Legal Medicine 84 (2026) 102885

Contents lists available at ScienceDirect

Legal Medicine

journal homepage: www.elsevier.com/locate/legalmed

Artificial intelligence in forensic age assessment: a systematic review☆

Valentina Bugelli [a,1], Francesco Calabrò [a,1], Laura Donato [b], Rossana Cecchi [c], Jessika Camatti [a,*], Marco Di Paolo [d], Lorenzo Franceschetti [e]

[a] University of Parma, Parma, Italy
[b] University of Rome Tor Vergata, Rome, Italy
[c] University of Modena and Reggio Emilia, Modena, Italy
[d] University of Pisa, Pisa, Italy
[e] University of Milano, Milano, Italy

ARTICLE INFO

Keywords:
Forensic age estimation
Artificial intelligence
Machine learning
Deep learning
Dental radiographs
DNA methylation

ABSTRACT

Forensic age assessment plays a crucial role in medico-legal contexts where an individual's chronological age cannot be reliably documented, such as immigration procedures. In recent years, artificial intelligence (AI) has increasingly been applied to automate and improve the accuracy of age estimation methods. However, the forensic applicability of these approaches, particularly in relation to legally relevant age thresholds, remains incompletely characterized.

A systematic review was conducted in accordance with the PRISMA 2020 guidelines. MEDLINE (via PubMed) and Scopus were searched from database inception to 1 March 2026. Studies were included if they applied AI techniques to age estimation in a forensic or medico-legal context and reported quantitative performance metrics.

The search identified 1,197 records, of which 48 studies met the inclusion criteria. Most studies investigated imaging-based approaches, particularly panoramic dental radiographs, while others explored skeletal imaging, DNA methylation markers, or emerging molecular biomarkers. Convolutional neural networks were the most frequently used modelling approach. AI models applied to dental radiographs commonly reported mean absolute errors between approximately 0.5 and 1.5 years, whereas DNA methylation–based models showed errors between 3 and 5 years. MRI-based skeletal models often achieved prediction errors close to one year. However, relatively few studies specifically evaluated classification performance around legally relevant thresholds such as 18 years.

AI shows considerable potential for improving the automation and reproducibility of forensic age estimation. Further research with larger and more diverse datasets, external validation, and standardized reporting is needed to ensure reliable implementation in medico-legal practice.

1. Introduction

Age assessment plays a critical role in forensic and medico-legal practice, particularly in contexts where the chronological age of an individual cannot be reliably documented. Determining whether a person has reached the age of legal majority is essential in several legal settings, including immigration procedures, asylum applications, criminal responsibility, and child protection systems. In such situations, forensic age estimation is used to support judicial authorities when documentary evidence is unavailable, unreliable, or disputed [1,2].

Traditional forensic age assessment relies primarily on the evaluation of biological maturation through medical imaging or dental examination. Commonly used methods include the analysis of dental development on orthopantomography, skeletal maturation of the hand and wrist, clavicular ossification on radiographs or computed tomography, and magnetic resonance imaging of growth plates. Molecular

☆ This article is part of a special issue entitled: 'Proceedings Congress Forensic Science' published in Legal Medicine.

* Corresponding author.
E-mail addresses: valentina.bugelli@unipr.it (V. Bugelli), francesco.calabro@unipr.it (F. Calabrò), donato.laura@gmail.com (L. Donato), rossana.cecchi@unimore.it (R. Cecchi), jessika.camatti@unipr.it (J. Camatti), marco.dipaolo@unipi.it (M. Di Paolo), lorenzo.franceschetti@unimi.it (L. Franceschetti).

[1] **These authors equally contributed as first author.**

https://doi.org/10.1016/j.legalmed.2026.102885
Received 14 May 2026; Received in revised form 22 May 2026; Accepted 24 May 2026
Available online 30 May 2026

approaches, such as DNA methylation analysis, have also been explored as potential biomarkers of biological age. However, these methods may be affected by interobserver variability, population-specific developmental differences, and uncertainties when applied near critical legal age thresholds, particularly around the age of 18 years [3–6].

In recent years, artificial intelligence (AI), including machine learning and deep learning techniques, has been increasingly applied to age estimation tasks. AI-based systems can analyse complex imaging or molecular datasets and automatically extract patterns associated with biological maturation. Several studies have reported promising performance in predicting chronological age using radiological images, dental radiographs, magnetic resonance imaging, or epigenetic data. In addition to regression-based age prediction, some AI models have been specifically designed to classify individuals according to legal age thresholds, which is particularly relevant in forensic and immigration contexts [7–9].

Importantly, many studies in this field were originally designed for chronological age prediction or biological maturation assessment rather than for direct medico-legal threshold classification. This distinction is highly relevant in forensic practice because medico-legal decision-making often focuses on whether an individual has reached a legally relevant age threshold, particularly 18 years of age, rather than on regression-based prediction accuracy alone.

Despite the growing number of studies in this field, the forensic applicability, methodological quality, and real-world reliability of AI-based age estimation systems remain incompletely characterized. The heterogeneity of datasets, imaging modalities, model architectures, and validation strategies makes it difficult to draw clear conclusions regarding their diagnostic performance and practical implementation [2,7].

The aim of this systematic review is therefore to evaluate the current evidence on the use of artificial intelligence for forensic age assessment. In particular, the review focuses on the accuracy of AI-based models for age estimation and their applicability in determining legal age thresholds, including the distinction between individuals below and above the age of majority.

2. Materials and methods

2.1. Study design and research question

This study was conducted as a systematic review evaluating the use of artificial intelligence for forensic age assessment. The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines [10].

The study protocol was prospectively registered on the Open Science Framework (OSF) to promote methodological transparency and reproducibility. A publicly accessible view-only version of the protocol is available at:

https://osf.io/pukaw/overview?view_only = 2b8def26d0f44610b1272da8172a6610.

The review aimed to assess the performance and forensic applicability of AI-based models for age estimation, with particular emphasis on methods used for determining legal age thresholds, including the distinction between individuals below and above the age of majority.

2.2. Eligibility criteria and search strategy

Studies were included if they met the following criteria: original research applying artificial intelligence, machine learning, or deep learning techniques to age estimation; conducted in a forensic or medico-legal context, or explicitly applicable to forensic age assessment; use of imaging data (e.g., dental radiographs, X-rays, MRI) or molecular/epigenetic data (e.g., DNA methylation); reporting quantitative performance metrics, such as mean absolute error (MAE), root mean square error (RMSE), sensitivity, specificity, or area under the receiver operating characteristic curve (AUC); full-text articles published in English.

Exclusion criteria were: studies focused exclusively on pediatric clinical growth assessment without forensic relevance; reviews, editorials, letters, or conference abstracts lacking full methodological details; studies without objective performance evaluation; studies primarily addressing personal identification rather than age estimation, as this topic is addressed in a dedicated companion review (Part I) within the same three-part systematic review series on artificial intelligence in forensic science.

Conference proceedings and abstracts were excluded due to limited methodological detail and the absence of full peer-reviewed publication.

MEDLINE (via PubMed) and Scopus were searched from database inception to the final search date (1 March 2026). The search was supplemented by backward and forward citation tracking. Database-specific search strategies are reported in Supplementary Appendix 1.

Additionally, manual screening of reference lists from included studies was performed to identify potentially relevant articles not captured in the initial database search.

2.3. Study selection and data extraction

All retrieved references were screened independently by two reviewers. After duplicate removal, studies were evaluated through the following stages: title and abstract screening; full-text eligibility assessment. Discrepancies between reviewers were resolved through consensus discussion.

A predefined standardized extraction form was used. The following variables were extracted: bibliographic information; study objective; sample characteristics and age range; imaging modality or biological material used; AI model architecture and training approach; validation method (internal or external); reference standard for chronological age; performance metrics (MAE, RMSE, classification accuracy, sensitivity/specificity); performance around legal age thresholds, when reported; reported limitations and potential sources of bias. Data extraction was conducted independently by two reviewers.

2.4. Risk of bias, quality assessment, and data synthesis

Risk of bias was assessed using: QUADAS-2, particularly for studies addressing classification of legal age status; PROBAST (AI-adapted) for predictive age estimation models.

Special attention was given to: dataset representativeness, potential data leakage, validation strategy (internal vs external), transparency and completeness of reporting. Given the anticipated heterogeneity in imaging modalities, age ranges, and evaluation metrics, results were synthesized narratively.

Studies were grouped into the following categories: dental age estimation (orthopantomography and third molar analysis); skeletal age estimation (hand, pelvis, and other radiographic methods); MRI-based age estimation; DNA methylation and molecular approaches; alternative or exploratory methods. Where methodological comparability allowed, quantitative comparison of error metrics across studies was considered.

Where studies reported comparable quantitative performance metrics, descriptive pooled estimates of mean absolute error (MAE) were calculated to provide an approximate summary of predictive accuracy across different methodological categories (e.g., dental imaging, MRI-based skeletal imaging, and DNA methylation approaches). Given the heterogeneity of study designs, datasets, and evaluation protocols, pooled values were interpreted as descriptive summary estimates rather than formal *meta*-analytic effect sizes.

3. Results

3.1. Study selection

The database search identified 1,197 records (PubMed/MEDLINE: 608; Scopus: 589). After removal of 274 duplicate records, 923 records remained and were screened based on titles and abstracts. Of these, 863 records were excluded.

A total of 60 reports were sought for retrieval, of which 3 could not be retrieved. Consequently, 57 full-text articles were assessed for eligibility. Following full-text evaluation, 9 articles were excluded for the following reasons: not related to age estimation (n = 2), no application of artificial intelligence or machine learning to age estimation (n = 3), AI used only for feature extraction with age estimation performed using classical non–machine learning models (n = 1), not focused on forensic age assessment (n = 1), and not reporting original data (n = 2).

Ultimately, 48 studies met the inclusion criteria and were included in the qualitative synthesis [11–58]. The study selection process is summarized in Fig. 1 (PRISMA flow diagram).

3.2. Study characteristics

The 48 studies included in this systematic review investigated the application of artificial intelligence techniques for forensic age estimation using a variety of biological and imaging data sources. The main characteristics of the included studies are summarized in Table 1, while detailed information on study design, datasets, AI models, validation strategies, and performance metrics is provided in Supplementary **Table S1**.

Overall, the majority of studies focused on dental imaging, particularly panoramic radiographs (orthopantomography), which represented the most frequently used modality for AI-based age estimation (Table 1). Other approaches investigated DNA methylation markers, magnetic resonance imaging (MRI) of skeletal structures, hand radiographs, cone-beam computed tomography (CBCT), and molecular biomarkers such as microRNA.

The sample sizes of the included studies varied substantially, ranging from fewer than 100 individuals to more than 20,000 participants, reflecting the heterogeneity of available datasets. Similarly, the age ranges investigated differed considerably across studies, with some

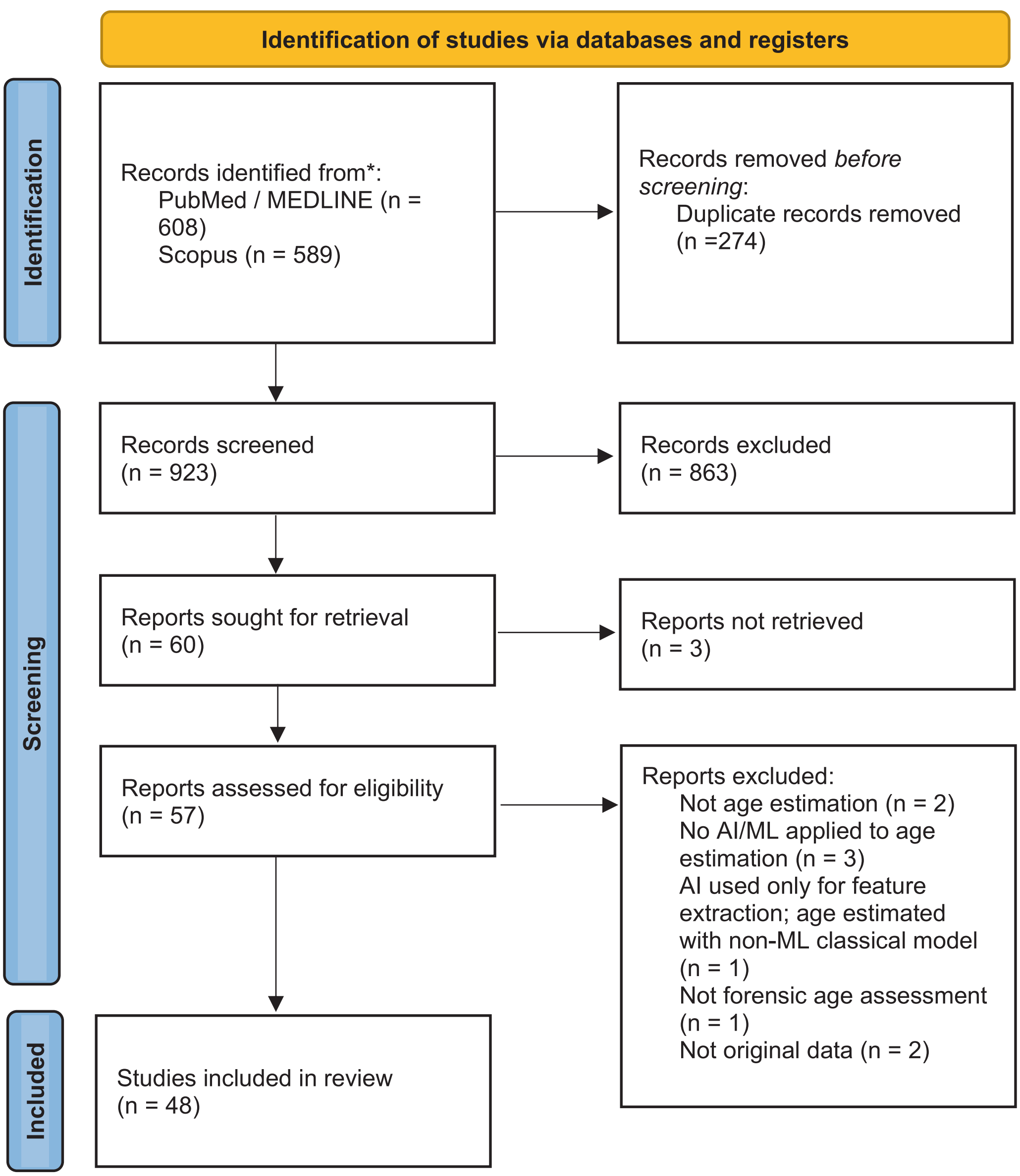

Fig. 1. PRISMA Flow diagram of the systematic review.

Table 1
Overview of the studies included in the systematic review (n = 48).

Characteristic	Category	Number of studies (%)
Data modality	Panoramic radiographs (OPG)	28 (58%)
	DNA methylation	9 (19%)
	MRI	3 (6%)
	CBCT	1 (2%)
	Hand radiographs	1 (2%)
	Blood miRNA	1 (2%)
	Other / combined approaches	5 (11%)
AI model type	Convolutional neural networks (CNN)	24 (50%)
	Machine learning regressors (RF, SVR, KNN, GBM etc.)	13 (27%)
	Neural networks (ANN / GRNN / MLP)	6 (13%)
	Support vector machine models	4 (8%)
	Other models	1 (2%)
Validation strategy	Train–test split	18 (38%)
	Cross-validation	19 (40%)
	Independent / external validation	6 (13%)
	Other / mixed approaches	5 (9%)
Legal age threshold analysis	Yes	22 (46%)
	No	26 (54%)

Abbreviations.
OPG – Orthopantomography (panoramic radiograph); CNN – Convolutional Neural Network; RF – Random Forest; SVR – Support Vector Regression; KNN – K-Nearest Neighbors; GBM – Gradient Boosting Machine; ANN – Artificial Neural Network; GRNN – Generalized Regression Neural Network; MLP – Multilayer Perceptron; SVM – Support Vector Machine; CBCT - Cone Beam Computed Tomography.

focusing on children or adolescents, while others included adult populations or wide lifespan cohorts (Supplementary **Table S1**).

3.3. Artificial intelligence models

A wide range of artificial intelligence approaches were applied across the included studies (Table 1). Convolutional neural networks (CNNs) represented the most frequently used models, particularly in studies analyzing radiological images such as panoramic dental radiographs or MRI scans.

In addition to deep learning architectures, several studies employed traditional machine learning algorithms, including random forests, support vector machines, k-nearest neighbors, gradient boosting models, and other regression-based approaches. In some cases, multiple algorithms were compared within the same study to identify the best-performing model for age prediction (Supplementary **Table S1**). These approaches enabled automated analysis of complex imaging or molecular datasets and facilitated the identification of patterns associated with biological maturation.

The main methodological workflow underlying AI-based forensic age estimation approaches identified in the included studies is summarized in Fig. 2.

3.4. Validation strategies

The methodological strategies used to evaluate AI models varied among the included studies (Table 1). The most common validation approaches consisted of train–test split designs and cross-validation procedures, particularly k-fold cross-validation.

A smaller number of studies reported the use of independent validation datasets, allowing for external testing of model performance (Supplementary **Table S1**). However, external validation across independent populations remained relatively limited, and many studies relied on single-center or population-specific datasets, which may restrict the generalizability of the proposed models.

3.5. Model performance

3.5.1. Chronological age prediction performance

Across the included studies, reported performance metrics varied depending on the data modality and modelling approach (Supplementary **Table S1**). Where comparable performance metrics were available, pooled estimates of MAE suggested that AI models based on dental imaging achieved an average error of approximately 1.1 years, while MRI-based skeletal models showed a pooled MAE of around 1.0 years. In contrast, molecular approaches based on DNA methylation demonstrated lower predictive accuracy, with a pooled MAE of approximately 3.7 years.

In studies based on dental radiographs, AI models frequently achieved mean absolute errors (MAE) in the range of approximately 0.5–1.5 years, particularly when deep learning architectures were applied.

Studies based on DNA methylation markers generally reported higher prediction errors, typically in the range of 3–5 years, although these approaches offer the advantage of applicability to biological samples rather than imaging data. MRI-based skeletal age estimation models also demonstrated promising performance, often achieving prediction errors close to 1 year.

3.5.2. Legal age threshold classification

A smaller but medico-legally important subgroup of studies specifically evaluated AI performance in relation to legally relevant age thresholds, particularly the distinction between individuals younger or older than 18 years. In these studies, classification-oriented metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC) were considered more informative for forensic applicability than regression-based error metrics alone (Table 1, Supplementary **Table S1**).

Several dental imaging studies reported high classification performance for threshold-based age categorization. Where reported, classification accuracies generally ranged between approximately 85% and 96%, depending on the imaging modality, dataset composition, and model architecture. For example, Guo et al. reported classification accuracies up to 95.9% using deep convolutional neural networks applied to orthopantomograms, while other studies evaluating third molar development or dental maturity classification also demonstrated promising performance around the 18-year threshold [14,20,45,51,53].

However, relatively few studies provided comprehensive reporting of classification-oriented performance measures such as sensitivity, specificity, calibration, or uncertainty around the decision boundary. Consequently, direct comparison of forensic applicability across studies remained limited.

3.6. Risk of bias assessment

Overall, several included studies demonstrated moderate methodological concerns or risk of bias, particularly regarding dataset representativeness, limited external validation, and insufficient reporting transparency. Many studies relied on single-center or population-specific datasets and internal validation procedures only, which may have resulted in optimistic estimates of model performance.

Additional concerns included potential risk of data leakage, incomplete reporting of preprocessing pipelines, and limited assessment of model calibration or uncertainty around legal decision thresholds. Studies specifically addressing legal age classification frequently showed concerns regarding generalizability across different populations and medico-legal contexts.

The overall risk-of-bias assessment based on QUADAS-2 and PROBAST evaluations is summarized in Supplementary **Fig. S1**.

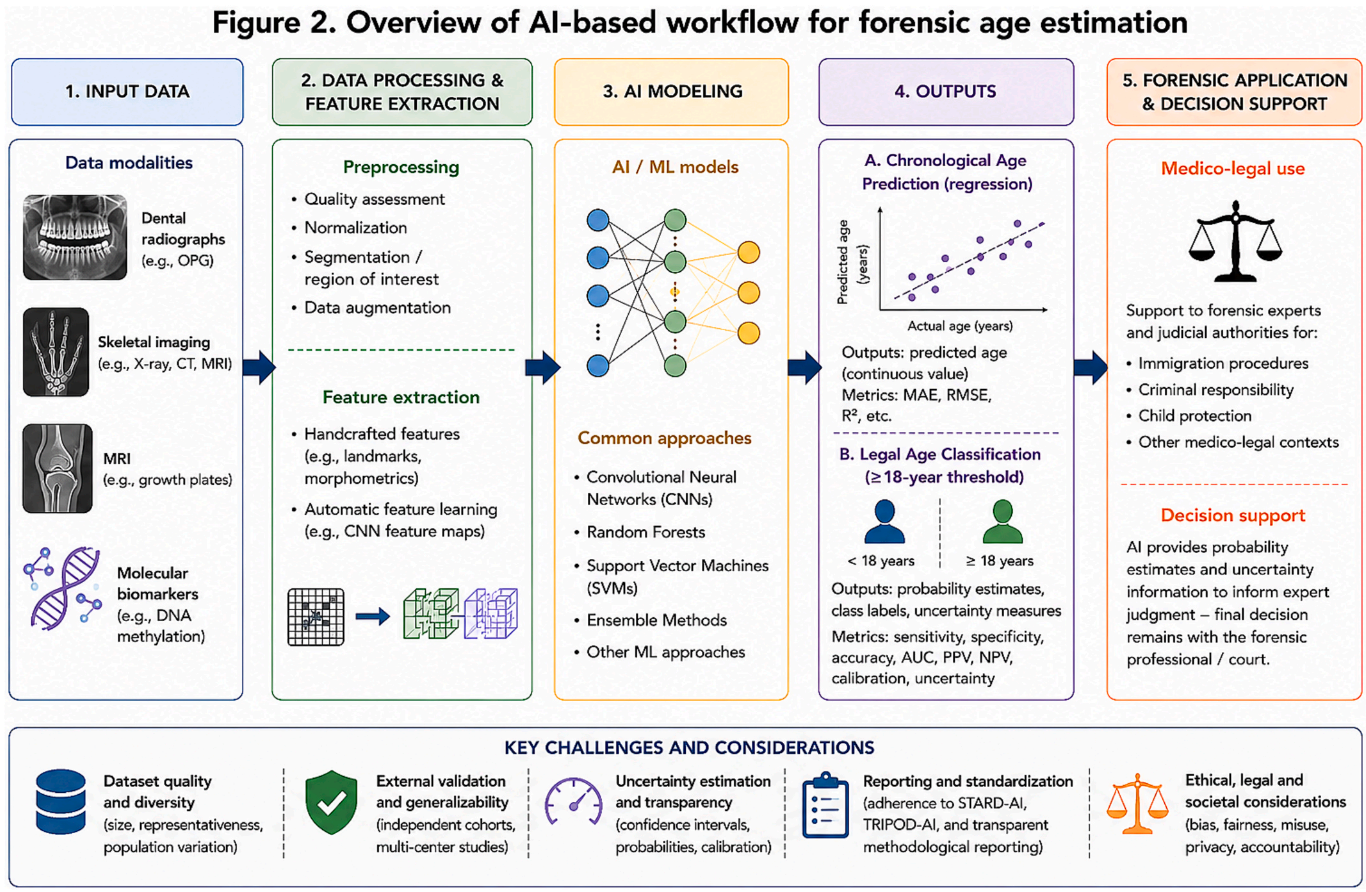

Fig. 2. Artificial intelligence workflow for forensic age estimation. Abbreviations: AI, artificial intelligence; AUC, area under the receiver operating characteristic curve; CNN, convolutional neural network; CT, computed tomography; MAE, mean absolute error; ML, machine learning; MRI, magnetic resonance imaging; NPV, negative predictive value; OPG, orthopantomogram; PPV, positive predictive value; RMSE, root mean square error; STARD-AI, Standards for Reporting Diagnostic Accuracy Studies–Artificial Intelligence; SVM, support vector machine; TRIPOD-AI, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis–Artificial Intelligence.

4. Discussion

This systematic review examined the current evidence on the use of artificial intelligence (AI) for forensic age assessment. A total of 48 studies were included, reflecting the rapid expansion of research exploring machine learning and deep learning approaches in medico-legal age estimation. The included studies investigated a variety of data sources, including dental radiographs, skeletal imaging, and molecular biomarkers, illustrating the multidisciplinary nature of AI-based age estimation research [11–58].

Among the different modalities, dental imaging represented the most frequently investigated approach. Numerous studies applied deep learning models to panoramic radiographs for automated age estimation or dental maturity assessment [14,15,19,23,26,39,42,52–54]. Other investigations focused specifically on third molar development or automated dental staging frameworks [20,30,55]. These findings confirm the central role of dental development in forensic age estimation and highlight how AI techniques are increasingly being used to automate the interpretation of dental imaging data.

Across the included studies, convolutional neural networks (CNNs) were the most commonly used modelling approach for image-based datasets [14,19,39,52–54]. In addition to deep learning architectures, several studies employed traditional machine learning algorithms such as random forests, support vector machines, and gradient boosting models, particularly in analyses involving structured data or molecular biomarkers [37,41,49,56,57]. In some studies, multiple algorithms were compared in order to identify the best-performing model for age prediction [41,49].

Most CNN-based models relied on transfer learning from pretrained architectures such as ResNet, DenseNet, or EfficientNet. Data augmentation techniques were frequently used to improve generalization in relatively small radiographic datasets. However, few studies reported model explainability approaches such as Grad-CAM or saliency maps, which may be important for medico-legal transparency.

Regarding predictive performance, several AI models applied to dental radiographs reported mean absolute errors (MAE) between approximately 0.5 and 1.5 years, indicating high predictive accuracy in image-based age estimation tasks [26,41,53,58]. Comparable performance was observed in studies using advanced deep learning architectures trained on large radiographic datasets [52–54]. These findings suggest that automated AI-based image analysis may achieve performance comparable to, or in some cases exceeding, traditional manual approaches to age estimation.

In contrast, DNA methylation–based models generally reported prediction errors in the range of 3–5 years [34,40,47,56]. Although these models are less precise than imaging-based approaches, they offer important advantages in forensic scenarios where only biological material is available. Several studies applied machine learning algorithms to methylation markers in blood samples to develop epigenetic clocks capable of predicting chronological age [18,34,40,47,56,57].

MRI-based skeletal age estimation showed potentially promising results in the limited number of available studies [38,51], with some models achieving prediction errors approaching approximately one year. However, the currently available evidence remains constrained by the relatively small number of studies, limited dataset diversity, and scarce external validation. Consequently, although MRI-based approaches may represent an attractive radiation-free alternative for medico-legal evaluation of living individuals, further large-scale and externally validated studies are required before their forensic applicability can be more definitively established.

Another important aspect emerging from the included studies concerns the evaluation of AI models in relation to legal age thresholds, particularly the distinction between individuals younger or older than 18 years. Several investigations specifically addressed this issue using dental radiographs or skeletal imaging datasets [14,20,38,45,51,53], highlighting the relevance of AI-based classification models for forensic decision-making in immigration and criminal justice contexts.

An important distinction emerging from this review is that many AI studies were primarily designed for chronological age prediction rather than for direct medico-legal threshold classification. Although regression-based metrics such as MAE are useful for evaluating average prediction accuracy, they may not fully capture forensic applicability in real-world medico-legal scenarios, where the key question often concerns whether an individual has reached a legally relevant threshold, particularly 18 years of age.

In these contexts, classification-oriented measures such as sensitivity, specificity, area under the receiver operating characteristic curve (AUC), predictive values, and uncertainty around the decision boundary may provide more clinically and legally meaningful information than regression error alone. However, relatively few studies comprehensively reported these metrics, limiting direct comparison of forensic applicability across different AI approaches.

This discrepancy highlights an important gap between the current technical development of AI-based age estimation systems and the practical requirements of medico-legal decision-making. Future research should therefore prioritize threshold-oriented validation strategies, transparent uncertainty reporting, and forensically interpretable classification frameworks specifically designed for forensic applications.

Although AI-based systems can achieve relatively low prediction errors, forensic age estimation should not be interpreted as a precise measurement of chronological age. In medico-legal practice, uncertainty intervals remain crucial, particularly around legally relevant thresholds such as 18 years. AI models should therefore be considered decision-support tools rather than definitive age determination methods.

Despite the promising performance reported in many studies, several methodological limitations were identified. One of the most important issues concerns the limited use of external validation datasets. Many studies relied exclusively on internal validation strategies such as train–test splits or cross-validation [39,53], which may lead to optimistic estimates of model performance.

Dataset representativeness also represents a significant limitation. Several studies were based on relatively small sample sizes or datasets derived from a single geographic population [37,41,50], which may limit the generalizability of the proposed models to different demographic or ethnic groups.

Another methodological concern relates to the potential risk of data leakage in machine learning pipelines, particularly when preprocessing or feature extraction steps are not clearly separated between training and testing datasets. Transparent reporting of model development and validation strategies is therefore essential to ensure reproducibility.

Another important limitation concerns the geographical concentration of available datasets. A large proportion of studies were conducted in Asian populations, particularly in China and Southeast Asia. Since dental and skeletal maturation patterns may vary between populations, models trained on single-population datasets may show limited external validity in forensic applications across different ethnic or geographic groups.

The increasing application of artificial intelligence in forensic age estimation also raises important ethical and medico-legal considerations. Age assessment outcomes may have significant legal implications, particularly in contexts such as immigration procedures, asylum applications, and criminal responsibility evaluations, where the determination of whether an individual has reached the age of majority may directly affect legal status and access to protection measures.

The use of AI-based systems should be approached with caution, particularly when algorithmic models are used to support decisions with potential legal consequences. Issues such as model transparency, explainability, and accountability are particularly relevant, especially for deep learning systems whose internal decision-making processes may be difficult to interpret.

Recent discussions in the medico-legal literature have highlighted the need for methodological frameworks capable of evaluating the role of artificial intelligence in healthcare decision-making and potential liability scenarios [59]. Although forensic age estimation represents a specific application domain, similar principles apply, including the importance of transparent model development, rigorous validation procedures, and careful expert interpretation of AI-assisted results.

For these reasons, AI-based age estimation systems should be considered decision-support tools rather than autonomous decision-makers, and their results should always be interpreted within a broader forensic and medico-legal framework.

Overall, the findings of this review indicate that AI-based methods have considerable potential to improve the automation and reproducibility of forensic age estimation, particularly through automated analysis of radiological images [14,26,41,53]. However, further research is required before these systems can be reliably implemented in forensic practice.

Future studies should prioritize external validation across diverse populations, the use of larger and more representative datasets, and the development of transparent methodological and reporting standards for AI-based forensic applications. Particular attention should also be given to the evaluation of model performance around legally relevant age thresholds, especially the distinction between individuals younger or older than 18 years.

5. Conclusions

Artificial intelligence is rapidly emerging as a valuable tool for forensic age estimation. The evidence reviewed in this study suggests that AI-based approaches, particularly deep learning models applied to dental radiographs and other imaging modalities, can achieve high predictive accuracy and may improve the automation and reproducibility of age estimation procedures.

Despite these promising results, important methodological challenges remain. Many studies rely on population-specific datasets and internal validation strategies, which may limit the generalizability of the proposed models. Furthermore, the heterogeneity of datasets, imaging modalities, and evaluation metrics highlights the need for standardized methodological frameworks and transparent reporting practices. Importantly, the currently available literature remains largely dominated by studies focused on chronological age prediction, whereas comparatively fewer investigations specifically address medico-legal threshold classification around legally relevant age boundaries.".

Future research should focus on large, diverse, and well-annotated datasets, as well as rigorous external validation across different populations. Particular attention should also be given to the performance of AI models around legal age thresholds, especially the distinction between individuals younger or older than 18 years, which remains a critical issue in forensic and medico-legal contexts.

Overall, while artificial intelligence offers promising opportunities to support forensic age assessment, further research is required to ensure that these systems are reliable, interpretable, and suitable for real-world medico-legal decision-making.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.legalmed.2026.102885. Supplementary Appendix 1 : database search strategies. Supplementary Table S1 : detailed characteristics of the studies included in the systematic review evaluating artificial intelligence approaches for forensic age estimation. Supplementary Figure S1 : summary risk-of-bias assessment.

References

[1] A. Schmeling, R. Dettmeyer, E. Rudolf, V. Vieth, G. Geserick, Forensic age estimation, Dtsch. Arztebl. Int. 113 (2016) 44–50, https://doi.org/10.3238/arztebl.2016.0044.

[2] B. Marcante, L. Marino, N.E. Cattaneo, A. Delicati, P. Tozzo, L. Caenazzo, Advancing forensic human chronological age estimation: Biochemical, genetic, and epigenetic approaches from the last 15 years: a systematic review, Int. J. Mol. Sci. 26 (2025) 3158, https://doi.org/10.3390/ijms26073158.

[3] R. Shedge, T. Kanchan, V. Warrier, S.G. Dixit, K. Krishan, Forensic age estimation using conventional radiography of the medial clavicular epiphysis: a systematic review, Med. Sci. Law 61 (2021) 138–146, https://doi.org/10.1177/0025802420988223.

[4] O. Lopatin, M. Barszcz, F. Bolechala, K.J. Wozniak, The fusion of ossification centers - a comparative review of radiographic and other imaging modalities of age assessment in living groups of children, adolescents, and young adults, Leg. Med. 61 (2023) 102185, https://doi.org/10.1016/j.legalmed.2022.102185.

[5] I.M. Suciyanie, F.P. Gultom, A.N. Hidayat, A.W. Suhartono, M. Yuniastuti, E. I. Auerkari, Accuracy of forensic age estimation using cementum annulation and dentin translucency in adult: a systematic review and meta-analysis, Int. J. Legal Med. 136 (2022) 1443–1455, https://doi.org/10.1007/s00414-022-02777-2.

[6] S.B. Balla, S.H. Ali, A.P. Mula, S.K. Tadakamadla, M.B. Rizky, A. Franco, N. Angelakopoulos, J. Tadakamadla, The effect of impaction on the mineralisation of third molars and forensic age estimation: a systematic review and meta-analysis, Forensic Sci. Int. 378 (2026) 112650, https://doi.org/10.1016/j.forsciint.2025.112650.

[7] R. Rokhshad, F. Nasiri, N. Saberi, R. Shoorgashti, S.S. Ehsani, Z. Nasiri, A. Azadi, F. Schwendicke, Deep learning for age estimation from panoramic radiographs: a systematic review and meta-analysis, J. Dent. 154 (2025) 105560, https://doi.org/10.1016/j.jdent.2025.105560.

[8] K. Pichetpan, P. Singsuwan, P. Mahakkanukrauh, Age estimation using medial clavicle by histomorphometry method with artificial intelligence: a review, Med. Sci. Law 64 (2024) 329–342, https://doi.org/10.1177/00258024241270779.

[9] F. Haugg, G. Lee, H. He, J. Johnson, A. Zapaishchykova, D.S. Bitterman, B.H. Kann, H.J.W.L. Aerts, R.H. Mak, Imaging biomarkers of ageing: a review of artificial intelligence-based approaches for age estimation, Lancet Healthy Longev. 6 (2025) 100728, https://doi.org/10.1016/j.lanhl.2025.100728.

[10] M.J. Page, J.E. McKenzie, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M.M. Lalu, T. Li, E.W. Loder, E. Mayo-Wilson, S. McDonald, L.A. McGuinness, L.A. Stewart, J. Thomas, A.C. Tricco, V.A. Welch, P. Whiting, D. Moher, The PRISMA 2020 statement: an updated guideline for reporting systematic reviews, BMJ 372 (2021) n71, https://doi.org/10.1136/bmj.n71.

[11] H.-M. Zhou, Z.-L. Zhou, Y.-H. He, T.-A. Liu, L. Wan, Y.-H. Wang, Forensic bone age assessment of hand and wrist joint MRI images in Chinese Han male adolescents based on deep convolutional neural networks, Int. J. Legal Med. 138 (2024) 2427–2440, https://doi.org/10.1007/s00414-024-03282-4.

[12] Q. Zheng, Z. Ge, H. Du, G. Li, Age estimation based on 3D pulp chamber segmentation of first molars from cone-beam computed tomography by integrated deep learning and level set, Int. J. Legal Med. 135 (2021) 365–373, https://doi.org/10.1007/s00414-020-02459-x.

[13] Z.-Y. Zhang, C.-X. Yan, Q.-M. Min, et al., Leverage effective deep learning searching method for forensic age estimation, Bioengineering 11 (2024) 674, https://doi.org/10.3390/bioengineering11070674.

[14] E. Yilmaz, C. Görürgöz, H.C. Kış, E.M. Canger, B. Öztaş, Forensic dental age estimation with deep learning: a modified Xception model for panoramic X-ray images, Forensic Sci. Med. Pathol. 21 (2025) 565–579, https://doi.org/10.1007/s12024-025-00962-4.

[15] B.S. Yavuz, O. Ekmekcioglu, H. Ankarali, Comparison of different dental age estimation methods with deep learning: Willems, Cameriere-European, London Atlas, Int. J. Legal Med. 139 (2025) 1661–1672, https://doi.org/10.1007/s00414-025-03452-y.

[16] T. Yamagishi, W. Sakurai, K. Watanabe, K. Toyomane, T. Akutsu, Development and comparison of forensic interval age prediction models by statistical and machine learning methods based on the methylation rates of ELOVL2 in blood DNA, Forensic Sci. Int. Genet. 69 (2024) 103004, https://doi.org/10.1016/j.fsigen.2023.103004.

[17] T.-J. Wu, C.L. Tsai, Y.-H. Huang, T.-Y. Fan, Y.-P. Chen, Efficacy of machine learning assisted dental age assessment in local population, Leg. Med. 59 (2022) 102148, https://doi.org/10.1016/j.legalmed.2022.102148.

[18] A. Vidaki, D. Ballard, A. Aliferi, T.H. Miller, L.P. Barron, D. Syndercombe Court, DNA methylation-based forensic age prediction using artificial neural networks and next generation sequencing, Forensic Sci. Int. Genet. 28 (2017) 225–236. Doi: 10.1016/j.fsigen.2017.02.009.

[19] W. Upalananda, C. Phisutphithayakun, P. Assawasuksant, P. Tanwattana, P. Prasatkaew, A fully automated deep learning framework for age estimation in adults using periapical radiographs of canine teeth, Int. J. Legal Med. 139 (2025) 2869–2881, https://doi.org/10.1007/s00414-025-03558-3.

[20] W. Upalananda, A. Charuakkra, S. Chaichulee, End-to-end vs. human-defined feature extraction: Comparing deep learning approaches for age classification using mandibular third molars, J. Forensic Odontostomatol. 43 (2025) 20–30, https://doi.org/10.5281/zenodo.17776415.

[21] M. Štepanovský, A. Ibrová, Z. Buk, J. Velemínská, Novel age estimation model based on development of permanent teeth compared with classical approach and other modern data mining methods, Forensic Sci. Int. 279 (2017) 72–82, https://doi.org/10.1016/j.forsciint.2017.08.005.

[22] M. Spólnicka, E. Pośpiech, B. Pepłońska, et al., DNA methylation in ELOVL2 and C1orf132 correctly predicted chronological age of individuals from three disease groups, Int. J. Legal Med. 132 (2018) 1–11, https://doi.org/10.1007/s00414-017-1636-0.

[23] M.B. Sivri, S. Taheri, R.G. Kırzıoğlu Ercan, Ü. Yağcı, Z. Golrizkhatami, Dental age estimation: a comparative study of convolutional neural network and Demirjian's method, J. Forensic Leg. Med. 103 (2024) 102679, https://doi.org/10.1016/j.jflm.2024.102679.

[24] T. Siahaan, A. Reckert, J. Becker, et al., Molecular and morphological findings in a sample of oral surgery patients: what can we learn for multivariate concepts for age estimation? J. Forensic Sci. 66 (2021) 1524–1532, https://doi.org/10.1111/1556-4029.14704.

[25] R. Saric, J. Kevric, N. Hadziabdic, et al., Dental age assessment based on CBCT images using machine learning algorithms, Forensic Sci. Int. 334 (2022) 111245, https://doi.org/10.1016/j.forsciint.2022.111245.

[26] L.T. Rocha, M.S. Ingold, F.K. Panzarella, et al., Applicability of Willems method for age estimation in Brazilian children: Performance of multiple linear regression and artificial neural network, Egypt. J. Forensic Sci. 12 (2022) 19, https://doi.org/10.1186/s41935-022-00271-9.

[27] M.R. Refn, M.-L. Kampmann, A. Vyöni, et al., Independent evaluation of an 11-CpG panel for age estimation in blood, Forensic Sci. Int. Genet. 76 (2025) 103214, https://doi.org/10.1016/j.fsigen.2024.103214.

[28] A.H.M. Razikin, N.N.A. Latip, M.S. Azlan, N.N. Azmi, A.A. Fauzi, P. Nambiar, Ascertaining the validity of a modified survey method for dental age estimations of Malaysian Malay and chinese juveniles, Anil Aggrawal's internet, J. Forensic Med. Toxicol. 25 (2024), https://doi.org/10.5281/zenodo.10019099.

[29] P.-L. Pröve, E. Jopp-van Well, B. Stanczus, et al., Automated segmentation of the knee for age assessment in 3D MR images using convolutional neural networks, Int. J. Legal Med. 133 (2019) 1191–1205, https://doi.org/10.1007/s00414-018-1953-y.

[30] P. Pintana, W. Upalananda, S. Saekho, U. Yarach, K. Wantanajittikul, Fully automated method for dental age estimation using the ACF detector and deep learning, Egypt. J. Forensic Sci. 12 (2022) 38, https://doi.org/10.1186/s41935-022-00314-1.

[31] V. Pinchi, I. Bianchi, F. Pradella, et al., AI-assisted age estimation in children based on a combination of bone and tooth maturity, Forensic Sci. Int. 378 (2026) 112688, https://doi.org/10.1016/j.forsciint.2025.112688.

[32] S.-J. Park, S. Yang, J.-M. Kim, et al., Automatic and robust estimation of sex and chronological age from panoramic radiographs using a multi-task deep learning network: a study on a south Korean population, Int. J. Legal Med. 138 (2024) 1741–1757, https://doi.org/10.1007/s00414-024-03204-4.

[33] C. Pan, S. Yi, C. Xiao, Y. Huang, X. Chen, D. Huang, The evaluation of seven age-related CpGs for forensic purpose in blood from Chinese Han population, Forensic Sci. Int. Genet. 46 (2020) 102251, https://doi.org/10.1016/j.fsigen.2020.102251.

[34] J. Naue, H.C.J. Hoefsloot, O.R.F. Mook, et al., Chronological age prediction based on DNA methylation: Massive parallel sequencing and random forest regression, Forensic Sci. Int. Genet. 31 (2017) 19–28, https://doi.org/10.1016/j.fsigen.2017.07.015.

[35] N. Mohammad, R. Ahmad, M.H.A. Gaus, A. Kurniawan, M.Y.P. Mohd Yusof, Accuracy of automated forensic dental age estimation lab (F-DentEst Lab) on large Malaysian dataset, Forensic Sci. Int. 361 (2024) 112150, https://doi.org/10.1016/j.forsciint.2024.112150.

[36] N. Mohammad, A.M. Muad, R. Ahmad, M.Y.P. Mohd Yusof, Reclassification of Demirjian's mandibular premolars staging for age estimation based on semi-automated segmentation of deep convolutional neural network, Forensic Imaging 24 (2021) 200440, https://doi.org/10.1016/j.fri.2021.200440.

[37] R. Merdietio Boedi, S. Shepherd, F. Oscandar, A.J. Franco, S. Mânica, Machine learning assisted 5-part tooth segmentation method for CBCT-based dental age estimation in adults, J. Forensic Odontostomatol. 42 (2024) 22–29, https://doi.org/10.5281/zenodo.11061543.

[38] M. Auf der Mauer, E. Jopp-van Well, J. Herrmann, et al., Automated age estimation of young individuals based on 3D knee MRI using deep learning, Int. J. Legal Med. 135 (2021) 649–663, https://doi.org/10.1007/s00414-020-02465-z.

[39] L. Matthijs, L. Delande, J. De Tobel, et al., Artificial intelligence and dental age estimation: Development and validation of an automated stage allocation

technique on all mandibular tooth types in panoramic radiographs, Int. J. Legal Med. 138 (2024) 2469–2479, https://doi.org/10.1007/s00414-024-03298-w.
[40] J.A. Mathew, G. Paul, J. Jacob, J. Kumar, N. Dubey, N.S. Philip, A new robust AI/ML based model for accurate forensic age estimation using DNA methylation markers, Forensic Sci. Med. Pathol. 21 (2025) 1145–1162, https://doi.org/10.1007/s12024-025-00985-x.
[41] X. Liang, C. Wang, D. Wen, et al., Combining machine learning algorithms to construct a new method for inferring dental age of children with missing teeth in Southern China, Int. J. Legal Med. 140 (2026) 357–374, https://doi.org/10.1007/s00414-025-03591-2.
[42] I. Kim, S. Yang, Y. Choi, H. Kwon, C. Lee, W. Park, Automated sex and age estimation from orthopantomograms using deep learning: a comparison with human predictions, Forensic Sci. Int. 374 (2025) 112531, https://doi.org/10.1016/j.forsciint.2025.112531.
[43] S.R. Hong, K.-J. Shin, S.-E. Jung, E.H. Lee, H.Y. Lee, Platform-independent models for age prediction using DNA methylation data, Forensic Sci. Int. Genet. 38 (2019) 39–47, https://doi.org/10.1016/j.fsigen.2018.10.005.
[44] M. Han, S. Du, Y. Ge, et al., With or without human interference for precise age estimation based on machine learning? Int. J. Legal Med. 136 (2022) 821–831, https://doi.org/10.1007/s00414-022-02796-z.
[45] Y.-C. Guo, M. Han, Y. Chi, et al., Accurate age classification using manual method and deep convolutional neural network based on orthopantomogram images, Int. J. Legal Med. 135 (2021) 1589–1597, https://doi.org/10.1007/s00414-021-02542-x.
[46] A. Galibourg, S. Cussat-Blanc, J. Dumoncel, N. Telmon, P. Monsarrat, D. Maret, Comparison of different machine learning approaches to predict dental age using Demirjian's staging approach, Int. J. Legal Med. 135 (2021) 665–675, https://doi.org/10.1007/s00414-020-02489-5.
[47] A. Freire-Aradas, L. Girón-Santamaría, A. Mosquera-Miguel, et al., A common epigenetic clock from childhood to old age, Forensic Sci. Int. Genet. 60 (2022) 102743, https://doi.org/10.1016/j.fsigen.2022.102743.
[48] L. Feng, F. Peng, S. Li, et al., Systematic feature selection improves accuracy of methylation-based forensic age estimation in Han chinese males, Forensic Sci. Int. Genet. 35 (2018) 38–45, https://doi.org/10.1016/j.fsigen.2018.03.009.
[49] A. Faragalli, L. Ferrante, N. Angelakopoulos, R. Cameriere, E. Skrami, Do machine learning methods solve the main pitfall of linear regression in dental age estimation? Forensic Sci. Int. 367 (2025) 112353 https://doi.org/10.1016/j.forsciint.2024.112353.
[50] C. Fang, X. Liu, J. Zhao, et al., Age estimation using bloodstain miRNAs based on massive parallel sequencing and machine learning: a pilot study, Forensic Sci. Int. Genet. 47 (2020) 102300, https://doi.org/10.1016/j.fsigen.2020.102300.
[51] F. Fan, H. Liu, X. Dai, et al., Automated bone age assessment from knee joint by integrating deep learning and MRI-based radiomics, Int. J. Legal Med. 138 (2024) 927–938, https://doi.org/10.1007/s00414-023-03148-1.
[52] F. Fan, W. Ke, X. Dai, et al., Semi-supervised automatic dental age and sex estimation using a hybrid transformer model, Int. J. Legal Med. 137 (2023) 721–731, https://doi.org/10.1007/s00414-023-02956-9.
[53] B. Büyükçakır, J. Bertels, P. Claes, D. Vandermeulen, J. de Tobel, P.W. Thevissen, OPG-based dental age estimation using a data-technical exploration of deep learning techniques, J. Forensic Sci. 69 (2024) 919–931, https://doi.org/10.1111/1556-4029.15473.
[54] Ž. Bizjak, T. Robič, DentAge: deep learning for automated age prediction using panoramic dental X-ray images, J. Forensic Sci. 69 (2024) 2069–2074, https://doi.org/10.1111/1556-4029.15629.
[55] N. Banar, J. Bertels, F. Laurent, et al., Towards fully automated third molar development staging in panoramic radiographs, Int. J. Legal Med. 134 (2020) 1831–1841, https://doi.org/10.1007/s00414-020-02283-3.
[56] A. Aliferi, S. Sundaram, D. Ballard, et al., Combining current knowledge on DNA methylation-based age estimation towards the development of a superior forensic DNA intelligence tool, Forensic Sci. Int. Genet. 57 (2022) 102637, https://doi.org/10.1016/j.fsigen.2021.102637.
[57] A. Aliferi, D. Ballard, M.D. Gallidabino, H. Thurtle, L. Barron, D. Syndercombe Court, DNA methylation-based age prediction using massively parallel sequencing data and multiple machine learning models, Forensic Sci. Int. Genet. 37 (2018) 215–226. Doi: 10.1016/j.fsigen.2018.09.003.
[58] A. Abuabara, T.V.P.M. do Nascimento, K.R. da Cruz, et al., Dental age estimation by comparing Demirjian's method and machine learning in Southeast Brazilian youth, Forensic Sci. Med. Pathol. (2025), in press. Doi: 10.1007/s12024-025-01042-3.
[59] R. Cecchi, F. Calabrò, J. Camatti, A.L. Santunione, M. Sperti, E.A. Zizzi, M.A. Deriu, Artificial intelligence in healthcare: Proposal for a new medico-legal methodology in medical liability, Leg. Med. 80 (2026) 102764, https://doi.org/10.1016/j.legalmed.2025.102764.