing Imaizumi et al.'s iliac crest results (MAE: 9.49 years). In the lower extremities, our approach that combined femur metrics with tibial measurements and cortical thickness analysis achieved an MAE of 2.74 years, reducing error compared to single-bone analyses of the femur (Pham et al.: 5.74 years; Imaizumi et al.: 8.98 years). The results highlight that integrating multiple skeletal indicators provides more accurate age estimates than relying on isolated features.

The technical superiority of our approach is based on three key points. First, we expanded anatomical coverage beyond previous studies, including both cranial and post-cranial systems. Second, our advanced feature extraction captures a wider range of age-related changes, from macrostructural to microstructural characteristics. Third, the XGBoost algorithm proved more effective for modeling our multidimensional skeletal data than earlier support vector regression methods.

Finally, all regions meet the forensic standard for age estimation (± 10 years) [48]. However, accuracy varies by anatomical region due to differences in developmental and degenerative markers. The skull remains the optimal choice for precise estimates, while the torso provides reliable secondary indicators. In contrast, lower extremities are less precise in adulthood.

4.2.1 Age analysis across different age groups

To understand the accuracy of age estimation, it is important to analyze results across various age brackets. The table below presents the mean absolute error (MAE) for age estimation in the skull, torso, and lower extremities. This analysis reveals how age-related changes impact the reliability of skeletal indicators, providing insights crucial for forensic applications.

Table 4:4 Mean Absolute Error (MAE) by age bracket and skeletal region.

Age Bracket	Skull MAE	Torso MAE	Lower extremities MAE
18-20	1.39	**0.68**	2.03
20-30	1.17	2.93	3.20
30-40	1.26	2.30	2.52
40-50	**0.97**	1.83	1.36
50-60	1.35	1.93	1.65
60-70	1.53	3.79	3.55
70+	1.82	4.37	4.88

The results demonstrate a clear relationship between age estimation accuracy and specific anatomical developmental stages across different skeletal regions. This is depicted in the figure below, which illustrates the evolution of mean absolute error (MAE) by age brackets. The data reveals how various bones provide reliable age markers at different life phases, with precision varying according to the distinct biological processes occurring in each age group.

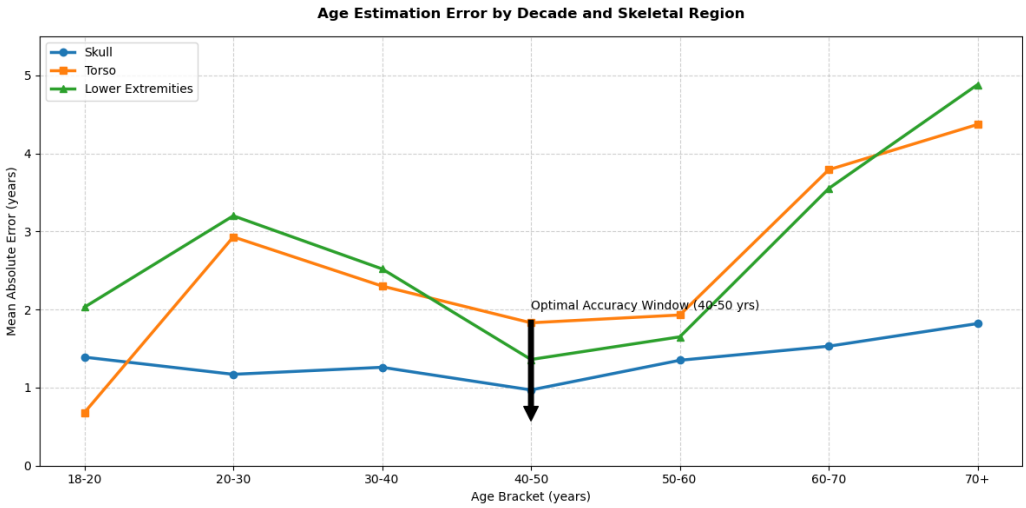

Figure 4:2 Mean Absolute Error in age estimation by decade and skeletal region.

For young adults (ages 18-30), the developmental period offers excellent age estimation potential through multiple skeletal indicators. Cranial analysis achieves strong results, with MAE values ranging from 1.39 to 1.17 years, due to the predictable fusion sequence of cranial features. For instance, the lambdoid suture typically completes fusion between the ages of 21 and 30. When examining an individual whose lambdoid suture appears nearly fused but retains partial openness, we can estimate an age of approximately 28-29 years. Similarly, the initiation of sagittal suture fusion around age 32 serves as another reliable indicator for this age group.

Postcranial elements show more variable performance during young adulthood. The torso demonstrates exceptional accuracy for the 18-20 age group, with an MAE of just 0.68 years. This high precision is primarily due to the medial epiphyseal fusion of the clavicles, which occurs between ages 22 and 25, creating a clear chronological marker. However, this precision decreases slightly after this fusion is complete, as subsequent changes in the sternal ribs occur more gradually.

In contrast, the lower extremities prove to be the least reliable, exhibiting an MAE ranging from 2.03 to 3.20 years. This is attributed to the transition from growth-based to degenerative-based markers, with femoral head fusion typically completing by age 18, limiting the availability of distinct indicators thereafter.

For middle-aged individuals, particularly those in the 40-50 age bracket, this period emerges as the most accurate for skeletal age estimation across all anatomical regions. This accuracy is largely due to the dataset containing the highest number of samples within this age range, providing our machine learning model with more training data. During this period, the skull stands out as the most precise indicator, achieving an impressive MAE of just 0.97 years. This remarkable accuracy can be attributed to the progressive fusion pattern of the coronal suture, which typically completes around age 50. The observable stages of fusion in the coronal suture serve as highly reliable age markers, especially when

combined with evidence from other sutures that fused earlier in life, creating a robust framework for precise age determination. This reliability is further evidenced by the skull's uniform performance across age groups, highlighted by the blue line in Figure 4:2.

After the age of 60, skeletal age estimation becomes increasingly unreliable due to the shift from developmental to degenerative changes in bone. While cranial sutures may still offer some consistency, postcranial elements exhibit much greater variability. This variability arises because degenerative markers depend not only on chronological age but are also heavily influenced by a wide range of biological and environmental factors. These include physical activity levels, nutritional history, body weight, systemic diseases, and genetic variation, all of which can significantly alter the rate and pattern of skeletal degeneration [49]. Consequently, two individuals of the same chronological age can display markedly different skeletal characteristics based on their life histories.

Such variability introduces significant sources of error in age estimation. Methods that rely on degenerative changes may overestimate age in individuals with physically demanding lifestyles or chronic illnesses, while underestimating it in those with lower activity levels or better health. Moreover, population differences in aging patterns complicate the application of standard reference data. Therefore, skeletal age estimation in older adults requires careful consideration of both intrinsic factors, such as genetics and health, and extrinsic factors, including labor history and nutrition, to reduce misinterpretation and improve accuracy.

Conclusion

The chapter analyzes the accuracy of sex and age estimation through skeletal remains, revealing how anatomical differences influence results. The skull consistently shows the highest accuracy for sex estimation, while the torso and lower extremities demonstrate greater variability in age estimation. These findings align with anatomical principles, emphasizing the unique features of each region. Such insights can be further explored in forensic science to improve estimation methods. Overall, the results highlight the importance of understanding skeletal anatomy in enhancing forensic practices and outcomes. This study lays a foundation for future research in this critical area.

Bibliography

[47] D. H. Ubelaker and C. M. DeGaglia, “Population variation in skeletal sexual dimorphism,” *Forensic Science International*, vol. 278, pp. 407.e1–407.e7, Sep. 2017, doi: https://doi.org/10.1016/j.forsciint.2017.06.012

[48] Mathew DG, Rajesh S, Koshi E, Priya LE, Nair AS, Mohan A. Adult forensic age estimation using mandibular first molar radiographs: A novel technique. J Forensic Dent Sci. 2013 Jan;5(1):56-9. doi: 10.4103/0975-1475.114552. PMID: 23960417; PMCID: PMC3746475.

[49] S. P. Nawrocki, "The nature and sources of error in the estimation of age at death from the skeleton," in *Age Estimation of the Human Skeleton*, K. E. Latham and M. Finnegan, Eds. Springfield, IL: Charles C. Thomas, 2010, pp. 79–101.

Conclusion

This Master's Thesis explored the integration of post-mortem computed tomography (PMCT) with machine learning techniques to estimate age and sex from skeletal remains, contributing significantly to the field of digital forensics. The study addressed a fundamental challenge in forensic science: the biological profile estimation of unidentified individuals. By developing a non-invasive, objective, and reproducible system, this research aimed to extract meaningful anatomical features from 3D reconstructions and use them in predictive models, overcoming the limitations of traditional forensic methods that often rely on subjective visual assessments and manual measurements.

The research focused on several major objectives. Firstly, it aimed to identify and quantify skeletal features indicative of age and sex. Secondly, it involved evaluating these features across multiple anatomical regions, including the skull, torso, and lower extremities. Lastly, the study aimed to apply and optimize XGBoost models to analyze the extracted features for accurate biological profile estimation. Using a comprehensive dataset of PMCT scans from the New Mexico Decedent Image Database (NMDID), the study ensured high-resolution volumetric data from a diverse demographic sample.

Methodologically, the study began with preprocessing of PMCT data, involving thresholding, filtering, and morphological operations to isolate bone structures. Precise 3D reconstructions using the NIfTI format ensured standardized and anatomically accurate representations of the skull, torso, and lower limbs. Feature extraction combined voxel-based measurements, anatomical landmark detection, morphological analysis, and texture features from the Gray-Level Co-occurrence Matrix (GLCM).

The machine learning component employed XGBoost, a gradient-boosted tree-based algorithm known for its robustness in handling structured data. Separate models were trained for regression (age estimation) and classification (sex determination), with performance evaluated using k-fold cross-validation to enhance generalizability. Results indicated that PMCT-based features from various anatomical regions offered differing predictive power, with cranial features yielding the highest accuracy for both age and sex estimation. Specifically, the skull achieved 90.5% accuracy for sex determination and a mean absolute error of 1.56 years for age estimation, underscoring the forensic significance of cranial anatomy. The findings showed that combining machine learning with PMCT imaging leads to objective and reproducible results that can match or exceed traditional forensic methods. This approach reduces subjective interpretation and uses detailed anatomical data.

Despite these advancements, the research encountered challenges, including the limited size and demographic imbalance of the dataset. This highlights the need for larger, more diverse PMCT databases. Additionally, while the segmentation and feature extraction processes were largely automated, expert oversight remained crucial for ensuring anatomical accuracy, especially in

overlapping regions. Another limitation is the difficulty of generalizing findings across different populations, as skeletal features can vary significantly by ancestry, nutrition, and environmental exposure. While this study focused on anatomically relevant regions, regional or population-specific models may be necessary to optimize performance in various contexts. Furthermore, legal and ethical considerations surrounding PMCT data access and use continue to pose challenges for broader adoption in routine forensic practice.

In conclusion, the Master's Thesis demonstrated that the combination of PMCT imaging and machine learning offers a promising solution to the challenges in forensic age and gender estimation. By developing and validating a robust analytical framework that uses advanced imaging, anatomical analysis, and artificial intelligence, this work represents a meaningful step toward modernizing forensic anthropology. Future research should focus on expanding datasets, refining anatomical segmentation techniques, and exploring advanced machine learning architectures. Ultimately, integrating these methods into forensic practice will support more accurate, ethical, and efficient identification processes, serving both scientific and societal needs.