People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University Echahid Cheikh Larbi Tebessi- Tebessa

Faculty of Exact Sciences and Natural and Life Sciences
Department of computer science
Domiciliation laboratory: laboratory of mathematics, informatics and systems (LAMIS)

Thesis

Presented and Publicly Defended for Obtaining the Diploma of Doctorate in Third Cycle

By:
Hazem Farah

Domain: Mathematics and computer science **Stream:** Computer science
Speciality: Artificial intelligence and its applications

Title

Contribution to the categorization of chest x ray radiographic

Defended on: 09/04/2026

Before the jury composed of:

Full name	Rank	University	
Pr.Bendjenna Hakim	Professor	Univ. of Tebessa	President
Pr.Akram Bennour	Professor	Univ. of Tebessa	Supervisor
Pr.Siam Abderrahim	Professor	Univ. of Khenchla	Examiner
Dr.Bendib Issam	MCA	Univ. of Tebessa	Examiner
Dr.Khediri Abderrazak	MCA	Univ. of Tebessa	Examiner

Academic year:2025-2026


بسم الله الرحمن الرحيم

Dedication

I dedicate this work to my beloved family, whose constant support, encouragement, and sacrifices have been the foundation of my journey.

To my parents, who instilled in me the values of perseverance, integrity, and the pursuit of knowledge, and whose unconditional love has always been my greatest strength.

To my brothers and sister, for their patience and encouragement during the long years of research.

To my friends and colleagues, who shared with me moments of challenge and inspiration.

Finally, I dedicate this thesis to all researchers and professionals striving to advance science and serve humanity through innovation and dedication.

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Professor BENNOUR Akram, from the Echahid ELchikh Laarbi Tbessi University in Tebessa, for his invaluable guidance and support throughout my years of research. His contribution to my project, his trust, and his encouragement were crucial to the completion of this work.

I would also like to extend my sincere appreciation also goes to Professor BENDJENNA Hakim, Director of the LAMIS laboratory and professor at Larbi Tébessi University – Tébessa, for kindly accepting to review and evaluate this work. His availability for enriching scientific discussions has greatly contributed to the advancement of this research.

I am equally grateful to Professor Bendib issam and doctor khediri abd errazak, from Larbi Tébessi University – Tébessa, and Professor Siam Abderahim, from Khenchla University, for their willingness to be part of this jury. It is truly a privilege to be able to share my work with them.

I would like to thank all my colleagues and friends from the LAMIS laboratory for the help they have provided.

I also wish to thank those who, during more difficult times, were there to read my work and help me move forward.

I extend my gratitude as well to my peers and to all those who contributed to improving my work through their reviews and constructive suggestions.

Last but not least, my heartfelt thanks go to my family and friends for their emotional support and encouragement, which have been a source of comfort in challenging moments.

الملخــص

تُعد صورة الأشعة السينية للصدر من أكثر وسائل التصوير الطبي استخدامًا في العالم، نظرًا لسهولتها، وانخفاض تكلفتها، وأهميتها التشخيصية في الممارسة السريرية. وبالإضافة إلى دورها الطبي، تلتقط صور الصدر بالأشعة السينية أيضًا البنى العظمية التي تتميز بالثبات النسبي مع مرور الزمن، مما يجعلها مرشّحة لتكون وسيلة جديدة للتعرّف البيومتري. وتبرز أهمية هذا التوجه بشكل خاص في الطب الشرعي والحالات ما بعد الوفاة، حيث تكون الوسائل البيومترية التقليدية مثل بصمات الأصابع، قزحية العين أو ملامح الوجه متدهورة أو غير متوفرة بسبب التحلل أو الإصابات. تبحث هذه الأطروحة في إمكانية استخدام صور الأشعة السينية للصدر كوسيلة بيومترية جديدة، وتقترح سلسلة من المساهمات لتطوير نظام متكامل للتعرّف على الهوية. المساهمة الأولى تقدم إثباتًا للمفهوم عبر استخدام تقنيات التعلم العميق، خصوصًا الشبكات العصبية السيامية المدربة بخوارزميةTriplet Loss ، لاستخراج خصائص مميزة من صور الصدر. وقد تم تقييم عدة نماذج عصبونية مدرَّبة مسبقًا (VGG16, VGG19, ResNet50/101, DenseNet) باعتبارها وحدات ترميز لاستخراج السمات.

المساهمة الثانية قدّمت شبكة الانتباه الذاتي المتبقي (SRAN) كعمود فقري معدل للشبكة، إذ تسمح هذه البنية بتركيز النموذج على الأجزاء التشريحية المهمة عبر آليات الانتباه المكاني والقنوي، مع تقليل تأثير العناصر غير المهمة، مما حسّن بشكل ملحوظ من قوة التمييز والموثوقية.

المساهمة الثالثة تناولت مشكلة الأنسجة الرخوة التي تتأثر بالأمراض وتتحلل بعد الوفاة. لمواجهة هذا التحدي، تم اعتماد نموذج Attention U-Net لعزل الهياكل العظمية (الأضلاع، الترقوة، القص، والفقرات)، وذلك لضمان أن النظام البيومتري يعتمد فقط على الخصائص الثابتة طويلة الأمد.

وأخيرًا، تجمع المساهمة الرابعة بين عملية التقسيم والتعلم المعتمد على الانتباه لتقديم إطار متكامل للتعرّف البيومتري المعتمد على الهيكل العظمي للصدر. يقوم هذا النظام بتوليد متجهات مدمجة تمكّن من عمليات التحقق (1:1) والتعرّف (1(N:، مما يجعله مناسبًا جدًا في البيئات الشرعية حيث قد تتوفر فقط البنى العظمية.

تم تقييم المنهج المقترح باستخدام قواعد بيانات عامة واسعة النطاق، مثل NIH ChestX-ray14 للتدريب وCheXpert للاختبار. وأظهرت النتائج أن صور الأشعة السينية للصدر، وخاصة عند التركيز على العظام، توفر خصائص هوية دقيقة وموثوقة. وبهذا، تقدم الأطروحة مساهمة علمية في مجال البيومترية المدعومة بالذكاء الاصطناعي، كما تقدم حلاً عمليًا في مجال الطب الشرعي، مما يفتح الطريق أمام أنظمة تحديد هوية قوية في الحالات ما بعد الوفاة.

الكلمات المفتاحية القياسات الحيوية، التعرّف القائم على الهيكل العظمي، التعرّف بعد الوفاة، علوم الأدلة الجنائية، التعلّم لأشعة السينية للصدر ، تقسيم الصور باستخدام شبكة العميق، الشبكة السيامية، شبكة الانتباه ذات البقايا الذاتية.

Abstract

Chest X-ray radiography is the most widely performed imaging modality worldwide, valued for its accessibility, affordability, and diagnostic relevance in clinical practice. Beyond its medical applications, chest X-rays also capture skeletal structures that are distinctive and relatively stable over time, which opens the possibility of using them as a novel biometric modality. This potential is particularly relevant in forensic and postmortem contexts, where traditional biometric traits such as fingerprints, iris, and facial features are often degraded or unavailable due to decomposition or trauma. This thesis investigates the feasibility of chest X-rays as a biometric trait and proposes a series of contributions toward developing a robust identification system.

The first contribution establishes a proof of concept by applying deep learning architectures, specifically Siamese networks trained with triplet loss, to chest radiographs. Several pre-trained convolutional neural networks (VGG16, VGG19, ResNet50/101, and Dense-Net) were evaluated as backbone encoders to extract discriminative features for person identification.

Building on this foundation, the second contribution introduces the Self-Residual Attention Network (SRAN), a modified backbone designed to improve feature representation. By integrating spatial and channel attention mechanisms with residual connections, SRAN enables the system to emphasize critical anatomical regions while suppressing irrelevant information, thereby enhancing robustness and discrimination.

The third contribution addresses the limitations posed by soft tissues, which are susceptible to pathological variations and postmortem degradation. To overcome this, an Attention U-Net segmentation model was employed to isolate skeletal structures such as ribs, clavicles, sternum, and vertebrae. This step ensures that the biometric system relies exclusively on stable anatomical features.

The final contribution integrates segmentation and attention-based feature learning into a complete skeletal-based chest X-ray biometric framework. This system generates compact embeddings that allow both verification (1:1) and identification (1:N), making it suitable for forensic scenarios where only skeletal information may be available.

The proposed approach was evaluated using large-scale public datasets, including NIH ChestX-ray14 for training and CheXpert for testing. The results demonstrate that chest X-rays, particularly when focused on skeletal structures, offer distinctive and reliable identity cues. By introducing chest radiographs as a new biometric modality, this work contributes to both the scientific field of AI-driven biometrics and the practical domain of forensic science, offering a pathway toward robust postmortem and medico-legal identification.

Key words: Chest X-ray (CXR), Biometrics**,** Skeletal-based identification**,** Postmortem identification**,** Forensic science**,** Deep learning**,** Siamese network**,** Triplet loss**,** Self-Residual Attention Network (SRAN)**,** Attention U-Net segmentation

Résume

La radiographie thoracique est la modalité d'imagerie la plus réalisée au monde, appréciée pour son accessibilité, son faible coût et sa valeur diagnostique en pratique clinique. Au-delà de ses applications médicales, la radiographie thoracique capture également des structures squelettiques distinctives et relativement stables dans le temps, ouvrant ainsi la voie à son utilisation comme modalité biométrique novatrice. Ce potentiel est particulièrement pertinent dans les contextes médico-légaux et post-mortem, où les traits biométriques classiques tels que les empreintes digitales, l'iris et les caractéristiques faciales sont souvent dégradés ou indisponibles en raison de la décomposition ou des traumatismes.Cette thèse explore la faisabilité de l'utilisation des radiographies thoraciques comme modalité biométrique et propose une série de contributions pour le développement d'un système d'identification robuste. La première contribution établit une preuve de concept en appliquant des architectures d'apprentissage profond, notamment des réseaux siamois entraînés avec la fonction de perte triplet, aux radiographies thoraciques. Plusieurs réseaux neuronaux convolutifs pré-entraînés (VGG16, VGG19, ResNet50/101 et DenseNet) ont été évalués comme encodeurs pour extraire des caractéristiques discriminatives pour l'identification des individus.

La deuxième contribution introduit le Self-Residual Attention Network (SRAN), une architecture modifiée visant à améliorer la représentation des caractéristiques. En intégrant des mécanismes d'attention spatiale et canal avec des connexions résiduelles, le SRAN permet au système de mettre en évidence les régions anatomiques critiques tout en supprimant les informations non pertinentes, améliorant ainsi la robustesse et la capacité de discrimination.

La troisième contribution traite des limites liées aux tissus mous, sensibles aux variations pathologiques et à la dégradation post-mortem. Pour y remédier, un modèle de segmentation Attention U-Net a été utilisé pour isoler les structures squelettiques telles que les côtes, les clavicules, le sternum et les vertèbres. Cette étape garantit que le système biométrique repose exclusivement sur des caractéristiques anatomiques stables.

Enfin, la quatrième contribution intègre la segmentation et l'apprentissage basé sur l'attention dans un cadre complet d'identification biométrique basé sur le squelette thoracique. Ce système génère des représentations compactes permettant la vérification (1:1) et l'identification (1:N), le rendant particulièrement adapté aux scénarios médico-légaux où seules les structures osseuses peuvent être exploitées.

L'approche proposée a été évaluée à l'aide de bases de données publiques à grande échelle, notamment NIH ChestX-ray14 pour l'entraînement et CheXpert pour les tests. Les résultats démontrent que les radiographies thoraciques, en particulier lorsqu'elles se concentrent sur les structures squelettiques, offrent des indices identitaires distinctifs et fiables. En introduisant la radiographie thoracique comme nouvelle modalité biométrique, ce travail contribue à la fois au domaine scientifique de la biométrie basée sur l'IA et au domaine pratique de la médecine légale, ouvrant la voie à une identification robuste en contexte post-mortem et médico-légal.

Les mots clés : Radiographie thoracique (CXR), Biométrie, Identification basée sur le squelette, Identification post-mortem, Médecine légale, Apprentissage profond Siamese network**,** Triplet loss**,** Self-Residual Attention Network (SRAN)**,** Attention U-Net segmentation

Contents

List of Tables

List of Figures

General Introduction

General Introduction

Medical imaging plays a central role in modern medicine, supporting physicians in diagnosis, treatment planning, monitoring disease progression, and research. Among the different imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), ultrasound, and positron emission tomography (PET) chest X-ray radiography (CXR) remains the most frequently used worldwide [1]. Its accessibility, relatively low cost, non-invasiveness, and speed of acquisition make it an indispensable diagnostic tool, especially in low- and middle-income countries where advanced modalities are less available. Chest radiographs are collected at an enormous scale, forming one of the largest digital medical imaging archives in existence. They are routinely performed to evaluate respiratory and cardiac conditions, including pneumonia, tuberculosis, lung cancer, heart enlargement, and more recently, COVID-19 [2,3].

Beyond their clinical significance, chest X-rays present unique properties from an anatomical perspective [4]. Unlike many other radiological images, chest radiographs capture both soft tissues and skeletal structures within the thoracic cavity. The skeletal components such as ribs, clavicles, sternum, and vertebrae are highly distinctive across individuals and maintain stability over long periods [5]. This makes chest radiographs particularly promising not only for clinical applications but also for biometric and forensic purposes, as explored in this thesis.

Artificial Intelligence in Chest X-Ray Analysis

Over the past decade, artificial intelligence (AI), and more specifically deep learning, has transformed medical image analysis [6,7]. Convolutional neural networks (CNNs) and their extensions have achieved human-level or even superhuman performance in image classification, segmentation, and detection tasks [8]. Within chest X-ray analysis, deep learning has been applied to:

- **Disease detection:** Identifying thoracic pathologies such as pneumonia, tuberculosis, lung nodules, cardiomegaly, and pleural effusion.
- **Multi-label classification:** Recognizing multiple disease patterns within a single radiograph.
- **Localization and segmentation:** Isolating lungs, heart, or lesions to aid diagnosis.
- **Image enhancement:** Reducing noise or reconstructing high-quality images from low-dose exposures.

Large public datasets such as NIH ChestX-ray14, CheXpert, and MIMIC-CXR have fueled this revolution, enabling reproducible and scalable research. These datasets provide not only raw images but also metadata and pathology labels, accelerating the development of AI-powered diagnostic systems.

However, most of these efforts remain focused on clinical pathology detection, overlooking another important potential application: person identification. While AI has demonstrated its power in recognizing diseases, its ability to capture identity-specific features embedded in radiographs has received little attention. This gap forms the foundation of the present research.

The Identity Problem and Biometrics

Establishing identity is a crucial requirement across many domains, from secure access control and law enforcement to healthcare and disaster victim identification (DVI). Traditional authentication methods rely on passwords, ID cards, or tokens, but these are vulnerable to theft, duplication, or loss [9]. To overcome these limitations, biometric systems have emerged, leveraging inherent physiological or behavioral traits of individuals. Common biometric modalities include fingerprints, facial recognition, iris scans, voice, and gait [10].

While these systems are widely used, they face significant challenges in forensic and postmortem scenarios. Soft tissues degrade rapidly after death due to decomposition, burning, or trauma, making fingerprints, iris, or facial recognition ineffective. Even DNA, while highly reliable, is time-consuming and expensive to analyze, limiting its usefulness in large-scale or time-sensitive investigations[11-13].

Forensic science therefore requires new biometric modalities that are both distinctive and resistant to postmortem changes. This is where chest X-rays, and particularly their skeletal features, provide a compelling solution.

Chest X-Rays as a Biometric Modality

Chest radiographs capture unique anatomical structures of the thoracic skeleton, such as rib cage morphology, clavicular shapes, and vertebral configurations [14]. These structures vary across individuals due to genetics, lifestyle, and health factors, but remain stable over time and resistant to decomposition. Unlike facial features or soft tissues, bones persist long after death, making them particularly suitable for postmortem identification [15] .

Moreover, chest radiographs are widely available in hospital archives and often linked to patient metadata, making them practical for retrospective identification in forensic cases. For instance, in scenarios such as airplane crashes, mass disasters, or unidentified remains, chest radiographs collected during an individual's lifetime can be matched against postmortem imaging to establish identity. This dual availability pre-mortem hospital records and post-mortem forensic radiographs offers a unique advantage not found in many other biometric modalities [16].

Nevertheless, leveraging chest X-rays for biometrics presents significant challenges. Radiographs contain a mixture of soft tissues and skeletal structures, with soft tissues prone to disease-related variations or postmortem degradation. To address this, specialized methods are required to isolate skeletal features and extract robust identity-specific representations.

Objectives of the Thesis

The central objective of this thesis is to investigate, design, and validate a chest X-ray based biometric system, with a specific focus on forensic and postmortem applications. More concretely, this research aims to:

1. Demonstrate the feasibility of chest radiographs as a new biometric modality.
2. Develop deep learning architectures capable of extracting discriminative identity features.
3. Introduce attention mechanisms and segmentation to enhance robustness and interpretability.
4. Provide a skeletal-based system resilient to decomposition and pathological variations.
5. Evaluate performance on large-scale public datasets and forensic-relevant settings.

Contributions of the Thesis

This thesis is structured around four main contributions that progressively address the challenges identified above:

1. **Proof of concept:** We designed a Siamese neural network trained with triplet loss to assess whether chest radiographs carry identity-specific information. Using pretrained CNN backbones such as VGG16, VGG19, ResNet50/101, and DenseNet, we demonstrated that even raw chest X-rays can achieve meaningful identification.
2. **Enhanced feature learning with attention:** To overcome the limitations of generic CNNs, we introduced the Self-Residual Attention Network (SRAN) as a modified backbone. This architecture integrates spatial and channel attention with residual connections, enabling the system to dynamically focus on critical thoracic structures while suppressing irrelevant regions.
3. **Bone segmentation:** Recognizing the instability of soft tissues in postmortem contexts, we employed an Attention U-Net to segment skeletal structures. By isolating ribs, clavicles, sternum, and vertebrae, we ensured the biometric system relies on stable and distinctive features.
4. **Final skeletal-based biometric framework:** We integrated segmentation into a complete identification system. The architecture produces compact embeddings that enable both verification (1:1) and identification (1:N). By focusing on skeletal anatomy, the system offers resilience against decomposition and is particularly suited for forensic applications.

Significance and Impact

The research presented in this thesis introduces chest X-rays as a novel biometric modality, extending the scope of medical imaging beyond clinical diagnosis. Its contributions are threefold:

- **Scientific significance:** It pioneers skeletal-based biometric identification using chest radiographs, a domain with little prior exploration.
- **Technical innovation:** It integrates deep learning, triplet loss, attention mechanisms, and segmentation into a coherent framework.
- **Forensic relevance:** It provides a practical solution for postmortem and medico-legal identification, where traditional biometrics fail.

By bridging the gap between healthcare archives and forensic needs, this research opens new pathways for identity verification in contexts where resilience, robustness, and reliability are paramount. The structure of this thesis is organized as follows. It begins with a General Introduction, which presents the context, motivations, problematics, objectives, and overall contributions of the research. Chapter I provides the scientific context by focusing on chest X-ray radiographs, outlining their principles, medical importance, limitations, and emerging uses beyond clinical diagnosis. Chapter II introduces the field of biometrics and reviews related works, discussing conventional biometric modalities, their limitations in forensic and postmortem scenarios, and recent approaches that leverage medical images for identification. The last two chapters constitutes the core of this thesis and presents its original scientific contributions. Chapter 3 focuses on deep learning based biometric identification using full chest X-ray radiographs. It introduces Siamese network architectures trained with metric learning strategies and presents the integration of self-residual attention mechanisms to enhance feature discrimination and robustness. This chapter establishes the feasibility of chest X-ray images as a biometric modality while highlighting the limitations of appearance-based representations.

Chapter 4 builds upon the findings of Chapter 3 by addressing its identified limitations through anatomically constrained modeling. It introduces a deep learning–based bone segmentation framework and proposes a skeletal-based biometric identification system designed to improve robustness, particularly in forensic and postmortem scenarios. Finally, the thesis concludes with a General Conclusion, which summarizes the findings, highlights the scientific and forensic contributions, and proposes perspectives for future research in this emerging field.

CHAPTER I :

Chest X-ray radiographs

Chapter I : Chest X-ray radiographs

1. Introduction

In the field of medical diagnostics, the use of medical image processing has played a pivotal role in supporting clinical decision-making. Across the wide spectrum of radiological imaging techniques ranging from ultrasound, X-ray, magnetic resonance imaging (MRI), computed tomography (CT), and positron emission tomography (PET) each modality offers unique methods for capturing internal body images[17,18]. Despite these technological advances, it remains a fact that only specific sections or regions within these radiological images provide clinically significant information for the consulting physician [19]. Moreover, there are several reasons why both pathologists and radiologists recognize that the images produced through such radiological examinations do not always deliver fully accurate or comprehensive diagnostic details[4]. In cases of minor or less critical diseases, such limitations may not drastically impact the diagnosis. However, in more severe or complex medical conditions, these shortcomings can potentially lead to significant clinical consequences.

Another important aspect to consider is the dual nature of chest X-rays: while they are indispensable in clinical practice, they also present intrinsic limitations. The images often display overlapping anatomical structures that may complicate interpretation. For example, ribs can obscure portions of the lung fields, while clavicles and sternum may conceal or distort other thoracic details. In female patients, breast tissue can introduce additional shadows, and in some cases, abdominal structures may also appear within the field of view. These overlapping elements create challenges for clinicians, as diagnostically relevant features may be partially hidden or confounded by surrounding anatomy [20].

Despite these limitations, chest radiography continues to be one of the most widely used diagnostic tools worldwide. Its advantages are clear: rapid acquisition, relatively low radiation dose compared to CT scans, affordability, and broad availability even in resource-limited settings. These qualities make chest X-rays the first-line imaging modality in many clinical workflows, whether in emergency medicine, primary care, or large-scale public health screenings. In addition to traditional diagnostic use, chest X-rays have also found their place in emerging technological applications, ranging from computer-aided detection of thoracic diseases to novel approaches in biometric person identification[21,22].

In this thesis, we place chest X-rays at the center of our investigation. Beyond their established clinical importance, we argue that their unique ability to capture both soft tissues and skeletal structures makes them a promising resource for identity identification and verification, especially in forensic and postmortem contexts. This chapter provides a comprehensive overview of chest radiography, highlighting not only its medical applications but also its potential in new research domains such as artificial intelligence and biometrics.

2. What Are Chest X-Ray Radiographs?

Chest X-ray radiography is a diagnostic imaging procedure that captures a two-dimensional representation of the chest's internal structures using ionizing radiation [23]. The technique involves exposing the chest area to a small, controlled amount of X-rays, which pass through the body and are absorbed by tissues at different rates. This difference in absorption creates a contrast image: bones, being denser, absorb more X-rays and appear white on the image, whereas air-filled structures such as the lungs allow more X-rays to pass through and thus appear darker [24].

A standard chest X-ray typically consists of two views:

- **Posteroanterior (PA) view**: This is the most common view, where the X-ray beam passes from the back (posterior) to the front (anterior) of the patient, illustrate in figure 1, who stands upright with the chest against the image receptor [25].

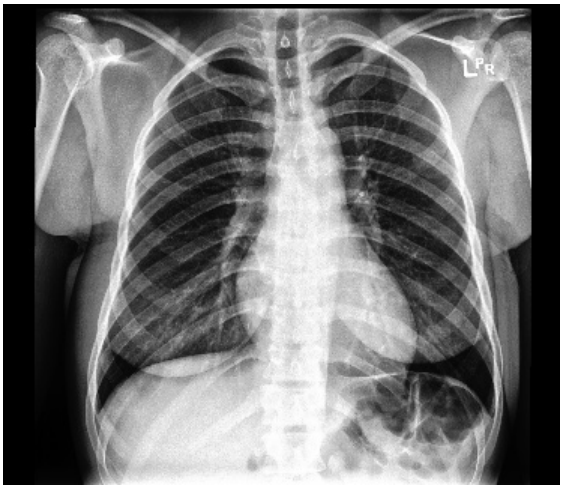


Figure 01 : Frontal chest x ray radiographic

- **Lateral view**: An additional side view is often taken to provide further anatomical detail and assist in localizing abnormalities [26], ateral chest view illustrate in figure 2.

Figure 02 : Ateral chest x ray radiographic

Other specialized views may be used based on clinical need, such as the anteroposterior (AP) view (often used when the patient cannot stand), lateral decubitus (lying down) views for evaluating fluid levels, or oblique views for detailed examination of certain structures.

2.1. Technical Considerations

The quality and diagnostic value of a chest X-ray depend on several technical factors. Radiographers adjust parameters such as exposure time, tube voltage (kVp), and current (mAs) to balance image clarity with minimizing radiation exposure. Modern imaging departments increasingly use digital radiography (DR) systems, replacing older analog film systems. Digital systems offer numerous advantages, including faster image acquisition, enhanced image processing capabilities, reduced need for repeat imaging, and easy integration with electronic medical records [27].

Digital chest X-rays also allow for advanced post-processing techniques, such as contrast adjustment, zooming, and highlighting of specific regions of interest, which can improve diagnostic accuracy. The widespread adoption of digital imaging has also facilitated the development of computer-aided diagnosis (CAD) systems and artificial intelligence (AI) tools, which are becoming more common in modern radiology practice [28].

3. What Can We Do with Chest X-Ray Radiographs?

Chest X-ray radiographs serve as a cornerstone in clinical diagnostics due to their versatility. The broad range of applications makes them an indispensable tool across various medical specialties. Their utility spans from routine health checkups to critical emergency assessments, reflecting their capacity to provide valuable information quickly and efficiently [29].

3.1. Diagnostic Purposes

One of the primary uses of chest X-rays is the diagnosis of diseases affecting the thoracic organs, particularly the lungs, heart, and chest wall. Physicians rely on chest X-rays to investigate symptoms such as cough, chest pain, shortness of breath, or unexplained weight loss. By analyzing the radiograph, clinicians can detect abnormalities like infections, tumors, or fluid accumulation [30].

3.2. Screening Programs

In many countries, chest X-rays are integral to population screening programs, particularly for infectious diseases such as tuberculosis (TB) [31]. They are also employed in pre-employment screening, immigration health checks, and periodic assessments for workers in industries with high exposure to respiratory hazards, such as mining or chemical manufacturing [32].

3.3 Monitoring and Follow-up

Chest X-rays are invaluable for monitoring disease progression and evaluating treatment responses. For instance, they are frequently used in follow-up assessments of pneumonia, tuberculosis, lung cancer, and congestive heart failure. Changes in radiographic findings over time help guide therapeutic decisions [33].

3.4 Emergency Medicine Applications

In emergency and critical care settings, chest X-rays are often the first imaging tool employed due to their speed and accessibility. They help rapidly identify life-threatening conditions such as pneumothorax (collapsed lung), hemothorax (blood in the pleural space), pulmonary embolism signs, and acute heart failure [34].

3.5 Non-Clinical and Research Applications

Beyond their traditional clinical roles, chest X-rays have emerged in research fields exploring forensic identification, age estimation, and gender determination. With the advancement of machine learning algorithms, chest X-ray images are being repurposed for non-diagnostic tasks, which opens up intriguing possibilities in security and forensic science [35].

4. Pathologies Detectable with Chest X-Ray Radiographs

Chest X-rays are capable of revealing a wide array of pathological conditions affecting different regions of the thoracic cavity. Below, we classify and elaborate on the most frequently encountered pathologies detected through chest radiography [36].

4.1 Pulmonary Diseases

- Pneumonia: Appears as localized or diffuse areas of increased opacity within the lung fields, often accompanied by air bronchograms.

- Tuberculosis (TB): Characterized by upper lobe infiltrates, cavitary lesions, and sometimes calcified nodules in chronic cases.
- Chronic Obstructive Pulmonary Disease (COPD): Shows signs such as hyperinflation of the lungs, a flattened diaphragm, and reduced vascular markings.
- Lung Cancer: May present as a solitary pulmonary nodule, mass, or more subtle signs like mediastinal widening or hilar lymphadenopathy.
- Pulmonary Edema: Associated with congestive heart failure, typically seen as bilateral perihilar "bat-wing" opacities.

4.2 Cardiac and Vascular Conditions

- Cardiomegaly: Identified by an enlarged cardiac silhouette, often quantified using the cardiothoracic ratio.
- Congestive Heart Failure: Manifests through an enlarged heart, pulmonary vascular congestion, interstitial edema, and pleural effusions.
- Aortic Aneurysm and Dissection: May be suspected from a widened mediastinum or abnormal contour of the aortic arch.

4.3 Pleural and Diaphragmatic Abnormalities

- Pneumothorax: Recognized by the absence of lung markings and a visible pleural line, indicating collapsed lung tissue.
- Pleural Effusion: Characterized by homogenous opacification and blunting of the costophrenic angle.
- Diaphragmatic Abnormalities: Elevation or irregular contour of the diaphragm can signify paralysis, hernia, or subdiaphragmatic pathology.

4.4 Skeletal Abnormalities

- Rib Fractures: Detected as discontinuities in the bony cortex of ribs, sometimes with associated soft tissue changes.
- Spinal Deformities: Scoliosis and kyphosis are apparent on chest X-rays through the curvature and angulation of the spine.
- Bone Lesions: Metastatic deposits from cancers elsewhere in the body may manifest as lytic or sclerotic lesions in ribs or vertebrae.

4.5 Mediastinal Masses

- Mediastinal Tumors or Lymphadenopathy: Enlarged lymph nodes or masses can widen the mediastinum and alter the normal contour of the heart and trachea.

5. Chest X-Ray Radiographs in Biometric Person Identification

In recent years, an exciting and innovative use of chest X-ray radiographs has emerged in the field of biometrics the science of identifying individuals based on unique physiological or anatomical features. Traditionally, biometric systems have relied on external traits such as

fingerprints, facial structure, iris patterns, or voice. However, these traits can be subject to changes due to age, injuries, cosmetic alterations, or deliberate spoofing attacks [37].

Chest X-rays offer a novel approach to biometrics by focusing on internal anatomical structures, particularly the rib cage, spinal column, clavicles, and overall thoracic skeletal configuration. These internal features are much more stable over time and less prone to intentional alteration or forgery [38]. Recent studies have demonstrated that machine learning and deep learning models can effectively extract unique skeletal patterns from chest radiographs, allowing for accurate identification of individuals. This approach holds significant potential for forensic science, disaster victim identification, and even secure access systems where higher resilience against identity fraud is required. Moreover, biometric identification through chest X-rays can be particularly useful in post-mortem identification scenarios, where external biometrics may be compromised due to trauma or decomposition, but skeletal structures remain largely intact. The growing integration of AI with chest X-ray imaging systems promises to further enhance the reliability and efficiency of such identification methods, paving the way for their future adoption in real-world forensic and security applications [39-42].

6. Datasets

The selection of appropriate datasets plays a critical role in medical science. To this end, two of the most widely recognized public chest radiograph collections were employed: the NIH ChestX-ray14 dataset and the CheXpert dataset.

6.1. NIH ChestX-ray14 Dataset

The NIH ChestX-ray14 collection is among the largest publicly available repositories of chest radiographs and has been extensively used in the medical imaging community. It contains 112,120 frontal-view X-ray images collected from 30,805 unique patients. On average, each patient is represented by three to four images, often including follow-up examinations. The images are stored in 8-bit PNG format with a resolution of 1024 × 1024 pixels.In addition to the radiographs, the dataset provides detailed metadata for each entry, including patient age, gender, anonymized patient ID, number of follow-up visits, projection view (posterior-anterior or anterior-posterior), and diagnostic labels covering up to 14 thoracic pathologies as well as the “no finding” category. This rigorous anonymization process makes the dataset suitable for large-scale public research [43].

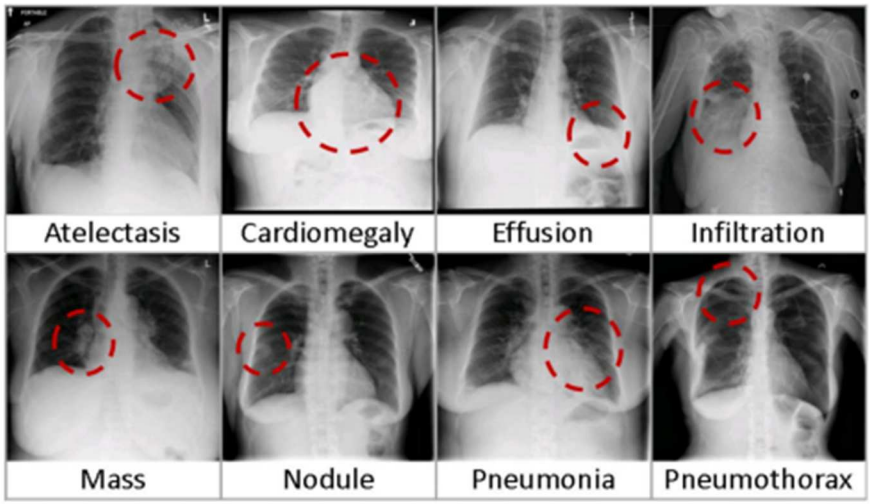


Figure 3: Eight visual examples of common thorax diseases [43]

The dataset's combination of large sample size, patient diversity, and longitudinal follow-up makes it a valuable resource for evaluating biometric recognition approaches, as it allows the system to be tested on repeated images of the same individual over time. Figure 3 shows a samples of radiographs with pathologies detected.

6.2. CheXpert Dataset

The CheXpert dataset is one of the largest and most widely used collections of chest radiographs for research in medical image analysis. Released in 2019 by Stanford University, it was specifically designed to advance the development of artificial intelligence systems in radiology. The dataset contains 224,316 chest radiographs from 65,240 unique patients, making it one of the most comprehensive publicly available resources in the field.Each radiograph is accompanied by labels for 14 thoracic observations, including common conditions such as pneumonia, cardiomegaly, pleural effusion, atelectasis, and lung opacity. Labels were automatically extracted from radiology reports using natural language processing (NLP) tools and later validated to ensure clinical relevance. In addition to disease labels, the dataset also includes patient identifiers, which are particularly valuable in tasks requiring longitudinal analysis or person-level recognition. One of the key strengths of CheXpert lies in its diversity: images were collected from a large and varied patient population, covering a wide range of ages, genders, and clinical scenarios. This makes the dataset robust for training models intended for real-world applications. Another important characteristic is the label uncertainty policy, where some findings are marked as uncertain (U) instead of strictly positive or negative. This better reflects clinical practice, where certain conditions are ambiguous and open to interpretation, and it challenges models to handle real-world uncertainty [44].

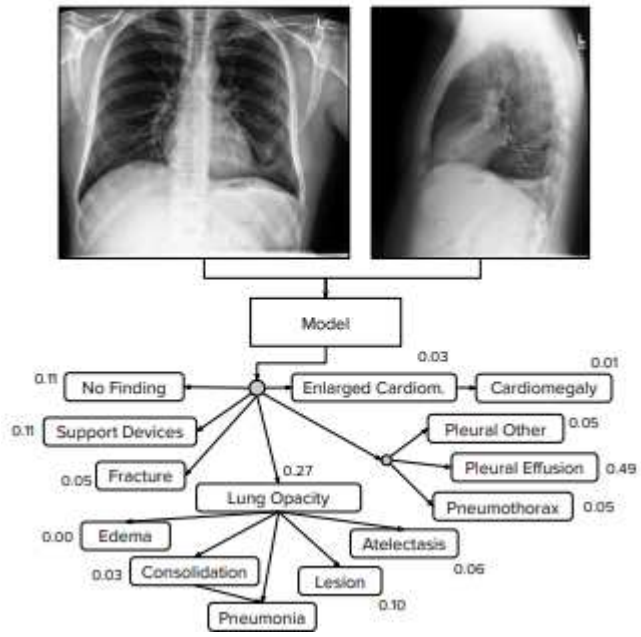


Figure 4: The CheXpert task is to predict the probability of different observations from multi-view chest radiographs [44].

6.2. MIMIC-CXR dataset:

The MIMIC-CXR dataset is a large-scale, publicly available collection of chest radiographs released as an extension of the Medical Information Mart for Intensive Care (MIMIC) initiative, with the aim of supporting advanced research in medical image analysis and clinical decision support. The dataset comprises more than 370,000 chest X-ray images acquired from over 65,000 patients, with each image paired with a corresponding free-text radiology report generated during routine clinical practice. Images are provided in both posteroanterior (PA) and anteroposterior (AP) views, capturing the variability inherent to real-world imaging conditions, particularly in critical care environments where portable imaging is frequently used. Disease labels were automatically extracted from radiology reports using natural language processing techniques, resulting in annotations for a wide range of thoracic findings while also introducing a degree of label uncertainty typical of large retrospective clinical datasets. Beyond image–label pairs, MIMIC-CXR includes rich metadata and is fully linked to the broader MIMIC-IV clinical database, enabling multimodal investigations that combine imaging data with longitudinal patient information. Owing to its scale, diversity, and clinical realism, MIMIC-CXR has become a reference benchmark for developing and evaluating deep learning models for chest X-ray analysis, including representation learning, transfer learning, and explainable artificial intelligence (XAI) methods. Despite its strengths, the dataset primarily reflects clinical diagnostic scenarios rather than forensic or post-mortem conditions; nevertheless, its extensive anatomical variability and large subject population make it a valuable resource for learning robust skeletal and structural representations from chest radiographs.

7. Conclusion

Chest X-ray radiography continues to be a foundational tool in modern medicine, offering invaluable insights into the structure and function of the thoracic cavity. Its widespread use is driven by its accessibility, affordability, and diagnostic versatility, making it indispensable

across clinical, emergency, and public health settings. From diagnosing common pulmonary infections like pneumonia and tuberculosis to detecting more complex cardiac, skeletal, and mediastinal abnormalities, chest X-rays remain a first-line imaging modality that plays a pivotal role in patient care.

In addition to its traditional medical applications, the technological evolution of digital imaging and the rise of artificial intelligence have broadened the scope of chest X-rays beyond conventional diagnostics. Chest radiographs are now being integrated into innovative research domains, including predictive analytics, automated disease detection, and forensic medicine. Notably, the recent exploration of chest X-rays in the realm of biometric person identification represents a significant leap forward. By focusing on stable internal skeletal structures, chest X-rays offer a robust alternative to external biometric traits, providing new opportunities for secure identification in both forensic and security applications. In summary, the chest X-ray remains one of the most important and versatile imaging tools in healthcare, with expanding possibilities driven by advancements in technology. As imaging techniques continue to evolve, the role of chest radiographs is poised to extend further, not only improving patient diagnosis and monitoring but also contributing to emerging fields like forensic identification and biometric security.

CHAPTER II :

Biometrics and State of the Art

Chapter II : Biometrics and State of the Art

1. Introduction

Biometric technologies have become foundational tools in forensic science, providing efficient, accurate, and often non-invasive means of identifying individuals[45]. Traditionally, biometric identification has relied on modalities such as fingerprint analysis, DNA profiling, and facial recognition, which have demonstrated high reliability in both civil and criminal applications. These technologies assist law enforcement in identifying victims, confirming identities of suspects, and solving complex crimes [46].

As global populations increase and violent incidents rise, the demand for rapid and robust forensic techniques continues to grow. This evolution is particularly critical in mass casualty events, war zones, and areas impacted by natural disasters, where traditional means of identification may be insufficient or unavailable [47]. In these contexts, novel biometric modalities such as chest radiographs are emerging as promising solutions for post-mortem identification.

Biometrics in forensic science are generally categorized into physiological, behavioral, and medical imaging modalities. Physiological traits such as fingerprints and iris scans remain relatively constant over time and provide high identification accuracy. Behavioral traits like gait or typing rhythm are more dynamic but offer continuous authentication in digital systems. However, the applicability of each modality varies significantly depending on whether the subject is living or deceased. Post-mortem identification imposes stringent requirements due to decomposition, trauma, and the unavailability of certain body parts or biometric data [48].

2. The Identification Problem in Modern Security Systems

The challenge of establishing reliable identity verification mechanisms extends far beyond digital services, reaching into critical domains such as forensic science, disaster victim identification (DVI), and postmortem investigations. In these scenarios, the primary requirement is to ensure that identification is both accurate and resistant to failure under extreme conditions. Traditionally, identity verification has relied on knowledge-based credentials (such as passwords or PINs) and possession-based tokens (such as ID cards, smart cards, or certificates). While effective in certain contexts, these approaches are inadequate in forensic and postmortem settings, where credentials are unavailable, lost, or irrelevant.

Conventional biometric systems based on fingerprints, iris patterns, facial features, or voice recognition have been developed to overcome these limitations by linking identity directly to intrinsic human traits. However, even these systems face significant constraints when applied to forensic science. Fingerprints may be destroyed by burning, mutilation, or decomposition. Facial features are subject to alteration due to trauma, disfigurement, or postmortem degradation. Iris and voice recognition likewise require cooperation or preservation of soft tissues, which are rarely available in mass disaster or long postmortem intervals. These vulnerabilities underline the urgent need for alternative biometric modalities that remain distinctive and stable in conditions where traditional methods fail [49].

In this context, chest X-ray–based skeletal biometrics emerge as a promising solution. Radiographs inherently capture the thoracic skeleton including ribs, clavicles, sternum, and vertebrae which possesses high distinctiveness across individuals and remarkable permanence over time [50]. Unlike soft tissues, skeletal structures resist decomposition and cannot be easily altered or forged, making them particularly valuable in postmortem identification and medico-legal investigations. Furthermore, chest X-rays are already widely collected in hospitals and medical archives, providing premortem references that can be matched against postmortem images in forensic cases [51].. The reliance on skeletal traits thus offers several advantages: resilience against spoofing and fraud, applicability in high-security environments such as borders and research facilities, and, most importantly, robustness in forensic contexts where conventional biometrics collapse. As governments, forensic agencies, and security institutions face rising challenges related to identity theft, disaster victim identification, and medico-legal verification, there is a growing emphasis on developing advanced biometric systems based on radiographic imaging. The work presented in this thesis directly addresses these challenges by exploring chest X-ray radiographs as a novel and reliable biometric modality for both clinical and forensic applications.

3. Definition

According to the International Organization for Standardization (ISO), biometrics also referred to as biometric recognition is defined as “the automated recognition of individuals based on their biological and behavioral characteristics” (ISO/IEC 2382-37, 2012). The term automated emphasizes the role of algorithms executed by machine systems to identify individuals, although human assistance may be used to enhance accuracy. The term recognition refers to the process of associating an identity with a person based on distinctive physical or behavioral characteristics, collectively known as identifiers or traits [52].

Physical characteristics are biological attributes that are inherently part of the human body, such as fingerprints, facial structure, iris patterns. In the context of this research, skeletal structures visible in chest X-ray radiographs. Behavioral characteristics, on the other hand, are patterns of action generated by the body, such as signatures, voice, or keystroke dynamics. Unlike traditional

identification methods which rely on what a person knows (e.g., passwords, PIN codes) or possesses (e.g., ID cards, smart cards, certificates) biometric systems authenticate based on what the individual is (biological traits) or does (behavioral traits). Because these identifiers are inseparably linked to the person, they cannot be forgotten, copied, or transferred, making them inherently more resistant to compromise. Biometric systems encompass a broad range of modalities, including fingerprint, face, ear shape, iris, retina, palmprint, vein patterns, voice, signature, gait, and even odor, showed in figure In this research, the focus is placed on skeletal-based biometrics using chest X-ray imagery a less conventional but highly stable modality offering significant advantages in forensic identification, high-security authentication, and scenarios where traditional biometrics are unavailable or unreliable [53].

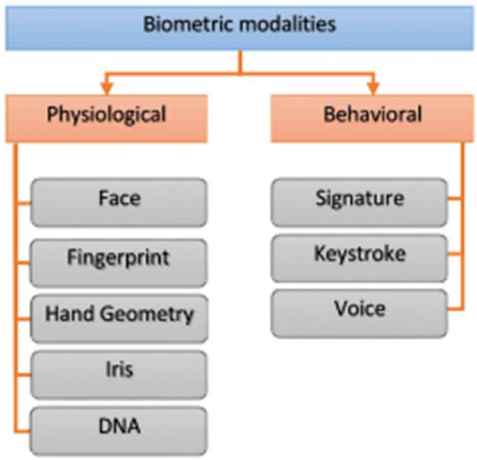


Figure 5: biometric modalities.

4. Biometric Characteristics Requirements

A biometric characteristic (or trait) is a measurable physical or behavioral attribute of an individual that can be used to distinguish them from others. It forms the basis of how an individual is recognized in an automated biometric system. One of the most critical steps in designing a practical biometric system is determining which characteristics to employ for reliable identification.

Each biometric trait has its strengths and weaknesses, and the choice often depends on the application domain, environmental constraints, and the characteristics of the population being identified. In certain fields such as forensic science the selection of traits is also driven by the condition of the subject, which may be damaged, decomposed, or otherwise unidentifiable through traditional biometric means. In such cases, internal and stable characteristics, like skeletal structures, become crucial for accurate identification.

According to [53] an effective biometric characteristic should meet the following requirements:

1. **Universality:** Every individual within the target population should possess the characteristic. For example, iris patterns cannot be used for blind individuals, and handwritten signatures are impractical in illiterate populations. In skeletal biometrics particularly in forensic contexts universality is high since all individuals possess bone structures, regardless of soft-tissue degradation.
2. **Uniqueness:** The trait should have sufficient variation between individuals to allow reliable distinction. Skeletal shapes and anatomical bone features, as visible in chest X-rays, exhibit unique patterns that remain identifiable even in postmortem forensic examinations.
3. **Permanence:** The characteristic should remain stable over time, at least for the duration of the recognition system's use. Skeletal structures, especially ribs and clavicles, undergo minimal changes in adulthood, making them valuable for long-term identification, including in forensic casework where remains may be decades old.
4. **Measurability:** The trait must be quantifiable and suitable for digital processing. In skeletal biometrics, radiographic imaging systems can reliably capture bone structures for analysis, whether in living subjects for security purposes or in forensic autopsies for victim identification.
5. **Performance:** The chosen biometric must allow for acceptable accuracy, processing speed, and computational efficiency in practical applications. For skeletal biometrics in forensic science, performance also depends on handling degraded or incomplete radiographs while still achieving reliable matching.
6. **Acceptability:** Users must be willing to participate in the acquisition process. While chest X-ray capture is more specialized than facial or fingerprint scanning, it is widely accepted in medical, forensic, and legal contexts where accurate identification is paramount.
7. **Circumvention**: Resistance the biometric should be difficult to forge or bypass. Skeletal biometrics are inherently resistant to common spoofing attacks and are particularly valuable in forensic science because they are immune to most alteration attempts.

It is rare for a single biometric trait to perfectly meet all these criteria. A practical biometric system should strike a balance offering high accuracy and speed, reasonable resource requirements, user safety, acceptance by the target population, and resilience against fraudulent attacks [55]. In forensic applications, skeletal-based biometrics provide the added advantage of functioning when external identifiers are unavailable, making them a critical tool for victim identification, disaster victim recovery, and postmortem human recognition.

5. Biometric System Architecture

A typical biometric recognition system consists of four main modules (Figure 6):

1. **Biometric Sensor**: Captures the biometric trait from the subject and converts it into a digital format for further processing. The quality of the captured data is crucial for system performance. In skeletal-based biometrics, this involves acquiring high-quality chest X-ray images, where resolution, noise levels, and positioning directly influence identification

accuracy. In forensic science, this step often uses postmortem radiographs to document skeletal structures for comparison with antemortem medical records.

2. **Enrollment:** The raw biometric data undergoes preprocessing to enhance quality (e.g., denoising, contrast adjustment). Discriminative features are then extracted to form a compact digital representation known as a template, which summarizes the essential biometric characteristics. In skeletal biometrics, feature extraction might include bone contour detection, rib curvature analysis, or clavicle geometry. In forensic identification, these templates can be matched against stored hospital or security records to confirm identity.
3. **Storage System:** The storage system can range from a simple file on a smartcard to a large-scale database managed by a database management system (DBMS). Alongside the generated biometric template, additional biographic information such as name, password, address, or case reference numbers may also be stored. Regardless of its size or complexity, the most critical aspect is template security. If a stored template is compromised, it may be possible to reconstruct the original biometric characteristic, posing a serious security threat. In the context of skeletal biometrics, protecting stored chest X-ray based templates is particularly important, as reconstructed anatomical structures could reveal sensitive personal or medical information. In forensic science, secure template storage ensures the integrity of evidence, maintaining its admissibility in legal proceedings.
4. **Matching Module:** During the operational phase, the biometric system performs matching to verify or identify a person. This process begins by extracting the discriminatory features from the newly acquired biometric sample called query features using the same feature extraction method employed during enrollment. The stored template is then retrieved and compared with the query features to determine whether they originate from the same individual. The result of this comparison is typically expressed as a similarity score ranging from 0 (no match) to 1 (perfect match), which guides the system in making the final identity decision.

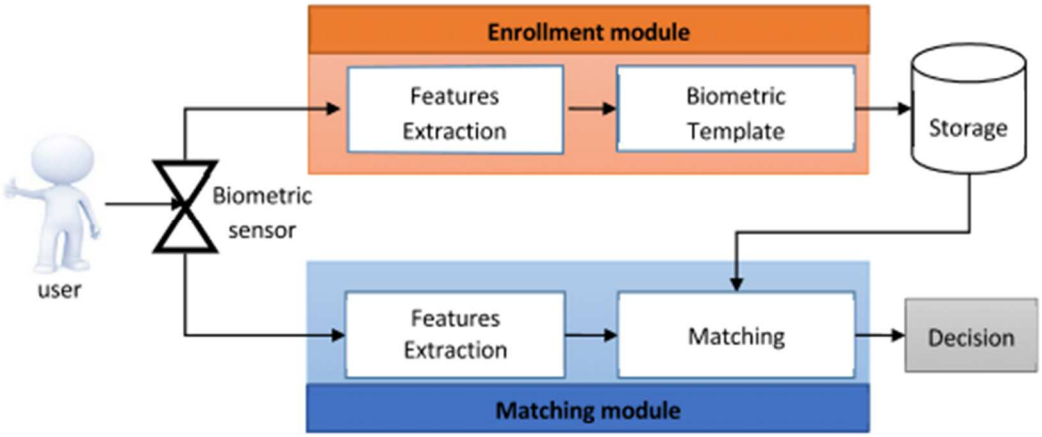


Figure 6: biometric system architecture.

Biometric systems can operate in two main modes:

1. **Verification Mode** (1:1 comparison): The system compares the query features against a single stored template, verifying a claimed identity. For example, in a secure medical archive, a person claiming to be a registered doctor could be verified by matching their skeletal biometric against the one stored in the database.
2. **Identification Mode** (1:N comparison): The system compares the query features against all templates in the database to determine whether the individual is already registered, and if so, to reveal their identity. This mode is particularly important in forensic science, where skeletal biometrics from a chest X-ray may be compared against large hospital or law enforcement databases to identify unknown remains or confirm the identity of disaster victims.

By integrating skeletal chest X-ray biometrics into these modules, both security applications and forensic investigations benefit from a method that is highly resistant to spoofing, remains stable over time, and retains effectiveness even when traditional surface-based biometric traits are unavailable.

6. Traditional Biometric Modalities

Traditional biometric modalities represent the long-established cornerstone of identity verification systems, utilizing distinctive physiological or behavioral characteristics inherent to individuals. These well-studied methods including fingerprints, facial recognition, iris patterns, voice characteristics, palm prints, and signature dynamics leverage unique biological traits or learned behaviors that are difficult to replicate or transfer. Due to their relative ease of capture, proven reliability, and extensive research history, these modalities form the backbone of security systems worldwide, integrated into applications ranging from smartphone authentication to border control and forensic databases. Their widespread adoption underscores their fundamental role in establishing and confirming personal identity across diverse sectors.

1. Behavioral Biometrics

Behavioral biometrics analyze patterns in human activity that are difficult to replicate. Common modalities include:

- **Gait**: Unique walking patterns influenced by physiology and movement. Although promising, gait is impractical in post-mortem scenarios.
- **Keystroke Dynamics**: Typing rhythms and patterns, useful in cybersecurity but irrelevant in forensic identification.
- **Signature & Handwriting**: Pressure, speed, and formation patterns analyzed in civil authentication systems.
- **Speech Recognition**: Tone, pitch, and vocal habits, primarily used in live authentication and inapplicable after death.

Although these modalities have found use in security systems, their reliance on live behavior renders them largely inapplicable in post-mortem identification scenarios [56].

2. Physiological and Biological Biometrics

These modalities focus on stable anatomical and biological features, making them more applicable in both ante-mortem and post-mortem scenarios.

a. Fingerprint Recognition Fingerprints are highly individualized and remain stable over time. Their widespread use is supported by extensive databases. However, in post-mortem contexts, fingerprint analysis can fail due to burns, decomposition, or skin damage. Recent studies have employed deep learning for enhanced fingerprint classification, achieving up to 97.83% accuracy using CNNs. Gender and finger-type classification models have also shown promising results, even under constrained datasets [57].

b. Facial Recognition Facial recognition is widely used in surveillance and criminal investigations. However, in post-mortem contexts, facial disfigurement or decomposition often renders these methods ineffective. To overcome such challenges, advanced systems like OpenFace, SeetaFace, and FaceNet have been integrated with probabilistic frameworks like Bayesian inference, improving performance even with degraded inputs [58].

c. Dental Analysis (Odontology) Teeth are resistant to decay and trauma, making dental analysis a cornerstone of post-mortem identification. Modern odontology uses radiographic comparisons, AI-based dental segmentation, and implant detection systems. AI advancements have enabled detection accuracies of up to 99.1% for natural teeth, with systems capable of identifying implants, root canals, and prosthetics. Automated comparison of antemortem and postmortem radiographs has reached accuracies near 85%, aiding rapid and accurate identification [58].

3. Limitations of Conventional Biometrics in Post-Mortem Cases

Despite their effectiveness, traditional biometric methods face severe limitations in post-mortem scenarios:

- **Fingerprint analysis** fails in cases of skin damage or decomposition.
- **Facial recognition** becomes unreliable with disfigurement or trauma.
- **DNA profiling** is hindered by contamination or biological degradation.
- **Iris and vein pattern analysis** are inapplicable once tissue degradation occurs.

Moreover, in high-casualty events such as natural disasters or military conflicts (e.g., Gaza, Ukraine), the volume and condition of remains necessitate faster, more resilient identification methods. The fragility of soft tissues and the logistical challenges associated with DNA testing delay identification and can lead to mismatches or inconclusive results. Therefore, imaging-based biometrics, particularly those involving skeletal or internal anatomical features, are being explored as alternative

Biometric Modality	Type	Key Characteristics	Strengths	General Limitations	Forensic Limitations	Applications
Fingerprint	Physiological	Ridge patterns on fingertips	High uniqueness, fast, mature technology	Can be spoofed, affected by cuts or skin damage	Often unusable in decomposed, charred, or waterlogged bodies; ridge detail may be destroyed	Law enforcement, border control, forensic identification
Face Recognition	Physiological	Facial geometry and texture	Non-intrusive, works from a distance	Affected by aging, lighting, facial coverings	Faces may be unrecognizable due to trauma, decomposition, or intentional disfigurement	Surveillance, security, missing persons
Iris Recognition	Physiological	Unique iris texture patterns	Extremely accurate, stable over lifetime	Requires close-up imaging, cooperation issues	Postmortem iris degrades rapidly, making it unusable beyond early death stages	Border security, high-security systems
Retina Scan	Physiological	Blood vessel patterns in retina	High accuracy, difficult to forge	Intrusive, requires close contact	Retinal tissue deteriorates quickly after death, unsuitable for delayed forensic identification	Military/government security, specialized forensic cases
Palmprint	Physiological	Ridge patterns and creases on palm	Larger area than fingerprint, good uniqueness	Needs controlled hand placement, skin issues	Palm ridge details also degrade in advanced decomposition or severe burns	Access control, forensic verification
Hand Geometry	Physiological	Shape and size of hand	Easy to acquire, stable after adulthood	Less unique than other traits	Limited forensic utility; hand geometry is not sufficiently distinctive for	Time & attendance, low-security applications

Biometric Modality	Type	Key Characteristics	Strengths	General Limitations	Forensic Limitations	Applications
					postmortem identification	
Vein Recognition	Physiological	Subcutaneous vein patterns (hand/finger)	Hidden from view, difficult to forge	Needs infrared imaging, affected by blood flow	Blood vessels collapse after death, making veins difficult to capture in forensic scenarios	Banking security, high-security authentication
DNA	Physiological	Genetic code unique to each person	Extremely unique, reliable forensics	Slow, expensive, privacy concerns	Strongest forensic tool but limited by sample degradation (fire, chemicals) or contamination	Criminal investigations, victim ID, paternity testing
Signature	Behavioral	Writing style, stroke dynamics	Familiar, legally recognized	Easily forged, variable over time	Impossible in forensic postmortem scenarios (requires conscious effort)	Banking, contracts, civil cases
Voice Recognition	Behavioral	Vocal pitch, tone, speech patterns	Can be captured remotely	Sensitive to noise, illness	Unusable postmortem; may be unusable if recording quality is poor	Call centers, forensic speaker recognition
Keystroke Dynamics	Behavioral	Typing rhythm and speed	Continuous monitoring possible	Affected by fatigue, needs data	No postmortem utility; requires live subject	Cybersecurity, behavioral forensics
Gait Recognition	Behavioral	Walking pattern	Works at a distance, unobtrusive	Affected by injury, clothing, surface	Not usable postmortem; forensic use limited to video surveillance	Surveillance, criminology
Ear Shape	Physiological	Outer ear	Stable over	Sensitive to	May be damaged in	Surveillance, forensic

Biometric Modality	Type	Key Characteristics	Strengths	General Limitations	Forensic Limitations	Applications
Recognition		geometry	lifetime	hairstyle/coverings	trauma, burns, or decomposition; forensic application limited	studies
Odor Recognition	Physiological	Chemical composition of body odor	Hard to forge	Requires advanced sensors, environment-dependent	Body odor changes after death due to decomposition, making it unreliable in forensics	Forensic tracking, experimental security
Chest X-ray Skeletal Biometrics	Physiological	Anatomical structures (ribs, clavicle, sternum, spine)	Highly stable, resistant to spoofing, unaffected by soft tissue decomposition	Requires radiographic imaging equipment; health exposure limits everyday use	Strong forensic utility: bones remain intact long after death, enabling postmortem identification from records	Forensic identification, disaster victim ID, high-security authentication, medical identity verification

Table 1 : biometric modalities.

7. Discussion on Biometric Modalities with Emphasis on Chest X-Ray Skeletal Biometrics in Forensic Science

Over the past few decades, biometric technologies have advanced significantly and are now integrated into many aspects of security and identity verification. Each biometric trait whether physical or behavioral offers unique benefits and constraints. From a forensic science perspective, however, their usefulness is closely linked to how well they endure after death, how resistant they are to environmental damage, and whether they can meet the evidentiary requirements of criminal and postmortem investigations.

Conventional physiological traits such as fingerprints remain the most established form of identification. Their uniqueness and the long history of investigative use make them highly reliable under normal conditions. Yet, they lose much of their utility when the body undergoes decomposition, burning, or immersion in water, as ridge details often deteriorate beyond recognition. Similarly, iris and retina scans, although highly accurate for living individuals, offer only a very narrow postmortem window. The soft ocular tissues degrade rapidly, sometimes within hours, which greatly limits their forensic relevance. By contrast, behavioral traits like signature, voice, typing rhythm, or walking style have no application in forensic or postmortem scenarios. Since they require voluntary human participation, they cannot be recorded after death. Even in the living, these modalities are prone to variability and manipulation, making them unsuitable for robust forensic investigations.

DNA analysis is often considered the gold standard for forensic identification. Its discriminatory power is unmatched, and it is central to resolving criminal cases, kinship testing, and victim identification. Nevertheless, DNA is not without problems. Sample collection and sequencing are expensive and time-intensive, and results can be delayed in time-sensitive investigations. DNA is also susceptible to contamination, degradation in adverse environments, or intentional destruction. Moreover, ethical debates around genetic databases present additional challenges. Other modalities such as palmprints, ear geometry, and vascular imaging also face limitations in forensic work. Palmprints, like fingerprints, degrade when skin tissue is damaged. The shape of the ear is fragile and can easily be altered by trauma. Vein pattern recognition, which relies on blood vessels, becomes ineffective after circulation ceases and tissues collapse postmortem. Even experimental methods such as odor-based biometrics fail in this context, since chemical body signatures change dramatically as decomposition progresses.

In light of these limitations, radiographic chest images and skeletal patterns represent a particularly promising alternative for forensic and postmortem identification. Bone structures such as the ribs, sternum, clavicles, and vertebral column are some of the most durable elements of the human body. They remain intact even in challenging conditions such as burial, fire exposure, or advanced decomposition. This resilience makes skeletal imaging especially valuable in disaster victim identification (DVI), where other biometric traits may have been destroyed.

One of the strongest advantages of chest X-ray biometrics lies in the fact that such records already exist for a large portion of the population. Chest radiographs are routinely taken in medical examinations, hospital procedures, immigration screenings, and occupational health checks. These archived medical images provide a readily available source of premortem data that can be directly compared to postmortem radiographs. This is rarely the case with other biometric traits, as very few people have premortem iris scans, vein maps, or odor profiles stored in official records.

From an identification standpoint, the skeletal features captured in chest radiographs are both stable and individual-specific. Variations in rib configuration, clavicle length, sternum shape, and spinal curvature create a unique anatomical signature that can be matched much like a fingerprint. Unlike DNA testing, which requires laboratory analysis, radiographic comparison can be performed relatively quickly, particularly when assisted by automated recognition algorithms. This speed is critical during large-scale disasters where rapid identification is necessary.

Chest X-ray biometrics also benefit from being extremely difficult to falsify. While fingerprints can be replicated using artificial molds and facial features can be altered digitally, the internal skeletal framework of a person is virtually impossible to imitate. This greatly enhances its credibility in both forensic science and high-security applications. Of course, chest X-ray identification is not free of drawbacks. Specialized imaging equipment is required, and exposure to radiation, although minimal, prevents its widespread everyday use. In forensic medicine, however, these concerns are outweighed by the clear advantages, as radiographic imaging is already standard practice in autopsy and medico-legal investigations.

In summary, while established biometric modalities remain important in living populations and in certain forensic settings, they each encounter significant limitations once the subject is deceased. Chest X-ray skeletal biometrics overcome many of these challenges by providing a durable, distinctive, and often pre-recorded identifier that retains its value long after death. Forensic investigators can therefore rely on this modality for victim identification in mass disasters, missing persons cases, and postmortem verification when conventional methods have failed.

1. Chest X-ray Biometrics in Post-Mortem Identification

Chest X-ray Biometrics in Post-Mortem Identification offers a valuable, often underutilized, method for identifying deceased individuals, particularly when traditional means like fingerprints or dental records are unavailable or compromised. This technique leverages the inherent uniqueness of anatomical structures visible in chest radiographs, such as the rib cage configuration, cardiac silhouette, mediastinal outline, clavicle shape, and patterns of costal cartilage calcification. By comparing post-mortem chest X-rays with available ante-mortem radiographs from medical records, forensic experts can perform a comparative analysis of these stable bony and soft-tissue landmarks [60]. The method proves especially crucial in mass disaster scenarios, severely decomposed remains, or cases involving extensive trauma, providing a relatively quick, non-invasive, and objective means to establish identity based on the individual's internal radiographic "fingerprint" captured during life. Its reliability hinges on the quality and comparability of the images and the distinctiveness of the visualized anatomical features.

2. Motivation and Advantages

Chest X-ray imaging has emerged as a robust, non-invasive modality in post-mortem identification, particularly due to its ability to visualize skeletal and anatomical structures that remain intact even after soft tissue deterioration.

Advantages include:

- **Resistance to decomposition**: Skeletal elements like the ribs and vertebrae remain intact even under extreme conditions.
- **Availability in medical records**: Many individuals have existing chest X-rays due to prior respiratory or cardiovascular conditions.
- **Distinctive anatomical features**: Chest morphology, bone density, and thoracic structure exhibit individual-specific patterns, making them suitable for biometric analysis.

These characteristics make chest X-ray-based identification a valuable alternative when other modalities are compromised. It is particularly useful in cases where only partial remains are found, or where facial and fingerprint recognition are not feasible.

3. Biometric Process Using Chest X-rays

The general framework for post-mortem identification using chest X-rays includes:

- **Image Acquisition**: Scanning of post-mortem chest radiographs.
- **Feature Extraction**: Deep learning models extract key anatomical features.
- **Embedding & Matching**: Similarity comparisons are conducted using Euclidean or cosine distances in the latent space.
- **Identification/Verification**: Matching results are validated against ante-mortem X-ray databases.

This process does not require the preservation of soft tissue, making it especially suitable in challenging forensic scenarios. Moreover, the ability to automate the matching process using artificial intelligence enhances its applicability in large-scale identification tasks.

8. Related Works on Chest X-ray Identification

This section synthesizes findings from a comprehensive literature review, pinpointing current methodologies, unresolved questions, and research opportunities in chest X-ray-based person identification and verification. It summarizes significant research leveraging artificial intelligence including deep learning, machine learning, and attention mechanisms and explores diverse architectural approaches driving progress in radiographic biometrics.

he use of chest radiographs for biometric identification has attracted increasing attention in recent years. The unique anatomical structures visible in these images, such as the ribcage,

clavicles, and heart silhouette, provide distinctive biometric traits that remain relatively stable over time, making them suitable for both healthcare security and forensic applications. Over the past two decades, researchers have explored both traditional image-processing techniques and modern deep learning approaches to exploit this modality.

8.1. Classical Image-Processing Approaches

Early efforts in chest radiograph-based identification largely relied on handcrafted feature extraction and statistical analysis. For example, [61] developed a stochastic image-processing framework for postmortem identification using both radiographs and CT scans. The method employed Bag-of-Words (BoW) and Histogram of Oriented Gradients (HOG) features, combined with Euclidean distance for similarity measurement. Despite its novelty, the system achieved relatively modest accuracies ranging from 44.4% to 63.0%**,** reflecting the limitations of handcrafted descriptors in capturing complex anatomical variability.

Similarly, [62] designed a four-stage identification process incorporating image preprocessing, feature extraction, boundary enhancement, and similarity calculation. Their approach utilized Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve image contrast and applied a two-dimensional Discrete Fourier Transform (DFT) for frequency-domain feature representation. When tested on radiographs collected from the Miyazaki University School of Medicine, the method achieved 74.07% accuracy for deceased individuals. While this represented an improvement over [61], the reliance on handcrafted features restricted scalability and robustness.

8.2. Deep Learning-Based Methods

The shift toward deep learning brought a significant transformation in biometric identification from medical images. Deep neural networks, with their ability to automatically learn hierarchical features, began to outperform classical handcrafted pipelines.

In [63], researchers explored chest X-ray re-identification using a Siamese network trained with contrastive loss**.** Their model, based on ResNet-50, achieved an AUC of 0.9940 and 95.55% accuracy on the ChestX-ray8 dataset. Importantly, this study demonstrated that chest radiographs could support long-term re-identification, even across imaging sessions separated by up to ten years.

Another contribution, [64], presented a verification framework based on a deep convolutional neural network (DCNN) with an EfficientNetV2-S backbone**.** The pipeline included image acquisition, feature extraction, and cosine similarity-based identification. When evaluated on a dataset of 1,000 chest X-rays, the model achieved 83.0% accuracy, illustrating the feasibility of deep learning–based verification systems for reducing patient misidentification in healthcare environments.

8.3. Advances in Attention Mechanisms

To further improve performance, attention mechanisms have been widely incorporated into deep neural networks. These modules enable models to prioritize salient image regions while suppressing irrelevant or noisy information.

Several landmark studies across the computer vision domain illustrate the importance of attention in identification tasks. For instance, SENet [65] introduced squeeze-and-excitation blocks to model channel interdependencies, significantly boosting classification performance.

Similarly, CBAM [66] combined channel and spatial attention to refine CNN feature selection, enabling more discriminative embeddings. In [67], a residual attention network employing an encoder-decoder structure was proposed for face recognition, yielding 98.3% accuracy on CASIA-WebFace and MS-Celeb-1M. Likewise, [68] introduced a multiscale deep supervision model with reverse attention to reduce feature loss in person re-identification, achieving 89.0% mAP and 95.5% accuracy on the Market-1501 dataset.

Beyond recognition, attention mechanisms have influenced other computer vision applications such as video classification [69,70], image generation [71], and scene segmentation [72]. Collectively, these studies confirm that integrating attention into deep learning pipelines substantially enhances model robustness, particularly when dealing with noisy or high-dimensional data like medical images.

8.4. Our works

While the aforementioned works laid essential groundwork, they often faced limitations such as reliance on soft-tissue features, inadequate focus on skeletal structures, or lack of interpretability. To address these gaps, our research [73–76] has advanced the field by introducing attention-driven, triplet loss optimized frameworks specifically tailored for chest radiograph biometrics.

In [73], we presented the first application of a triplet loss-trained Siamese network with triple image inputs for chest radiograph identification. Leveraging transfer learning with pretrained models, the system effectively captured intricate skeletal and anatomical patterns. On the NIH ChestX-ray14 dataset, it achieved 97% accuracy**,** 95.3% precision, and 98.4% recall, establishing a strong benchmark for deep learning–based identification.

Expanding this approach, [74] integrated ResNet-50 with a spatial attention mechanism to highlight key anatomical regions such as the heart, lungs, and ribs. By combining spatial attention with triplet loss optimization, the model enhanced discriminability in challenging forensic contexts, including postmortem or disaster scenarios. This framework achieved 95.8% accuracy on the NIH ChestX-ray14 dataset, underscoring the value of attention for chest radiograph biometrics.

In [75], we introduced the Self-Residual Attention Network (SRAN)**,** which fused channel and spatial attention within a ResNet-50 Siamese backbone. By dynamically emphasizing skeletal features while minimizing noise, SRAN generated highly separable feature embeddings. The model achieved 98.3% accuracy on NIH ChestX-ray14 and 96.1% on CheXpert, outperforming prior methods and demonstrating the scalability of attention-enhanced architectures.

Finally, [76] explored a complementary approach using a VGG-16–based Siamese network for feature extraction. Optimized with triplet loss, the framework emphasized healthcare safety applications, aiming to reduce patient misidentification errors. The model achieved 99% training accuracy with only 0.001% loss, further validating the reliability of chest radiographs as a biometric modality.

8.5. Summary

Taken together, our works [73–76] have advanced chest radiograph biometrics by:

- Introducing triplet loss with triple image inputs for robust discriminative learning.
- Incorporating spatial and self-residual attention mechanisms to emphasize skeletal structures.
- Demonstrating state-of-the-art accuracy across large-scale datasets (ChestX-ray14 and CheXpert).
- Exploring multiple deep architectures (ResNet-50, VGG-16) to validate generalizability.

By addressing the shortcomings of earlier handcrafted and deep learning methods, our research contributes significantly to the development of reliable and interpretable chest radiograph identification systems, with applications in both healthcare and forensic domains.

Work	Method	Application	Modalities	Dataset	Accuracy
[75]	**Siamese NN Self-residual attention N Triplet loss**	**Person identification**	**Chest X-ray**	**ChestXray14 dataset and CheXpert**	**98**
[73]	**Siamese NN Triplet loss Transfer learning**	**Person identification**	**Chest X-ray**	**NIH ChestX-ray14 dataset**	**97**
[74]	**Siamese NN ResNet-50 Spatial attention N Triplet loss**	**Person identification**	**Chest X-ray**	**NIH ChestXray14 datase**	**95.8**
[63]	Siamese NN Contrastive loss ResNet50	Person reidentificatio n	Chest X-ray	ChestX-ray8 dataset	95.55
[64]	EfficientNet Cosine distance	Person identification	Chest X-ray	NIH ChestX-ray14 dataset	83.0
[62]	CLAHE Two-dimensional discrete Fourier transform (DFT)	Person identification	Chest X-ray	collected from data stored in the database of Miyazaki University School of Medicine	74.07
[61]	HOG BOW classifier Euclidean distance	Person identification	Chest X-ray	collected from data stored in the database of Miyazaki University School of Medicine	44.4 to 63.0
[77]	Siamese network with triplet loss CED, adaptive margin-based hard negative mining Generative domain-specific features	Palm vein detection	Palm	------	------
[78]	CNN for key point extraction KNN for classification Siamese network and triplet loss	Facial detection	Face	------	------
[79]	Triplet loss Deep Siamese network K-way face recognition network	Face recognition	Face	LFW dataset	94.8
[80]	Deep Siamese network	Face recognition	Face	------	91

	Triplet loss				
[81]	Attention feature learning ResNet50 Triplet loss	Face recognition	Face	Market-1501 dataset	95.5
[67]	Self-residual attention N	Face recognition	Face	CASIA-WebFace and MS-Celeb-1M	98.3

Table 2: related works.

9. Conclusion

Biometric technologies have transformed forensic science, yet their reliability in post-mortem conditions is limited. The use of chest X-ray radiographs presents a promising alternative, offering resilience against decomposition and trauma. Studies demonstrate high identification accuracy, especially when leveraging deep learning, attention mechanisms, and large-scale radiographic datasets.

As conflicts and mass casualty events increase globally, the integration of chest X-ray biometrics into forensic protocols is not only necessary but urgent. Further research should focus on standardizing datasets, improving cross-domain matching, and developing real-time identification systems to support field investigations. This approach represents a significant advancement in forensic science, offering new tools for rapid, ethical, and reliable identification in the most challenging circumstances.

.

CHAPTER III :

Deep Learning Based Chest X-ray Biometric Identification

Chapter III : Deep Learning Based Chest X-ray Biometric Identification

1. Introduction

The development of reliable biometric systems remains a central challenge in security and identity recognition. Traditional biometric modalities such as fingerprints, facial images, and iris patterns have achieved high accuracy in controlled civilian environments; however, their performance may degrade in unconstrained or extreme conditions, including variations in acquisition settings, physiological changes, and image quality. These limitations motivate the exploration of alternative biometric traits that can complement or extend existing modalities.

Chest X-ray radiographs represent a widely available medical imaging modality that captures rich anatomical information of the thoracic region. Recent advances in deep learning have enabled the extraction of discriminative representations from complex medical images, opening new possibilities for biometric identification beyond conventional traits. This chapter investigates the feasibility of using chest X-ray images as a biometric modality through deep learning–based representation learning, without imposing explicit anatomical constraints. The contributions presented in this chapter focus on learning identity-discriminative embeddings directly from full chest radiographs using Siamese neural networks and metric learning strategies. The first contribution establishes a baseline approach using triplet loss and pretrained convolutional encoders, demonstrating the potential of chest X-ray images for biometric identification. The second contribution extends this framework by introducing self-residual attention mechanisms to enhance feature discrimination and robustness.

Together, these contributions form the foundational stage of this thesis, highlighting both the capabilities and limitations of appearance-based deep learning approaches for chest X-ray biometric identification. The insights gained in this chapter motivate the anatomically constrained strategies developed in the subsequent chapter.

2. System Overview

The biometric framework proposed in this chapter follows a standard deep learning–based recognition pipeline adapted for chest X-ray images, as illustrated in Figure 7. The system operates on full radiographic images and is designed to learn identity-discriminative representations without explicit segmentation or anatomical isolation.

4.2.1 Image Acquisition

Chest X-ray radiographs are collected from clinical imaging repositories and serve as the primary input to the biometric system. These images may vary in acquisition conditions, resolution, and subject posture, reflecting realistic operational scenarios.

4.2.2 Image Preprocessing

Prior to feature extraction, the images undergo basic preprocessing steps, including resizing, normalization, and intensity standardization, to ensure compatibility with deep learning architectures and reduce variability unrelated to identity.

4.2.3 Representation Learning

The preprocessed images are passed through Siamese neural networks employing pretrained convolutional backbones. Transfer learning is used to leverage knowledge from large-scale visual datasets. Metric learning with triplet loss is applied to learn compact embeddings that minimize intra-subject variability while maximizing inter-subject separation. In the second contribution, attention mechanisms are integrated to guide the network toward more informative regions of the radiograph.

4.2.4 Template Generation and Storage

The learned embeddings are stored as biometric templates in a database. Each template represents the identity-specific features extracted from a chest X-ray image and may be associated with identity labels or reference identifiers.

4.2.5 Matching and Decision

During verification, a query embedding is compared to a claimed identity template (1:1 matching). During identification, the query is compared against all enrolled templates (1:N matching). Similarity scores are computed using distance metrics, and the system outputs an identity decision based on predefined thresholds.

While this framework demonstrates the feasibility of chest X-ray–based biometric identification, it remains influenced by global appearance variations. These limitations motivate the anatomically constrained skeletal-based approach introduced in Chapter 5.

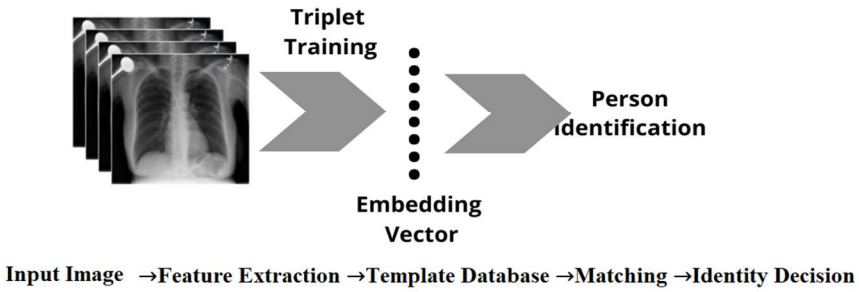


Figure 7: general system overview.

Contribution 1:

Triplet Network for Chest X-ray Based Biometric Identification

The first stage of our research introduces a deep learning framework tailored to the challenges of biometric recognition from chest radiographs. This contribution establishes the feasibility of chest X-rays as a biometric modality, particularly in forensic and postmortem scenarios where traditional identifiers (e.g., fingerprints, iris, or facial features) may be unavailable.

1. Methodology

To achieve this goal, we adopted a Siamese neural network architecture trained with triplet loss as illustrate in figure 8. The network is designed to learn highly discriminative embeddings by encoding three input images an anchor, a positive (same identity), and a negative (different identity) into a shared feature space. By enforcing compactness between anchor positive pairs while maximizing the separation between anchor negative pairs, the system develops robust identity representations from radiographic anatomy.

A key aspect of our design is the shared encoder, which processes all three inputs using identical weights. This ensures consistent feature extraction and facilitates meaningful comparisons between samples. To strengthen the encoder, we leveraged transfer learning from several widely used convolutional neural networks VGG16, VGG19, ResNet50, ResNet101, and Dense-Net originally pre-trained on large-scale image datasets. These backbones were integrated into the Siamese structure and extended with additional dense and normalization layers to adapt them to medical imaging.

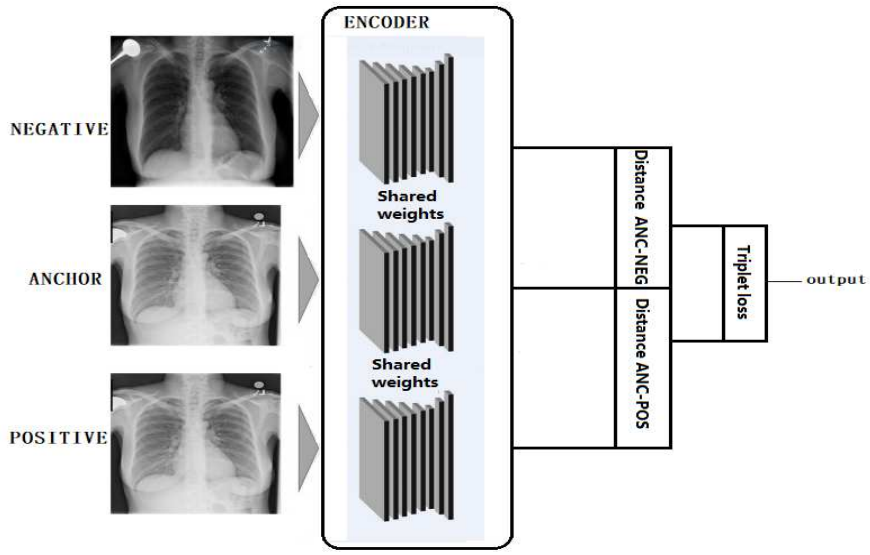


Figure 8: Proposed method.

1.1.Siamese Network with Triplet Loss

A Siamese neural network, which processes two inputs with identical weights, compares the output vectors. Consequently, a twin neural network is another term for it. Initially, it was used for signature verification and facial recognition [65] . The neural network receives two input images that are employed in the network architecture. In our experiment, we build a Siamese network that encodes three input images anchor, positive, and negative to their feature vectors using an encoder that contain a pre-trained model with some added layers . The model then produces three vectors that depict chests. The distance between these vectors is then produced by using the distance function to calculate the similarity between the three feature maps. Using the triplet loss function, the network learns to place greater distance between dissimilar images and less distance between identical images.

In the context of Siamese networks with triplet loss in our case, the distance formula between the anchor f(A) and positive f(P) embeddings, as well as between the anchor f(A) and negative f(N) embeddings, can be expressed in this way.

Let: f(A)=(a1,a2,...,an) be the anchor embedding, f(P)=(p1,p2,...,pn) be the positive embedding and f(N)=(n1,n2,...,nn) be the negative embedding. The Euclidean distance between the anchor f(A) and positive f(P) embeddings is given by the equation 1. Similarly, the Euclidean distance between the anchor f(A) and negative f(N) embeddings is given by: equation 2.

$$distance\_ap=\sum i=1n(ai-pi) \quad (1)$$

$$distance\_an=\sum i=1n(ai-ni)2 \quad (2)$$

The goal is to learn embeddings in such a way that the distance between the anchor and positive embeddings is minimized, while the distance between the anchor and negative embeddings is maximized using triplet loss . This loss function is mathematically expressed as follows :

The triplet loss for a triplet (A,P,N), where A is the anchor image, P is the positive image, and N is the negative image, is defined by equation 3:

$$loss(A,P,N)= \max(ap\_distance-an\_distance+margin,0.0) \quad (3)$$

Where :

- $\|\cdot\|\|\cdot\|$ represents the Euclidean norm.
- f(A), f(P), and f(N) are the embeddings of the anchor, positive, and negative images, respectively.
- margin is a hyperparameter that specifies the minimum margin or separation between the distances of the anchor and positive embeddings f(A)−f(P)‖2) and the anchor and negative embeddings f(A)−f(N)‖2).

The goal during training is to minimize this triplet loss. The loss is zero when the distances between the anchor and positive embeddings are smaller than the distance between the anchor

and negative embeddings by at least the specified margin. Otherwise, the loss becomes the positive difference between the squared distances. [70]

This formulation encourages the model to learn embeddings in which positive pairs are closer together than negative pairs by at least the margin. Adjusting the margin allows you to control the trade-off between embedding separation and model robustness.

1.2.Transfer Learning and Encoder Design

Creating chest x ray identification system from scratch requires a significant investment of time and money. Utilizing transfer learning from trained models is an additional method. A wide range of pre-trained models and ideas are presented by the current machine learning community[66][69]. We are able to construct our own model on top of pre-trained models. This can be accomplished by training only recently added layers that need the least amount of time and resources to train, and by freezing existing model layers. There are several models, such as VGG19, VGG16 , ResNet50, ResNet101 and DenseNet . In our work , we use encoder for converting the input images into their feature vectors, figure 9 represent the structure of this encoder. We are using pre-trained models to extract feature , our goal is to evaluate each pre-trained model separately within the same encoder architecture for the same task. The use of transfer learning can significantly reduce the training time and size of the dataset. The Model is connected to Fully Connected (Dense) layers and the last layer normalizes the data using L2 Normalization. (L2 Normalization is a method that modifies the dataset values in a way that in each row the sum of the squares will always be up to 1) . By utilizing the those pre-trained models ,we harness its ability to capture intricate hierarchical features from images and conclude the best performance . To ensure the preservation of these learned features, all layers of the pre-trained models are set as non-trainable. Subsequently, we construct a tailored encoder model by extending the pre-trained architecture. This includes the addition of dense layers for further abstraction, batch normalization for regularization, and a final L2 normalization layer for feature vector normalization. Those features are passed to a distance layer which computes the distance between (anchor, positive) and (anchor, negative) pairs. By analyzing the performance of each pre-trained model to gain insights into their strengths and weaknesses for this task. This understanding can guide future model selection or improvements.

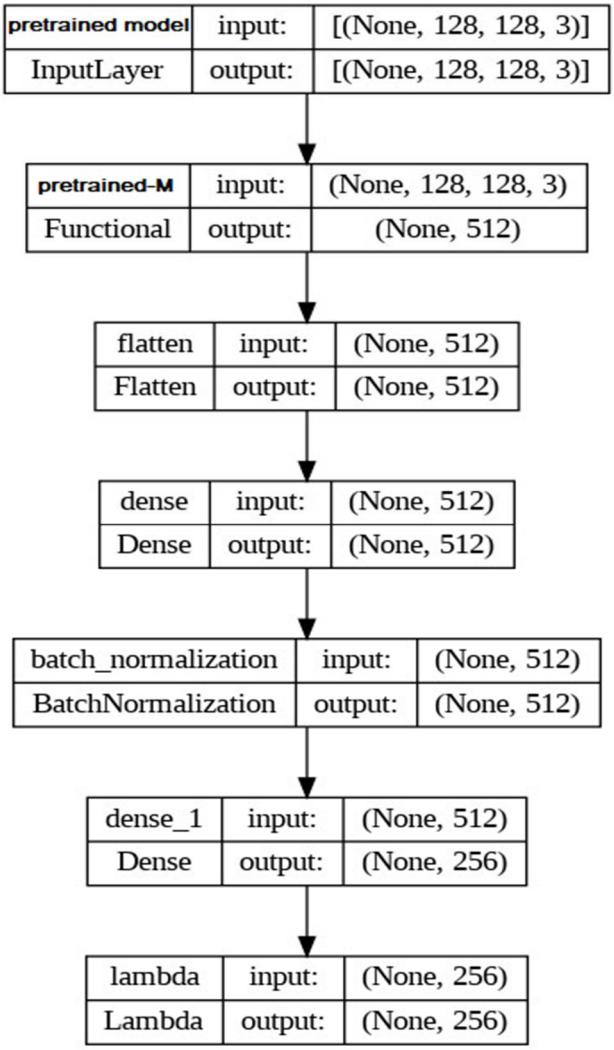


Figure 9: Encoder.

.

1.3.Data Preparation

Experiments were conducted using the NIH ChestX-ray14 dataset, one of the largest publicly available chest radiograph collections. To construct triplets, we ensured that each selected patient had at least two radiographs. For each anchor image, a positive sample (same patient ID) and a negative sample (different patient ID) were randomly chosen. Images were resized to 128 × 128 × 3, normalized, and encoded in RGB. To prevent data leakage, the dataset was partitioned into 90% training and 10% testing, with patient-disjoint splits ensuring that no individual appeared in both sets. This design allowed a fair evaluation of the network's generalization ability.

1.4.Results

To evaluate the performance of our proposed person identification system, we employ various assessment metrics tailored to assess the accuracy of identification tasks. The confusion matrix is one of the important measures used to assess classification metrics. For this purpose, we used it to provide a comprehensive imagine of a classification model's performance in distinguishing between dissimilar and similar images. The matrix reveals that the model accurately predicted dissimilar images (True Negatives) and similar images (True Positives).

However, there were instances that dissimilar images were erroneously classified as similar (False Positives), and similar images were mistakenly categorized as dissimilar(False Negatives). Precision denotes that the model's high accuracy in predicting similar images among its positive predictions. The recall, indicating the model's ability to identify all actual similar images .The F1 score, a balanced metric between precision and recall, encapsulating the model's effectiveness in handling both dissimilar and similar image classifications .

$$\text{Precision [P]} = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall [R]} = \frac{TP}{TP+FN} \quad (5)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

$$\text{F1Score} = 2\frac{P*R}{P+R} \quad (7)$$

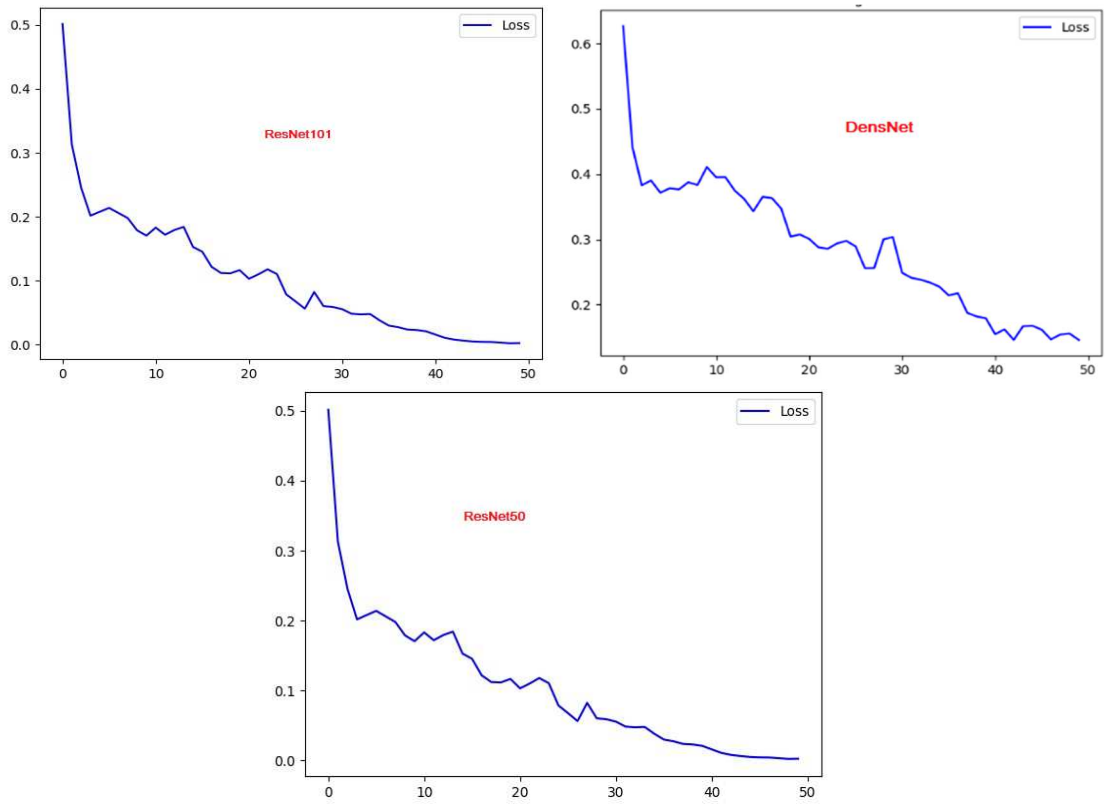
To evaluate the identification performance of our trained model, several metrics are computed. Training loss , test accuracy and other metrics .The identification model is trained using the triplet loss function to enhance the model's ability to learn effective representations for tasks such as identification , image similarity, and other tasks where understanding the similarity or dissimilarity between inputs is important, we evaluate our model with training loss to ensure that our model is learning by prove that our loss is decreasing over time, figure 3 show the training loss curb .In the proposed model we get a training accuracy of 99.9% and loss of 0.004% , else in testing we get an accuracy of 97% using the pre-trained model VGG19. That's prove that our proposed method have a very powerful performance in learning and testing in term of accuracy and loss and in other metrics shoed in table 3 . We trained our model for over 50 epochs.

Pre-trained Model	**TN**	**TP**	**FP**	**FN**	**Recall**	**precision**	**F1score**	**Accuracy**
VGG16	49.63	46.97	0.7	3.03	93.94	98.5	96.1	96.4
VGG19	**49.67**	**47.24**	**2.33**	**0.76**	**98.4**	**95.3**	**96.8**	**97**
ResNet50	43.89	46.91	3.09	6.11	88.4	93.8	91.01	90.8
ResNet101	48.64	46.98	3.02	1.36	97.1	93.9	95.5	94.9
DensNet	45.92	42.78	7.22	4.08	91.2	85.5	88.3	88.7

Table 3 : Confusion matrix and performance metrics.

In this comparative analysis of pre-trained models for feature extraction in our encoder, our experimental results provide valuable insights into their respective performances. VGG16 exhibited an impressive accuracy of 96.4%, showcasing a high precision of 98.5% and a robust recall of 93.94%. VGG19 demonstrated an even higher accuracy of 97%, driven by a well-balanced precision of 95.3% and a remarkable recall of 98.4%. ResNet50, while

achieving an accuracy of 90.8%, displayed a slightly lower precision of 93.8% but maintained a commendable recall of 88.4%. ResNet101 emerged as a strong contender with a balanced accuracy of 95.6%, combining a high recall of 97.1% and a precision of 93.9%. DensNet, although showing a lower accuracy of 88.7%, delivered a balanced precision of 85.5% and a recall of 91.2%. The best evaluation model is VGG19, with the highest overall accuracy, precision, and recall. Its impressive performance makes it the optimal choice for feature extraction in our experimental setup, emphasizing its efficacy in accurately identifying positive instances while minimizing false positives and false negatives. The comprehensive evaluation presented here guides the selection of the most suitable pre-trained model for the given task, setting the stage for improved outcomes in related applications. In conclusion, while all models demonstrate strong performance overall, there are nuances in their abilities to balance precision, recall, and accuracy. VGG16 and ResNet101 stand out for their high precision and recall rates, making them suitable for tasks where minimizing both false positives and false negatives is crucial. ResNet50, although exhibiting slightly lower recall, still maintains high accuracy, indicating its reliability in making correct identification. VGG19 and DenseNet, while performing well, may require additional attention to minimize false positives and negatives, respectively. These insights can inform the selection and optimization of pre-trained models based on specific research objectives and application requirements.



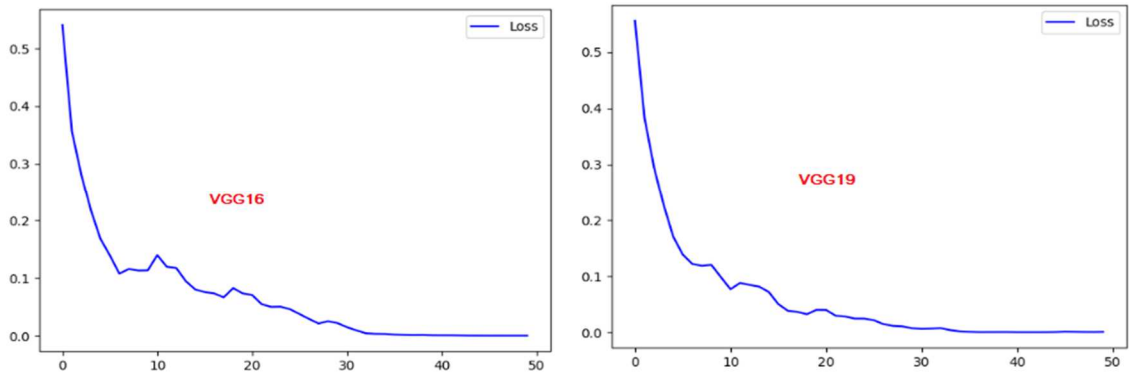


Figure 10: Trainings losses of different pre-trained models over epochs.

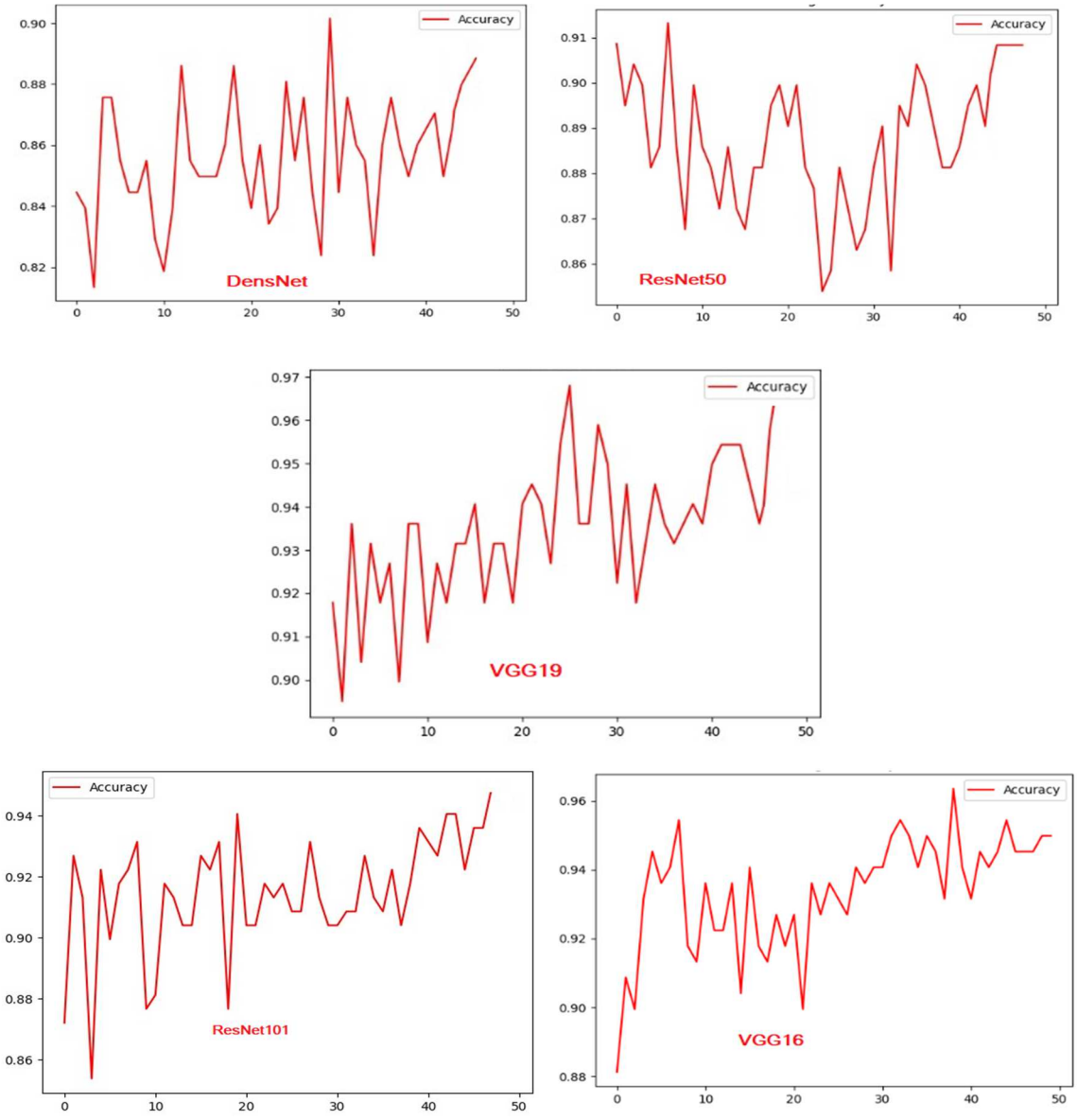


Figure 11: Testing accuracy of different pre-trained models over epochs.

2. Summary of Contribution

This contribution demonstrates the first use of Siamese networks with triplet loss for biometric identification from chest radiographs. By integrating transfer learning from multiple pretrained CNNs and a carefully designed encoder, the framework effectively extracts discriminative skeletal and anatomical features. While this stage establishes a solid baseline, it also highlights limitations related to soft-tissue variability and decomposition, motivating the subsequent steps in our research, where attention mechanisms and bone segmentation are introduced.

2.1.Limitations and Motivation for Advancement

Although the initial Siamese triplet framework demonstrated that chest radiographs contain distinctive biometric information, several limitations became evident:

- **2.1.1. Sensitivity to Soft-Tissue Variability.** Since the system processed full chest X-rays, embeddings were often influenced by lung textures, medical conditions, or soft-tissue changes. This reduced stability, especially in forensic or postmortem scenarios where soft tissue is degraded.
- **2.1.2. Generic Feature Extraction.** The pretrained backbones (ResNet, VGG, DenseNet) were originally optimized for natural image recognition rather than radiographic interpretation. As a result, they sometimes emphasized irrelevant patterns such as background intensity or imaging artifacts instead of bone structure.
- **2.1.3. Absence of Attention Mechanisms.** The network lacked the ability to prioritize the most discriminative skeletal features. Without a selective focus, the embeddings included noise and non-essential details, which weakened inter-class separation.

These shortcomings indicated that while deep metric learning with triplet loss is effective, a more specialized backbone was needed one capable of selectively emphasizing skeletal regions while suppressing irrelevant cues. To overcome these challenges, we proposed an improved architecture in the next stage of our work, introducing a Self-Residual Attention Network (SRAN) within the Siamese framework. This design integrates residual learning with self-attention modules to enhance focus on stable, identity-preserving structures, ultimately improving robustness in both clinical and forensic applications.

Contribution 2:

Siamese Network with Self-Residual Attention Backbone

The first contribution demonstrated that Siamese networks with triplet loss and pretrained CNNs (VGG, ResNet, DenseNet) could extract discriminative embeddings from chest radiographs for biometric identification. However, three main limitations were identified. To overcome these drawbacks, we proposed a Self-Residual Attention Network (SRAN) as the backbone of the Siamese triplet framework. SRAN combines residual learning with channel spatial attention mechanisms, enabling the model to selectively focus on identity-preserving skeletal traits while suppressing redundant or noisy regions.

1. Methodology

The system follows the Siamese triplet paradigm: three inputs (anchor, positive, negative) are passed through identical encoders, producing embeddings that are compared with Euclidean distance and optimized with triplet loss.

The novelty lies in the encoder backbone, which integrates ResNet-50 with multiple attention-enhanced modules:

a) **ResNet-50 backbone**: extracts initial low- and mid-level features using residual connections, mitigating vanishing gradients and ensuring robust general feature representation.
b) **Self-Residual Attention Blocks (SRABs)**: refine features by applying channel and spatial attention while preserving identity information through residual skip connections.
c) **Attention Block**: amplifies critical signals already enhanced by SRABs, ensuring finer discrimination.
d) **Reverse Attention Block**: highlights regions that might have been neglected by previous modules, preventing loss of subtle but informative skeletal details.
e) **Dense + Normalization layers**: flatten and transform feature maps into a 256-dimensional embedding vector normalized via L2, ensuring uniform representation.

This embedding generator forms the basis of both training and testing.

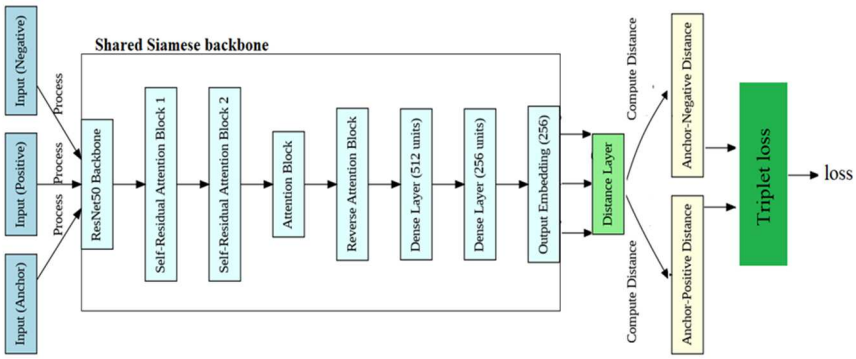


Figure 12 . Proposed Architecture.

As you can see in figure 12, the model begins with three input images anchor, positive and negative processed through the ResNet50 backbone to extract features. The output then flows into the Self-Residual Attention Block 1, where feature refinement occurs through convolution, batch normalization, and Softplus activation, complemented by channel and spatial attention mechanisms that emphasize important features while using residual connections for gradient flow. The refined output is further enhanced by Self-Residual Attention Block 2, which applies similar operations for deeper feature refinement. Next, the Attention Block enhances the feature maps using additional attention mechanisms before passing them to the Reverse Attention Block, which focuses on highlighting crucial features through reverse attention strategies. After flattening the enhanced feature maps, the model processes them through dense layers, yielding a final output embedding of 256 units. In order to help the model learn and optimize the triplet loss for identification and verification tasks, this embedding is then sent into the Distance Layer, which calculates the Euclidean distances between the anchor and both positive and negative embeddings. By making the changes mentioned above, we could reduce the information duplication across channels and identify the most crucial aspect of chest x-ray images.

1.1.ResNet-50:

ResNet50, or Residual Network with 50 layers, is a deep convolutional neural network architecture that helps address the vanishing gradient problem common in deep networks. It incorporates skip (or residual) connections, allowing the network to learn identity mappings, which eases the flow of information through layers without degrading performance as layers increase. ResNet50 has 50 convolutional layers, organized into bottleneck building blocks, and is known for its robust feature extraction capabilities, making it highly effective in computer vision tasks. In this work, ResNet50 serves as the head backbone of our Siamese neural network for person identification and verification, processing input images (anchor, positive, and negative) to extract initial feature embeddings. These embeddings are then refined through additional attention and residual blocks to ensure the model focuses on identity-distinguishing features, critical for accurate identification.

1.2.Self-Residual Attention Block (SRA Block) :

The Self-Residual Attention Block is a key building block that combines convolutional operations with both channel attention and spatial attention mechanisms to refine features. The input tensor X is first processed through a convolutional layer that applies a filter to extract meaningful features:

$$X_{conv}=Conv2D(X) \quad (8)$$

This is followed by batch normalization and a Softplus activation to stabilize the outputs and introduce non-linearity:

$$X_{conv\_bn}=BatchNormalization(X_{conv}) \quad (9)$$

$$X_{conv\_act}=Softplus(X_{conv\_bn}) \quad (10)$$

a. Channel Attention:

The channel attention mechanism helps the model focus on important channels by applying global pooling operations:

- Global Average Pooling (GAP):

$$GAP(X)=1/(H\times W)\ \Sigma_{i=1}^{H}\Sigma_{j=1}^{W}X(i,j) \quad (11)$$

- Global Max Pooling (GMP):

$$GMP(X)=\max_{i,j} X(i,j) \quad (12)$$

These pooled features are concatenated and passed through a dense layer with a sigmoid activation:

$$CA(X)=\sigma(Dense([GAP(X_{conv\_act}),GMP(X_{conv\_act})])) \quad (13)$$

Finally, the attention weights are reshaped and multiplied element-wise with the input tensor to reweight the channels:

$$X_{channel\_att}=X_{conv\_act}\times CA(X) \quad (14)$$

b.Spatial Attention:

Spatial attention highlights important spatial regions by reducing the channel dimension:

- Average Pooling across channels:

$$AvgPool(X)=1/C\ \Sigma_{k=1}^{C}X(:,:,k) \quad (15)$$

- Max Pooling across channels:

$$MaxPool(X)=\max_{k} X(:,:,k) \quad (16)$$

The concatenated features are passed through a convolutional layer with a sigmoid activation:

$$SA(X)=\sigma(Conv2D([AvgPool(X),MaxPool(X)])) \quad (17)$$

The spatial attention weights are multiplied element-wise with the channel-attended output:

$$X_{spatial\_att} = X_{channel\_att} \times SA(X) \quad (18)$$

c. Residual Connection :

A residual connection is added to preserve gradient flow and ensure stable training. If the input tensor X has a different number of channels than the output, it is reshaped using a 1×1 convolution:

$$X_{res} = Conv2D(X) \quad (19)$$

The final output is the sum of the spatially and channel-attended features and the residual connection:

$$X_{out} = X_{spatial\_att} + X_{res} \quad (20)$$

1.3.Attention Block :

The Attention Block focuses on applying attention to the already refined features, using both channel and spatial attention mechanisms. It performs similar global pooling operations, but without the residual component of the previous blocks, the attention block architecture illustrated in figure 13 .

a. Channel Attention:

Similar to the SRA Block, the global average pooling and max pooling are computed:

$$CA(X) = \sigma(Dense([GAP(X), GMP(X)])) \quad (21)$$

The output from the dense layer is reshaped and multiplied element-wise with the input tensor:

$$X_{channel\_att} = X \times CA(X) \quad (22)$$

b. Spatial Attention:

The spatial attention works similarly to the one in the SRA Block. Channel averages and maxima are computed and concatenated:

$$SA(X) = \sigma(Conv2D([AvgPool(X), MaxPool(X)])) \quad (23)$$

The spatial attention weights are multiplied element-wise with the channel-attended output:

$$X_{spatial\_att} = X_{channel\_att} \times SA(X) \quad (24)$$

Thus, the output of the Attention Block is:

$$X_{out} = X_{spatial\_att} \quad (25)$$

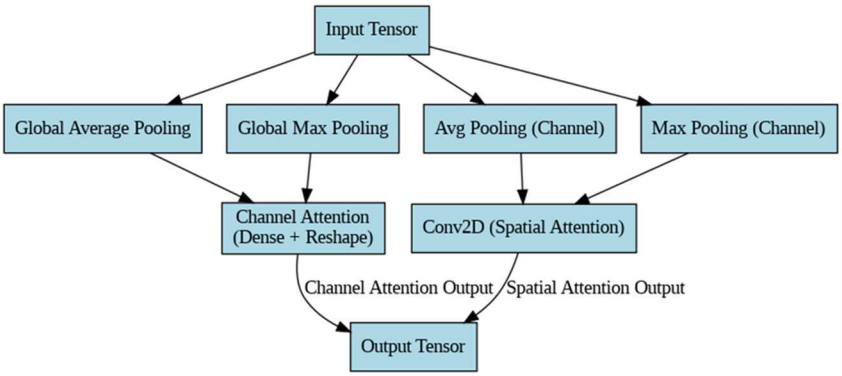


Figure 13. Attention Block architecture.

1.4. Reverse Attention Block:

The Reverse Attention Block is applied to reverse the attention from prior steps and highlight regions that might have been overlooked. This ensures comprehensive feature extraction. Like the previous blocks, the input tensor is first processed through a convolutional layer and Batch normalization and Softplus activation, showed in figure 14 .

a. Reverse Channel Attention:

Global average pooling and max pooling are applied similarly to before, but the reverse attention mechanism helps the model focus on regions previously down-weighted:

$$RA(X)=\sigma(\text{Dense}([GAP(X_{conv\_act}),GMP(X_{conv\_act})])) \quad (26)$$

The reverse attention weights are then multiplied with the output:

$$X_{reverse\_att}=X_{conv\_act}\times RA(X) \quad (27)$$

Thus, the Reverse Attention Block output is:

$$X_{out}=X_{reverse\_att} \quad (28)$$

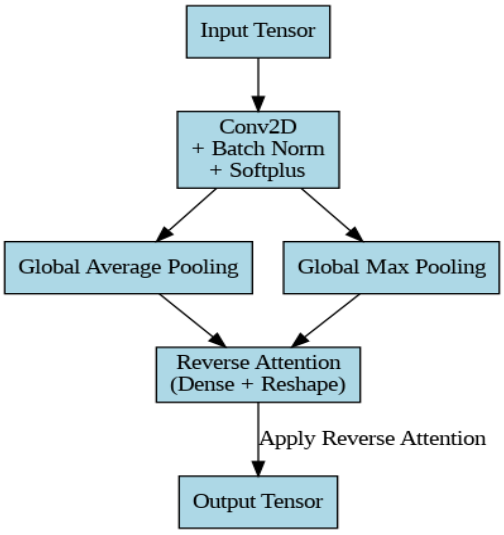


Figure 14. Reverse Attention Block architecture.

1.5.Distance Layer:

In order to calculate triplet loss, this layer calculates the Euclidean distance between the anchor, positive, and negative image embeddings.

Given two vectors, A (anchor) and P (positive), the Euclidean distance between them is:

$$D\_AP = \sqrt{([\sum(A-P)]^2)} \quad (29)$$

Similarly, the Euclidean distance between the anchor and negative embedding N is:

$$D\_AN = \sqrt{([\sum(A-N)]^2)} \quad (30)$$

1.6. Triplet Loss:

Triplet Loss is a crucial training tool that ensures the model learns to differentiate between individuals by structuring the embedding space. It works by bringing similar images closer together and pushing dissimilar images farther apart. Specifically, Triplet Loss minimizes the distance between the anchor (A) and a positive sample (P) from the same class, while maximizing the distance between the anchor (A) and a negative sample (N) from a different class. In this setup, A represents the anchor image, P is the positive sample, N is the negative sample, and the margin hyperparameter ensures that the anchor-negative distance exceeds the anchor-positive distance by a fixed amount, reinforcing separation in the embedding space.

$$L(A,P,N)=\max(0 , D\_AP - D\_AN + \text{margin}) \quad (31)$$

1.7.Data Preparation:

The images are read in RGB format after being preprocessed to a consistent size of (200, 200). To guarantee that every patient has at least two samples available, we choose one image at random to serve as the anchor for each patient. Next, we randomly select negative instances from various patients and create positive examples for each patient who has the same patient ID as the anchor image. For uniformity, the RGB images are downsized to 200x200. Using images from the ChestX-ray14 dataset, we choose an anchor image and generate positive pairs (same patient ID) and negative pairs (different patient IDs) in order to train Siamese network architectures using triplet loss. Additionally, we use CheXpert dataset just for testing, we construct actual image pairs for each subset, excluding patients who have only one radiograph, by using the patient ID labels to form positive and negative pairs.

2. Testing phase:

The embedding serves as the foundation for our distance-based verification and identification tasks, relying on the learned similarity between image embeddings to establish accurate identification. To comprehensively evaluate the effectiveness of our system, we assess both our trained model and the same architecture without training, as a baseline, for comparison. The results of these analyses are detailed in the following two sections, which present identification and verification outcomes individually. We use two datasets for testing: the NIH ChestX-ray14 dataset and the CheXpert dataset. The inference time of the trained model is optimized for practical applications, with embedding generation taking approximately 3 hours for 43000 patients. The subsequent identification process requires 9 minutes, while the verification process completes in just 1.6 minutes, demonstrating the system's efficiency and readiness for real-world deployment. For Inference we need a GPU

such as an NVIDIA T4, V100 or Nvidia RTX 4090 is preferred to accelerate the embedding generation and ensure efficient computation during the identification and verification processes. Else Google Colab (Pro+ for faster GPUs) could be more efficient and at least 16 GB of RAM is required for managing large image datasets and intermediate computations. For storage we need a minimum of 500 GB of storage is recommended to store the model, datasets (NIH ChestX-ray14 and CheXpert), embeddings, and temporary files generated during inference. Our point of power that we don't need to retrain our model in cases of new datasets or patients, we generate just embeddings and store them using our embedding generator.

2.1.Identification:

In the identification process showed in figure 16, the system processes an anchor image (a reference chest X-ray) along with multiple comparison images by passing each image through the embedding generator. The model extracts a 256-dimensional embedding for each image, capturing identity-specific features. To determine identity, the Euclidean distance between the embedding of the anchor image and each comparison image is calculated. The system then ranks the comparison images by distance and selects the top 5 closest matches showed in figure 15 example of identifying a person. This top-5 ranking approach enhances identification accuracy by allowing for slight variations between images while prioritizing those with highly similar embeddings. To integrate new patients into the system, first, generate embeddings for each radiograph of the new patients using the proposed model the we store these embeddings in a secure database alongside patient identifiers, when a new query radiograph would be processed by the model, generating a fresh embedding, which would then be matched against stored embeddings using our distance-based approach. The top-5 similarity search method would help identify the patient by retrieving the closest matches. We will also evaluate the performance of this identification process using the non-trainable architecture to compare the effectiveness of our trained model against the baseline in table 1.

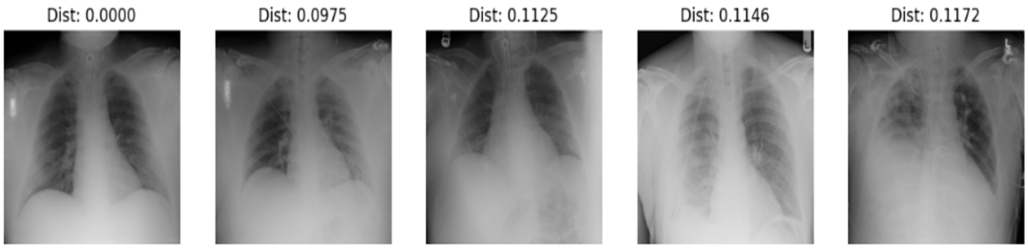


Figure 15. Example of testing simple for identification.

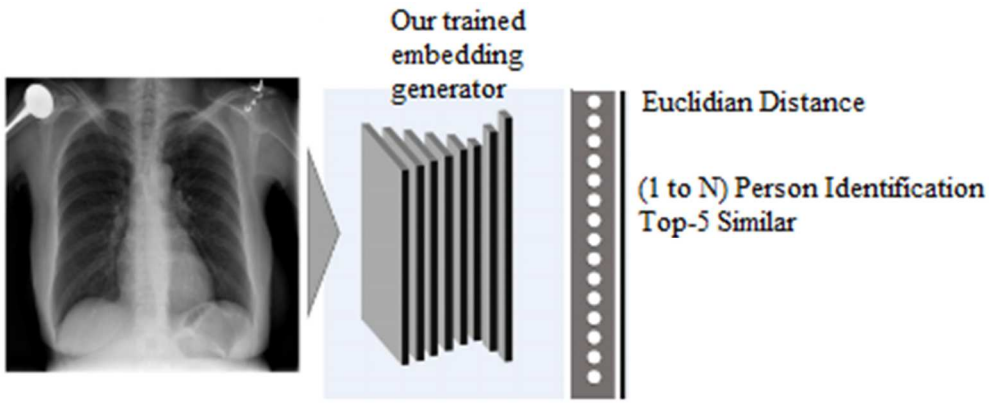


Figure 16. Identification System.

2.2.Verification:

In the verification process, the system assesses whether two given images correspond to the same individual. Both images are passed through the embedding generator of the Siamese model, which extracts their 256-dimensional embeddings. The Euclidean distance between these two embeddings is calculated. If the distance is small (below the specified threshold), the system confirms that the two images represent the same person. Conversely, if the distance is large, the images are classified as belonging to different individuals. The proposed patient verification system, which operates on a one-to-one basis, is illustrated in Figure 17. We will also evaluate this verification process using the non-trainable architecture to establish a baseline for comparison against our trained model in table 1.

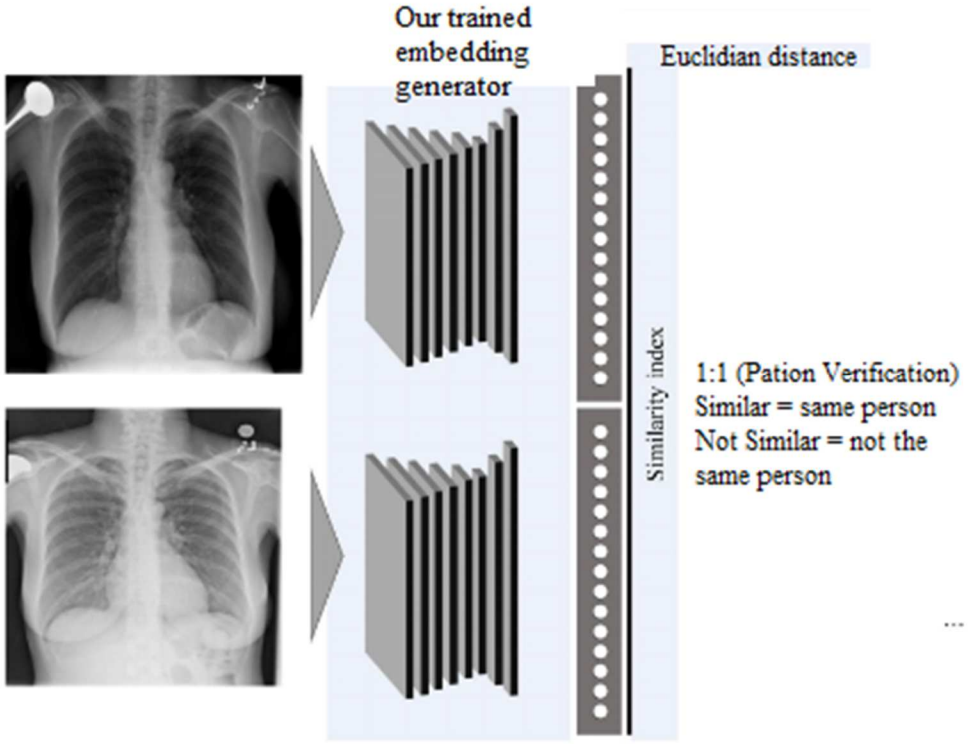


Figure 17 : Verification System.

3. Results and Discussion

In this section, we evaluate our trained Siamese model through multiple analyses, including the illustration of loss over 40 and 80 epochs in figure 18, that's proving that loss stable at 0.000. Additionally, we assess the model's ability to generate embeddings effectively and examine its verification and identification performance. For a comprehensive comparison we find that ResNet50, as a standalone backbone, typically trains 20–25% faster due to reduced architectural complexity. we also test the model architecture without training to observe its baseline performance. The findings can be compared with existing literature and theoretical frameworks to highlight similarities, differences, or new insights. This section reviews the empirical data obtained from the previously described experimental protocols. The primary goal is to present the results objectively, often utilizing tables, figures, and statistical analyses for clarity. By contrasting the results with prior research or theoretical expectations, we aim to uncover patterns, discrepancies, or novel discoveries. Our proposed method is implemented using the Keras library in Python, else we test our embedding generator in verification and identification using NIH and CheXpert Datasets, results showed in table 4.

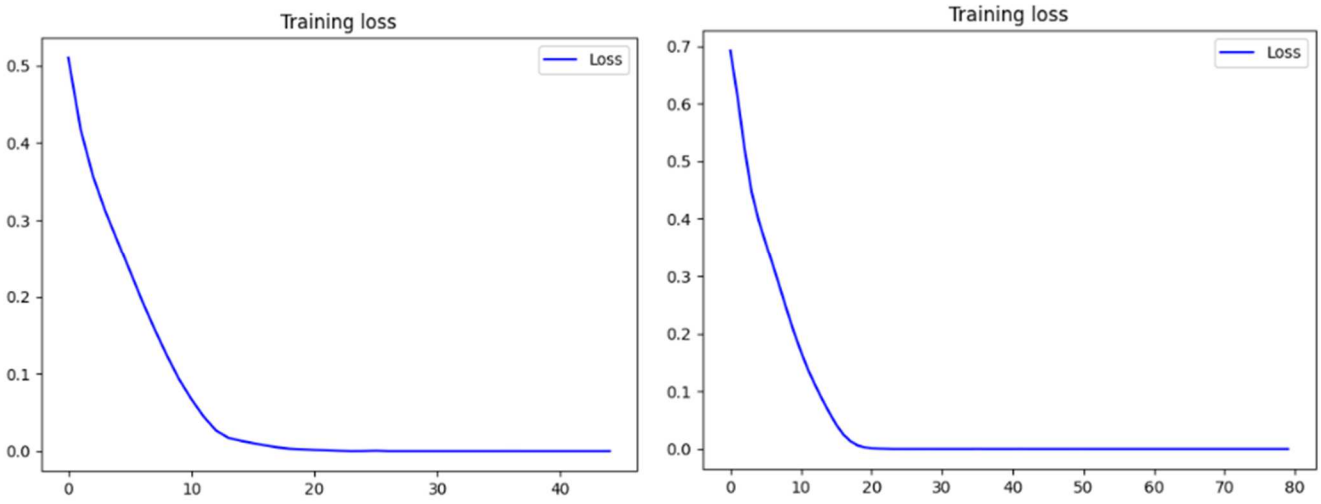


Figure 18: Training loss against epochs 40 and 80.

Model	Task	Accuracy NIh dataset	Accuracy CheXpert Dataset
Our Trainable Model	Identification	98.3	96.1
	Verification	91	90.8
The Same Architecture without training	Identification	56	61.2
	Verification	59	57

Table 4: Performance of our system.

4. Strengths of the SRAN Contribution

This contribution improves upon the baseline Siamese triplet approach by:

a) **Focusing on skeletal identity cues**: attention blocks emphasize ribs, clavicles, and vertebrae while reducing sensitivity to soft tissue.
b) **Improving robustness**: reverse attention ensures that overlooked features are recovered, enhancing generalization.
c) **Reducing feature redundancy**: channel reweighting prevents duplication across embeddings, ensuring efficient identity representation.
d) **Better forensic applicability**: unlike soft-tissue based methods, SRAN embeddings retain discriminative power in forensic or postmortem cases.

5. Limitations and Transition

Despite strong improvements, SRAN still processes full chest radiographs, meaning that soft tissues continue to influence embeddings, albeit to a lesser degree. In forensic scenarios where decomposition or pathology removes soft tissues, this may lead to performance degradation.

To address this limitation, our third contribution introduces a segmentation-based skeletal biometric system, where only bones (ribs, clavicles, sternum, spine) are isolated and used for recognition. This step further strengthens permanence and reliability for postmortem identification.

6. Conclusion

This chapter presented two deep learning–based biometric identification approaches relying on chest X-ray images as a whole. The first contribution demonstrated the feasibility of learning identity-specific representations using Siamese networks and triplet loss, confirming that chest X-ray images contain discriminative biometric information. However, the observed sensitivity to intra-subject variability and background activations highlighted the limitations of purely global representations. The second contribution addressed these limitations through the integration of self-residual attention mechanisms, enabling the model to focus more effectively on informative regions and improving identification and verification performance. The results confirm that attention-guided architectures enhance embedding discriminability and robustness.

Despite these improvements, both approaches remain influenced by non-anatomical variations and soft-tissue changes. These observations motivate the transition toward anatomically constrained representations, forming the basis of the skeletal-focused contributions developed in the next chapter.

CHAPTER IV :

Skeletal Representation and Bone-Based Biometric Identification

Chapter VI : Skeletal Representation and Bone-Based Biometric Identification

1. Introduction

The appearance-based biometric approaches presented in the previous chapter demonstrate that chest X-ray images contain identity-discriminative information. However, full radiographs include both skeletal structures and soft tissues, the latter being susceptible to variations caused by pathology, aging, imaging conditions, and postmortem changes. Such variability may reduce robustness and limit forensic applicability.

To address these limitations, this chapter focuses on skeletal representation and bone-based biometric identification using chest X-ray radiographs. Skeletal structures such as ribs, clavicles, the sternum, and the vertebral column exhibit high anatomical stability and inter-subject uniqueness, making them particularly suitable for biometric and forensic identification, including postmortem scenarios.

This chapter introduces two tightly connected contributions. The first presents a deep learning–based framework for automatic bone segmentation from chest X-ray images, establishing a dedicated anatomical preprocessing stage. The second contribution integrates the segmented skeletal representations into a biometric identification system, evaluating their discriminative capability, robustness, and interpretability.

By explicitly separating segmentation and identification, this chapter enables a detailed analysis of segmentation accuracy, error propagation, and the role of skeletal anatomy in identity recognition, thereby advancing chest X-ray biometrics toward forensic-grade reliability.

2 System Overview

The skeletal-based biometric framework extends the general recognition pipeline by incorporating a dedicated bone segmentation stage, as illustrated in Figure 8. The system is designed to restrict identity recognition to anatomically stable skeletal structures.

2.1 Image Acquisition

Chest X-ray radiographs are acquired from clinical archives or forensic examinations. Both premortem and postmortem images may be used, depending on the application context.

2.2 Bone Segmentation and Preprocessing

A segmentation stage based on an Attention U-Net architecture is applied to isolate skeletal structures, including ribs, clavicles, sternum, and the vertebral column. This step suppresses soft tissue information and ensures that subsequent processing relies only on stable anatomical features.

2.3 Skeletal Representation Learning

The segmented bone regions are processed by deep learning models employing transfer learning and attention-enhanced Siamese architectures. Metric learning with triplet loss is used to generate compact skeletal embeddings that capture individual-specific anatomical characteristics.

2.4 Template Generation

The resulting skeletal embeddings are stored as biometric templates in a secure database. These templates may be associated with identity labels or forensic case identifiers.

2.5 Matching and Identity Decision

In verification mode (1:1), the query skeletal template is compared against a claimed identity. In identification mode (1:N), it is matched against all enrolled templates. Similarity scores are computed, and the system outputs the most likely identity or validates the claim based on predefined criteria.

By focusing exclusively on skeletal structures, the proposed framework enhances robustness against soft-tissue variability and supports reliable biometric identification in both security and forensic scenarios.

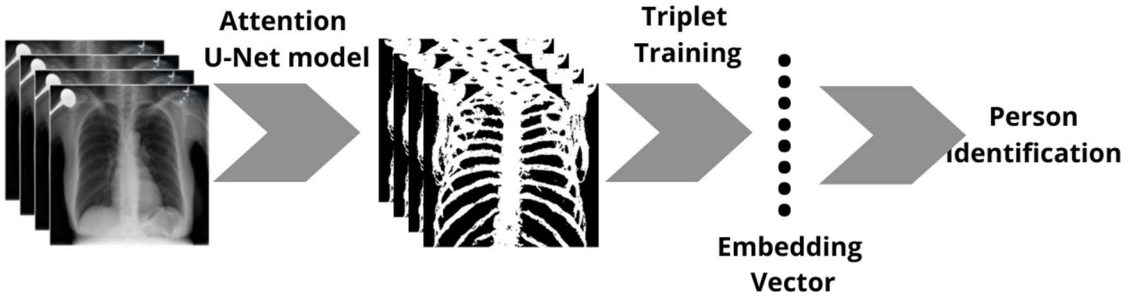


Figure 19: general system overview using bone.

Contribution 3:

Bone Segmentation from Chest X-ray Radiographs

One of the major challenges in developing a chest X-ray–based biometric system lies in the presence of soft tissue. Lungs, heart, and other soft tissues not only change significantly over time due to pathology, disease, or decomposition, but also introduce variations in appearance across radiographs. This reduces the permanence and reliability of soft-tissue features for long-term or forensic identification.

To overcome this limitation, we introduce bone segmentation as a separate contribution. The aim is to isolate the skeletal structures (ribs, sternum, clavicles, vertebrae), which are more stable and distinctive over time. By focusing exclusively on bones, we lay the foundation for a robust biometric system that remains reliable even in postmortem scenarios, where soft tissues are typically degraded.

1. Methodology

For the segmentation task, we adopt Attention U-Net, an advanced variant of the traditional U-Net architecture. Attention U-Net enhances segmentation performance by integrating attention gates into the skip connections, allowing the network to automatically focus on the most relevant anatomical regions and suppress irrelevant structures.

- **Input**: Raw chest X-ray radiographs.
- **Encoder**: Convolutional layers with downsampling capture multi-scale features.
- **Attention Gates**: Guide the network to highlight skeletal regions while reducing focus on soft tissue.
- **Decoder**: Upsampling with skip connections reconstructs bone masks at the original resolution.
- **Output**: Binary bone mask that highlights ribs, clavicles, sternum, and vertebral column.

Our study introduces a deep learning-driven approach for bone segmentation in chest radiographs, leveraging a U-Net architecture augmented with attention mechanisms. The subsequent sections elaborate on the data preparation strategy, architectural design, training methodology, and dataset specifications. Our approach is designed to enhance segmentation accuracy and robustness, offering valuable support in forensic and clinical applications, particularly in postmortem identification scenarios.

1.1.Dataset

In line with the methodology described in [82], this work makes use of the LIDC-IDRI dataset. To isolate bone structures from the CT scans, a segmentation strategy based on Hounsfield Units (HU) was applied. Voxels corresponding to bone were defined within the range of 300–700 HU, whereas non-bone regions such as air and soft tissue were assigned a value of -1024 HU, reflecting the standard CT representation of air. Using these segmented bone voxels, synthetic two-dimensional "bone X-ray" images were produced, which acted as ground-truth masks for model training. This process allowed the generation of a comprehensive synthetic dataset composed of paired digitally reconstructed radiographs (DRRs) and their corresponding bone masks. Overall, the dataset includes 386 images alongside their respective masks, providing a solid foundation for training the deep learning model aimed at bone segmentation and enhancement. Figure 19 illustrates a sample chest X-ray image alongside its corresponding mask.

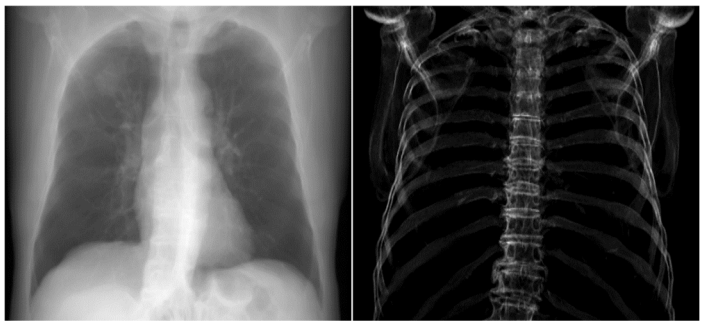


Figure 20: Sample of chest x ray image with mask

1.2.Data Preprocessing

During the preprocessing stage of bone segmentation from chest X-ray images, the Digitally Reconstructed Radiograph (DRR) dataset is first examined to verify the consistency and reliability of the input data. The bone structures are then delineated using the masks generated through the method described in [82]. To refine these masks, histogram equalization is applied, which enhances image contrast and emphasizes key structural features, as illustrated in Figure 20. This preprocessing pipeline plays a vital role in preparing high-quality input data for the segmentation network, thereby improving both the training efficiency and the accuracy of the final segmentation results.

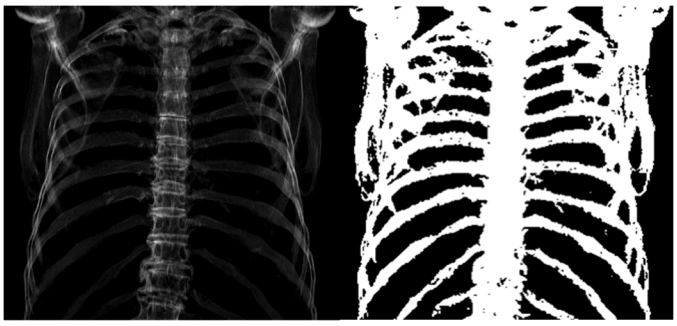


Figure 10: Preprocessing Results.

1.3.Bone Segmentation

In this study, we propose a deep learning framework for bone segmentation that builds upon an attention-augmented U-Net architecture. The design directly tackles the difficulties of extracting bone structures from X-ray images, where low contrast with surrounding soft tissues, imaging noise, and anatomical variability present major challenges. Accurate identification of both cortical and trabecular regions is especially important for reliable diagnosis and treatment planning [83,84].

Our model adopts a symmetric encoder–decoder structure with skip connections, further enhanced by attention mechanisms that highlight the most relevant bone features. The encoder progressively reduces spatial resolution while capturing hierarchical feature representations, enabling the network to learn both global context and fine details. Although skip connections preserve spatial information, they may also introduce redundant or irrelevant features. To address this, we integrate attention gates into the skip pathways. These gates generate spatial attention maps by combining encoder features with contextual information from deeper layers, thereby amplifying critical bone structures while suppressing background noise. This targeted filtering allows the model to delineate complex bone boundaries with greater precision. The inclusion of attention gates is particularly beneficial in clinical contexts, such as orthopaedic surgery, trauma evaluation, and diagnostic imaging, where reliable bone segmentation is essential [84]. Additionally, the framework's ability to robustly capture skeletal characteristics opens opportunities for biometric applications, where bone structure can serve as a distinctive identifier for identity verification. In essence, the attention-enhanced U-Net offers a strong and adaptable solution for bone segmentation from X-ray data, supporting both improved clinical workflows and emerging applications in medical imaging and biometrics.

1.4.Network Architecture

The proposed architecture comprises three primary components Encoder Blocks, Decoder Blocks, and Attention Gates as illustrated in Figure 21

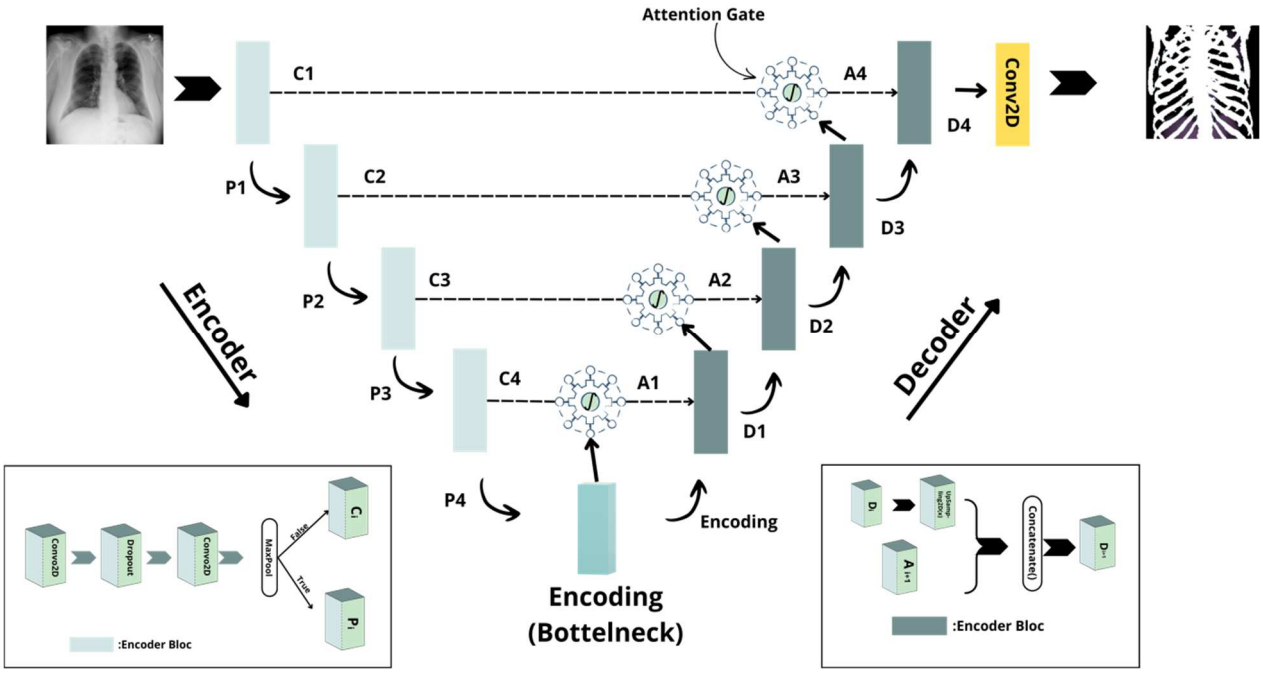


Figure 21: The proposed architecture

A. Encoder Blocks

Each encoder block is responsible for gradually capturing hierarchical features while reducing the spatial resolution of the input image. Structurally, the block is composed of two consecutive convolutional layers with 3×3 kernels, each followed by ReLU activations. To prevent overfitting, a dropout layer is placed between the convolutions. In addition, a max-pooling operation can be applied after the convolutions to further downsample the feature maps. The encoder block produces two outputs: a downsampled feature map that continues through the encoding pathway and a pre-pooled feature map that is preserved for use in the skip connections.

B. Decoder Blocks

The decoder pathway rebuilds the segmentation map from the compressed features produced by the encoder. Each decoder block begins with an up-sampling step to restore spatial resolution. The up-sampled output is then merged with the corresponding skip connection from the encoder, ensuring that both contextual and spatial details are retained. This combined feature set is passed through a convolutional block structured similarly to the encoder blocks but without pooling to generate a refined output. The up-sampling factor is carefully selected so that the resolution aligns with the corresponding encoder feature map.

C. Attention Gates

To enhance the effectiveness of skip connections, attention gates are incorporated into the architecture. These gates act as filters that highlight bone structures regions of interest while suppressing irrelevant background information. Each gate takes two inputs: a gating signal from deeper layers of the network and a skip connection from the encoder. The gating signal is processed through a convolutional layer with ReLU activation, while the skip connection is down-sampled using a convolution with stride 2. The resulting feature maps are combined

and passed through another convolutional layer with a sigmoid activation, producing an attention mask. This mask is then up-sampled to match the spatial size of the original skip connection and applied via element-wise multiplication to refine the features. Optionally, batch normalization can be included to further stabilize the output. An illustration of the attention gate is provided in Figure 22.

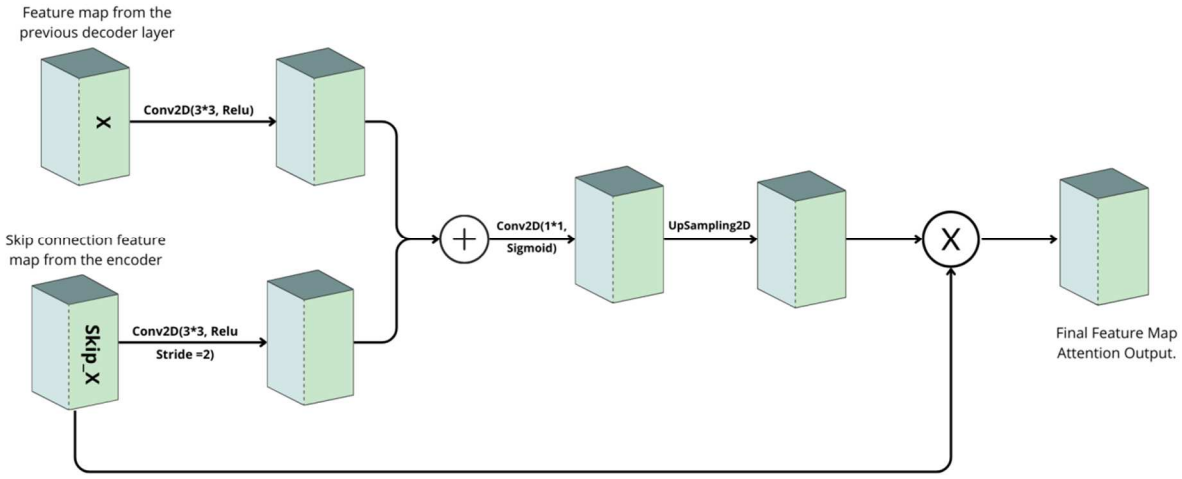


Figure 22: Attention gate architecture.

2. Overall Pipeline

The full network begins by processing the input X-ray image through a sequence of encoder blocks that progressively down-sample the data while capturing multi-scale features. At the deepest stage, a bottleneck layer implemented without pooling provides a compact yet comprehensive representation of the input. During decoding, spatial resolution is gradually restored through up-sampling, while attention gates refine the skip connections by suppressing irrelevant features and emphasizing bone structures. The refined encoder features are concatenated with the decoder outputs and passed through convolutional layers, ultimately leading to a 1×1 convolution with sigmoid activation that generates the final bone segmentation mask.

This attention-guided U-Net is particularly well-suited for bone segmentation in X-ray imaging, where accurate delineation of fine skeletal structures is essential for diagnosis and treatment planning. Beyond clinical use, its ability to robustly capture bone characteristics also opens promising applications in biometrics, such as identity verification. As future directions, we plan to benchmark this framework against conventional U-Net variants and investigate the inclusion of additional imaging modalities to further enhance segmentation performance.

3. Results

To rigorously evaluate the proposed attention-based U-Net for bone segmentation, a structured experimental framework was established. The model was implemented in Python using the Keras library and executed on Google Colab. The dataset, derived from LIDC-IDRI,

contained 386 X-ray images with corresponding bone masks, which were divided into 80% for training and 20% for validation to maintain balance and reproducibility.

An ablation study was designed to examine the specific contributions of the attention mechanism and data augmentation. Four experimental setups were considered:

- Baseline U-Net : a standard U-Net trained on the original dataset without attention or augmentation, used as a performance reference.
- U-Net with Attention : integration of the attention mechanism while keeping the original dataset unchanged, isolating its direct effect on segmentation accuracy.
- U-Net with Data Augmentation : application of augmentation methods such as rotation, scaling, contrast adjustment, and horizontal flipping, which expanded the dataset from 386 to 3,088 images, without incorporating attention. This tested the role of augmentation in improving generalization.
- Proposed Model : a combination of both the attention mechanism and data augmentation, aimed at assessing the synergy between these enhancements.

Alongside quantitative evaluation, qualitative comparisons were performed by overlaying predicted bone masks on the original X-ray images. These visual analyses highlighted both segmentation accuracy and anatomical consistency across experiments. Overall, this structured ablation study provided a clear understanding of how attention mechanisms and data augmentation contribute individually and jointly to the performance of the proposed model.

1.1.Performance Evaluation Metrics

To thoroughly evaluate the effectiveness of the proposed bone segmentation framework, we employ a set of well-established metrics commonly used in image segmentation tasks. These include the loss function, Dice coefficient (Dice), and precision, each providing complementary perspectives on model performance. The loss function implemented as either binary cross-entropy or Dice loss measures the difference between the predicted segmentation and the ground truth, guiding the optimization process during training. While loss is an essential indicator of convergence, it does not fully reflect segmentation accuracy, making additional metrics necessary. The Dice coefficient, expressed in Equation (32), quantifies the degree of overlap between predicted masks and reference annotations. With values ranging from 0 to 1, it is particularly well suited for medical imaging applications, as it effectively handles class imbalance. A higher Dice score corresponds to better alignment between the model's predictions and the ground truth.

$$dice = \frac{2 \times | X \cap Y |}{| \mathrm{x} | + | \mathrm{y} |} \qquad (32)$$

where X represents the set of predicted pixels, Y represents the ground truth pixels, and $|X \cap Y|$ denotes their intersection.

In addition, precision, defined in Equation (33), measures the proportion of correctly identified positive pixels out of all pixels predicted as positive. This metric highlights the model's ability to reduce false positives, which is especially important in medical imaging

tasks. High precision ensures that detected structures correspond to true anatomical regions, thereby reducing the risk of misdiagnosis or unnecessary clinical procedures.

$$precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad (33)$$

To provide a holistic evaluation of the model's performance, we analyze loss, Dice coefficient, and precision across multiple iterations. The graphical representations of evaluation loss and precision coefficient over training iterations are presented in Figures 3, and 4, respectively. By leveraging these metrics, we ensure a rigorous assessment of our approach, validating its capability to achieve accurate and reliable bone segmentation in chest radiographs.

1.2.Discussion

The performance of the proposed method was examined across several experimental setups to analyze the individual and combined effects of attention gates (AG) and data augmentation (DA) on bone segmentation. The results are summarized in Table 2, which highlights how each component contributes to improving accuracy and robustness. The baseline U-Net, used as a reference model, obtained a Dice coefficient of 83% and a precision of 87%, consistent with typical outcomes from conventional segmentation approaches. Incorporating data augmentation improved the model's generalization ability, raising the Dice score to 89% and lowering segmentation loss from 0.21 to 0.19. When only attention gates were introduced, the Dice score increased further to 92%, demonstrating their effectiveness in refining feature selection by suppressing irrelevant background regions. The best results were achieved when AG and DA were combined. In this configuration, the model reached a Dice score of 94%, the highest among all experiments, along with a precision of 98% and a segmentation loss of 0.06. These findings underline the complementary nature of AG and DA: attention gates enhance the model's focus on relevant bone structures, while data augmentation strengthens its ability to generalize across variations in imaging conditions. Together, they enable the generation of accurate, well-defined segmentation maps with superior reliability.

Experiment	U-net		U-net + AG	
	Without DA	+DA	Without DA	+ DA
Dice	83%	89%	92%	94%
Precision	87%	92%	97%	98%
Loss	0.21	0.19	0.1	0.06

Table 5: Segmentation Performance Across Experimental Configurations

A closer examination of the segmentation outputs shows that models equipped with attention gates (AG) achieve markedly better boundary delineation, especially in anatomically complex

areas where bones overlap with surrounding soft tissues. The baseline U-Net, by comparison, often produces over-segmented or fragmented contours in these regions. By selectively enhancing bone structures and suppressing irrelevant background information, AG improves spatial awareness and leads to more consistent segmentation results. The considerable decrease in Dice score from 94% with the full model to 83% without AG and data augmentation (DA) further highlights the importance of these components in ensuring robust performance. Moreover, the high precision obtained with AG demonstrates a clear reduction in false positives, an essential factor in medical imaging where misclassifications may adversely affect clinical decision-making.

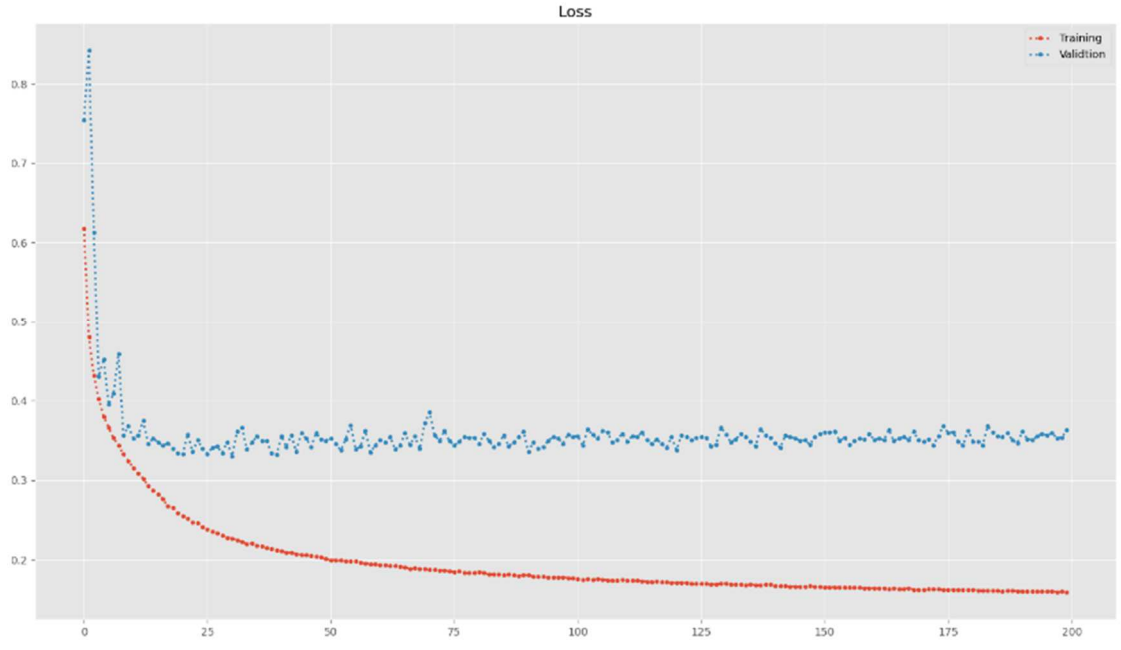


Figure 21: Training and validation Loss Curve.

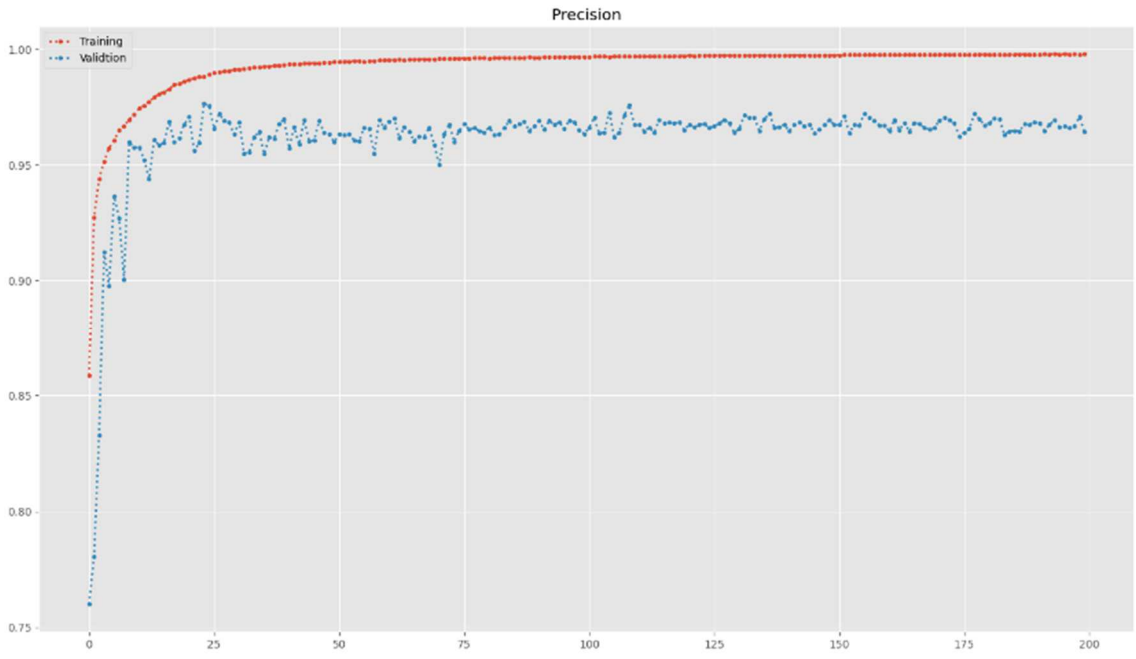


Figure 22: Training and validation Precision Curve.

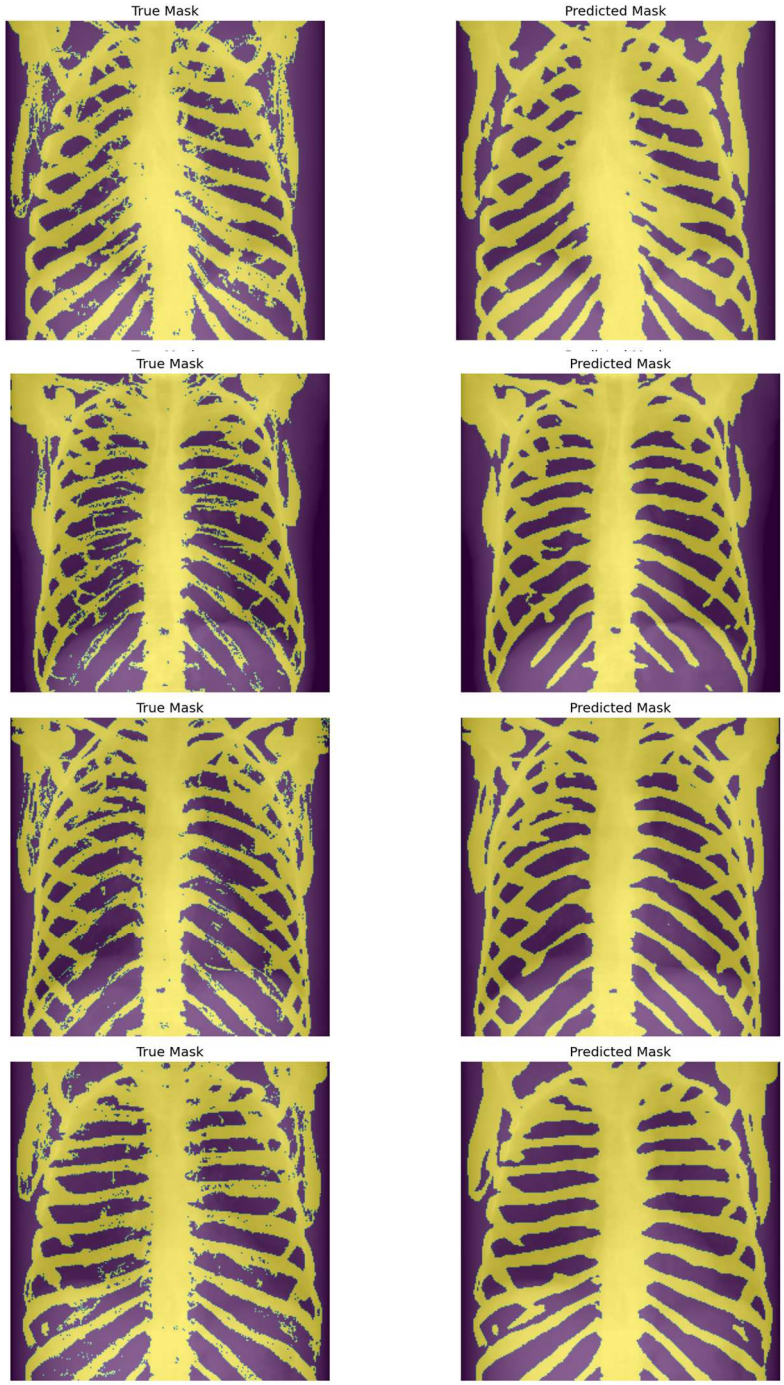


Figure 23: Comparaison of Ground truth and predicted Segmentation Masks.

Figures 23 and 24 illustrate the training loss and precision curves, respectively. The loss curve shows a steady decline throughout training, converging at lower values for the attention-enhanced U-Net compared to the baseline. Similarly, the precision curve demonstrates the stability of the proposed model, with validation precision remaining consistently high across epochs. A qualitative comparison is provided in Figure 25, where predicted segmentation masks are overlaid with ground truth annotations. These visual results confirm that the integration of attention gates improves segmentation quality by reducing noise and sharpening structural boundaries, particularly in challenging ribcage regions.

Contribution 4:

Skeletal-Based Biometric Identification Using Chest X-ray Bones

The previous contributions demonstrated that chest X-ray radiographs can serve as a promising biometric modality. However, they also highlighted key limitations. Approaches based on the full radiograph (Contributions 1 and 2) achieved strong performance but were inherently dependent on soft-tissue structures such as the lungs and heart. These regions are prone to pathological variations during life and are subject to rapid decomposition after death, making them less reliable for long-term or forensic identification. Contribution 3 addressed this limitation by introducing a segmentation strategy that isolates the skeletal structures, which are anatomically stable and resistant to postmortem changes.

Building upon this foundation, the fourth contribution presents a complete skeletal-based biometric recognition framework that relies exclusively on chest bones as the primary source of identity information. This transition from soft tissue to bone transforms chest X-ray biometrics into a modality suitable not only for clinical and security applications but also for forensic science and disaster victim identification.

1. Methodology

The proposed skeletal-based biometric system integrates the segmentation module with advanced deep feature learning to ensure robustness and discriminability. The pipeline proceeds as follows.

a) **Image Acquisition**

Chest X-rays are collected either from large-scale clinical repositories (NIH, CheXpert) in pre-mortem cases, or from forensic imaging setups in postmortem contexts.

b) **Bone Segmentation**

Attention U-Net, developed in Contribution 3, isolates ribs, clavicles, sternum, and vertebral column. The segmented masks are combined with the original images to yield skeletal-only radiographs.

c) **Feature Extraction**

The skeletal images are passed through a Siamese neural network enhanced with Resnet-50 backbone. This architecture emphasizes subtle skeletal traits while reducing redundancy and noise.

d) **Embedding Generation**

Each skeletal image is encoded into a 256-dimensional normalized feature vector. These embeddings capture person-specific anatomical patterns that remain stable across time and conditions.

e) **Matching and Decision Making**

Identification (1:N): Matches a query embedding against the database to find the closest identity.

Our proposed method employed a deep learning-based approach to segment skeletal bone from chest X-ray radiographs, enabling robust biometric person identification, illustrate in figure 26. We utilized the pre-trained model ResNet50 as the backbone of a triplet network, which was trained with triplet loss to generate discriminative embeddings for person identification. The segmented bone structures served as the primary input for our embedding generator, which produced unique 256-dimensional feature representations for each image. This feature extraction process enhanced the robustness and accuracy of our identification framework.

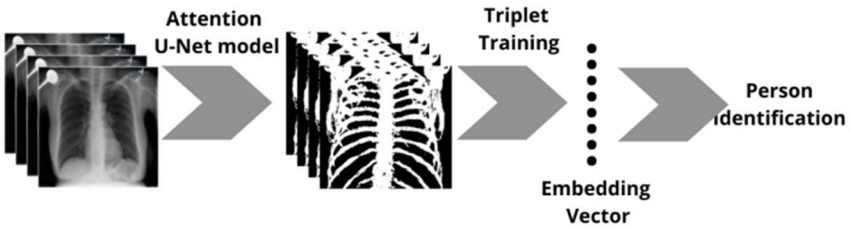


Figure 26: Global architecture for person identification.

1.1.Bone segmentation

In this work, chest X-ray images are preprocessed to isolate skeletal bone structures for use in person identification. The dataset includes X-ray scans accompanied by metadata such as patient ID, age, gender, and view position. To maintain consistency, all images are resized and normalized before being processed by the deep learning framework. For segmentation, an attention-enhanced U-Net is employed to effectively separate skeletal features from soft tissue and background noise. The model produces binary masks that emphasize the key bone structures, which serve as the foundation for building biometric identification models. By focusing on intrinsic skeletal characteristics, the system ensures robustness and reliability in person recognition.

The preprocessing pipeline is specifically designed to deliver high-quality segmentation, which is essential for accurate identification. As illustrated in Figure 21, the framework is built on a symmetric encoder–decoder architecture with skip connections, augmented by attention mechanisms that highlight the most relevant bone features. The encoder progressively reduces spatial resolution while extracting hierarchical representations that capture both fine details and global context. Although skip connections help preserve spatial information, not all transferred features are equally useful for bone delineation. To refine this process, attention gates are embedded within the skip connections. These gates generate spatial attention masks by combining deeper contextual signals with encoder features, selectively amplifying bone structures while suppressing irrelevant background information. This targeted refinement enhances the model's ability to accurately capture complex skeletal boundaries.

1.2. Person identification

We propose a deep learning framework for person identification from medical imaging, designed with a particular focus on challenging post-disaster and postmortem scenarios. Situated at the intersection of healthcare and security, our approach utilizes chest X-ray

radiographs as a reliable biometric source. At the core of the system lies a Siamese neural network, well suited for biometric applications due to its capacity to compare paired inputs through shared weights. The network maps anchor, positive, and negative chest X-ray images into discriminative feature vectors, thereby enabling effective representation learning.

As shown in Figure 27, all three inputs are processed by a common encoder with identical weights to ensure consistent feature extraction. For this purpose, we adopt ResNet-50 as the backbone architecture, leveraging its strong representational power through transfer learning. The resulting embeddings are passed to a distance layer optimized with a triplet loss function, which calculates Euclidean distances between the anchor and both the positive and negative samples. This setup enforces compact clustering of embeddings from similar identities while pushing apart those belonging to different individuals, ensuring robust and discriminative performance for person identification.

1.2.1. Siamese Neural Network

A Siamese neural network is designed to process multiple inputs with shared weights and compare their resulting feature representations, which is why it is often referred to as a twin network. Initially developed for applications such as signature verification and face recognition, this architecture has proven highly effective for similarity-based tasks [85–87]. In our work, we extend this concept by implementing a triplet-based Siamese network that encodes three inputs anchor, positive, and negative chest X-ray images into feature vectors. The encoder, built upon a pre-trained backbone with additional layers, generates three distinct feature representations. Similarity is then assessed by computing distances between these vectors, with the triplet loss function guiding the optimization. This ensures that embeddings from the same individual are pulled closer together, while those from different individuals are pushed farther apart, enabling robust identity discrimination.

1.2.2. Transfer Learning

Developing a chest X-ray identification model entirely from scratch demands extensive time and significant computational resources. To overcome these challenges, we adopt transfer learning, integrating pre-trained models to strengthen feature extraction. This strategy takes advantage of the broad and generalized representations acquired from large-scale datasets such as ImageNet, which is especially valuable in medical imaging tasks where annotated data is often limited. In this study, ResNet-50 is selected as the backbone encoder because of its strong residual learning framework and its well-documented success in medical image analysis. By introducing skip connections, ResNet-50 effectively mitigates the vanishing gradient issue, allowing the network to learn identity mappings more efficiently:

$$Y=F(x,\{Wi\})+x \tag{34}$$

Here, $F(x,\{Wi\})$ is the residual mapping, and x is the input. This structure allows gradients to flow directly through the identity connection, making deep networks easier to optimize. ResNet-50 consistently demonstrates higher accuracy and robustness when extracting discriminative features from grayscale chest X-ray images. Its depth (50 layers) strikes a balance between complexity and efficiency, capturing both local and global features effectively. In our setup, ResNet-50 is used as a frozen feature extractor, preserving its pre-

learned convolutional weights. The extracted feature maps are then processed by additional custom layers tailored to the identification task:

Fully Connected Layer: Projects the high-dimensional output to a lower-dimensional embedding:

$$Z=\sigma(W\cdot f+b) \quad (35)$$

where f is the feature vector from ResNet-50, W and b are trainable weights, and σ (sigma) is a non-linear activation function (e.g., ReLU).

Batch Normalization: Improves convergence and generalization by normalizing each mini-batch:

$$X^{(k)} = \frac{x^{(k)}-\mu_B^k}{\sqrt{(\sigma_B^{(k)})^2+\epsilon}} \quad (36)$$

L2 Normalization: Ensures all embeddings lie on a hypersphere, which stabilizes similarity comparisons:

$$V = \frac{z}{\|z\|_2} \quad (37)$$

The final normalized feature vector v is used to compute distances between samples using Euclidean in the distance layer. Empirically, ResNet-50 achieves faster convergence, higher top-1 accuracy, and better generalization compared to other architectures, making it an optimal choice for chest X-ray identification tasks.

1.2.3. Euclidean Distance and Triplet Loss

In a Siamese network with triplet loss, the similarity between embeddings is quantified using Euclidean distance. The distance between the anchor f(A) and positive f(P) embeddings, as well as between the anchor f(A) and negative f(N) embeddings, is computed as follows:

$$distance\_ap = \sum i = 1n(ai - pi)2 \quad (38)$$

$$distance\_an = \sum i = 1n(ai - ni)2 \quad (39)$$

Where f(A),f(P), and f(N) represent the feature embeddings of the anchor, positive, and negative images, respectively. The objective is to learn an embedding space where anchor-positive pairs are closer together while anchor-negative pairs are farther apart. This is achieved using the triplet loss function, which is formulated as:

$$loss(A, P, N) = max(ap\_distance - an\_distance + margin, 0.0) \quad (40)$$

Here, the margin is a hyper-parameter that enforces a minimum separation between similar and dissimilar embeddings. The loss is minimized when the anchor-positive distance remains smaller than the anchor-negative distance by at least the specified margin [88]. By tuning this margin, we can balance embedding separation and model robustness, ensuring that the learned representations effectively distinguish individuals based on chest X-ray features.

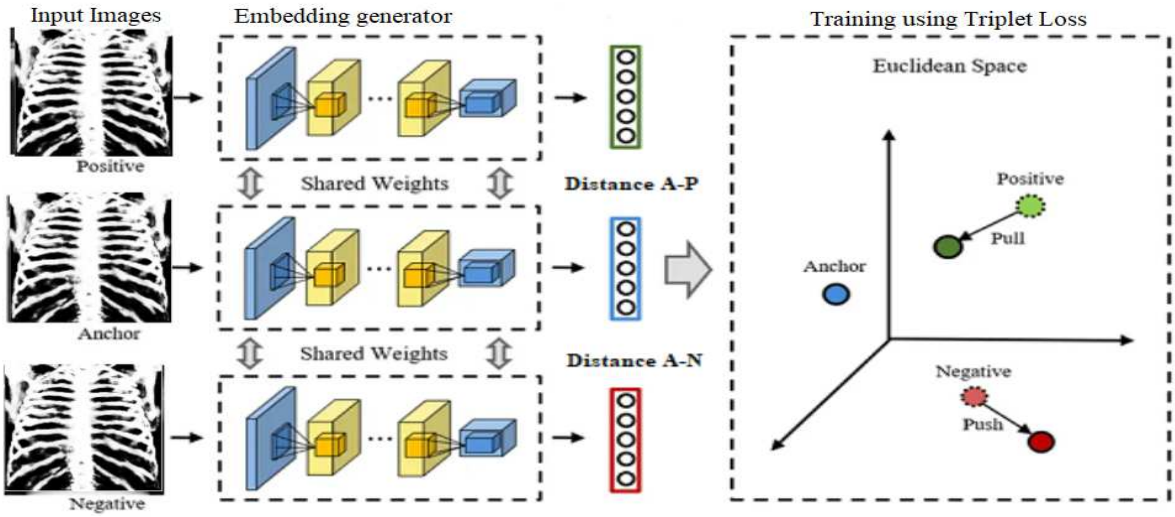


Figure 27: Triplet architecture.

2. Expremental results and disscusion

The obtained results are examined in light of existing studies and theoretical expectations to highlight both consistencies and deviations, as well as to uncover new insights. This section presents an objective analysis of the empirical data generated through the experimental procedures described earlier. The outcomes are reported using tables, figures, and statistical evaluations to ensure clarity and reproducibility. Our proposed framework was implemented using the Keras library in Python. The dataset consists of more than 20,000 chest X-ray images corresponding to 10,000 individuals, with 18,000 images allocated for training and 2,000 reserved for evaluation. Figure 28 illustrates the evaluation loss across iterations, alongside the training accuracy curve, providing an overview of the model's convergence behavior and performance stability over time.

Figure 28: Evaluation curbs.

In addition to evaluating quantitative performance, we further analyze the model's decision-making process and the structure of the learned feature space. To achieve this, we employ dimensionality reduction techniques such as UMAP and t-SNE, which allow for effective visualization of the embedding space. These methods demonstrate how the model organizes feature representations, revealing well-formed clusters for individual identities and clear separations between different subjects. This confirms the model's capability to capture discriminative skeletal characteristics for reliable identification.

To complement this analysis, interpretability methods are integrated to validate the relevance of the learned features. Specifically, Gradient-weighted Class Activation Mapping (Grad-CAM) is applied to highlight the image regions that most influence the model's predictions. By examining these activation maps, we confirm that the network bases its decisions on meaningful skeletal structures rather than background artifacts or irrelevant details. Together, the combination of embedding visualization and explainability techniques not only strengthens trust in the model's predictions but also provides valuable insights into its internal decision-making process.

2.1.Identification system

During the identification phase, as shown in Figure 29, the system processes an anchor image together with multiple candidate images, passing each through the embedding generator. For every input, the model produces a 256-dimensional feature embedding that encodes identity-specific skeletal characteristics. To determine the best match, the Euclidean distance is calculated between the anchor's embedding and those of the candidate images. The comparison set is then ranked according to these distances, with the image closest to the anchor selected as the predicted match. When new patients are enrolled in the system, their chest X-ray radiographs are first processed through the segmentation model to isolate the skeletal structures. The proposed identification model then generates embeddings for each radiograph, which are securely stored in a database together with the corresponding patient identifiers. This embedding-based representation ensures efficient and scalable retrieval during future identification tasks.

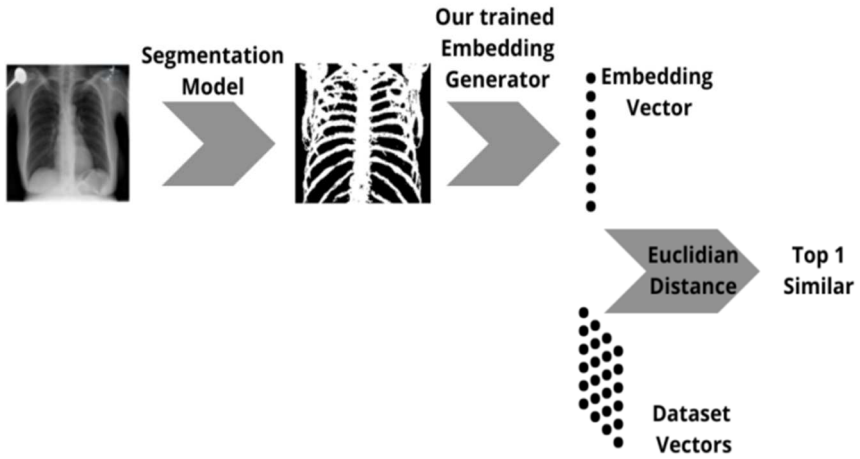


Figure 29: Identification system.

When a new query radiograph is processed, a fresh embedding is created and compared to the stored embeddings using the distance-based approach. The performance of this identification process is evaluated using accuracy and loss metrics shown in Table 2, as well as UMAP and t-SNE visualizations to assess the effectiveness of our trained model and architecture.

Accuracy	Loss	Recall	F1-scor
97.3	0.002	98	97.1

Table 6: Performance Evaluation.

2.2.Validating Discriminative Power and Interpretability

This section provides a detailed analysis of the embedding space produced by our identification system, using t-SNE and UMAP visualizations to validate both discriminative ability and interpretability. By projecting high-dimensional embeddings into two-dimensional space, we illustrate how the system effectively separates individual identities, characterized by large inter-class distances and tight intra-class clustering [89,90]. Employing both methods adds methodological rigor: t-SNE emphasizes local and global structure preservation, while UMAP focuses on topological relationships. Their consistency confirms that the observed separations are intrinsic to the learned embeddings rather than artifacts of a single technique. Beyond quantitative performance, these visualizations strengthen explainability by making it possible to trace how the model distinguishes between individuals, which is essential in high-stakes domains such as security or access control.The projections reveal the robustness of the learned features. In the t-SNE visualization, embeddings spread across a broad range (approximately –40 to 40 on the axes), forming distinct, well-separated clusters that correspond to unique individuals. Similarly, UMAP embeddings (ranging roughly from 3.5 to 8.0) display compact intra-class clusters while maintaining substantial separation between identities, as shown in Figure 30. This pronounced segregation minimizes overlap, drastically reducing the likelihood of misidentification. The clarity of the visual clusters not only confirms the model's ability to capture discriminative skeletal features but also reinforces transparency: stakeholders can easily interpret how the system groups semantically similar samples and separates dissimilar ones, thereby supporting accountability in decision-making.

To further validate the embeddings, Figure 31 presents a UMAP projection with a highlighted anchor sample. Three types of points are displayed: the anchor embedding (blue), its closest neighbor (red), and all other dataset embeddings (gray). The anchor and its nearest neighbor are tightly clustered in the upper-left region, clearly separated from other points scattered across UMAP Dimension 1 (≈75–100) and Dimension 2 (≈60–80). This proximity reflects a strong similarity in the high-dimensional space, confirming the model's ability to cluster related instances together. At the same time, the dispersion of gray points shows the diversity of unrelated embeddings, reinforcing that the model effectively distinguishes different identities.

From a research perspective, this visualization serves two purposes. First, it demonstrates that the system preserves meaningful relationships when reducing dimensionality: similar instances remain close, while dissimilar ones are pushed apart. Second, by labeling the anchor and its closest match, the visualization provides an intuitive illustration of how the model ranks identity candidates during verification or retrieval. The evident gap between the anchor–

closest pair and the rest of the dataset highlights the robustness of the learned features and the system's ability to minimize confusion across identities.

Overall, the complementary use of t-SNE and UMAP not only validates the system's accuracy and discriminative power but also enhances its interpretability, ensuring that it meets the demands of operational environments where precision, robustness, and transparency are paramount..

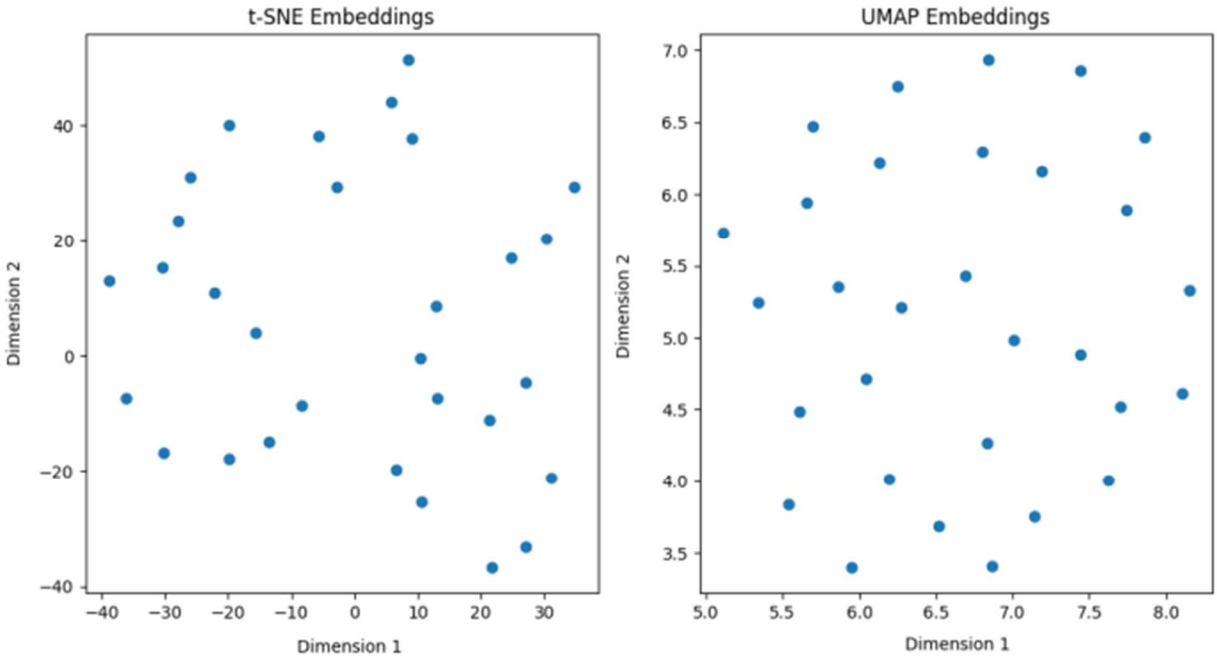


Figure 30: UMAP and t-SNE visualizations of the learned embedding space. Every point represents a unique identity.

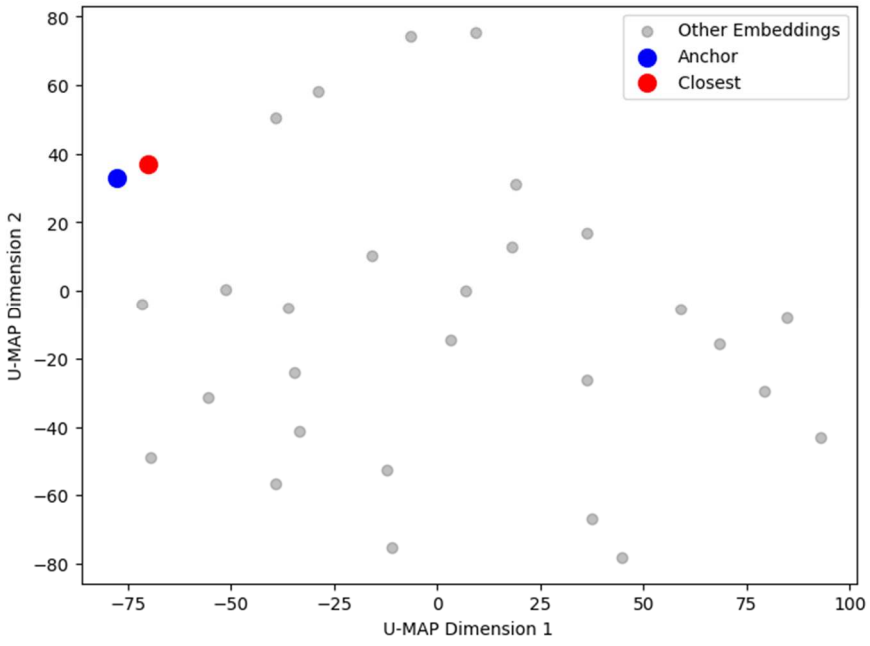


Figure 31: UMAP projection showing the position of a query image (anchor: blue) and its most similar sample (closest match: red) within the learned embedding space. All other samples are shown in gray. Clear proximity between anchor and match indicates accurate identity retrieval.

2.3.Analysis of t-SNE Embeddings for Identity Identification

The t-SNE visualization in figure 32 illustrates the embedding space of our identification system when tested with a sample image (anchor) to verify its capability in accurately identifying individuals. The anchor image, positioned within a dense cluster of its closest matches (nearest neighbors), demonstrates strong intra-class cohesion, with minimal dispersion among semantically similar instances. Notably, the t-SNE Dimension 1 (ranging from -60 to 40) and Dimension 2 (spanning -60 to 40) reveal a clear spatial separation between the anchor cluster and other embeddings, which are distributed distantly across the projection. This spatial segregation underscores the system's discriminative power, embeddings of the target individual remain tightly grouped, while those of unrelated identities are pushed far apart in the latent space. Such behavior ensures reliable identification, as the model minimizes false positives by maintaining large inter-class distances. The visual alignment between the anchor and its closest matches further validates the system's precision in retrieving relevant instances, even under high-dimensional complexity. By testing the system with a representative sample, we confirm its robustness in real-world scenarios, where accurate person identification is critical. The t-SNE projection not only quantifies this performance but also enhances interpretability, offering stakeholders a transparent view of how feature relationships govern retrieval outcomes. This dual focus on accuracy and explainability positions our system as a trustworthy solution for applications demanding both technical rigor and operational clarity.

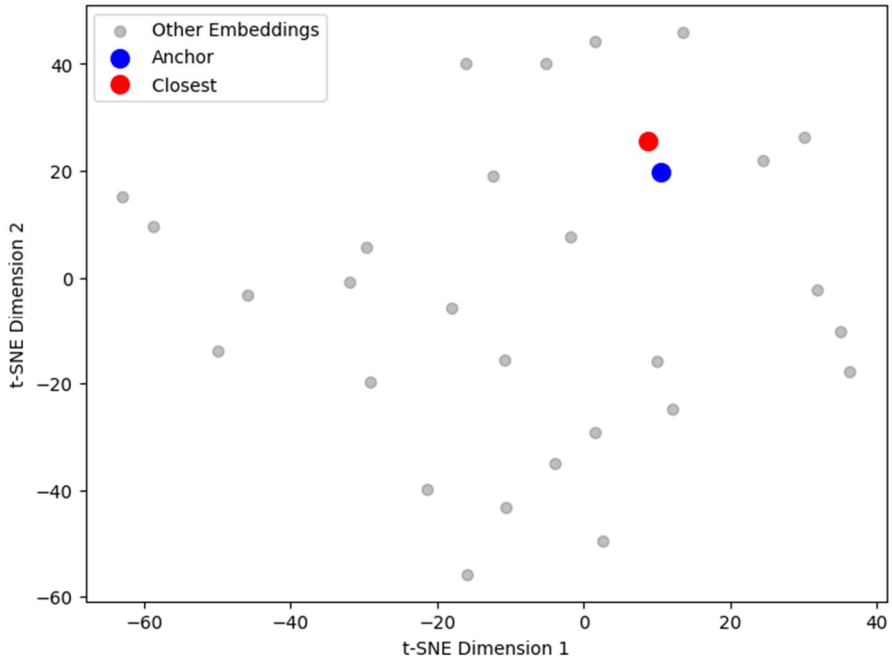


Figure 32: t-SNE projection of the embedding space. The query (anchor: blue) and its most similar sample (closest match: red) within the learned embedding space. while samples from other classes are spatially well-separated, supporting strong inter-class discrimination.

2.4.Explainability and Grad-CAM

Explainable AI (XAI) is a rapidly growing field dedicated to making artificial intelligence systems more transparent and interpretable [91,92]. As deep learning models become more complex, their decision-making processes often resemble "black boxes," making it difficult to understand how they arrive at a particular prediction. This lack of interpretability poses significant challenges, especially in high-stakes applications such as medical imaging, biometrics, and security. One of the most widely used techniques in XAI is Gradient-weighted Class Activation Mapping (Grad-CAM), which provides visual explanations for a model's predictions [93]. Grad-CAM generates heatmaps that highlight the most important regions in an image that influence the model's decision. Unlike traditional AI models that offer only numerical outputs, Grad-CAM allows researchers to see exactly which areas contribute to classification, segmentation, or detection tasks [94]. This technique is particularly valuable in medical imaging and biometric identification, where understanding model reasoning can enhance reliability, trust, and performance. In our research, we introduce a groundbreaking approach to person identification based on skeletal structures extracted from chest X-rays. Unlike conventional biometric methods such as fingerprints, iris recognition, or facial recognition, our approach focuses on unique skeletal patterns that can serve as identifiers. This innovation offers a highly secure, non-intrusive, and robust identification method, especially in medical and forensic applications. To ensure that our model makes accurate and explainable predictions, we apply Grad-CAM to visualize the critical skeletal regions it uses for identification. The figure 33 illustrates heatmaps generated using Grad-CAM, where bright yellow and red regions highlight the most influential areas in the model's decision-making process. From these visualizations, we observe that the model primarily focuses on the central spinal column and rib structures, reinforcing the idea that skeletal features are distinctive enough to differentiate individuals.

Our research marks a pioneering advancement in biometric identification. To the best of our knowledge, we are the first to discover and document the specific skeletal regions used for person identification through XAI-driven analysis. Prior studies in medical imaging and biometrics have explored bone structures, but no existing research has identified precisely which skeletal features are most crucial for individual recognition.

To empirically validate the faithfulness of the Grad-CAM explanations, we conducted a comparative masking experiment. In this experiment, we identified the top-k% most salient regions in the input images using Grad-CAM and evaluated the identification accuracy after masking these regions. We then compared the results with a baseline where an equal-sized random region was masked. The experiment was conducted at multiple saliency levels (k = 10%, 20%, 30%) across the test dataset. The results are summarized in table 3.

Masking Condition	Top-10%	Top-20%	Top-30%
Random Region Masking	92.1%	89.3%	85.7%
Saliency-Based Masking	83.4%	74.6%	66.2%
Accuracy Drop (Δ)	8.7%	14.7%	19.5%

Table 7:comparative masking experiment

The identification accuracy dropped significantly more when salient regions were masked, with a maximum drop of 19.5% at k = 30%. A paired t-test confirmed that the difference in performance between the saliency-based and random masking conditions was statistically significant ($p < 0.01$) at all masking levels. These findings indicate that the regions highlighted by Grad-CAM play a causal and informative role in the model's decision-making, thereby supporting the reliability and faithfulness of the explanations provided.

Our research marks a pioneering advancement in biometric identification. To the best of our knowledge, we are the first to discover and document the specific skeletal regions used for person identification through XAI-driven analysis. Prior studies in medical imaging and biometrics have explored bone structures, but no existing research has identified precisely which skeletal features are most crucial for individual recognition. By leveraging XAI and Grad-CAM, we unlock new insights into biometric identification, offering a scientifically validated approach that was previously unknown. This discovery not only enhances the credibility of our model but also lays the foundation for future research in secure and explainable AI-based identification systems. Our findings demonstrate that skeletal structures can serve as reliable biometric markers, expanding the possibilities for forensic science, medical identification, and security applications.

By leveraging XAI and Grad-CAM, we unlock new insights into biometric identification, offering a scientifically validated approach that was previously unknown. Despite the advantages of using Grad-CAM for visual explanation, it is important to acknowledge its limitations. In particular, the quality of Grad-CAM heatmaps can be affected by segmentation inaccuracies or noisy input boundaries, especially in medical imaging where anatomical structures may overlap or appear faint. These imperfections can lead to ambiguous or diffuse saliency maps that partially highlight irrelevant regions. While such limitations are inherent to many gradient-based attribution methods, they are especially pronounced when explanations are derived from post-segmentation features. Nevertheless, the results of our comparative masking experiment provide empirical support for the functional relevance of the Grad-CAM highlighted regions. The significant drop in identification accuracy when salient areas are occluded relative to random masking suggests that despite potential noise, Grad-CAM successfully captures critical regions used by the model. This balance between interpretability limitations and quantitative validation helps ground the explanatory power of Grad-CAM in

our skeletal-based identification context. This discovery not only enhances the credibility of our model but also lays the foundation for future research in secure and explainable AI-based identification systems. Our findings demonstrate that skeletal structures can serve as reliable biometric markers, expanding the possibilities for forensic science, medical identification, and security applications.

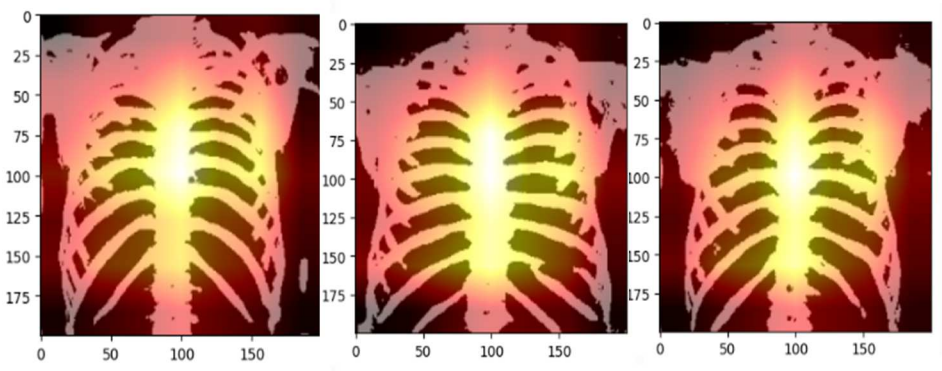


Figure 33: Grad cam visualization.

3. Discussion

The contributions presented in this thesis represent a progressive refinement of chest X-ray based biometric identification, evolving from baseline deep learning approaches toward an advanced, interpretable skeletal-focused pipeline. Each stage of this progression responds to specific limitations identified in the literature and introduces methodological innovations designed to enhance robustness, reliability, and forensic applicability.

From Baseline Experiments to Backbone Optimization

The first series of experiments explored the feasibility of person identification using Siamese and triplet architectures combined with multiple pretrained backbones. These baseline studies confirmed that chest radiographs carry sufficient discriminative information to serve as a biometric modality. Among the tested models, VGG19 with triplet loss achieved an accuracy of 97.0% on the NIH dataset, slightly outperforming ResNet-based approaches previously reported in related work. This finding not only validated earlier research showing the potential of convolutional architectures in capturing patient-specific radiographic traits but also highlighted the importance of backbone selection for maximizing discriminability.

Advancing with Self-Residual Attention Networks (SRAN)

Building upon this foundation, we introduced the Self-Residual Attention Network (SRAN)**,** which integrates channel and spatial attention mechanisms within a residual learning framework. This design emphasizes salient anatomical structures while suppressing irrelevant background noise, thereby improving both discriminability and generalization. With SRAN, identification accuracy reached 98.3% on NIH ChestX-ray14 and maintained strong cross-dataset performance on CheXpert (96.1%)**,** marking an improvement of nearly two percentage points over conventional ResNet50 with spatial attention. This step demonstrated that residual

attention mechanisms enhance feature localization and provide greater robustness under varying imaging conditions, addressing one of the key challenges identified in earlier work.

Segmentation as a Pathway to Skeletal Biometrics

A critical innovation of this thesis was the integration of Attention U-Net for bone segmentation, which laid the foundation for skeletal-based biometric identification. By isolating ribcage, clavicles, sternum, and vertebral structures, the system effectively eliminated soft-tissue dependencies that are prone to degradation in forensic and postmortem settings. Ablation studies confirmed the value of attention gates and augmentation, with Dice coefficients improving from 83% (baseline U-Net) to 94% (Attention U-Net + augmentation)**.** These results exceeded those of earlier U-Net based segmentation methods. For example, a DenseNet multitasking framework reported 88.38%**,** while the strongest prior approach reached 93% on the much smaller JSRT dataset. Our results therefore demonstrate the effectiveness of attention-enhanced segmentation when applied to larger, more diverse datasets such as DRRs derived from LIDC-IDRI.

Toward Skeletal-Only Identification

The combination of segmentation with SRAN produced a skeletal-only identification system, which achieved 97.3% accuracy and an F1-score of 97.1. Notably, these results approach or even surpass soft-tissue based methods, while offering clear advantages in forensic and postmortem applications where soft tissues are absent or degraded. Importantly, this system circumvents the vulnerabilities inherent to soft-tissue dependent models, such as sensitivity to pathology, trauma, or decomposition. By focusing solely on bones, it provides a persistent and tamper-resistant biometric source, making it particularly valuable for medico-legal and disaster victim identification.

Interpretability and Anatomical Insights

To ensure the reliability of this novel biometric paradigm, interpretability was integrated into the pipeline through Grad-CAM analysis. The visualizations consistently revealed that the network's most salient regions corresponded to skeletal structures, particularly the spinal column and ribcage. Comparative masking experiments provided causal evidence of this reliance, with identification accuracy dropping by up to 19.5% when these regions were occluded. A key novel insight emerged: the spinal column consistently acted as the dominant anatomical marker**,** with vertebral alignment, curvature, and spacing patterns proving uniquely discriminative across individuals. This is the first time such a precise anatomical basis for skeletal biometrics has been scientifically demonstrated. The integration of Grad-CAM not only validated the model's decisions but also strengthened its forensic credibility by grounding predictions in meaningful, stable anatomy.

Limitations and Future Directions

Despite these advances, several limitations remain. First, the segmentation stage relied on synthetic DRR data, which may not fully reflect the variability of real-world radiographs. Future validation with manually annotated chest X-rays and postmortem images is required to establish generalizability. Second, while identification performance was strong, verification accuracy (~91%) lags behind, suggesting the need for threshold optimization or advanced metric learning techniques to improve one-to-one matching. Third, potential dataset biases

related to age distribution, pathological variations, or imaging protocols should be systematically addressed to ensure fairness and applicability across diverse populations.

Comparative Context and Contribution to the Field

When compared to existing methods, the trajectory of this thesis clearly demonstrates added value:

- **Classical handcrafted methods** (e.g., CLAHE + Fourier, HOG + BoW) achieved only 44–74% accuracy**,** underlining their inadequacy for biometric tasks.
- **Soft-tissue based deep learning models** reached 95–98%**,** but their forensic utility is limited by reliance on degradable anatomical features.
- **Our skeletal-based pipeline** matched or exceeded these performances (97.3%) while delivering permanence, interpretability, and robustness against decomposition.

By bridging gaps in segmentation quality, backbone design, and forensic applicability, this thesis introduces **a** novel skeletal biometric paradigm. It not only improves state-of-the-art performance but also establishes, for the first time, a biologically grounded explanation of skeletal identity markers, positioning the spinal column as a central anatomical contributor. In summary, the contributions of this thesis mark a significant advancement in the field of chest X-ray biometrics. Through progressive innovations from backbone optimization and residual attention mechanisms to skeletal segmentation and interpretability the research demonstrates that chest X-rays, particularly skeletal structures, constitute a powerful biometric modality. The findings carry both scientific and practical implications, paving the way for robust, interpretable, and forensic-ready biometric systems**.**

4. Conclusion

This chapter investigated skeletal-based biometric identification by introducing a bone segmentation framework and integrating it into a deep learning identification pipeline. The segmentation results demonstrate that bone structures can be reliably extracted from chest X-ray images using deep neural networks, despite inherent annotation and anatomical challenges.

Building upon these segmented representations, the proposed skeletal-based identification system confirms that bone structures carry discriminative identity information. Embedding analyses, similarity measurements, and explainability experiments reveal that skeletal regions contribute meaningfully to identity separation while improving interpretability compared to global image-based approaches. Nevertheless, the results also highlight a dependency between segmentation quality and identification performance, as segmentation inaccuracies may affect downstream discriminative features. These findings emphasize the importance of anatomical preprocessing and provide a realistic assessment of skeletal-based biometric systems. Overall, this chapter validates skeletal representation as a viable and interpretable direction for chest X-ray biometric identification.

General Conclusion & Perspective

General Conclusion & Perspective

The present thesis has investigated the problem of biometric identification using chest radiographs, with a particular focus on developing methods that remain effective under forensic and postmortem conditions. Traditional biometric modalities, such as fingerprints, iris, or facial recognition, face significant challenges in cases of decomposition, trauma, or medical degradation. Soft-tissue based features extracted from chest X-rays, while discriminative, are similarly limited due to their sensitivity to pathology, age, and physiological changes. Against this background, this thesis proposed skeletal-based identification as a permanent and robust biometric modality, leveraging advances in deep learning, attention mechanisms, and explainable AI.

The research contributions followed a progressive and systematic trajectory. First, a comprehensive baseline was established using Siamese networks with triplet loss and multiple pre-trained convolutional backbones (VGG16, VGG19, ResNet50, ResNet101, and DenseNet). These experiments demonstrated that chest radiographs indeed carry unique identity information, with VGG19 achieving the highest identification accuracy of 97.0%. This finding confirmed the feasibility of radiograph-based identification and provided a strong foundation for further refinement.

Building on this baseline, the Self-Residual Attention Network (SRAN) was introduced to enhance feature extraction. By incorporating residual attention blocks, the model was able to focus more effectively on discriminative structures while suppressing irrelevant background information. The SRAN achieved 98.3% identification accuracy on the NIH ChestX-ray14 dataset and maintained high generalization with 96.1% accuracy on the CheXpert dataset. These results established SRAN as a state-of-the-art model for chest radiograph biometrics, outperforming existing attention-based ResNet approaches by nearly two percentage points.

Recognizing the importance of permanence in biometric identifiers, the thesis then addressed the challenge of skeletal isolation through bone segmentation. An Attention U-Net was developed and evaluated on a dataset of Digitally Reconstructed Radiographs (DRRs) derived from LIDC-IDRI. Ablation studies demonstrated the effectiveness of attention gates and data augmentation, with segmentation performance improving from a Dice coefficient of 83% (baseline U-Net) to 94% (U-Net with attention gates and augmentation). These results surpassed several related works and provided a reliable pipeline for generating skeletal masks, a crucial step toward skeletal-based identification.

The final contribution of this thesis combined bone segmentation with the Resnet-50 backbone to develop a skeletal-only biometric system. This approach achieved 97.3% identification accuracy, 98% recall, and a 97.1 F1-score, demonstrating that skeletal structures alone are sufficiently distinctive for person identification. Importantly, this performance was nearly equivalent to soft-tissue–based methods, yet offered superior robustness in forensic contexts where soft tissues may degrade. To strengthen the interpretability of the approach, Grad-CAM explainability was employed, revealing that the model relied primarily on the spinal column and rib cage. Masking experiments further validated this finding, as occluding salient skeletal regions caused accuracy drops of up to 19.5%, thus providing causal evidence of skeletal discriminability.

Collectively, these contributions establish a new direction in medical biometrics: skeletal-based identification using chest radiographs. They not only demonstrate high accuracy and robustness but also enhance interpretability, thereby supporting forensic admissibility. Moreover, the thesis contributes to practical deployment considerations by reporting computational performance, inference times, and GPU requirements, which are critical for transitioning from research to real-world applications.

Despite these promising results, several limitations were identified. First, the bone segmentation stage relied on synthetic DRR masks, which, while useful for training, may not fully capture the variability of real-world X-rays. Second, the evaluation datasets (NIH and CheXpert) consist of clinical radiographs, and although cross-dataset validation was performed, true postmortem radiographs remain to be tested. Third, verification accuracy (~91%) was lower than identification accuracy, indicating that threshold calibration and advanced metric learning are still required to optimize one-to-one matching scenarios. Finally, demographic and pathological biases in the datasets may influence results, underscoring the need for broader population studies.

In conclusion, this thesis demonstrates that chest radiographs, and in particular skeletal structures extracted from them, constitute a powerful biometric modality. Through baseline experiments, novel attention-based architectures, advanced segmentation techniques, and explainability analyses, it establishes skeletal-based identification as a robust, interpretable, and forensically relevant pathway. The findings significantly advance the state of the art and provide a foundation for future research and application in forensic science, clinical identity management, and secure biometric systems.

Future Works

While the results obtained in this thesis are encouraging, they also open several directions for future research:

1. **Validation on Real Postmortem Radiographs**

   To fully establish forensic applicability, the proposed skeletal-based system should be validated on datasets containing real postmortem radiographs. Such validation would

confirm the robustness of skeletal biometrics under conditions of decomposition, trauma, or tissue loss.

2. **Annotated Bone Segmentation Datasets**

   The segmentation stage would benefit greatly from training and evaluation on large-scale, manually annotated bone mask datasets derived from real chest X-rays. Collaboration with radiologists to build such datasets would allow more reliable benchmarking and improve generalization beyond synthetic DRRs.

3. **Enhanced Verification Strategies**

   While identification accuracy was near-perfect, verification performance remains a challenge. Future work should investigate advanced metric learning approaches, such as hard triplet mining, N-pair losses, proxy-based embeddings, or contrastive self-supervised pretraining, to strengthen one-to-one decision reliability.

4. **Cross-Dataset and Cross-Population Studies**

   The generalizability of skeletal biometrics should be assessed across different imaging devices, clinical populations, and age groups. Evaluating demographic and pathological variability will ensure fairness, robustness, and applicability to diverse populations.

5. **Integration with Multimodal Biometrics**

6. Skeletal biometrics could be combined with other modalities, such as dental radiographs, CT-based skeletal features, or even residual soft-tissue features, to form a multimodal system. This would enhance accuracy and provide redundancy for cases with incomplete data.

7. **Operational Deployment Studies**

   Beyond algorithmic development, future research should explore deployment pipelines for forensic and clinical use. This includes optimizing runtime efficiency, establishing standardized forensic protocols, ensuring compliance with legal and ethical frameworks, and addressing data privacy and security concerns.

8. **Explainability and Forensic Acceptance**

   The integration of explainable AI should be further refined to generate case-level visual reports that can be used in forensic investigations or court proceedings. Improving the transparency and interpretability of skeletal biometrics will support their acceptance in real-world applications.

Final Statement:
Through these contributions and future directions, this thesis lays the groundwork for a new generation of biometric systems based on skeletal features in chest radiographs. The results demonstrate not only technical feasibility but also significant forensic and clinical potential.

With further validation, expansion, and operational integration, skeletal-based biometrics can evolve into a reliable, permanent, and legally admissible identity verification tool, addressing critical gaps in both medical and forensic domains.

Bibliographie

Bibliography

1. Yang, Z., Xu, X., Zhang, J., Wang, G., Kalra, M. K., & Yan, P. (2025). Chest X-ray Foundation Model with Global and Local Representations Integration. *arXiv preprint arXiv:2502.05142*.
2. Hussain, S., Mubeen, I., Ullah, N., Shah, S. S. U. D., Khan, B. A., Zahoor, M., ... & Sultan, M. A. (2022). Modern diagnostic imaging technique applications and risk factors in the medical field: a review. *BioMed research international*, *2022*(1), 5164970.
3. Akhter, Y., Singh, R., & Vatsa, M. (2023). AI-based radiodiagnosis using chest X-rays: A review. *Frontiers in big data*, *6*, 1120989.
4. Båth, M., Håkansson, M., Börjesson, S., Hoeschen, C., Tischenko, O., Kheddache, S., ... & Månsson, L. G. (2005). Nodule detection in digital chest radiography: effect of anatomical noise. *Radiation protection dosimetry*, *114*(1-3), 109-113.
5. Zarshenas, A., Liu, J., Forti, P., & Suzuki, K. (2019). Separation of bones from soft tissue in chest radiographs: Anatomy-specific orientation-frequency-specific deep neural network convolution. *Medical physics*, *46*(5), 2232-2242.
6. Tsuneki, M. (2022). Deep learning models in medical image analysis. *Journal of Oral Biosciences*, *64*(3), 312-320.
7. Eli, A. A., & Ali, A. (2024). Deep Learning Applications in Medical Image Analysis: Advancements, Challenges, and Future Directions. *arXiv preprint arXiv:2410.14131*.
8. Sarvamangala, D. R., & Kulkarni, R. V. (2022). Convolutional neural networks in medical image understanding: a survey. *Evolutionary intelligence*, *15*(1), 1-22.
9. Windley, P. J. (2005). *Digital Identity: Unmasking identity management architecture (IMA)*. " O'Reilly Media, Inc.".
10. Ross, A., & Jain, A. K. (2004, December). Biometrics: when identity matters. In *Chinese Conference on Biometric Recognition* (pp. 1-2). Berlin, Heidelberg: Springer Berlin Heidelberg.
11. Sauerwein, K., Saul, T. B., Steadman, D. W., & Boehnen, C. B. (2017). The effect of decomposition on the efficacy of biometrics for positive identification. *Journal of forensic sciences*, *62*(6), 1599-1602.

12. Bhatt, S., Sehrawat, J. S., & Gupta, V. (2025). A systematic review of iris biometrics in forensic science: applications and challenges. *Egyptian Journal of Forensic Sciences*, *15*(1), 12.
13. Cornett, D. C., Bolme, D. S., Steadman, D. W., Sauerwein, K. A., & Saul, T. B. (2019, September). Effects of postmortem decomposition on face recognition. In *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)* (pp. 1-8). IEEE.
14. Stephan, C. N., D'Alonzo, S. S., Wilson, E. K., Guyomarc'h, P., Berg, G. E., & Byrd, J. E. (2018). Skeletal identification by radiographic comparison of the cervicothoracic region on chest radiographs. *New perspectives in forensic human skeletal identification*, 277-292.
15. Watamaniuk, L., & Rogers, T. (2010). Positive personal identification of human remains based on thoracic vertebralmargin morphology. *Journal of forensic sciences*, *55*(5), 1162-1170.
16. Sony, R., Isaac, C. V., Vanbaarle, A., Devota, C. J., Fenton, T., Hefner, J. T., & Ross, A. (2025). Automatic Comparative Chest Radiography Using Deep Neural Networks. *IEEE Access*.
17. Fu, S., Zhang, M., Mu, C., & Shen, X. (2018). Advancements of medical image enhancement in healthcare applications. *Journal of healthcare engineering*, *2018*, 7035264.
18. Deheyab, A. O. A., Alwan, M. H., Rezzaqe, I. K. A., Mahmood, O. A., Hammadi, Y. I., Kareem, A. N., & Ibrahim, M. (2022, December). An overview of challenges in medical image processing. In *Proceedings of the 6th international conference on future networks & distributed systems* (pp. 511-516).
19. Demner-Fushman, D., Kohli, M. D., Rosenman, M. B., Shooshan, S. E., Rodriguez, L., Antani, S., ... & McDonald, C. J. (2015). Preparing a collection of radiology examinations for distribution and retrieval. *Journal of the American Medical Informatics Association*, *23*(2), 304-310.
20. Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S., Karthikesalingam, A., King, D., ... & Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ digital medicine*, *4*(1), 65.
21. Gambato, M., Scotti, N., Borsari, G., Zambon Bertoja, J., Gabrieli, J. D., De Cassai, A., ... & Causin, F. (2023). Chest X-ray interpretation: detecting devices and device-related complications. *Diagnostics*, *13*(4), 599.
22. Singh, S. (2024). Computer-aided diagnosis of thoracic diseases in chest X-rays using hybrid cnn-transformer architecture. *arXiv preprint arXiv:2404.11843*.
23. DUPONT, Mackenzie, CASTRO, Robert, KIK, Sandra V., *et al.* Computer-Aided Reading of Chest Radiographs for Pediatric Tuberculosis: Current Status and Future Directions. *medRxiv*, 2024, p. 2024.10. 08.24314837.
24. Veldkamp, W. J., Kroft, L. J., & Geleijns, J. (2009). Dose and perceived image quality in chest radiography. *European journal of radiology*, *72*(2), 209-217.
25. Siela, D. (2002). Using chest radiography in the intensive care unit. *Critical Care Nurse*, *22*(4), 18-27.

26. Van Gelderen, F., & Van Gelderen, F. (2004). *Understanding X-rays*. Springer Berlin Heidelberg.
27. Park, B., & Sung, D. W. (2010). A comparative study of image quality and radiation dose with changes in tube voltage and current for a digital chest radiography. *Journal of the Korean Society of Radiology*, *62*(2), 131-137.
28. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, *18*(8), 500-510.
29. Irmici, G., Cè, M., Caloro, E., Khenkina, N., Della Pepa, G., Ascenti, V., ... & Cellina, M. (2023). Chest x-ray in emergency radiology: What artificial intelligence applications are available?. *Diagnostics*, *13*(2), 216.
30. Gambato, M., Scotti, N., Borsari, G., Zambon Bertoja, J., Gabrieli, J. D., De Cassai, A., ... & Causin, F. (2023). Chest X-ray interpretation: detecting devices and device-related complications. *Diagnostics*, *13*(4), 599.
31. Macpherson, L., Miller, C., Hamada, Y., Rangaka, L., Ruhwald, M., Falzon, D., ... & Esmail, H. (2025). Policies, practices, opportunities and challenges for tuberculosis screening: a global survey of national tuberculosis programmes. *BMJ Global Health*, *10*(7).
32. Gibikote, S., & Kirupakaran, H. (2016). The routine pre-employment screening chest radiograph: Should it be routine?. *Indian Journal of Radiology and Imaging*, *26*(03), 402-404.
33. Wielpütz, M. O., Heußel, C. P., Herth, F. J., & Kauczor, H. U. (2014). Radiological diagnosis in lung disease: factoring treatment options into the choice of diagnostic modality. *Deutsches Ärzteblatt International*, *111*(11), 181.
34. Irmici, G., Cè, M., Caloro, E., Khenkina, N., Della Pepa, G., Ascenti, V., ... & Cellina, M. (2023). Chest x-ray in emergency radiology: What artificial intelligence applications are available?. *Diagnostics*, *13*(2), 216.
35. Dutta, S., Keerthana, R., Shekar, D. C., & Kumar, S. S. (2024, December). Biological Profile Estimation of the Human Skeleton Based Forensic Identification Using Deep Learning Image Classification. In *2024 International Conference on IoT Based Control Networks and Intelligent Systems (ICICNIS)* (pp. 1369-1375). IEEE.
36. Arthur, R. (2001). The neonatal chest X-ray. *Paediatric respiratory reviews*, *2*(4), 311-323.
37. Morishita J, Katsuragawa S, Kondo K, Doi K. An automated patient recognition method based on an image-matching technique using previous chest radiographs in the picture archiving and communication system environment. Med Phys. 2001 Jun;28(6):1093-7. doi: 10.1118/1.1373403. PMID: 11439478
38. Danaher, L. A., Howells, J., Holmes, P., & Scally, P. (2011). Is it possible to eliminate patient identification errors in medical imaging?. Journal of the American College of Radiology, 8(8), 568-574.
39. Morishita, J., Watanabe, H., Katsuragawa, S., Oda, N., Sukenobu, Y., Okazaki, H., ... & Doi, K. (2005). Investigation of misfiled cases in the PACS environment and a solution to prevent filing errors for chest radiographs1. Academic radiology, 12(1), 97-103.

40. Ueda, Y., Morishita, J., Kudomi, S., & Ueda, K. (2016). Usefulness of biological fingerprint in magnetic resonance imaging for patient verification. Medical & biological engineering & computing, 54, 1341-1351.
41. Singh, H. (2013). 101 Chest X-Ray Solutions. JP Medical Ltd.
42. Walz, C., Schwarz, C. S., Imdahl, K., Steffan, C., & Germerott, T. (2023). Comparison of the quality of clinical forensic examination of victims of physical violence conducted by clinicians and forensic examiners. International journal of legal medicine, 137(6), 1777-1786.
43. Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2097-2106).
44. Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., ... & Ng, A. Y. (2019, July). Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 33, No. 01, pp. 590-597).
45. Saini, M., & Kapoor, A. K. (2016). Biometrics in forensic identification: applications and challenges. *J Forensic Med*, *1*(108), 2.
46. Dey, A., Tharmavaram, M., Pandey, G., Rawtani, D., & Mustansar Hussain, C. (2020). Conventional and emerging biometrics techniques in forensic investigations. *Technology in Forensic Science: Sampling, Analysis, Data and Regulations*, 175-197.
47. Emergency Care Research Institute: Patient identification errors. 2016.Availableat https://www.ecri.org/Resources/HIT/Patient%20ID/Patient_Identification_Evidence_Based_Literature_final.pdf. Accessed Sep 28, 2021.
48. Jain, A. K., & Ross, A. (2015). Bridging the gap: from biometrics to forensics. Philosophical Transactions of the Royal society B: Biological sciences, 370(1674), 20140254.
49. Saini, M., & Kapoor, A. K. (2016). Biometrics in forensic identification: applications and challenges. J Forensic Med, 1(108), 2.
50. Singh, H. (2013). 101 Chest X-Ray Solutions. JP Medical Ltd.
51. Walz, C., Schwarz, C. S., Imdahl, K., Steffan, C., & Germerott, T. (2023). Comparison of the quality of clinical forensic examination of victims of physical violence conducted by clinicians and forensic examiners. International journal of legal medicine, 137(6), 1777-1786.
52. Singhal, Z., Gupta, P., & Garg, K. (2012). Biometric recognition: Personal identification technique. *International Journal of Computational Engineering & Management*, *15*(3), 6-10.
53. Barbole, A., & Godase, M. (2015). Advancements in Bio-Metrics Provides Additional Security. Available at SSRN 4607940.
54. Jain, A. K., Flynn, P., & Ross, A. A. (Eds.). (2007). Handbook of biometrics. Springer Science & Business Media.
55. Maltoni, D., Maio, D., Jain, A. K., & Prabhakar, S. (2009). Handbook of fingerprint recognition. London: Springer London.

56. Shadman, R., Wahab, A. A., Manno, M., Lukaszewski, M., Hou, D., & Hussain, F. (2025). Keystroke dynamics: Concepts, techniques, and applications. *ACM Computing Surveys*, *57*(11), 1-35.
57. EFRIZONI, L., ARMOOGUM, S., & ZAKARIA, M. (2024). DEEP LEARNING INNOVATIONS IN FINGERPRINT RECOGNITION: A COMPARATIVE STUDY OF MODEL EFFICIENCIES. *INTERNATIONAL JOURNAL*, *1*(1), 28-35.
58. Cornett, D. C., Bolme, D. S., Steadman, D. W., Sauerwein, K. A., & Saul, T. B. (2019, September). Effects of postmortem decomposition on face recognition. In 2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS) (pp. 1-8). IEEE.
59. Khanagar, S. B., Vishwanathaiah, S., Naik, S., Al-Kheraif, A. A., Divakar, D. D., Sarode, S. C., ... & Patil, S. (2021). Application and performance of artificial intelligence technology in forensic odontology–A systematic review. Legal Medicine, 48, 101826.
60. Beckett, R. G., & Conlogue, G. J. (Eds.). (2020). *Advances in Paleoimaging: Applications for Paleoanthropology, Bioarcheology, Forensics, and Cultural Artifacts*. CRC Press, Taylor & Francis Group.
61. Ishigami, R., Zin, T. T., Shinkawa, N., & Nishii, R. (2017, March). Human identification using X-Ray image matching. In Proceedings of the International Multi Conference of Engineers and Computer Scientists (Vol. 1).
62. Cho, H.; Zin, T.T.; Shinkawa, N.; Nishii, R. Post-mortem human identification using chest X-ray and ct scan images. Int. J. Biomed.Soft Comput. Hum. Sci. 2018, 23, 51–57.
63. Packhäuser, K.; Gündel, S.; Münster, N.; Syben, C.; Christlein, V.; Maier, A. Deep learning-based patient re-identification is able to exploit the biometric nature of medical chest X-ray data. Sci. Rep. 2022, 12, 14851. [CrossRef]
64. Ueda, Y.; Morishita, J. Patient Identification Based on Deep Metric Learning for Preventing Human Errors in Follow-up X-Ray Examinations. J. Digit. Imaging 2023, 36, 1941–1953. [CrossRef]
65. Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).
66. Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In Proceedings of the European conference on computer vision (ECCV) (pp. 3-19).
67. Ling, H., Wu, J., Wu, L., Huang, J., Chen, J., & Li, P. (2019). Self residual attention network for deep face recognition. *IEEE Access*, *7*, 55159-55168.
68. Ma, D., Li, S., Zhang, X., & Wang, H. (2017). Interactive attention networks for aspect-level sentiment classification. *arXiv preprint arXiv:1709.00893*.
69. Wang, X.; Girshick, R.; Gupta, A.; He, K. Non-local neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 7794–7803.
70. Buades, A.; Coll, B.; Morel, J.-M. A Non-Local Algorithm for Image Denoising. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20 June 2005; Volume 2, pp. 60–65.

71. Zhang, H.; Goodfellow, I.; Metaxas, D.; Odena, A. Self-attention generative adversarial networks. arXiv 2018, arXiv:1805.08318.
72. Fu, J.; Liu, J.; Tian, H.; Fang, Z.; Lu, H. Dual attention network for scene segmentation. arXiv 2018, arXiv:1809.02983.
73. Hazem, F., Akram, B., Mekhaznia, T., Ghabban, F., Alsaeedi, A., & Goyal, B. (2024). Beyond Traditional Biometrics: Harnessing Chest X-Ray Features for Robust Person Identification. Acta Informatica Pragensia, 13(2), 234-250.
74. Hazem, F., Akram, B., mekhaznia, T., Al-Sarem, M., Shukla, P. K., & Khalaf, O. I. (2024, June). X-Ray Insights: A Siamese with CNN and Spatial Attention Network for Innovative Person Identification. In International Conference on Intelligent Systems and Pattern Recognition (pp. 240-253). Cham: Springer Nature Switzerland.
75. Farah, H., Bennour, A., Kurdi, N. A., Hammami, S., & Al-Sarem, M. (2024). Channel and Spatial Attention in Chest X-Ray Radiographs: Advancing Person Identification and Verification with Self-Residual Attention Network. Diagnostics, 14(23), 2655.
76. Hazem, F., Akram, B., Khalaf, O. I., Sikder, R., & Algburi, S. (2023, December). X-ray insights: Innovative person identification through Siamese and Triplet networks. In IET Conference Proceedings CP870 (Vol. 2023, No. 39, pp. 40-49). Stevenage, UK: The Institution of Engineering and Technology.
77. Thapar, D.; Jaswal, G.; Nigam, A.; Kanhangad, V. PVSNet: Palm Vein Authentication Siamese Network Trained Using Triplet Loss and Adaptive Hard Mining by Learning Enforced Domain Specific Features. In Proceedings of the 2019 IEEE 5th International Conference on Identity, Security, and Behavior Analysis (ISBA), Hyderabad, India, 22–24 January 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–8.
78. Kumar, C.R.; Saranya, N.; Priyadharshini, M.; Gilchrist, D. Face recognition using CNN and siamese network. Measurement. Sensors 2023, 27, 100800.
79. Lai, S.C.; Lam, K.M. Deep Siamese Network for Low-Resolution Face Recognition. In Proceedings of the 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Tokyo, Japan, 14–17 December 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1444–1449.
80. Kumar, K.V.; Teja, K.A.; Bhargav, R.T.; Satpute, V.; Naveen, C.; Kamble, V. One-Shot Face Recognition. In Proceedings of the 2023 2nd International Conference on Paradigm Shifts in Communications Embedded Systems, Machine Learning and Signal Processing (PCEMS), Nagpur, India, 5–6 April 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 1–6.
81. Wu, D., Wang, C., Wu, Y., Wang, Q. C., & Huang, D. S. (2021). Attention deep model with multi-scale deep supervision for person re-identification. IEEE Transactions on Emerging Topics in Computational Intelligence, 5(1), 70-78.
82. Singh, A., Lall, B., Panigrahi, B. K., Agrawal, A., Agrawal, A., Thangakunam, B., & Christopher, D. J. (2022). Semantic segmentation of bone structures in chest X-rays including unhealthy radiographs: A robust and accurate approach. International Journal of Medical Informatics, 165, 104831.
83. Dodamani PS, Danti A. Grey wolf optimization guided non-local means denoising for localizing and extracting bone regions from X-ray images. Biomed Pharmacol J. 2023;16(2):935–46. doi:10.13005/bpj/2676. [Google Scholar] [CrossRef]

84. 30.Xie Y, Chen L, Zhou L, Liu Y, Han Z, Li Y, et al. Attention mechanisms in medical image segmentation: a survey. arXiv:2305.17937. 2023. [Google Scholar]
85. Chicco, D. (2021). Siamese neural networks: An overview. *Artificial neural networks*, 73-94.
86. Bromley, J., Guyon, I., LeCun, Y., Säckinger, E., & Shah, R. (1993). Signature verification using a" siamese" time delay neural network. *Advances in neural information processing systems*, *6*.
87. De Rosa, G. H., & Papa, J. P. (2022). Learning to weight similarity measures with Siamese networks: a case study on optimum-path forest. In *Optimum-Path Forest* (pp. 155-173). Academic Press.
88. Guo, J., Wang, H., Cheng, Z., Zhang, X., & Yan, D. M. (2020). Learning local shape descriptors for computing non-rigid dense correspondence. *Computational Visual Media*, *6*(1), 95-112.
89. McInnes, L., Healy, J., & Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv:1802.03426.
90. Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W., Ng, L. G., ... & Newell, E. W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. Nature biotechnology, 37(1), 38-44.
91. Samek, W., Wiegand, T., & Müller, K. R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. arXiv preprint arXiv:1708.08296.
92. Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., ... & Ranjan, R. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. ACM Computing Surveys, 55(9), 1-33.
93. Suara, S., Jha, A., Sinha, P., & Sekh, A. A. (2023, November). Is grad-cam explainable in medical images?. In International Conference on Computer Vision and Image Processing (pp. 124-135). Cham: Springer Nature Switzerland.
94. Chien, J. C., Lee, J. D., Hu, C. S., & Wu, C. T. (2022). The usefulness of gradient-weighted cam in assisting medical diagnoses. Applied Sciences, 12(15), 7748.