Convolutional neural network of age-related trends digital radiographs of medial clavicle in a Thai population: a preliminary study

Phisamon Kengkard[1], Jirachaya Choovuthayakorn[1], Chollada Mahakkanukrauh[1], Nadee Chitapanarux[1], Pittayarat Intasuwan[2], Yanumart Malatong[2], Apichat Sinthubua[2,3], Patison Palee[4], Sakarat Na Lampang[5], Pasuk Mahakkanukrauh[2,3]

[1]Faculty of Medicine, Chiang Mai University, Chiang Mai, [2]Department of Anatomy, Faculty of Medicine, Chiang Mai University, Chiang Mai, [3]Excellence in Osteology Research and Training Center, Chiang Mai University, Chiang Mai, [4]College of Arts, Media and Technology, Chiang Mai University, Chiang Mai, [5]Department of Oral Biology and Diagnostic Sciences, Faculty of Dentistry, Chiang Mai University, Chiang Mai, Thailand

Abstract: Age at death estimation has always been a crucial yet challenging part of identification process in forensic field. The use of human skeletons have long been explored using the principle of macro and micro-architecture change in correlation with increasing age. The clavicle is recommended as the best candidate for accurate age estimation because of its accessibility, time to maturation and minimal effect from weight. Our study applies pre-trained convolutional neural network in order to achieve the most accurate and cost effective age estimation model using clavicular bone. The total of 988 clavicles of Thai population with known age and sex were radiographed using Kodak 9000 Extra-oral Imaging System. The radiographs then went through preprocessing protocol which include region of interest selection and quality assessment. Additional samples were generated using generative adversarial network. The total clavicular images used in this study were 3,999 which were then separated into training and test set, and the test set were subsequently categorized into 7 age groups. GoogLeNet was modified at two layers and fine tuned the parameters. The highest validation accuracy was 89.02% but the test set achieved only 30% accuracy. Our results show that the use of medial clavicular radiographs has a potential in the field of age at death estimation, thus, further study is recommended.

Key words: Clavicle, Age estimation, Radiography, Convolutional neural network

Introduction

Over the past decades, crimes and pandemics have posted a real challenge for experts to establish biological profiles of human remains. Over 1,800 bodies from the Indian ocean Tsunami in 2004 remain unidentified even 7 months after the incidence [1]. Apart from that, there are growing cases of murdered victims yet to be given justice. According to Thai ministry of justice, for the past 17 years (2002–2019), over 4,000 unnamed bodies have been reported [2]. The European union and The US also face the same problem. The EU has more than 1,500 accumulated bodies [3] while the US has 4,000 bodies reported per year [4]. In all country, money and resources are spent storing and managing the bodies especially in Thailand where those bodies have to be kept up to 20 years [2]. However, the major portion of those bodies are of the people who have been reported missing. In order

Corresponding author:
Pasuk Mahakkanukrauh
Department of Anatomy & Excellence in Osteology Research and Training Center, Chiang Mai University, Chiang Mai 50200, Thailand
E-mail: pasuk034@gmail.com

to improve the situation, a more effective way of constructing one's identity is necessary and this is where age plays a crucial role.

Bone has long been recognized as one of the five key elements in identification process [5]. Its relationship with age and persistency makes it the most preserved part that remain in the crime scene. Various bone parts have been used by forensic anthropologists as an age indicator for example: pelvis, cranium, femur and clavicle [6-8]. Because of the characteristic of clavicles: low weight bearing, the longest time to maturate, easy to access, they are considered as the best candidate for accurate age estimation. The statement is correlate with Marera and Satyapal [9] study which found that estimating age from clavicle yields the most accurate result. However, methods used nowadays mainly rely on observing epiphyseal plate fusion of clavicles which could only predict age of young adult and adolescent. In addition, it requires specialists to interpret the result which leads to the problem of subjectivity and time consumption [10]. An alternative method which is widely used in bone age estimation other than epiphyseal plate union is bone mineral density (BMD). However, Botha et al. [11] stated that BMD in blacks does not significantly correlate with age which disagrees with previous research. Other methods used in this field: cortical thickness and histomorphometry are both impractical in real life application due to their methodology inconsistency and sample variety. Apart from this, histomorphometry is destructive to the specimen and should be an alternative way to forensic human identification.

Imaging techniques used to estimate bone age are conventional radiograph, digital X-ray radiograph (DXR), dual-energy X-ray absorptiometry (DXA), computed tomography (CT), magnetic resonance imaging (MRI) [12, 13]. DXA, CT, and MRI are considered hard - to – obtain and not cost efficient in our country. Consequently, digital radiography, which is easy to obtain and cost efficient, was chosen. Furthermore, DXR, compared to conventional radiograph, provides clearer and more accurate pictures. The concept of imaging and age estimation is based on the knowledge that bone density will deteriorate overtime due to bone modeling and remodeling mechanism. As we age, the remodeling ability diminishes, accompanied by changes in hormonal response and reduced production of calcitriol, results in increase of bone porosity [14].

Artificial Intelligence is an ability of a computer or robots to do tasks that are usually done by human. Deep learning is a type of artificial intelligence that imitates the way humans gain certain type of knowledge. There are 2 main types of deep learning: Artificial Neural Network (ANN) and Convolutional Neural Network (CNN). ANN refers to a network that resembles human neural network. After receiving signal from data, neurons weigh their data which is passed on from layer to another. If the system detects any defect, it will change the weight of that layer. CNN refers to a system that learns mainly by pattern recognition and is widely used in images recognition – classification and object detection. It has been used for analyzing medical images in various field such as prediction of blood glucose level from electrocardiogram [15], Not repeated in the main text. Please spell it out. evaluation of lung nodules on chest CT and classification of ear abnormality.

Thus, we aim to embrace the help of CNN (GoogLeNet; Google Inc., Mountain View, CA, USA) with digital radiograph to create a more convenient, refined, and non-subjective method to estimate age from medial clavicle in a Thai population, which has not yet been done. This would promote precise personal identification and facilitate the forensic team to return justice to the dead.

Materials and Methods

Human dry-bone specimens of Thai national (Mongoloid) were provided by the Forensic Osteology Research Center, Faculty of Medicine, Chiang Mai University, Thailand. The study sample encompasses 374 female clavicles and 641 male clavicles. All samples were collected between 2006–2018 with the age at death ranging from 20–100. The cause of death was mostly natural. Individuals with health condition affecting bone mass density for example Thalassemia or with visible bone deformities such as fractures and erosions were excluded. Further biomedical information such as weight, lifestyle and personal medication was not available. The study was approved by the Research ethics committee, Faculty of Medicine, Chiang Mai University, Thailand (Research ID: ANA2563-07823).

Digital X-ray protocol

The samples were placed in an anteroposterior position with distal ends placed on the fixed clay and medial ends elevated 1 cm above the x-ray plate. Each radiograph comprised of 8 pairs of clavicles and the main focus is on the medial end of the clavicles (Fig. 1). Radiographs of the specimens

were obtained using Kodak 9000 Extraoral Imaging System (Eastman Kodak, Rochester, NY, USA) with the criteria: 60 kV, 2 mA, DC 1 seconds, 124 cm of focus-film distance at the Oral Radiology Clinic of Faculty of Dentistry, Chiang Mai University, Thailand. The radiographs were then stored as bitmap files, organized by collection number, sex and age.

Image pre-processing

The radiographs of the clavicular sample were processed using Adobe Photoshop 2019 (version 19.1.6, Adobe Systems Inc., San Jose, CA, USA). The region of interest was selected at the medial end of each clavicle using a template size 3×3 cm^2 (261×261 pixels) with a reference point being at the sternal border of the clavicle (Fig. 1). Images containing artifacts such as remnant of clay were excluded. The number of total samples used was 625 and 363 for male and female respectively, allowing them to be distributed in 8 age groups (Table 1).

Deep learning algorithm

CNN which excels at image classification were used in this study [16]. A pre-trained network called GoogLeNet was adopted to minimize training duration and reduce the need for large amount of training data [17]. The loss3-classifier and output classification layer of the network were substituted with 'Fully connected layer' and 'The final class output'. Other parameters were adjusted based on data set characteristics. MATLAB R2021b (MathWorks, Natick, MA, USA) were used as the main platform of this study.

Data augmentation

To prevent overfitting due to small sample size, all images were randomly augmented during training. The samples could be reflected horizontally, rotated from –60 to 60, rescaled from 0.5 to 1.5 and translated horizontally and vertically from –10 to 10.

Initial experiment

The network was tested to evaluate initial performance of GoogLeNet model. Two sets of experiment containing different datasets were carried out using the same parameters (InitialLearnRate 0.001, Epochs 30 and MiniBatch Size 8) to evaluate the effect of sex on age estimation. The validation accuracy of male-only dataset and male-and-female dataset were 31.89% and 25.91% respectively. The results imply that gender specific data yield better outcome despite having smaller size of training data. Thus, shifting this research focus to development of gender-specific network. Hence, only male data were used in this research.

Sample modification

From the initial experiment, the model was found to be inadequately trained due to small dataset and unequally distributed data. After careful consideration, the category 20–29 was excluded from the study since its number of samples were significantly smaller than others. From each remaining

Table 1. Number of total samples

Age range (yr)	Total male	Total female	Total
20–29	6	5	11
30–39	20	13	33
40–49	64	33	97
50–59	113	63	176
60–69	155	74	229
70–79	141	90	231
80–89	95	72	167
90–100	31	13	44
Total	625	363	988

Table 2. Samples of both original and generated clavicular images

Age range (yr)	Total male	20% (validation)	80%	GAN	Discarded	Total
30–39	20	4	16	500	43	473
40–49	64	12	52	500	24	528
50–59	113	22	91	500	34	557
60–69	155	30	125	500	55	570
70–79	141	28	113	500	42	571
80–89	95	18	77	500	56	521
90–100	31	6	25	500	61	464
Total	619	120	499	3,500	315	3,684

GAN, Generative Adversarial Network.

Fig. 1. The radiographs of 8 pairs clavicular samples, red box=3×3 cm^2 region of interest.

7 categories, 20% of samples were randomly separated as a test set. Then, Generative Adversarial Network (GAN) was used to generate 500 more samples per category with Epochs 5,000 and MiniBatchSize 16. Newly generated images that did not resemble clavicular structure were excluded. The total 3,999 figure samples of both original and generated clavicular images were randomly allocated to training and validation set with the ratio of 75:25 (Table 2).

Blind case testing

After the training process with the highest validation accuracy and lowest loss training, the test set was assessed by the network to evaluate the performance of the model. The result of the trained network was exported into the workspace using MATLAB. Commands were used to loaded pictures of clavicles (120 blind images) into the current folder to predict the age of the test set. The code for prediction age was below the transfer learning with deep network designer live editor, initially separated during sample modification, to the network in a one-by-one manner.

Results

The total of 3,684 images of male samples, after excluding from 3,999 images that did not resemble the clavicular structure, were used for this model with an age range at death
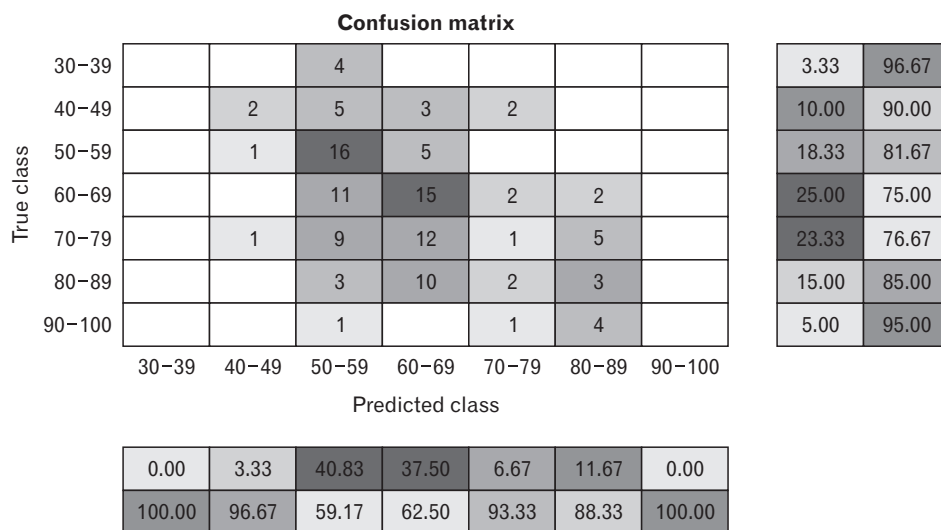
Confusion matrix

True class \ Predicted class	30-39	40-49	50-59	60-69	70-79	80-89	90-100		
30-39			4					3.33	96.67
40-49		2	5	3	2			10.00	90.00
50-59		1	16	5				18.33	81.67
60-69			11	15	2	2		25.00	75.00
70-79		1	9	12	1	5		23.33	76.67
80-89			3	10	2	3		15.00	85.00
90-100			1		1	4		5.00	95.00
	0.00	3.33	40.83	37.50	6.67	11.67	0.00		
	100.00	96.67	59.17	62.50	93.33	88.33	100.00		

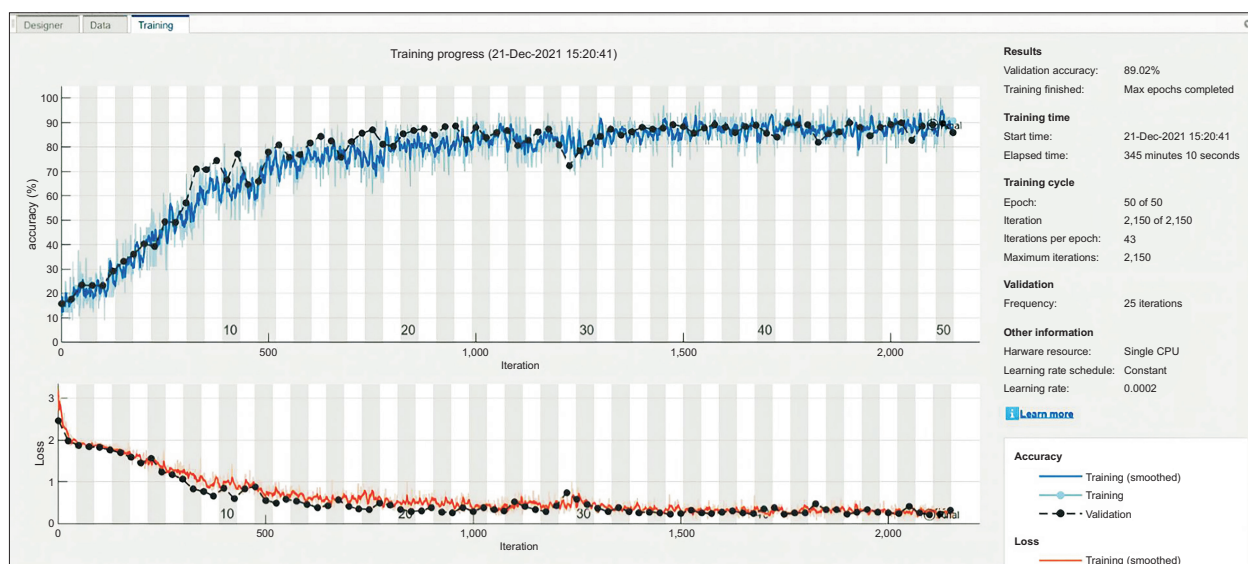**Fig. 2.** Confusion matrix of the predicted age range.

Fig. 3. Graph of training accuracy and loss for the final CNN model. CNN, convolutional neural network.

from 30 to 100 years. A training dataset was used for CNN models, and a validation dataset was used to avoid overfitting. The original dataset was randomly divided into 75% for training and 25% for validation. Fig. 2 shows the confusion matrices from the training and validation dataset of models. The correct age prediction for each class was 96.67%, 59.17%, 62.5%, 93.33%, and 88.33% for 40–49, 50–59, 60–69, 70–79, and 80–89 classes, respectively. Because the age group of 30–39 and 90–100 have a small sample size, the confusion matrix that the auto-random test was auto-excluded those two groups for testing (Fig. 2).

Fig. 3 demonstrates training progress, accuracy, and loss after training of the final CNN model, the highest validation accuracy was 89.02% when applied hyperparameters of the age estimation model as follows: InitialLearnRate 0.0002, L2regularization 0.0002, MaxEpochs 50, MiniBatchSize 64, Momentum 0.99 and Validation frequency 25 (Fig. 3).

The testing image dataset was assessed by the network to evaluate the performance of the trained model. When testing dataset with 120 hidden images in a one-by-one manner was tested by the CNN model, the correct age was achieved with 30.0% accuracy (correctly assigned 36 out of 120 cases).

Discussion

Clavicle is considered the bone of choice in age estimation due to its universal reliability compared to other bone parts. It has been proved to be the most reliable indicator for age estimation in radiological studies [17]. The medial part of the clavicle is widely used in forensic field since it was less affected by weight than the body and acromion part. In postmortem condition, medial clavicle is also the least damaged part while the acromion part is highly fragile especially in elderly [18]. The body is markedly affected by force and is composed with different cortical components and properties compared with acromion and sternal end. Accordingly, it is the most frequent (70%–80%) location of fracture [9].

In the current study, the area of 3×3 cm from sternal end of the medial clavicle was selected as a region of interest from the reason that the medial end is the most preserved part and would accurately depicts bone cortex and trabecular representing bone porosity. Physiologically, our bodies balance between bone degenerative process and remodeling process which the latter decreases as we age [13]. This results in increase porosity and decrease bone mass in elderly [19]. In our settings, using anatomical landmarks on radiographs

to mark region of interest such as costoclavicular ligament [18] was not accomplished due to technical limitations and radiographs quality.

Currently, clavicle is used in age estimation by observing gross appearance of epiphyseal plate closure separating into 4 stages (Webb and Suchey classification) by using completeness of epiphyseal union. However, even though it is very accurate, it could only estimate age in adolescents and young adult (15–29 years old in Thai population) [19]. The Study Group on Forensic Age Diagnostics Arbeitsgemeinschaft für Forensische Altersdiagnostik (AGFAD), recommended that the medial clavicular epiphyses using the radiographic images are one of the evidence to determine for adulthood legal age estimation [18, 20, 21] while estimation in other age range is still in debate [22]. Thus, clavicles aged between 30–100 years old were chosen, seeking an alternative way to accurately estimate age at this range and clavicles aged under 30 years were excluded from this study.

According to Marera and Satyapal [9], age assessment using clavicle epiphyseal plate closure was more precise in males than in females. This correlates with the fact that female is affected from osteoporosis more and faster than male from smaller bone size and the process of estrogen (bone resorption) as the age advanced [23]. The effect of estrogen mainly acts on trabecular bone [24] thus, male, and female individuals show a different pattern of osteoporosis and should yield different pattern. From this reason, male and female bones was not categorized as a same data set to train deep learning.

Our study also shows that when combined with female, the accuracy of age estimation using deep learning were 25.9% which is lower than male alone (31.9%). Besides, our Forensic Osteology Research Center houses bone collection of 278 Thai mongoloid males which is twice as much of female. Consequently, unlike previous studies, male is our main focus.

Regarding the knowledge mentioned before that bone porosity increases as age advances, nowadays a lot of research is conducted using various imaging techniques to detect bone porosity or bone mass for age estimation. These methods could widely expand age range of interest and are more practical compared to the epiphyseal method. According to Benito et al. [18], grey average together with cortical thickness of clavicle from radiograph show significant negative correlation with age. Similar results have been obtained by Botha et al. [11] and Navega et al. [25] which used BMD from

DXA as an indicator. However, these two studies were conducted using femur instead of clavicle.

The advancement of Deep learning in recent years, especially the development of a certain type of neural network called CNN, has benefited medical field through the use of pattern recognition [23]. In Forensic works, CNN has mainly been studied as a tool for age and sex assessment [26]. Several modalities of images from X-ray, DXA, MRI and CT are applied to various networks, newly developed or pre-trained, to obtain the highest performing model [27]. Navega et al. [25] accomplished in combining deep learning with DXA imaging modality of femur to estimate age at death in Caucasian female. In Mongoloid, this type of studies is still limited.

In this study, left and right clavicles were computed altogether since author wanted deep learning to be able to estimate age at death from both sides. This would allow applying this deep learning in real life situations where any sides of the clavicle could be left in crime scene.

The results from our first few experiments using solely original male clavicular images yield only 68% validation accuracy despite achieving 100% training accuracy. This indicates overfitting which, in layman's terms, means that AI model has learned in a manner that is only applicable to the training sample and is no longer generalizable to the overall population. Ideally, to overcome this problem, more training data should be collected and general basic starting point of training of 1,000 images per category should be reached [28]. However, in medical field, this task is merely impossible due to various reasons such as ethical privacy, high cost of obtaining data, requirement of data labelling from specialists and rare incidence of diseases [29-31]. Although there are several open access databases offering medical images for researchers, our study, which focus on Thai population, may not benefit from them. To tackle the insufficient data, our study used data augmentation, a technical solution widely adopted in medical imaging field [31-33].

GAN, a machine learning framework composing of a generator network and discriminator network, are extensively used for medical image synthesis [32]. Many studies have evaluated the quality of data generated from GAN by using it for various tasks through training of other CNNs. Frid-Adar, whose study focused on image classification, reported a performance boosted from 88.4% to 92.4% specificity. Another study on lung cytology images also reported a better outcome after using GAN with a rise of 4.3% accuracy. Other studies have shown promising results as well [32, 34,

35]. These illustrate GAN's generalizability thus can be apply with clavicular radiographs. The result from our study is in concordance with the mentioned studies as we found a 5% increase in validation accuracy of testing data.

Pattern recognition is the core concept of image classification carried out by CNN, implying that image with better quality would enable distinguish features to be drawn out more accurately [33]. In his research, Samuel Dodge found noise and blur to have a negative impact on neural networks' performance, including GoogLeNet's. However, training CNN with low quality image may not always decrease the network performance. If the trained network is tested on images of the approximate quality, its performance is likely to be improved. On the contrary, if the test set contains higher resolution data, the network's validation accuracy would have a tendency to plummet [33]. Our study was situated in this scenario which the majority of images (3,186 of 3,684 or 86.5%) used for training were derived from GAN with apparent lower resolution while the test set comprised of only original images.

This could explain the wide gap between our training and testing validation accuracy, a sharp decrease from 87% to 30% accuracy respectively. Our problem could be solved utilizing GAN with optimal fine tuning for better resolution, but due to hardware and time limitation, this was not possible. Another explanation of our finding would be the unequally distributed data [36]. From the confusion matrix (Fig. 2), it showed that our network performed best in the category with the most images and most of the false positive falls into that same category. Additional information such as socioeconomic, nutrition, occupation, and other factors which were not included in the assessment would be beneficial to our study.

Our initial hypothesis was that decreasing trabecular density, a known indicator for increasing age, would correlate with overall grey level of clavicular radiographs allowing the network to recognize difference pattern of each age group. However, with our imaging protocol, digital radiograph may not be able to capture minimal variation in grey level of each age group, resulting in poor categorization [18]. Further studies using other imaging techniques such as MRI or CT should be considered as it would allow CNN to be trained with higher quality images.

Our preliminary study has given an insight into what CNN could achieve in the field of age at death estimation. Even though the accuracy of the test set is quite low, the

results from this study show the various possibilities to apply the CNN model for age estimation in a Thai population sample and to decrease the subjectivity as well as errors in the measurement. The result shows possibility of using CNN as a part of identification tool although, facing major limitation of small dataset. The accuracy of network is expected to increase as more data and resources are provided. In the future, collaboration with other bone collections to expand our samples and to develop more accurate results is expected. Apart from that, a more extensive network modification with advanced technique would result in an even more practical model. The success of further development will present an opportunity for real life application of the technology.

ORCID

Phisamon Kengkard: https://orcid.org/0000-0002-0209-304X
Jirachaya Choovuthayakorn:
https://orcid.org/0000-0001-5507-7855
Chollada Mahakkanukrauh:
https://orcid.org/0000-0001-5022-6129
Nadee Chitapanarux: https://orcid.org/0000-0002-0695-0061
Pittayarat Intasuwan: https://orcid.org/0000-0002-3008-7974
Yanumart Malatong: https://orcid.org/0000-0003-1191-4206
Apichat Sinthubua: https://orcid.org/0000-0002-5349-4214
Patison Palee: https://orcid.org/0000-0003-4538-4266
Sakarat Na Lampang: https://orcid.org/0000-0001-5825-1468
Pasuk Mahakkanukrauh:
https://orcid.org/0000-0003-0611-7552

Author Contributions

Conceptualization: PM. Data acquisition: PK, JC. Data analysis or interpretation: CM, NC. Drafting of the manuscript: PK, JC, CM, NC. Critical revision of the manuscript: PI, YM, AS, PP, SNL. Approval of the final version of the manuscript: all authors.

Conflicts of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgements

References

1. Morgan OW, Sribanditmongkol P, Perera C, Sulasmi Y, Van Alphen D, Sondorp E. Mass fatality management following the South Asian tsunami disaster: case studies in Thailand, Indonesia, and Sri Lanka. PLoS Med 2006;3:e195.

2. Srinak N. The study of unclaimed and unidentified bodies management in Thailand and other countries. J Thai Justice Syst 2020;13:139-52.

3. Evert L. Unidentified bodies in forensic pathology practice in South Africa: demographic and medico-legal perspectives [thesis]. Hatfield: University of Pretoria; 2011.

4. National Missing and Unidentified Persons System. The nation's silent mass disaster [Internet]. Washington, D.C.: National Institute of Justice [cited 2022 Oct 29]. Available from: https://namus.nij.ojp.gov/.

5. Manzoor Mughal A, Hassan N, Ahmed A. Bone age assessment methods: a critical review. Pak J Med Sci 2014;30:211-5.

6. Hermetet C, Saint-Martin P, Gambier A, Ribier L, Sautenet B, Rérolle C. Forensic age estimation using computed tomography of the medial clavicular epiphysis: a systematic review. Int J Legal Med 2018;132:1415-25.

7. Shirley NR. Age and sex estimation from the human clavicle: an investigation of traditional and novel methods [PhD dissertation]. Knoxville: University of Tennessee; 2009.

8. Schulz R, Mühler M, Mutze S, Schmidt S, Reisinger W, Schmeling A. Studies on the time frame for ossification of the medial epiphysis of the clavicle as revealed by CT scans. Int J Legal Med 2005;119:142-5.

9. Marera DO, Satyapal KS. Fusion of the medial clavicular epiphysis in the South African and Kenyan populations. Int J Morphol 2018;36:1101-7.

10. Milenkovic P, Djukic K, Djonic D, Milovanovic P, Djuric M. Skeletal age estimation based on medial clavicle--a test of the method reliability. Int J Legal Med 2013;127:667-76.

11. Botha D, Lynnerup N, Steyn M. Age estimation using bone mineral density in South Africans. Forensic Sci Int 2019;297:307-14.

12. Kranioti EF, Bonicelli A, García-Donas JG. Bone-mineral density: clinical significance, methods of quantification and forensic applications. Res Rep Forensic Med Sci 2019;9:9-21.

13. Rowe P, Koller A, Sharma S. Physiology, bone remodeling. StatPearls. Treasure Island: StatPearls Publishing; 2022.

14. Seeman E, Delmas PD. Bone quality--the material and structural basis of bone strength and fragility. N Engl J Med 2006;354:2250-61.

15. Li Y, Huang Z, Dong X, Liang W, Xue H, Zhang L, Zhang Y, Deng Z. Forensic age estimation for pelvic X-ray images using deep learning. Eur Radiol 2019;29:2322-9.

16. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.

17. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. Paper presented at: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015 Jun 7-12; Boston, USA. p. 1-9.

18. Benito M, Sánchez JA, Codinha S. Age-at-death estimation based on radiological and image analysis methods in clavicle in a current Spanish population. Int J Legal Med 2014;128:523-33.

19. Chantharawetchakun T, Vachirawongsakorn V. Age estimation in the Thai male population using epiphyseal union of the medial clavicle. Chiang Mai Med J 2021;60:149-55.

20. Schulz R, Mühler M, Reisinger W, Schmidt S, Schmeling A. Radiographic staging of ossification of the medial clavicular epiphysis. Int J Legal Med 2008;122:55-8.

21. Schmeling A, Grundmann C, Fuhrmann A, Kaatsch HJ, Knell B, Ramsthaler F, Reisinger W, Riepert T, Ritz-Timme S, Rösing FW, Rötzscher K, Geserick G. Criteria for age estimation in living individuals. Int J Legal Med 2008;122:457-60.

22. Subramanian S, Viswanathan VK. Bone age. StatPearls. Treasure Island: StatPearls Publishing; 2022.

23. Fourcade A, Khonsari RH. Deep learning in medical image analysis: a third eye for doctors. J Stomatol Oral Maxillofac Surg 2019;120:279-88.

24. Ott SM. Cortical or trabecular bone: what's the difference? Am J Nephrol 2018;47:373-5.

25. Navega D, Coelho JD, Cunha E, Curate F. DXAGE: a new method for age at death estimation based on femoral bone mineral density and artificial neural networks. J Forensic Sci 2018;63:497-503.

26. Thurzo A, Kosnáčová HS, Kurilová V, Kosmeľ S, Beňuš R, Moravanský N, Kováč P, Kuracinová KM, Palkovič M, Varga I. Use of advanced artificial intelligence in forensic medicine, forensic anthropology and clinical anatomy. Healthcare (Basel) 2021;9:1545.

27. Guo YC, Han M, Chi Y, Long H, Zhang D, Yang J, Yang Y, Chen T, Du S. Accurate age classification using manual method and deep convolutional neural network based on orthopantomogram images. Int J Legal Med 2021;135:1589-97.

28. Mutasa S, Sun S, Ha R. Understanding artificial intelligence based radiology studies: What is overfitting? Clin Imaging 2020;65:96-9.

29. Chartrand G, Cheng PM, Vorontsov E, Drozdzal M, Turcotte S, Pal CJ, Kadoury S, Tang A. Deep learning: a primer for radiologists. Radiographics 2017;37:2113-31.

30. Zhang P, Zhong Y, Li X. ACCL: adversarial constrained-CNN loss for weakly supervised medical image segmentation. *arXiv*. 2005.00328 [Preprint]. 2020 [cited 2022 Dec 6]. Available from: https://doi.org/10.48550/arXiv.2005.00328.

31. Guan S, Loew M. Breast cancer detection using synthetic mammograms from generative adversarial networks in convolutional neural networks. J Med Imaging (Bellingham) 2019;6:031411.

32. Wang Y, Zhou L, Wang M, Shao C, Shi L, Yang S, Zhang Z, Feng M, Shan F, Liu L. Combination of generative adversarial network and convolutional neural network for automatic subcentimeter pulmonary adenocarcinoma classification. Quant Imaging Med Surg 2020;10:1249-64.

33. Dodge S, Karam L. Understanding how image quality affects deep neural networks. Paper presented at: 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX); 2016 Jun 6-8; Lisbon, Portugal. p. 1-6.

34. Aggarwal A, Mittal M, Battineni G. Generative adversarial network: an overview of theory and applications. Int J Inf Manag Data Insights 2021;1:100004.

35. Kazeminia S, Baur C, Kuijper A, van Ginneken B, Navab N, Albarqouni S, Mukhopadhyay A. GANs for medical image analysis. Artif Intell Med 2020;109:101938.

36. Buda M, Maki A, Mazurowski MA. A systematic study of the class imbalance problem in convolutional neural networks. Neural Netw 2018;106:249-59.