

Gender Estimation from Panoramic Dental X-ray Images using Deep Convolutional Networks

Ilić, Ivan; Vodanović, Marin; Subašić, Marko

Source / Izvornik: **IEEE EUROCON 2019 - 18th International Conference on Smart Technologies, 2019, 1 - 5**

Conference paper / Rad u zborniku

Publication status / Verzija rada: **Published version / Objavljena verzija rada (izdavačev PDF)**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:127:352991>

Rights / Prava: [Attribution-NonCommercial-NoDerivatives 4.0 International/Imenovanje-Nekomercijalno-Bez prerada 4.0 međunarodna](#)

Download date / Datum preuzimanja: **2025-04-02**



Repository / Repozitorij:

[University of Zagreb School of Dental Medicine Repository](#)



Gender Estimation from Panoramic Dental X-ray Images using Deep Convolutional Networks

Ivan Ilić¹, Marin Vodanović², and Marko Subašić¹

¹Faculty of Electrical Engineering and Computing, University of Zagreb

²School of Dental Medicine, University of Zagreb

Abstract—Current techniques for gender estimation from X-ray images, except being time-consuming, require a highly experienced expert to perform the process. Deep convolutional neural networks have shown to be a very successful technique in many computer vision tasks, mainly because of high accuracy, stability, and processing speed. In this paper, we propose a new method to the gender estimation from panoramic dental X-ray images based on analysis of images using deep convolutional neural networks trained to perform a binary classification. Detailed insight is provided into architecture, hyperparameters and training procedure of our best performing model obtaining an accuracy of 94.3% on a test set. Further experiments have been performed to get a better understanding of anatomical structures which carry the most important information for gender estimation. The presented method requires no special knowledge or equipment to be used, and besides high accuracy, it is also extremely fast with only 18ms of processing time per image on a dedicated GPU.

Index Terms—Deep convolutional networks, image classification, gender estimation, partial identification.

I. INTRODUCTION

In their everyday work, stomatologists estimate many different stomatological parameters about a patient from various information sources. One such source are panoramic X-ray images of teeth and surrounding bone structures. Among many other, information about a person's dental status, age and gender can be obtained from such images. One possible use of such data is in obtaining a partial identification of a living or dead person whose identity is unknown.

In this paper, we present a new approach to gender estimation from panoramic X-ray images of teeth. The usual approach to gender estimation uses different forensic evaluation procedures based on the morphologic characterization of cranial and postcranial bones. It is considered that gender identification with 80% accuracy is possible from the adult skull alone. Accuracy is increasing up to 90% if the lower jaw is included. In case that cranial and postcranial bones like long bones, vertebral bones, scapula, and pelvis are used for gender determination, the accuracy increases above 95%. All of these methods are time-consuming and require an experienced expert (forensic anthropologists, pathologist or similar) familiar with different gender determination techniques [1]. Our approach is based on the application of deep learning techniques for analysis of dental X-ray images, namely deep convolutional neural networks (CNN) which have demonstrated to be very successful at various image analysis tasks. The main advantage

of deep convolutional neural networks is the ability to learn features and patterns in data without any prior knowledge or context, through parameter optimization. On the other hand, training of the network has to be backed up by a sample image database containing all possible image variations. This makes the quality and the volume of the image database a significant factor when designing such systems.

To the best of our knowledge, there is no other published research on deep learning methods to automate gender estimation from panoramic dental X-ray images.

II. CONVOLUTIONAL NETWORKS FOR CLASSIFICATION

In the field of machine learning, deep convolutional neural networks represent a class of artificial neural networks [2] that are most commonly used for image recognition and video analysis, but they perform very well in many other fields such as natural language processing [3] and speech recognition [4]. The classical form of a deep convolutional neural network includes convolutional, max-pooling and fully connected layers. Along with these, our network utilizes batch normalization and dropout layers which are also quite common in CNN based solutions.

One of the problems that deep convolutional networks are very good at solving is image classification [5]. Classification is a task to categorize data samples into a set of discrete classes, based on their characteristics or semantics. A typical example is determining the species of the animal in the image or, given a set of sentences written in different languages, categorize each sentence based on the language it was written in. More formally, classification is a task of finding function f that maps input x to a set of discrete classes y_i , where $i = 1, \dots, n$ and n is the number of classes.

Deep CNNs can perform classification with very high accuracy and speed by estimating the function f through supervised learning [6]. Usually, when multiple classes are involved, classes are encoded using *one-hot* encoding [7]. The number of neurons in the output layer matches the number of different classes n in data, and each neuron corresponds to one class. For classification, a *softmax* [7] activation function is usually utilized to obtain probabilities in the output layer. Other network parameters like the number of layers, types of layers and their parameters have to be adapted to the problem at hand to obtain optimal performance, but some of the parameters can only be optimized through extensive

experiments. Once the network architecture is determined, the appropriately chosen training set is used to train the network. The training procedure has some additional parameters like loss function and a training algorithm that have to be selected carefully. After the training process is finished, we can feed new images through the network to obtain a class estimation for new images.

To estimate the gender, we have built, trained and evaluated multiple deep convolutional neural networks. The architecture of our best performing model is shown in Figure 1. To get the best results we utilize transfer learning through pretrained VGG16 convolutional neural network [8], which we use as a feature extractor.

III. IMAGE DATASET

To train our deep convolutional neural models, we use the dataset containing 4155 images provided by School of Dental Medicine, University of Zagreb. All images were in *gray-scale* JPG format with dimensions of 3256 by 1536 pixels, 8 bits per pixel, with known patient age and gender for each image. Images were recorded for the regular dentist examinations through a period of several years using various orthopantomograph devices, and the database was collected for a different study. All images were anonymized. Unfortunately, the age distribution of our image dataset is unbalanced and is shown in Figure 2. To overcome that problem, we have used popular strategies that will be discussed later. Female gender takes up 60% of our image dataset, while male gender takes up the remaining 40%.

Examples of images from the dataset are shown in Figure 3. Subfigures 3a and 3c present panoramic X-ray images of two males, and subfigures 3b and 3d present panoramic X-ray images of two females. As expected, there are no strong and easy to spot visual features that can be used to distinguish two genders. Various specific features that vary among genders and are visible in the panoramic X-ray images are mentioned in the forensic literature like shape parameters and structural differences of the lower jaw in [9], though the intra-gender variations are very small and the focus of the paper is on the age estimation. Another paper [10] reports that a specific angle of the lower jaw differs among genders, but the difference is small and depends on the age. However, a more recent study [11] argues that a shape feature of the lower jaw that is commonly used for the gender estimation cannot be used reliably for the purpose. These findings suggest that there is no easy way to spot the clues for gender estimation from panoramic X-ray images of teeth. Furthermore, various age-related conditions like fillings, crowns, implants and missing teeth are present in our database, and our gender estimation should function regardless of these conditions. Some of these conditions are visible in subfigures 3c and 3d which come from two elderly patients.

IV. RESULTS AND DISCUSSION

Since our dataset is fairly small and, as we mentioned earlier, age-wise skewed we decide to use a stratified 10-fold

cross-validation technique to ensure that we obtain unbiased results. We use stratification with respect to the age of the patients to ensure equal distribution of age across all train and test splits. Each of the 10 splits contains 3740 and 415 images in the train and the test set, respectively. Input images are non-uniformly scaled to the new dimensions of 512 by 512 pixels to reduce complexity and training time, but still retain important details in the images. Following network hyperparameters have been used: batch size 25, *sigmoid* activation function in output layer and *relu* activation function in every other layer, categorical cross-entropy loss, Adam optimizer [12] with initial learning rate set to 10^{-4} . To avoid overfitting, we use batch normalization layer [13] together with dropout layer after the dense layer. We trained the model on each split for 75 epochs on an Nvidia TitanXp GPU. The code was written in Python using the Keras [14] library.

After training, the model estimates gender with an average accuracy of 94.3% on the test set. Averaged accuracies for age groups are shown in Table I. The model predicts gender with higher accuracy for patients that are between 20 and 60 years of age, which was expected due to the age distribution shown in Figure 2.

TABLE I: Experimental results for age groups

Age group	Accuracy
Less than 20 years old	92.6%
Between 20 and 30 years old	94.1%
Between 30 and 40 years old	95.0%
Between 40 and 50 years old	96.6%
Between 50 and 60 years old	95.4%
Between 60 and 70 years old	93.5%
Between 70 and 80 years old	91.4%
Above 80 years old	50.0%

We were interested in finding out if any specific anatomical structures contribute to the successful gender estimation, so we decide to run four additional experiments to gain a better understanding of how our model infers gender from images. In each experiment, we generate the new test set from an existing test set by replacing a specific part of the image with random noise with a uniform distribution over the closed interval with the minimum of 25 and maximum of 230. Then we test our already trained model on these four versions of the original test set to see how accuracy changes depending on which part of the image is masked by noise. This can provide us with some idea whether the masked region is an essential source of information for gender estimation.

In the first experiment, we apply noise over the teeth region in each image. In the second experiment, we increase the region from the previous experiment by one-third of the leftover space, in each direction. In the third experiment, we calculate the new mask region as the negative of the region from the first experiment, and we apply noise again. Finally, in the fourth experiment we do the same as in the third, but this time we use the inverted mask from the second experiment. For each experiment, mask size and position are kept constant. Examples of images for these four scenarios are shown in

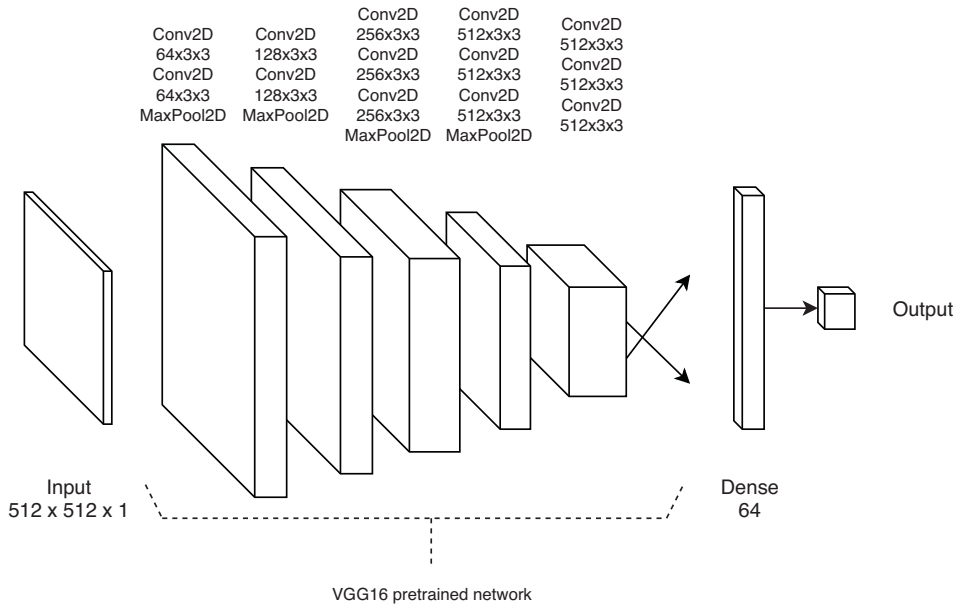


Fig. 1: Neural network model architecture

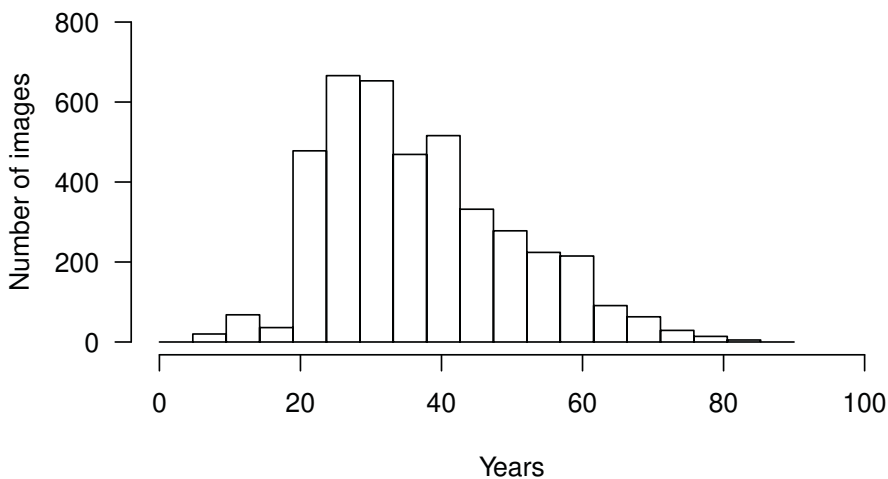


Fig. 2: Age distribution histogram of the database

Figure 4, and test results for these four scenarios are shown in Table II.

TABLE II: Results for masking experiments

Experiment	Accuracy
First experiment	86.0%
Second experiment	77.8%
Third experiment	80.6%
Fourth experiment	83.9%

The accuracy drop was most significant when the area sur-

rounding the teeth was masked out. This indicates that teeth are probably not a critical part of the image for gender estimation by our model, but the region surrounding the teeth might carry more important information. This is in accordance with the anthropological approach to gender determination. The most common osteological features of skull which distinguish male skull from female skull are the skull size, size of muscular ridges on cranial bones, size of frontal sinuses and supraorbital ridges, roundness of the upper margin of the orbit, size of the palate, robustness of the lower jaw, gonial angle etc. [15]. Size



(a) Male



(b) Female

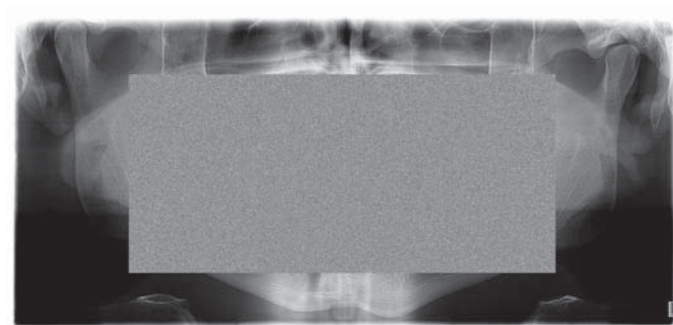


(c) Male

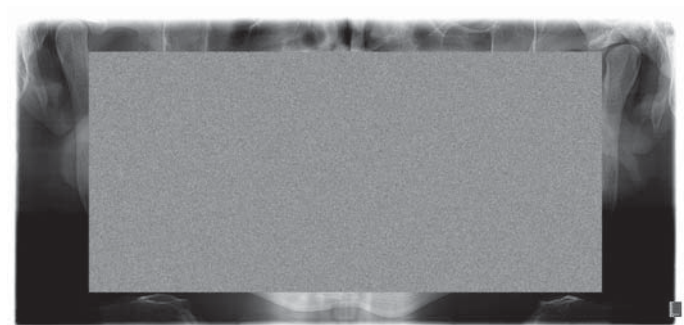


(d) Female

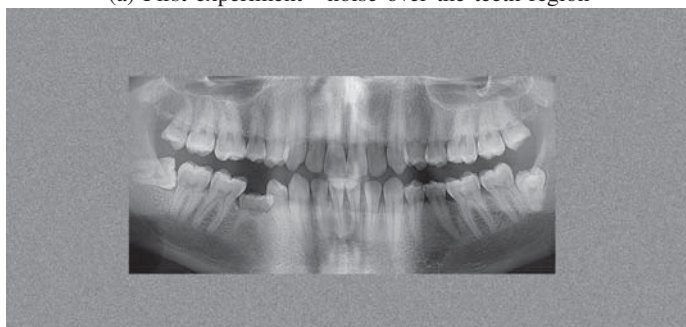
Fig. 3: Panoramic X-ray image examples



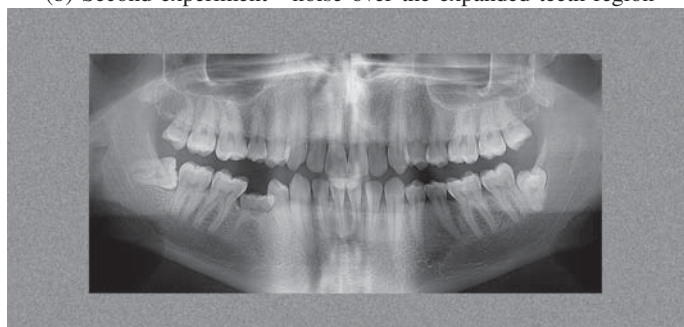
(a) First experiment - noise over the teeth region



(b) Second experiment - noise over the expanded teeth region



(c) Third experiment - noise over the inverted teeth region



(d) Fourth experiment - noise over the inverted expanded teeth region

Fig. 4: Image masking examples

of the teeth is of less importance for gender determination, especially if panoramic dental X-ray images are used. Some morphological and histological characteristics of teeth show some level of gender dimorphism, but they are also useless for gender determination based on panoramic X-ray images.

Memory requirements and execution time are crucial practical aspects of each method. On average, it takes 18 milliseconds for our model to process a single image. The model uses approximately 31.5 million parameters which use around 120 megabytes of GPU memory.

V. CONCLUSION

Our experimental results demonstrate that gender can be estimated from panoramic dental X-ray images using deep convolutional neural networks with accuracy exceeding the current state of the art. Along with achieved high accuracy, our approach is high-speed due to relatively simple network architecture. From the applicative point of view, our method requires no special equipment or knowledge to be utilized. Our experiments aimed at identifying image regions that provide useful data for gender estimation suggest the teeth themselves are less important for the estimation. It seems that, like in the osteological-anthropological approach, cranial bones and structures visible on the panoramic X-ray like maxilla, mandible, sinuses, and vertebrae are more important for our model.

ACKNOWLEDGMENT

This research has been supported by the European Regional Development Fund under the grant KK.01.1.1.01.0009 (DAT-ACROSS). We also gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

REFERENCES

- [1] M. İşcan and K. Kennedy, *Reconstruction of life from the skeleton*. Liss, 1989. [Online]. Available: https://books.google.hr/books?id=0_jaAAAAMAAJ

- [2] R. J. Schalkoff, *Artificial neural networks*. McGraw-Hill New York, 1997, vol. 1.
- [3] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proceedings of the 25th international conference on Machine learning*. ACM, 2008, pp. 160–167.
- [4] O. Abdel-Hamid, A.-r. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional neural networks for speech recognition," *IEEE/ACM Transactions on audio, speech, and language processing*, vol. 22, no. 10, pp. 1533–1545, 2014.
- [5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [6] S. B. Kotsiantis, I. Zaharakis, and P. Pintelas, "Supervised machine learning: A review of classification techniques," *Emerging artificial intelligence applications in computer engineering*, vol. 160, pp. 3–24, 2007.
- [7] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006.
- [8] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [9] D. Bhardwaj, "Radiographic evaluation of mandible to predict the gender and age," *JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH*, 2014. [Online]. Available: <https://doi.org/10.7860/jcdr/2014/9497.5045>
- [10] R. Upadhyay, J. Upadhyay, P. Agrawal, and N. Rao, "Analysis of gonial angle in relation to age, gender, and dentition status by radiological and anthropometric methods," *Journal of Forensic Dental Sciences*, vol. 4, no. 1, pp. 29–33, 2012. [Online]. Available: <http://www.jfds.org/article.asp?issn=0975-1475;year=2012;volume=4;issue=1;spage=29;epage=33;aulast=Upadhyay;t=6>
- [11] A. Peregrina, S. Azer, E. E. Tao, and W. Johnston, "Radiographic blind test of curvature of the posterior border of the mandibular ramus as a morphological indicator of gender: Mandibular ramus curvature as indicator of gender," *Journal of Prosthodontics*, vol. 25, 06 2016.
- [12] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
- [13] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.
- [14] F. Chollet *et al.*, "Keras," <https://keras.io>, 2015.
- [15] D. R. Brothwell, *Digging up bones: the excavation, treatment and study of human skeletal remains*, 3rd ed. Cornell University Press, 1981.