DOI: https://doi.org/10.17816/fm16110

The potential of synthetic minority oversampling technique to enhance the precision of gender prediction: an investigation of artificial neural networks with cephalometry

Vitria Wuri Handayani[1, 2], Ahmad Yudianto[3, 4], Mieke Sylvia M.A.R.[5], Riries Rulaningtyas[6, 7], Muhammad Rasyad Caesarardhi[8], Ramadhan Hardani Putra[9]

[1] Medical Faculty, Universitas Airlangga, Surabaya, Indonesia;
[2] Nursing Department, Pontianak Polytechnic Health Ministry, Pontianak, Indonesia;
[3] Department of Forensics and Medicolegal, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia;
[4] Magister of Forensic Sciences, Postgraduate School, Universitas Airlangga, Surabaya,Indonesia;
[5] Forensic Odontology Department, Dental Medical Faculty, Univesitas Airlangga, Surabaya, Indonesia;
[6] Physics Department, Sains and Technology Faculty, Universitas Airlangga, Surabaya, Indonesia;
[7] Biomedical Department, Sains and Technology Faculty, Universitas Airlangga, Surabaya,Indonesia;
[8] Department of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia;
[9] Department of Dentomaxillofacial Radiology, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, Indonesia

ABSTRACT

BACKGROUND: When creating models utilizing artificial neural networks, the quantity of training data and the distribution of data need to be considered, particularly when making gender predictions.

AIM: This study seeks to determine the potential impact of using the synthetic minority oversampling technique (SMOTE) on gender prediction using the artificial neural networks model.

MATERIALS AND METHODS: The current study utilized a dataset consisting of 297 cephalometric measurements from Indonesian patients, comprising 229 samples from females and 68 samples from males. WebCeph is used to measure certain parameters, such as Sella-Nation-Point A (SNA) angle, mandibular length, mandibular angle, Sella-Glabella-Point A (SGA) angle, and diagnosis. Data processing and artificial neural networks model creation were conducted using Python.

RESULTS: The gender identification accuracy of the artificial neural networks model is 87% for females and 0% for males, resulting in an overall average accuracy of 78%. When using SMOTE, the accuracy is 22%, with 0% for females and 37% for males. However, when using SMOTE and normalization, the accuracy increases to 71%, with 82% for females and 30% for males. The accuracy of normalization without SMOTE is 76%, with 86% for females and 14% for males.

CONCLUSIONS: This research has proven the efficacy of SMOTE in improving the classification of male matrices. Nevertheless, this study reveals that the overall accuracy results of SMOTE are suboptimal in comparison to the absence of SMOTE and normalization. The application of data balancing strategies is necessary to achieve optimal accuracy in gender prediction when artificial neural networks, and other parameters must be applied.

Keywords: artificial neural networks; cephalometry; gender determination; SMOTE.

To cite this article:
Handayani VW, Yudianto A, Mieke Sylvia MAR, Riries R, Rasyad M, Putra R. The potential of synthetic minority oversampling technique to enhance the precision of gender prediction: an investigation of artificial neural networks with cephalometry. *Russian Journal of Forensic Medicine.* 2024;10(2):139–151. DOI: https://doi.org/10.17816/fm16110

Received: 09.01.2024 **Accepted:** 15.04.2024 **Published online:** 07.06.2024

DOI: https://doi.org/10.17816/fm16110

Потенциал метода синтетической передискретизации меньшинства для повышения точности определения пола: исследование искусственных нейронных сетей с помощью цефалометрии

V.W. Handayani[1, 2], A. Yudianto[3, 4], MAR Sylvia Mieke[5], R. Riries[6, 7], M.R. Caesarardhi[8], R. Putra[9]

[1] Medical Faculty, Universitas Airlangga, Surabaya, Indonesia;
[2] Nursing Department, Pontianak Polytechnic Health Ministry, Pontianak, Indonesia;
[3] Department of Forensics and Medicolegal, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia;
[4] Magister of Forensic Sciences, Postgraduate School, Universitas Airlangga, Surabaya,Indonesia;
[5] Forensic Odontology Department, Dental Medical Faculty, Univesitas Airlangga, Surabaya, Indonesia;
[6] Technology Faculty, Universitas Airlangga, Surabaya, Indonesia;
[7] Biomedical Department, Sains and Technology Faculty, Universitas Airlangga, Surabaya,Indonesia;
[8] Department of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia;
[9] Department of Dentomaxillofacial Radiology, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, Indonesia

АННОТАЦИЯ

Обоснование. При создании моделей, использующих искусственные нейронные сети, необходимо учитывать количество обучающих данных и их распределение, в частности, при прогнозировании пола.

Цель исследования — определить потенциальную эффективность метода синтетической передискретизации меньшинства (synthetic minority oversampling technique, SMOTE) при определении пола умерших с помощью искусственной нейронной сети.

Материалы и методы. В данном исследовании использовали набор данных, состоящий из 297 цефалометрических измерений индонезийских пациентов (229 женщин и 68 мужчин). Для измерения определённых параметров, таких как угол SNA (Sella-Nation-Point A), длина нижней челюсти, угол нижней челюсти, угол SGA (Sella-Glabella-Point A), и диагностики использовали программу WebCeph. Обработку данных и создание искусственной нейронной сети выполняли на языке программирования Python.

Результаты. Точность определения пола с помощью искусственной нейронной сети составляет 87% для женщин и 0% для мужчин (в среднем 78%). При использовании SMOTE-алгоритма точность определения пола составляет 22% (0% для женщин, 37% для мужчин). Однако при использовании SMOTE-алгоритма в сочетании с нормализацией данных точность возрастает до 71% (82% для женщин, 30% для мужчин). Точность модели при нормализации данных без применения SMOTE составляет 76% (86% для женщин, 14% для мужчин).

Заключение. Данное исследование доказало эффективность SMOTE в улучшении классификации мужских матриц. Тем не менее результаты общей точности недостаточно оптимальны по сравнению с результатами, полученными без применения метода SMOTE и нормализации данных. Для достижения оптимальной точности в определении пола при использовании искусственных нейронных сетей и других параметров необходимо применение стратегий балансировки данных.

Ключевые слова: искусственные нейронные сети; цефалометрия; определение пола; SMOTE.

Как цитировать:
Handayani V.W., Yudianto A., Mieke Sylvia MAR., Riries R., Caesarardhi M.R., Putra R. Потенциал метода синтетической передискретизации меньшинства для повышения точности определения пола: исследование искусственных нейронных сетей с помощью цефалометрии // *Судебная медицина.* 2024. Т. 10, № 2. С. 139–151. DOI: https://doi.org/10.17816/fm16110

Рукопись получена: 09.01.2024 **Рукопись одобрена:** 15.04.2024 **Опубликована online:** 07.06.2024

DOI: https://doi.org/10.17816/fm16110

合成少数超采样技术提高性别鉴定准确性的潜力：利用头颅测量学对人工神经网络的研究

Vitria Wuri Handayani[1, 2], Ahmad Yudianto[3, 4], Mieke Sylvia M.A.R.[5], Riries Rulaningtyas[6, 7], Muhammad Rasyad Caesarardhi[8], Ramadhan Hardani Putra[9]

[1] Medical Faculty, Universitas Airlangga, Surabaya, Indonesia;
[2] Nursing Department, Pontianak Polytechnic Health Ministry, Pontianak, Indonesia;
[3] Department of Forensics and Medicolegal, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia;
[4] Magister of Forensic Sciences, Postgraduate School, Universitas Airlangga, Surabaya,Indonesia;
[5] Forensic Odontology Department, Dental Medical Faculty, Univesitas Airlangga, Surabaya, Indonesia;
[6] Physics Department, Sains and Technology Faculty, Universitas Airlangga, Surabaya, Indonesia;
[7] Biomedical Department, Sains and Technology Faculty, Universitas Airlangga, Surabaya,Indonesia;
[8] Department of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia;
[9] Department of Dentomaxillofacial Radiology, Faculty of Dental Medicine, Universitas Airlangga, Surabaya, Indonesia

摘要

论证。在使用人工神经网络创建模型时，有必要考虑训练数据的数量及其分布，尤其是在预测性别时。

研究目的是确定合成少数超采样技术（synthetic minority oversampling technique, SMOTE）在使用人工神经网络确定死者性别方面的潜在有效性。

材料和方法。本研究使用的数据集包括对印度尼西亚患者（229 名女性和 68 名男性）进行的297次头颅测量。WebCeph 软件用于测量某些参数，如 SNA 角（Sella-Nation-Point A）、下颌长度、下颌角、SGA 角（Sella-Glabella-Point A）和诊断。数据处理和人工神经网络的创建使用 Python 编程语言进行。

结果。使用人工神经网络进行性别鉴定的准确率为：女性 87%，男性 0%（平均 78%）。当使用 SMOTE 算法时，性别确定的准确率为 22%（女性为 0%，男性为 37%）。然而，当 SMOTE 算法与数据归一化结合使用时，准确率提高到 71%（女性为 82%，男性为 30%）。在不使用 SMOTE 算法的情况下，使用数据归一化的模型准确率为 76%（女性为 86%，男性为 14%）。

结论。这项研究证明了 SMOTE 在改进男性矩阵分类方面的有效性。然而，与不使用 SMOTE 和数据归一化的结果相比，总体准确度结果还不够理想。为了在使用人工神经网络和其他参数时实现性别确定的最佳精度，需要应用数据平衡策略。

关键词：人工神经网络；头颅测量；性别鉴定；SMOTE。

引用本文：
Handayani VW, Yudianto A, Mieke Sylvia MAR, Riries R, Rasyad M, Putra R. 合成少数超采样技术提高性别鉴定准确性的潜力：利用头颅测量学对人工神经网络的研究. *Russian Journal of Forensic Medicine.* 2024;10(2):139–151. DOI: https://doi.org/10.17816/fm16110

收到：09.01.2024 **接受：**15.04.2024 **发布日期：**07.06.2024

BACKGROUND

Radiography is an essential component of forensic odontology because it is a straightforward, cost-effective, and noninvasive method of identification, particularly for identifying corpses by comparing premortem and postmortem radiographs [1]. Dental radiographs can provide reliable information, such as comparison data of the anatomical shape of teeth, periapical anomalies, fillings, cavities, outlines, and positions of impacted teeth, among others [1]. Broadbent invented lateral cephalometry in 1931, which has the benefit of gaining a comprehensive visual representation of the cranial structure and soft tissue contour [2]. In addition, lateral cephalometry permits the evaluation of several anatomical components, such as the nasal bones, frontal sinuses, sphenoid sinuses, and other images that assist in the gender identification process [3-5].

Patil et al, used artificial neural networks (ANN) to predict gender using mandibular radiograph [6]. ANN is a type of computing system that utilizes machine learning (ML) to generate intricate information (7). An ANN is a computational model that draws inspiration from the biological nervous system of the brain [7]. The ANN can obtain and retain knowledge (information-based) and can be characterized as a collection of processing units represented by artificial neurons, interconnected by numerous interconnections (artificial synapses), and implemented by vectors and matrices of synaptic weights [8,9]. The main advantage ANN is its capacity to acquire knowledge and maintain stability even when confronted with minor errors by generating information through the learning process [7,8,10].

A previous study conducted by Handayani et al demonstrated how well both of these models predict the gender of adult Indonesians by utilizing 297 lateral cephalometry ANN models with 80% of the samples for training, 10% of the samples for testing, and 10% of the samples for validating. Although an overall accuracy of 80% is achieved, data imbalance reduces the accuracy of predicting the male gender. Data imbalance occurs when the number of people in each class is unequal. In other words, one class is not well represented (minority class), whereas the other class has more cases (majority class). In many real-life situations, the class mismatch problem occurs [11]. The most well-known method to deal with uneven data is the synthetic minority oversampling technique (SMOTE). SMOTE creates new fake data patterns by combining samples from the minority class with their K closest neighbors in a straight line [11]. This study uses an ANN model to predict gender and examines the possible effects of applying SMOTE.

AIM

This study seeks to determine the potential impact of using SMOTE on gender prediction using the ANN model.

MATERIALS AND METHOD

A total of 297 cephalometry images were taken from the medical records of patients at the Airlangga University Dental and Mouth Hospital (RSGMP Unair) in Surabaya, part of this data is taken from a previous study [12]. Subsequently, the sample will be partitioned into three subsets, i.e., 70% of the cephalometry images will be allocated for training purposes, 15% will be reserved for validation tests, and the remaining 15% will be utilized for testing. Training, development, and analysis of the data will be conducted by cephalometric algorithms using ANN in Python. The following inclusion criteria will be employed:

1) Cephalometric images are obtained from preexisting cephalometric photographs at the RSGMP FKG Airlangga University in Surabaya.

2) Cephalometric pictures are acquired from individuals of Indonesian descent.

3) Cephalometric photographs are captured using standardized instruments and identical equipment.

4) Cephalometric photographs are in a satisfactory state, with no evidence of superimposition.

5) Cephalometric photographs exhibit a satisfactory state of preservation, devoid of any discernible distortions.

6) Cephalometric photographs are captured by operators possessing a minimum of D3 radiology education, together with a minimum of 1 year of practical experience operating cephalometric equipment.

7) Cephalometric images of individuals between the ages of 18 and 40 years. The cephalometric photograph exhibits a comprehensive dentition in both the mandibular and maxillary regions, except for the third molar.

8) Cephalometric photographs of individuals who lack either orthodontic intervention or prior records of orthodontic intervention are taken.

9) Cephalometric photographs are obtained from individuals lacking prior records of orthognathic surgery.

10) Cephalometric photographs of individuals who do not possess any documented instances of jaw injuries are taken.

This study employs the following instruments:

1) The cephalometric photographs were captured using equipment that adheres to defined protocols.

2) The cephalometric device employed in this study is the ZULASSUNG THA/HV-GEN Type THA100.

3) Operators of cephalometric equipment possess standardized skills.

4) A computer system equipped with a minimum of two 8-GB random access memory modules and a solid-state drive with a storage capacity of 1 TB.

5) The computer is equipped with the NVIDIA GeForce RTX 3060 graphics processing unit.

6) The utilization of Python in web development and Google Collaboration.

7) The utilization of Pandas.

8) The utilization of the SHapley Additive exPlanation (SHAP).

9) The utilization of WebCeph application.

Variables and Definition

The following variables are used in this study:

I. Dependent variables: The precision of cephalometric analysis utilizing the ANN techniques, considering the variables listed in Table 1.

II. Independent variable: Gender-based variations in cephalometric data employing the artificial intelligence computational analysis method, ANN.

III. Control variables: Age, cephalometric instrument, operator skill, Indonesian people.

RESULTS

Cephalometric radiographs were obtained from patients who sought treatment at Airlangga University Dental Hospital in Surabaya, Indonesia, as part of this research endeavor. This study has passed the ethics review with ethical number 316/HRECC.FODM/III/2023 conducted by Airlangga University Faculty of Dental Medicine Health Research Ethical Clearance Commission. We choose a suitable shape for people between the ages of 18 and 40 years. The collection comprises 297 cephalometric photographs, including 229 images in the female category and 68 images in the male category. The observed disparity in data between males and females can be attributed to the predominant representation of female patients seeking orthodontic therapy, as indicated by the majority of cephalometry data collected. The Python programming language is utilized to partition the photos into three distinct segments, i.e., 80% for training, 15% for validation, and 15% for testing, as depicted in Table 2.

A. Measurement of the variables

We employed the backpropagation method in this investigation. In this methodology, two computational processes, i.e., sophisticated computations to ascertain the disparity between the output of the ANN and the intended goal, are executed. The next step entails a reverse calculation that utilizes the obtained errors to modify the weights of all of the neurons in the system.

The ANN cephalometric parameters were measured using the end web application, which is a medical record application for orthodontic care. This application is utilized because of its cost-free nature and its capability to reduce errors associated with manual procedures. The visual representation of the measurement obtained using the front-end web application is depicted in Fig. 1.

After obtaining the value of the parameters searched using the front-end web application, the results of the four parameters were analyzed using Pandas in Python, and the ANN was trained based on the diagnosis of malocclusion of the patient as listed in Table 3. Notably, the sample ratio is in Class I malocclusion angle with normal occlusion.

The average values of the four parameters were determined by analyzing them using Pandas in Python. The parameters obtained from the cephalometric photographs were used as a gender reference in the ANN algorithm. The results are shown in Table 4.

B. Network training and validation

Python is used to develop the ANN architecture. The first step involves training and validating the model before conducting the tests. This study employs four distinct scenarios. The first scenario involves the absence of SMOTE and normalization. The second scenario entails the presence of

Table 1. Definition of the operational variables

Parameter	Definition
Age	The length of life of the respondent is seen from the age shown in the cephalometric photo. The ratio is 18–40 years
Cephalometric photo	The cephalometric photo was taken at RSGMP Unair by an operator with the same level of skill-using a standard device. There was no superimposition, and the teeth were complete, except for the third molar. The photo was taken from a patient who has never undergone jaw and maxilla surgery, or orthodontic treatment, or has never experienced any jaw and maxilla trauma
Instrument	The standardized tool used to capture cephalometric photos at RSGMP Unair is the ASAHI Hyper-XCM Panoramic Cephalometric Type D.052 SB
Operator skill	The required qualifications for the operator position include a minimum of a D3 degree in Radiography and at least 1 year of experience operating cephalometric radiology equipment
Indonesian people	The cephalometric photo was taken from a Mongoloid race, which was generalized from cephalometric photo results originating from Indonesian citizens. Photo selection was performed by an operator from RSGMP Unair
SNA angle	Angles formed from the points SellaNasion–A (Wyle)
SGA angle	Angles formed from the points Sella_Glabella_A on the cephalometric photographs (Johnson)
Mandibular length	A lining of the lower mandibular boundary that is projected to the most posterior point from the head of the condyle to the more anterior point of the jaw (Jarabak)
Mandibular angle	Angles formed from the Gonion_Gnathion line (Go_Gn) together with the base of the anterior skull (S_N; Steiner)
Diagnosis	The differential diagnosis of the skeletal pattern and the dental classification of malocclusion

Table 2. Distribution of sample frequency based on gender

	Training (70%)	Validation (15%)	Testing (15%)	Total sample
Female	153	38	38	229
Male	46	11	11	68
Total sample	199	49	49	297

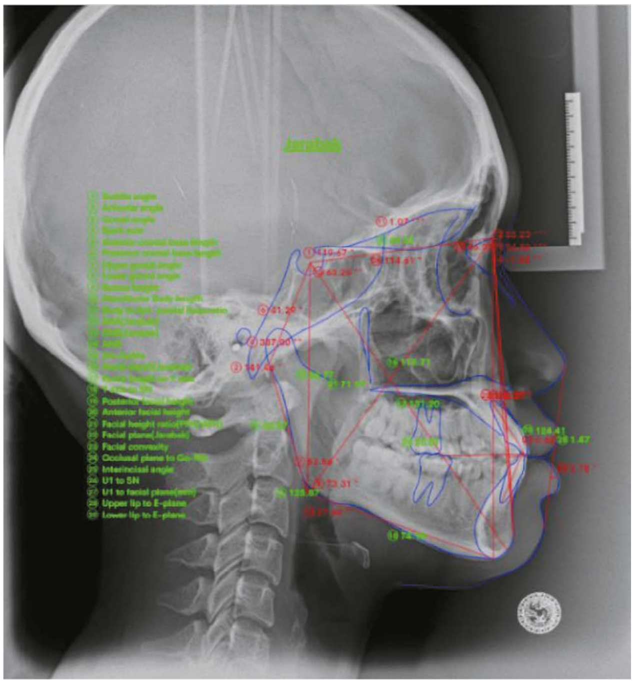

Fig. 1. Analysis of cephalometric parameters using the front-end web application [13].

SMOTE without normalization. The third scenario involves the combination of SMOTE and normalization. The fourth scenario involves the presence of normalization without SMOTE. The training and validation methods of photocephalometry are exemplified in Fig. 2 and 3. The observation of a decreasing loss graph in both the training and validation stages provides additional support for this claim.

C. ANN model analysis

The results of accuracy, precision, recall, f1 score, and support of the ANN model under four different scenarios are presented in Table 5. The first scenario involves developing the model without incorporating SMOTE and normalization techniques. The accuracy outcome for the first scenario is 0.78, or 78%, with a male f1 score of 0.00 and a female f1 score of 0.87. The second scenario of the ANN model involves the use of SMOTE without normalization. The accuracy score obtained is 0.22, which corresponds to 22%. The male f1 score is 0.00, whereas the female f1 score is 0.37. The third ANN model scenario involves the use of SMOTE and normalization. The accuracy result is 0.71, which corresponds to a 71% accuracy rate. The male f1 score is 0.30, whereas the female f1 score is 0.82. In the final scenario, the absence of SMOTE with normalization leads to an accuracy result of 0.76, or 76%. The male f1 score is 0.14, whereas the female f1 score is 0.86.

A matrix classification provides a concise representation of the classification accuracy of a classifier in relation to a given set of test data. A two-dimensional matrix is created, with one dimension containing the true class of an object and the other dimension containing the class assigned by the classifier. This study employed a two-class design, with one class representing females and the other class representing males, which were classified as the positive class and the negative class, respectively. Within this framework, the four cells of the matrix are classified as true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), as specified in Table 6.

In this study, the ANN model utilizes four cephalometric factors as the foundation for gender prediction. However, certain parameters may undermine the accuracy of the prediction. The SHAP model was employed in this work to determine the most relevant parameters in predicting gender. Fig. 4 illustrates that the mandibular angle had the strongest positive impact on gender prediction, whereas the SGA angle had the most negative impact on gender prediction.

Table 3. Distribution of sample frequency based on diagnosis

Diagnosis	Gender		Total sample
	Female	Male	
Class I malocclusion angle with normal occlusion	45	17	62
Class I malocclusion angle with maxilla protrusion	40	3	43
Class I malocclusion angle with open bite	0	1	1
Class II malocclusion angle with normal occlusion	57	23	80
Class II malocclusion angle with maxilla protrusion	64	16	80
Class III malocclusion angle with normal occlusion	11	7	18
Class III malocclusion angle with maxilla protrusion	10	1	11
Class III malocclusion angle with maxillary retrusion	2	0	2
Total sample	229	68	297

DOI: https://doi.org/10.17816/fm16110

Table 4. Average values of the cephalometric parameters based on gender and diagnosis

Diagnosis	Gender	Mean of the parameter			
		SNA angle	Mandibular length (mm)	Mandibular angle	SGA angle
Class I malocclusion angle with normal occlusion	Male	84.78	83.46	30.45	72.81
	Female	82.75	78.32	32.5	72.46
Class I malocclusion angle with maxilla protrusion	Male	88.7	83.18	28.43	74.4
	Female	87.79	81.11	30.69	75.02
Class I malocclusion angle with open bite	Male	80.47	72.73	32.95	70.5
	Female	–	–	–	–
Class II malocclusion angle with normal occlusion	Male	84.24	80.53	31.47	72.71
	Female	82.51	75.74	37	71.45
Class II malocclusion angle with maxilla protrusion	Male	89.47	78.94	26.90	76.04
	Female	87.52	79.23	31.91	75.76
Class III malocclusion angle with normal occlusion	Male	84.29	57.85	49.573	71.1
	Female	82.3	79.12	30.38	68.98
Class III malocclusion angle with maxilla protrusion	Male	88.96	81.22	20.52	64.9
	Female	86.94	84.98	25.49	75.01
Class III malocclusion angle with maxillary retrusion	Male	–	–	–	–
	Female	74.37	73	35.35	71.35

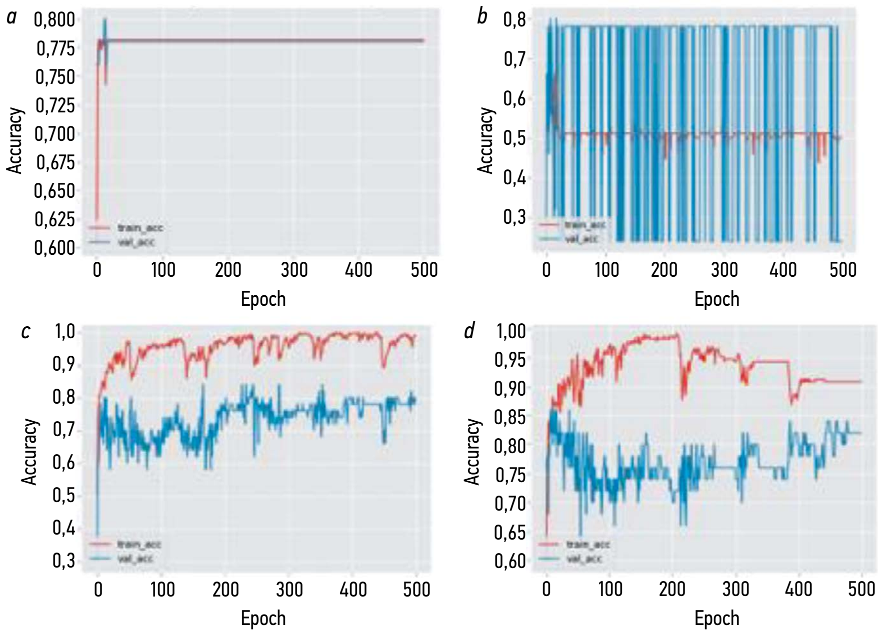

Fig. 2. Graph of the success of the training and validation of ANN: *a* — training and validation accuracy without SMOTE and normalization; *b* — training and validation accuracy with SMOTE and without normalization; *c* — training and validation accuracy with SMOTE and normalization; *d* — training and validation accuracy with normalization and without SMOTE.

DISCUSSION

The creation of a machine or autonomous mechanism with intelligence has long been a cherished aspiration of researchers across several scientific and technical disciplines, and ANN can be utilized to address various technical and scientific difficulties [7]. Although established more than 50 years ago, ANN has garnered substantial interest and research in the early 1990s and still shows great promise for further exploration [9]. The applications of intelligent systems cover a wide range of areas, including the field of forensic odontology [9]. This study utilizes a combined approach of parameter measurement using deep learning, e.g., WebCeph, and ANN to conduct gender determination.

The accurate interpretation of the results of an ANN prediction model is crucial. This interpretation provides

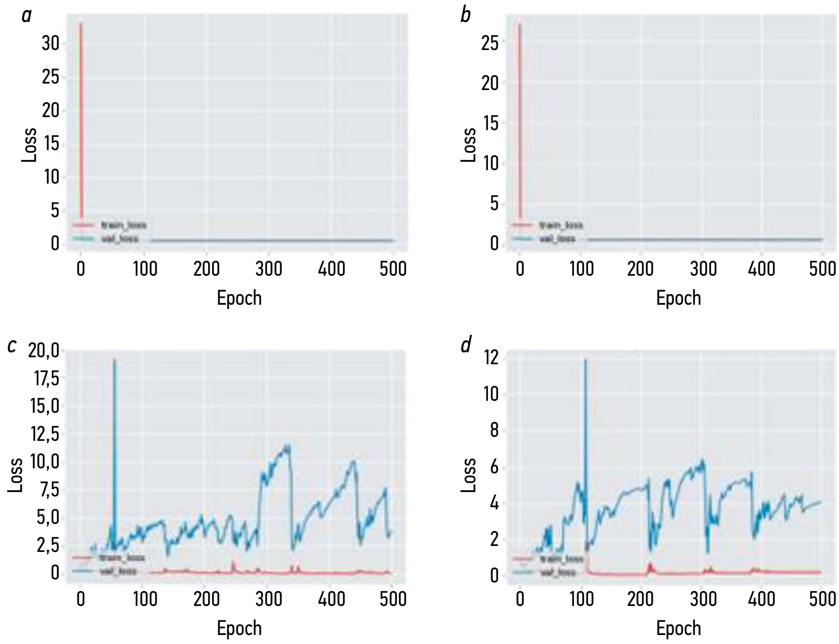

Fig. 3. Graph of the failure of the training and validation of ANN: *a* — training and validation loss without SMOTE and normalization; *b* — training and validation loss with SMOTE and without normalization; *c* — training and validation loss with SMOTE and normalization; *d* — training and validation loss with normalization and without SMOTE.

Table 5. Accuracy, precision, recall, f1 score, and support of the ANN model

Scenario		Precision	Recall	f1 score	Support
Without SMOTE and normalization	Female	0.78	1.00	0.87	38
	Male	0.00	0.00	0.00	11
	Accuracy	0.78	49		
	Macro average	0.39	0.50	0.44	49
	Weighted average	0.60	0.78	0.68	49
With SMOTE and without normalization	Female	0.00	0.00	0.37	39
	Male	0.22	1.00	0.00	11
	Accuracy			0.22	49
	Macro average	0.11	0.50	0.18	49
	Weighted average	0.05	0.22	0.08	49
With SMOTE and normalization	Female	0.80	0.84	0.82	38
	Male	0.33	0.27	0.30	11
	Accuracy	0.71	49		
	Macro average	0.57	0.56	0.56	49
	Weighted average	0.70	0.71	0.70	49
With normalization and without SMOTE	Female	0.78	0.95	0.86	38
	Male	0.33	0.09	0.14	11
	Accuracy			0.76	49
	Macro average	0.90	0.57	0.57	49
	Weighted average	0.88	0.80	0.74	49

dependable user confidence, valuable insights for model enhancement, and an improved understanding of the modeling process [14]. However, the effectiveness of the ANN technique is dependent on a significant volume of acquired data, as demonstrated in previous research conducted by Vathsala Patil et al., Tanvi et al., Deeepthi Bharadwaj et al., Naveen et al., and Usha Jambunath et al. The investigations conducted had a sample size of over 500 radiographs [6, 15–17]. By assessing the precision rate, the ANN demonstrated effectiveness and accuracy in detecting gender, regardless

Table 6. Matrix classification results of males and females using ANN

	Without SMOTE and normalization	
n=49	Predicted: YES	Predicted: NO
Actual: NO	TP=38	TN=0
Actual: YES	FP = 11	FN=0
	38	11
	With SMOTE and without normalization	
n=49	Predicted: YES	Predicted: NO
Actual: NO	TP=0	TN=38
Actual: YES	FP = 0	FN=11
	38	11
	With SMOTE and normalization	
n=49	Predicted: YES	Predicted: NO
Actual: NO	TP=32	TN=6
Actual: YES	FP = 8	FN=3
	38	11
	With normalization and without SMOTE	
n=49	Predicted: YES	Predicted: NO
Actual: NO	TP=36	TN=2
Actual: YES	FP = 10	FN=1
	38	11

Note. TP — true positives; TN — true negatives; FP — false positives; FN — false negatives.

Fig. 4. SHapley Additive exPlanation values attribute to each feature the change in the expected model prediction when conditioning on that feature.

of the specific diagnosis of the sample used. These systems can forecast future values of a certain process by considering several past samples within its domain [9]. As a result, ANN shows promise in the field of forensic odontology, particularly in the prediction identification of gender [18].

Lateral cephalometry has long been an essential tool for diagnosing orthodontic issues and creating treatment programs; it is now also applicable in the field of forensic odontology [18,19]. In a prior investigation, Bao discovered that automatic analysis software conducts cephalometric measurements with a level of effectiveness that is nearly sufficient for clinical application [19].

A sizable dataset is required to create a ML model [20]. Handayani et al. determined that collecting data on cephalometry presents a greater challenge than panoramic radiography, as RSGMP Unair exclusively uses cephalometric radiography for patients undergoing orthodontic therapy [13]. In addition, the statistics between males and females exhibit a significant discrepancy, which can be attributed to the fact that women are more aware of dental esthetics than men. Handayani et al. also discovered data imbalances that prevent the accuracy of the results from exceeding 90%. The composition of the ANN model is 80% for training, 10% for validation, and 10% for testing, with precision values for males and females being 25% and 88%, respectively, in that study [13].

Table 5 summarizes the four scenarios of the ANN model. The first scenario is the ANN model without SMOTE and normalization, which achieves a precision rate of 78% in predicting gender. The second scenario is the ANN model with SMOTE and without normalization, which leads to a decrease in accuracy by 22%. The third scenario is the ANN model with SMOTE and normalization, which achieves a better accuracy result of 71%. The fourth scenario is the ANN model with normalization and without SMOTE, which yields a better result, of 76%. The discrepancy in these precision values can be attributed to the variations in the distribution of the number of samples for male and female cephalometry. In cases where the sample numbers are imbalanced, the accuracy value tends to favor larger sample sizes [21]. This finding is corroborated by the metrics of accuracy, precision, and recall obtained from ANN [22].

SMOTE is a resampling strategy that seeks to augment the number of samples in the minority class by generating synthetic samples in that class. SMOTE is utilized to stabilize datasets with a significantly imbalanced ratio. Upon analyzing the results of this study, the matrix classification presented in Table 6 shows that SMOTE can enhance the accuracy of male predictions but fall short in predicting females. Overall, the absence of normalization in SMOTE does not contribute to the improvement in accuracy and reduces the accuracy of the model. The findings of Elreedy and Atiya and Duan et al. indicated that the implementation of SMOTE can enhance the accuracy of samples in the minority class [23,24]. However, when normalization is applied, SMOTE exhibits improved performance, which can be attributed to its capability to mitigate the problem of overfitting. In addition, the procedure for generating fresh synthetic samples differs from that of the multiplication method.

Cephalometric radiographs are examined to assess the SNA angles, SGA angles, and mandibular length parameters. We utilize the SHAP method, which is a Python-based metric, to measure the level of integration of the cephalometric components used in the ANN and determine their importance. The SHAP approach is utilized to aid ANN models in acquiring precise parameters to identify a novel category of parameter measurements [25]. Moreover, the SHAP approach aids in determining whether there exists a singular answer in this category that possesses a collection of desirable attributes. Hence, using the SHAP methodology in our research, we have ascertained that the mandibular angle derived from cephalometric radiographs emerges as the most dependable parameter for assessing gender prediction in the Indonesian population. This research pertains to the studies conducted by Sikka et al. and Patil et al, which showed significant variations in the mandible between males and females. This research shows that the selected characteristics, particularly the SGA angle, pose a problem in accurately predicting gender [6,26].

ANNs possess a notable advantage in that they do not require any prior familiarity with system models [27]. This attribute is highly advantageous when used for the processing of files that exhibit missing or corrupted data. However, because of the exclusive operation of ANNs on the tasks for which they have been trained, a substantial amount of data needs to be used. To accomplish any objective, the data need to be retrained, and the augmentation of the image count within the cephalometry picture collection has the potential to exhibit improved performance [27, 28]. This study is hampered by the shortage of male samples in comparison to female samples, as well as the limited parameters employed, which can negatively impact the accuracy of gender prediction and do not measure the density of the cranium.

CONCLUSION

This study proposes the use of SMOTE as a solution to address the issue of imbalanced data in an ANN model used for gender identification. The efficacy of SMOTE in augmenting male samples has been demonstrated, although this study did not observe any improvement in accuracy. Among the four selected characteristics, the mandibular angle has a significant influence on gender prediction. To enhance performance, future research could involve expanding the dataset and conducting a thorough analysis of gender sample distribution. Furthermore, researchers should evaluate additional characteristics and parameters, such as skull density, that can be utilized in the study of unidentified cranium and mass disasters.

ADDITIONAL INFORMATION

Funding source. This research was funded by Indonesian Health Ministry.

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. V.W. Handayani — data collecting, editing the article; A. Yudianto — supervisor; M. Sylvia M.A.R. — cephalometric manual data measures; R. Rulaningtyas — algorithm development; M.R. Caesarardhi — algortithm revised; R.H. Putra — image data filtering.

Acknowledgements. We express our gratitude to the Indonesian Health Ministry for providing financing for this project, as well as to Airlangga University for their cooperation in conducting this study.

REFERENCES

1. Tahir H. Book of abstracts: The 4th Indonesia International Symposium of Forensic Odontology "Incorporating recent advances and new technologies for delivering good evidence in forensic odontology". Amerta Media; 2023. P. 45.

2. Subramanian AK, Chen Y, Almalki A, et al. Cephalometric analysis in orthodontics using artificial intelligence: A comprehensive review. *Biomed Res Int.* 2022;2022:1880113. EDN: VCAUJB doi: 10.1155/2022/1880113

3. Ruth MS. *Sefalometri radiografi dasar.* Surabaya: Sagung Seto; 2013.
4. Indra Sukmana B, Rijaldi F. Buku Ajar Kedokteran Gigi Forensik [Internet]. Vol. VI. 2022. P. 1–79. Available from: https://idndentist.com/article/93
5. Taner L, Gursoy G, Uzuner F. Does gender have an effect on craniofacial measurements? *Turkish J Orthod.* 2019;32(2):59–64. doi: 10.5152/TurkJOrthod.2019.18031
6. Patil V, Vineetha R, Vatsa S, et al. Artificial neural network for gender determination using mandibular morphometric parameters: A comparative retrospective study. *Cogent Eng.* 2020;7(1):1–12. doi: 10.1080/23311916.2020.1723783
7. Chen M, Chalita U, Saad W, et al. Artificial neural networks-based machine learning for wireless networks: A tutorial. *IEEE Commun Surv Tutorials.* 2019;21(4):3039–3071. doi: 10.1109/COMST.2019.2926625
8. Dastres R, Soori M. Artificial neural network systems. *Int J Imaging Robot.* 2021;2021(2):13–25.
9. Da Silva IN, Hernane SD, Andrade FR, et al. *Artificial neural networks: A practical course.* Springer Nature: Switzerland; 2017. 252 p.
10. Wu Y, Feng J. Development and application of artificial neural network. *Wirel Pers Commun.* 2018;102(5):1645–1656. doi: 10.1007/s11277-017-5224-x
11. Elreedy D, Atiya AF, Kamalov F. A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Mach Learn.* 2023;113(7):4903–4923. EDN: KFIOLR doi: 10.1007/s10994-022-06296-4
12. Handayani V, Yudianto A, Mieke Sylvia MAR, Rulaningtyas R. Classification of Indonesian adult forensic gender using cephalometric radiography with VGG16 and VGG19: A preliminary research. *Acta Odontol Scand.* 2024;(83):308–316. doi: 10.2340/aos.v83.40476
13. Handayani VW. *Cephalometry radiology based on rrtificial intelligence model for predict gender determination in unidentified cranium.* Universitas Airlangga; 2024.
14. Hapsari RK, Miswanto M, Rulaningtyas R, et al. Modified gray-level haralick texture features for early detection of diabetes mellitus and high cholesterol with iris image. *Int J Biomed Imaging.* 2022;2022:5336373. EDN: BVEODY doi: 10.1155/2022/5336373
15. Satish BN, Moolrajani C, Basnaker M, Kumar P. Dental sex dimorphism: Using odontometrics and digital jaw radiography. *J Forensic Dent Sci.* 2017;9(1):43. doi: 10.4103/jfo.jfds_78_15
16. Arab MA, Khankeh HR, Mosadeghrad AM, Farrokhi M. Developing a hospital disaster risk management evaluation model. *Risk Manag Healthc Policy.* 2019;(12):287–296. doi: 10.2147/RMHP.S215444
17. Vahanwala S. Assessment of the effect of dimensions of the mandibular ramus and mental foramen on age and gender using digital panoramic radiographs: A retrospective study. *Contemp Clin Dent.* 2019;9(3):343–348. doi: 10.4103/ccd.ccd_26_18
18. Tahir H. Book of abstracts: The 4th Indonesia International Symposium of Forensic Odontology "Incorporating recent advances and new technologies for delivering good evidence in forensic odontology". Amerta Media; 2023. P. 36–37.
19. Bao H, Zhang K, Yu C, et al. Evaluating the accuracy of automated cephalometric analysis based on artificial intelligence. *BMC Oral Health.* 2023;23(1):191. doi: 10.1186/s12903-023-02881-8
20. Ramezanzade S, Laurentiu T, Bakhshandah A, et al. The efficiency of artificial intelligence methods for finding radiographic features in different endodontic treatments: A systematic review. *Acta Odontol Scand.* 2022;81(6):422–435. doi: 10.1080/00016357.2022.2158929
21. Shung KP. *Accuracy, precision, recall or F1?* [Internet]. Towards Data Science. 2018 [cited 2023 Sep 2]. Available from: https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9. Accessed: 15.04.2024.
22. Jeong SH, Yun JP, Yeom HG, et al. Deep learning based discrimination of soft tissue profiles requiring orthognathic surgery by facial photographs. *Sci Rep.* 2020;10(1):16235. doi: 10.1038/s41598-020-73287-7
23. Elreedy D, Atiya AF. A comprehensive analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance. *Inf Sci (Ny).* 2019;(505):32–64. doi: 10.1016/j.ins.2019.07.070
24. Duan F, Zhang S, Yan Y, Cai Z. An oversampling method of unbalanced data for mechanical fault diagnosis based on meanradius-SMOTE. *Sensors.* 2022;22(14):5166. EDN: PURYZO doi: 10.3390/s22145166
25. Zhang K, Zhang Y, Wang M. A unified approach to interpreting model predictions scott. *Nips.* 2012;16(3):426–430. doi: 10.48550/arXiv.1705.07874
26. Sikka A, Jain A. Sex determination of mandible: A morphological and morphometric analysis. *Int J Contemp Med Res.* 2016;3(7):1869–1872.
27. Smart Energy and Electric Power Systems [Internet]. Pavithra V, Jayalakshmi V. Smart energy and electric power system: Current trends and new intelligent perspectives and introduction to ai and power system. Elsevier; 2023. P. 19–36. doi: 10.1016/B978-0-323-91664-6.00001-2
28. Fan DP, Zhang J, Xu G, et al. Salient objects in clutter. *IEEE Trans Pattern Anal Mach Intell.* 2023;45(2):2344–2366. doi: 10.1109/TPAMI.2022.3166451

СПИСОК ЛИТЕРАТУРЫ

1. Tahir H. Book of abstracts: The 4th Indonesia International Symposium of Forensic Odontology "Incorporating recent advances and new technologies for delivering good evidence in forensic odontology". Amerta Media, 2023. P. 45.
2. Subramanian A.K., Chen Y., Almalki A., et al. Cephalometric analysis in orthodontics using artificial intelligence: A comprehensive review // Biomed Res Int. 2022. Vol. 2022. P. 1880113. EDN: VCAUJB doi: 10.1155/2022/1880113
3. Ruth M.S. Sefalometri radiografi dasar. Surabaya: Sagung Seto, 2013.
4. Indra Sukmana B, Rijaldi F. Buku ajar kedokteran gigi forensik [Internet]. 2022. Vol. VI. P. 1–79. Available from: https://idndentist.com/article/93
5. Taner L., Gursoy G., Uzuner F. Does gender have an effect on craniofacial measurements? // Turkish J Orthod. 2019. Vol. 32, N 2. P. 59–64. doi: 10.5152/TurkJOrthod.2019.18031
6. Patil V., Vineetha R., Vatsa S., et al. Artificial neural network for gender determination using mandibular morphometric parameters: A comparative retrospective study // Cogent Eng. 2020. Vol. 7, N 1. P. 1–12. doi: 10.1080/23311916.2020.1723783
7. Chen M., Chalita U., Saad W., et al. Artificial neural networks-based machine learning for wireless networks: A tutorial. // IEEE Commun Surv Tutorials. 2019. Vol. 21, N 4. P. 3039–3071. doi: 10.1109/COMST.2019.2926625
8. Dastres R., Soori M. Artificial neural network systems // Int J Imaging Robot. 2021. Vol. 21, N 2. P. 13–25.
9. Da Silva I.N., Hernane S.D., Andrade F.R., et al. Artificial neural networks: A practical course. Springer Nature: Switzerland, 2017. 252 p.
10. Wu Y., Feng J. Development and application of artificial neural network // Wirel Pers Commun. 2018. Vol. 102, N 5. P. 1645–1656. doi: 10.1007/s11277-017-5224-x

11. Elreedy D., Atiya A.F., Kamalov F. A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning // Mach Learn. 2023. Vol. 113, N 7. P. 4903–4923. EDN: KFIOLR doi: 10.1007/s10994-022-06296-4

12. Handayani W.V., Yudianto A., Mieke Sylvia MAR., Rulaningtyas R. Classification of Indonesian adult forensic gender using cephalometric radiography with VGG16 and VGG19: A preliminary research // Acta Odontol Scand. 2024. N 83. P. 308–316. doi: 10.2340/aos.v83.40476

13. Handayani V.W. Cephalometry radiology based on rrtificial intelligence model for predict gender determination in unidentified cranium. Universitas Airlangga, 2024.

14. Hapsari R.K., Miswanto M., Rulaningtyas R., et al. Modified gray-level haralick texture features for early detection of diabetes mellitus and high cholesterol with iris image // Int J Biomed Imaging. 2022. Vol. 2022. P. 5336373. EDN: BVEODY doi: 10.1155/2022/5336373

15. Satish B.N., Moolrajani C., Basnaker M., Kumar P. Dental sex dimorphism: Using odontometrics and digital jaw radiography // J Forensic Dent Sci. 2017. Vol. 9, N 1. P. 43. doi: 10.4103/jfo.jfds_78_15

16. Arab M.A., Khankeh H.R., Mosadeghrad A.M., Farrokhi M. Developing a hospital disaster risk management evaluation model // Risk Manag Healthc Policy. 2019. N 12. P. 287–296. doi: 10.2147/RMHP.S215444

17. Vahanwala S. Assessment of the effect of dimensions of the mandibular ramus and mental foramen on age and gender using digital panoramic radiographs: A retrospective study // Contemp Clin Dent. 2019. Vol. 9, N 3. P. 343–348. doi: 10.4103/ccd.ccd_26_18

18. Tahir H. Book of abstracts: The 4th Indonesia International Symposium of Forensic Odontology "Incorporating recent advances and new technologies for delivering good evidence in forensic odontology". Amerta Media, 2023. P. 36–37.

19. Bao H., Zhang K., Yu C., et al. Evaluating the accuracy of automated cephalometric analysis based on artificial intelligence // BMC Oral Health. 2023. Vol. 23, N 1. P. 191. doi: 10.1186/s12903-023-02881-8

20. Ramezanzade S., Laurentiu T., Bakhshandah A., et al. The efficiency of artificial intelligence methods for finding radiographic features in different endodontic treatments: A systematic review // Acta Odontol Scand. 2022. Vol. 81, N 6. P. 422–435. doi: 10.1080/00016357.2022.2158929

21. Shung K.P. Accuracy, precision, recall or F1? [Internet]. Towards Data Science. 2018 [2023 Sep 2]. Режим доступа: https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9. Дата обращения: 15.04.2024.

22. Jeong S.H., Yun J.P., Yeom H.G., et al. Deep learning based discrimination of soft tissue profiles requiring orthognathic surgery by facial photographs // Sci Rep. 2020. Vol. 10, N 1. P. 16235. doi: 10.1038/s41598-020-73287-7

23. Elreedy D., Atiya A.F. A comprehensive analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance // Inf Sci (Ny). 2019. N 505. P. 32–64. doi: 10.1016/j.ins.2019.07.070

24. Duan F., Zhang S., Yan Y., Cai Z. An oversampling method of unbalanced data for mechanical fault diagnosis based on meanradius-SMOTE // Sensors. 2022. Vol. 22, N 14. P. 5166. EDN: PURYZO doi: 10.3390/s22145166

25. Zhang K., Zhang Y., Wang M. A unified approach to interpreting model predictions scott // Nips. 2012. Vol. 16, N 3. P. 426–430. doi: 10.48550/arXiv.1705.07874

26. Sikka A., Jain A. Sex determination of mandible: A morphological and morphometric analysis // Int J Contemp Med Res. 2016. Vol. 3, N 7. P. 1869–1872.

27. Smart Energy and Electric Power Systems [Internet]. Pavithra V., Jayalakshmi V. Smart energy and electric power system: Current trends and new intelligent perspectives and introduction to ai and power system. Elsevier, 2023. P. 19–36. doi: 10.1016/B978-0-323-91664-6.00001-2

28. Fan D.P., Zhang J., Xu G., et al. Salient objects in clutter // IEEE trans pattern anal mach intell. 2023. Vol. 45, N 2. P. 2344–2366. doi: 10.1109/TPAMI.2022.3166451

AUTHORS' INFO

*** Ahmad Yudianto,** MD, PhD, Professor;
address: Jl. Prof. DR. Moestopo No.47, Pacar Kembang, Kec. Tambaksari, Surabaya, Jawa Timur 60132, Indonesia;
ORCID: 0000-0003-4754-768X;
e-mail: ahmad-yudianto@fk.unair.ac.id

Vitria Wuri Handayani, MD;
ORCID: 0000-0002-5076-0118;
e-mail: vitriawuri@gmail.com

MAR Mieke Sylvia, MD, PhD, Professor;
ORCID: 0000-0001-8821-0157;
e-mail: mieke-s-m-a-r@fkg.unair.ac.id

Rulaningtyas Riries, MD;
ORCID: 0000-0001-7058-1566;
e-mail: riries-r@fst.unair.ac.id

Muhammad Rasyad Caesarardhi, MD;
ORCID: 0000-0002-6986-0346;
e-mail: mrasyadc@gmail.com

Ramadhan Putra, MD;
ORCID: 0000-0002-0622-3892;
e-mail: ramadhan.hardani@fkg.unair.ac.id

ОБ АВТОРАХ

*** Yudianto Ahmad,** MD, PhD, Professor;
адрес: Jl. Prof. DR. Moestopo No.47, Pacar Kembang, Kec. Tambaksari, Surabaya, Jawa Timur 60132, Indonesia;
ORCID: 0000-0003-4754-768X;
e-mail: ahmad-yudianto@fk.unair.ac.id

Handayani Vitria Wuri, MD;
ORCID: 0000-0002-5076-0118;
e-mail: vitriawuri@gmail.com

Mieke Sylvia MAR, MD, PhD, Professor;
ORCID: 0000-0001-8821-0157;
e-mail: mieke-s-m-a-r@fkg.unair.ac.id

Riries Rulaningtyas, MD;
ORCID: 0000-0001-7058-1566;
e-mail: riries-r@fst.unair.ac.id

Muhammad Rasyad Caesarardhi, MD;
ORCID: 0000-0002-6986-0346;
e-mail: mrasyadc@gmail.com

Putra Ramadhan, MD;
ORCID: 0000-0002-0622-3892;
e-mail: ramadhan.hardani@fkg.unair.ac.id

* Corresponding author / Автор, ответственный за переписку