birpublications.org/dmfr

RESEARCH ARTICLE A fully automated method of human identification based on dental panoramic radiographs using a convolutional neural network

Young Hyun Kim, Eun-Gyu Ha, Kug Jin Jeon, Chena Lee and Sang-Sun Han

Department of Oral and Maxillofacial Radiology, Yonsei University College of Dentistry, Seoul, South Korea

Objectives: This study aimed to develop a fully automated human identification method based on a convolutional neural network (CNN) with a large-scale dental panoramic radiograph (DPR) data set.

Methods: In total, 2760 DPRs from 746 subjects who had 2–17 DPRs with various changes in image characteristics due to various dental treatments (tooth extraction, oral surgery, prosthetics, orthodontics, or tooth development) were collected. The test data set included the latest DPR of each subject (746 images) and the other DPRs (2014 images) were used for model training. A modified VGG16 model with two fully connected layers was applied for human identification. The proposed model was evaluated with rank-1, -3, and -5 accuracies, running time, and gradient-weighted class activation mapping (Grad-CAM)-applied images. **Results:** This model had rank-1, -3, and -5 accuracies of 82.84%, 89.14%, and 92.23%, respectively. All rank-1 accuracy values of the proposed model were above 80% regardless of changes in image characteristics. The average running time to train the proposed model was 60.9 s per epoch, and the prediction time for 746 test DPRs was short (3.2 s/image). The Grad-CAM technique verified that the model automatically identified humans by focusing on identifiable dental information.

Conclusion: The proposed model showed good performance in fully automatic human identification despite differing image characteristics of DPRs acquired from the same patients. Our model is expected to assist in the fast and accurate identification by experts by comparing large amounts of images and proposing identification candidates at high speed. *Dentomaxillofacial Radiology* (2022) **51**, 20210383. doi: 10.1259/dmfr.20210383

Cite this article as: Kim YH, Ha E-G, Jeon KJ, Lee C, Han S-S. A fully automated method of human identification based on dental panoramic radiographs using a convolutional neural network. *Dentomaxillofac Radiol* (2022) 10.1259/dmfr.20210383.

Keywords: Artificial intelligence; Deep learning; Forensic dentistry; Human identification; Panoramic radiography

Introduction

Dental panoramic radiography (DPR) is widely used for dental diagnosis, planning, treatment, and evaluation.¹ A single DPR contains anatomical information

Correspondence to: Sang-Sun Han, E-mail: sshan@yuhs.ac

on the oral and maxillofacial region, as well as an individual's dental treatment history,² such as prostheses, endodontic treatment, tooth extraction, orthodontic treatment, or dental surgery; this information can play an important role in the human identification process.^{3,4} In the forensic science field, ante- and post-mortem DPRs have been compared in criminal cases and in the aftermath of large-scale disasters.⁵

Previous investigators have manually compared anteand post-mortem panoramic radiographs to identify

Received 24 August 2021; revised 19 November 2021; accepted 19 November 2021; published online 02 December 2021

The authors Young Hyun Kim and Eun-Gyu Ha contributed equally to the work.

Young Hyun Kim and Eun-Gyu Ha have contributed equally to this study and should be considered as co-first authors.

matching dental features.⁶ However, manual comparisons are labor- and time-intensive.⁷ Furthermore, expert experience, training duration, and education have been reported as factors affecting identification accuracy.⁶ DPRs are taken as the X-ray source rotates around the subject, making overlapping anatomical features or blurring unavoidable.⁸ These geometric limitations can be exacerbated by the panoramic machines and head position of the patient,⁸ often leading to inconsistent results depending on the skill of the operator who manually compares DPRs.

The development of automated methods for human identification is being attempted with the goal of overcoming these limitations.⁹ Before the advent of artificial intelligence (AI), various computer-aided identification methods using comparisons between DPRs have shown high identification performance, but still require manual processing.^{4,5,10–14} As the recent trend in imagebased AI research, a few studies have introduced DPRbased automatic human identification models using AI algorithms.^{15,16} These previous researches reported that AI methods have a potential as efficient automation methods vs the cumbersome manual processes, but the methods were developed using only a small data set or some manual work was still involved.

This study aimed to propose a fully automated human identification method using a convolutional neural network (CNN) with a large-scale DPR data set. The CNN model automatically identifies the same person by comparing the latest acquired DPR and other DPRs without any manual processes. This model provides complete automation of the entire identification process, and has been validated on the basis of large-scale DPRs reflecting various clinical situations without any constraints.

Methods and materials

This study was approved by the Institutional Review Board of Yonsei University Dental Hospital (no. 2-2020-0050), and the requirement for informed consent was waived by this committee due to the retrospective nature of the study. All aspects of the study workflow were conducted in accordance with its relevant guidelines and regulations (Declaration of Helsinki). All DPRs were taken from January 2016 to April 2020 and were downloaded anonymously from the picture archiving and communication system of Yonsei University Dental Hospital to protect personal information.

Data preparation

A total of 2760 DPRs from 746 subjects (386 females and 360 males; age range, 13–89 years) were collected for this study. The number of DPRs per person ranged from a minimum of 2 to a maximum of 17, all with different acquisition times. Panoramic images with different image characteristics in various DPRs of the

Dentomaxillofac Radiol, 51, 20210383

 Table 1
 Differences in image characteristics between dental panoramic radiographs collected from the same subjects

Number of subjectsa
252
35
355
164
16
146

^aDuplicate subjects were counted when one or more image characteristics had changed between the training and test images of a single subject.

same patient were used due to oral condition changes. The data with changes in the oral condition were grouped into five subtypes according to the type of treatment: tooth extraction, oral surgery (resection of maxilla or mandible, fracture reduction, orthognathic surgery), prosthetics (implantation, crown or bridge restoration), orthodontics, or different tooth development stages (Table 1). All DPRs were obtained with five panoramic machines from four different manufacturers: a Cranex 3 + Ceph panoramic apparatus (Soredex Co, Helsinki, Finland), FCR XG5000 cassette reader (Fuji film Co, Tokyo, Japan), RAYSCAN α + (Ray Co. Ltd, Hwaseong-si, Korea), and PaX-i and PaX-i3D Green (Vatech Co., Hwaseong, Korea).

The overall workflow of the proposed method for fully automated human identification, the configuration of the data set, and examples with different image characteristics are shown in Figure 1. To simulate the DPR-based human identification process, the test data set included the latest DPR of each subject (746 images), and the other DPRs (2014 images) were used as the training data set of the model. The training data set was augmented using various methods, such as zooming, rotation, shearing, brightness, and contrast, to improve the performance and generalization of the model (Figure 2). As a result, the training data set (2014 images) was augmented six-fold (one original and five augmented images of each DPR), resulting in a total of 12.084 DPRs that were used for model training. The image brightness and contrast levels were adjusted by an oral radiologist to emphasize the high density of dental prostheses.

Proposed model

To develop a fully automated human identification model, we used the VGG16 algorithm pre-trained with ImageNet, a huge data set spanning 1000 classes and containing over 1 million images. Our model started training with the parameters of the pre-trained VGG16 to overcome the local minimum problem induced by the small dataset.¹⁷ The original VGG16 model consisted of 13 convolutional layers and three fully connected (FC) layers, but considering our data set, we only constructed



Figure 1 Design of this study. Multiple DPRs with different image characteristics were collected as a data set. The white numbers in the lower right corner of each DPR indicate the acquisition date (yymmdd). The proposed human identification model was developed using the VGG16 algorithm, and the Grad-CAM technique was applied to visualize where the model focused for prediction DPR, dental panoramic radiograph.

two FC layers after the last convolutional layer. The rectified linear unit (ReLU) was used as the activation function for 14 layers (13 convolutional layers and 1 FC layer) except the last FC layer. L2 regularization and the

dropout method were applied to make the cost function less affected by local noise and to reduce overfitting.

Each DPR was automatically resized to 512 (width)

× 256 (height) pixels before being input into the



Figure 2 Examples of original and augmented dental panoramic radiographs.

birpublications.org/dmfr

proposed model, and all pixel values of the input image were rescaled from 0 to 255 to a range from 0 to 1 to improve processing speed. The network was trained by extracting features from input images and provided a confidence score, which refers to the probability that the input image belongs to one patient's image data. If the panoramic radiograph with the highest confidence score belonged to the actual corresponding patient in the input data, it was considered to have been correctly identified. In order to find suitable parameters for achieving high performance of the model, we conducted hands-on tuning while changing the initial learning rate (0.0002–0.000002) and batch size (32, 64, 128) in various combinations. An initial learning rate of 0.00002 and batch size of 128 led to the best model performance.

We used an open-source deep learning library (TensorFlow 2.0) on Windows 10 to implement this CNN algorithm. The computing hardware included an AMD EPYC 8-core processor and NVIDIA Titan RTX.

Evaluation of the proposed model

Rank-1, -3, and -5 accuracies were computed for the proposed model. Rank-n accuracy is mainly used when there are many class labels, and the DPR data set in this study consisted of 746 classes. Rank-*n* accuracy is defined as the percentage of finding the *n*th matching target DPR among the training DPRs when the test DPR is input to the model.^{15,18} For instance, rank-5 accuracy is the percentage of cases in which the correct subject is included when the proposed model predicts the top five most likely subjects. The accuracy of the presented model according to differences in image characteristics was evaluated only for the rank-1 values. The running time performance was measured as the average training and test time for the proposed model.

For a visual evaluation of the model, the gradientweighted class activation mapping (Grad-CAM) technique was applied to the DPRs. The Grad-CAM technique visualizes the region of interest that influences the prediction of the CNN model.¹⁹ Therefore, applying the Grad-CAM technique allows a better understanding of the output results.

Results

Table 2 indicates the rank-1, -3 and -5 accuracies of the proposed model according to whether augmentation was applied. The model with training data augmentation had a rank-1 accuracy of 82.84%, which was higher

 Table 2
 Objective performance evaluation metrics of the test results according to data augmentation

Accuracy	Without augmentation	With augmentation
Rank-1	80.56%	82.84%
Rank-3	86.86%	89.14%
Rank-5	88.61%	92.23%

 Table 3
 Rank-1 accuracy of the model according to image characteristics

Rank-1 accuracy (Proposed model)
80.16%
82.86%
85.92%
84.76%
81.25%
81.51%

than the accuracy of the model without augmentation (80.56%). Of a total of 746 subjects, 688 were successfully identified within 5 candidates (rank-5) with the data augmentation, corresponding to a rate of 92.23%.

All accuracy values from the proposed model were above 80%, regardless of changes in image characteristics (Table 3). DPRs with prosthetics showed the highest identification rank-1 accuracy at 85.92%, and DPRs with tooth extraction had the lowest at 80.16%.

The average running time to train the proposed model using a total of 12,084 DPRs from 746 subjects was 60.9 s per epoch. After training, the prediction time for the 746 test DPRs was short, at 3.2 s per image.

Figure 3 shows the cases correctly identified as the same patient and provides a test DPR and its Grad-CAM image, as well as a matching DPR from the training data set. In the Grad-CAM images, some identifiable features such as an edentulous area, sinus shape, and prostheses were visualized as red and yellow colors. The red color indicates the area that had the greatest impact on the judgment of the proposed model, whereas the blue area indicates the opposite.

Figure 4 shows several cases where the input test DPR was incorrectly identified as a different subject. Although the model did not predict the correct subject, there was a high level of visual similarity between the test DPR and the DPR of the mismatched subject. In the Grad-CAM images of the test DPRs, the red and yellow areas appear to span a slightly wider range than the specific identifiable features shown in Figure 3.

Discussion

Human identification is a difficult, but necessary task in unexpected accidents such as natural disasters or terrorist acts that cause mass casualties.²⁰ Since DPRs are repeatedly taken—with a minimum of one time and a maximum of several times—during dental treatment, human identification is widely performed using anteand post-mortem DPRs.²¹ However, the process of manually comparing multiple DPRs is time-consuming and the expert's skill level may affect the reliability of the results.^{6,22} In addition, these shortcomings can be exacerbated if the victims to be identified are multinational or numerous. Therefore, we developed a fully

CNN model for human identification using DPRs Kim *et al*



Figure 3 Visual evaluation in cases correctly identified as the same subject. Grad-CAM images show the areas of interest indicated by the proposed model in different colors: red areas indicate the areas of highest interest, followed in order by the yellow and blue areas. DPR, dental panoramic radiograph; Grad-CAM, gradient-weighted class activation mapping.

automated AI-based identification model that did not require any manual intervention. The proposed model can present identification candidates in about 3s per image by comparing the input DPR with the trained DPRs in the database. Considering that the accuracy rate was 92.23% within the top five candidates predicted by the model (rank-5), this CNN-based automation model has the potential to help experts make final decisions quickly by reducing the range of candidates to be identified.

In the current study, we constructed a DPR database with a wide range of dental treatment histories, from endodontic treatment of single teeth to dental surgery such as hemi-mandibular resection, to simulate the various changes between ante- and post-mortem DPRs. Differences in the patient's head position and variability in the panoramic machines used to acquire the DPR were also considered. Based on the data set configured as close as possible to the real-world data, we obtained an 82.84% rank-1 identification accuracy for the proposed model. In fact, these values are lower than the accuracy presented in previous studies that proposed CNN-based automatic human identification models: Fan et al¹⁵ proposed DENT-net and reported an accuracy of 85.16% with a test dataset of 173 subjects (499 DPRs), and Matsuda et al¹⁶ reported 100% identification success with a test dataset of 30 subjects (30 DPRs) using fine-tuned VGG16 algorithm. However, we would carefully suggest that the large-scale volume and complexity of the DPR database used in the present study may have affected the identification accuracy.

A single DPR contains individually identifiable features such as anatomical structure and dental treatment status.^{2,3,5} Several researchers have attempted to perform automatic human identification based on DPRs analyzing tooth contours,^{4,10,11} dental codes,^{12,13} and dental work.^{5,14} Heinrich et al²² applied the speeded up robust features computer vision algorithm to find the correspondence points between two ante- and post-mortem DPRs and obtained robust identification results. Application of an appropriate image preprocessing method can improve the accuracy of identification by providing information about various matching points.²² Since prostheses such as implants and crowns are powerful identifiable points,²³ we included brightness and contrast-adjusted images highlighting identifiable features in the training dataset using image augmentation. Through image augmentation, the model achieved better rank-1 accuracy.

In the current study, we used the VGG16 algorithm, which performed well in the ImageNet Large Scale Visual Recognition Challenge.²⁴ This model also recorded the best performance when applied for DPR-based human identification problems compared with six other well-known CNN algorithms.¹⁶ The VGG16 model used in this study was modified by including two FC layers at



Figure 4 Visual evaluations in cases that were incorrectly identified. The left column is the test DPRs input to the proposed model, and the right column is the subject's DPR incorrectly predicted by the proposed model. Grad-CAM images show the areas of interest indicated by the proposed model in different colors: red areas indicate the areas of highest interest, followed in order by the yellow and blue areas. DPR, dental panoramic radiograph; Grad-CAM, gradient-weighted class activation mapping.

the end of the architecture, unlike the original design consisting of three FC layers. The model was modified because the number of images and classes in our DPR data set was smaller than the ImageNet data set used for pre-training the VGG16 model. The proposed model can quickly identify individuals with high accuracy, similar to that reported in previous studies. Therefore, it will be especially useful for identifying victims of mass casualty events or when the specific features of reference DPRs are difficult to compare.

Deep learning models are considered to be a "black box" because the user cannot understand how the model generates its output.²⁵ This means that even if a deep learning model performs well, it might focus on the wrong area to make predictions. Among the several techniques proposed to handle this problem, we applied the Grad-CAM technique to generate visual explanations for the predictions of the proposed model.²⁶ We expected that the regions provided by Grad-CAM would allow us to check whether the proposed model focuses on similar areas 'as humans see'. Through a visual analysis of Grad-CAM images, we confirmed that our model focused on meaningful features in DPR to identify humans, even if the predictions were incorrect. The areas shown in red in Figure 3 had identifiable dental information that deserved focus even if an expert manually compared the DPRs. By comparison, although Figure 4 did not predict the correct subject, both the test and mismatched DPRs may appear to have been acquired from the same subject. Due to the high visual similarity of the two images, red and yellow areas appeared over a wide range within the DPR. Since Grad-CAM visualizes the locations that support the output on the DPR, the observer can obtain some insights into the interpretation of the results.

There are some limitations of the current study. Since DPR is sensitive to image distortion according to the patient's head position,²⁷ verification using actual DPRs obtained from cadavers would be necessary. Although we could not verify the model with real post-mortem DPRs, we constructed and used a data set including various panoramic machines and head positions, resembling

what could occur when taking panoramic images of cadavers. Secondly, we were only able to split the data set into two parts (the training and test data sets), as in a previous study.¹⁶ We are aware of the data partitioning method for deep learning models, as represented by training, validation, and test data sets. However, since the CNN-based architecture we designed focused on finding similar candidates by comparing training and test DPRs, it was not feasible to use the general data set split method. In order to overcome this limitation, we tried to correct data-related bias by excluding images with exactly the same oral conditions among the DPRs from each person. The last limitation of this study is that the proposed model identifies humans by comparing only features extracted from images, so it matches DPRs with highly similar oral conditions to the same subject. However, these hurdles can also arise when experts are faced with cases where they need to identify humans using only image-based information. The proposed model can be an efficient tool to help the expert make a final decision when large-scale human identification is required.

In future work, a human identification model should be developed based on various image data sources such

REFERENCES

- Kantor ML, Slome BA. Efficacy of panoramic radiography in dental diagnosis and treatment planning. J Dent Res 1989; 68: 810–2. doi: https://doi.org/10.1177/00220345890680051101
- Morse DR, Esposito JV, Kessler HP, Gorin R. Age estimation using dental periapical radiographic parameters. A review and comparative study of clinically based and regression models with the operation desert storm victims. *Am J Forensic Med Pathol* 1994; 15: 303–18.
- Krishan K, Kanchan T, Garg AK. Dental evidence in forensic identification - an overview, methodology and present status. *Open Dent J* 2015; 9: 250–6. doi: https://doi.org/10.2174/ 1874210601509010250
- Zhou J, Abdel-Mottaleb M. A content-based system for human identification based on bitewing dental X-ray images. *Pattern Recognit* 2005; 38: 2132–42. doi: https://doi.org/10.1016/j.patcog. 2005.01.011
- Ajaz A, Kathirvelu D. Dental biometrics: computer aided human identification system using the dental panoramic radiographs. 2013 international conference on communication and signal processing: IEEE 2013: 717–21.
- Pinchi V, Norelli G-A, Caputi F, Fassina G, Pradella F, Vincenti C. Dental identification by comparison of antemortem and postmortem dental radiographs: influence of operator qualifications and cognitive bias. *Forensic Sci Int* 2012; 222(1-3): 252–5. doi: https://doi.org/10.1016/j.forsciint.2012.06.015
- Ito K, Nikaido A, Aoki T, Kosuge E, Kawamata R, Kashima I. A dental radiograph recognition system using phase-only correlation for human identification. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* 2008; E91-A: 298–305. doi: https://doi.org/10.1093/ietfec/e91-a. 1.298
- Rondon RHN, Pereira YCL, do Nascimento GC. Common positioning errors in panoramic radiography: a review. *Imaging Sci Dent* 2014; 44: 1–6. doi: https://doi.org/10.5624/isd.2014.44.1.1
- 9. Heo M-S, Kim J-E, Hwang J-J, Han S-S, Kim J-S, Yi W-J, et al. Artificial intelligence in oral and maxillofacial radiology: what is

as DPR, bitewing radiographs, and periapical radiographs. Such an attempt would be useful in the absence of an antemortem DPR or when there is a limit to the imaging modalities that can be acquired post-mortem.

Conclusions

The proposed model showed good performance in fully automatic human identification, even with various DPRs from the same patients that showed different image characteristics due to dental treatment. From a forensic point of view, our model is expected to be a useful tool to assist in rapid and accurate identification when it is necessary to compare large amounts of images by proposing identification candidates at high speed.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1A2C1007508).

currently possible? *Dentomaxillofac Radiol* 2021; **50**: 20200375. doi: https://doi.org/10.1259/dmfr.20200375

- Shamsafar F. A new feature extraction method from dental X-ray images for human identification. 2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP): IEEE 2013: 397– 402.
- Jain AK, Chen H. Matching of dental X-ray images for human identification. *Pattern Recognit* 2004; 37: 1519–32. doi: https:// doi.org/10.1016/j.patcog.2003.12.016
- Yuniarti A, Nugroho AS, Amaliah B, Arifin AZ. Classification and numbering of dental radiographs for an automated human identification system. *TELKOMNIKA* 2012; 10: 137. doi: https:// doi.org/10.12928/telkomnika.v10i1.771
- Nassar DE, Abaza A, Xin Li, Ammar H. Automatic construction of dental charts for postmortem identification. *IEEE Transactions* on Information Forensics and Security 2008; 3: 234–46. doi: https:// doi.org/10.1109/TIFS.2008.922452
- Hofer M, Marana AN. Dental biometrics: human identification based on dental work information. XX Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2007): IEEE 2007: 281–6.
- Fan F, Ke W, Wu W, Tian X, Lyu T, Liu Y, et al. Automatic human identification from panoramic dental radiographs using the convolutional neural network. *Forensic Sci Int* 2020; 314: 110416. doi: https://doi.org/10.1016/j.forsciint.2020.110416
- Matsuda S, Miyamoto T, Yoshimura H, Hasegawa T. Personal identification with orthopantomography using simple convolutional neural networks: a preliminary study. *Sci Rep* 2020; 10: 13559. doi: https://doi.org/10.1038/s41598-020-70474-4
- Hon M, Khan NM. Towards Alzheimer's disease classification through transfer learning. 2017 IEEE International conference on bioinformatics and biomedicine (BIBM): IEEE 2017: 1166–9.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* 2014;.

- Lee K-S, Jung S-K, Ryu J-J, Shin S-W, Choi J. Evaluation of transfer learning with deep convolutional neural networks for screening osteoporosis in dental panoramic radiographs. J Clin Med 2020; 9: 392. doi: https://doi.org/10.3390/jcm9020392
- Graham EAM. Disaster victim identification. Forensic Sci Med Pathol 2006; 2: 203–7. doi: https://doi.org/10.1007/s12024-006-0011-0
- Alqerban A, Storms A-S, Voet M, Fieuws S, Willems G. Early prediction of maxillary canine impaction. *Dentomaxillofac Radiol* 2016; 45: 20150232. doi: https://doi.org/10.1259/dmfr. 20150232
- Heinrich A, Güttler FV, Schenkl S, Wagner R, Teichgräber UK-M. Automatic human identification based on dental X-ray radiographs using computer vision. *Sci Rep* 2020; 10: 1–13. doi: https:// doi.org/10.1038/s41598-020-60817-6
- 23. Heinrich A, Güttler F, Wendt S, Schenkl S, Hubig M, Wagner R, et al. Forensic odontology: automatic identification of persons comparing antemortem and postmortem panoramic radiographs

using computer vision. *Rofo* 2018; **190**: 1152–8. doi: https://doi. org/10.1055/a-0632-4744

- Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging* 2018; 9: 611–29. doi: https://doi.org/10.1007/ s13244-018-0639-9
- McGovern A, Lagerquist R, John Gagne D, Jergensen GE, Elmore KL, Homeyer CR, et al. Making the black box more transparent: understanding the physical implications of machine learning. *Bull Am Meteorol Soc* 2019; 100: 2175–99. doi: https:// doi.org/10.1175/BAMS-D-18-0195.1
- 26. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE international conference on computer vision* 2017: 618–26.
- Tronje G. Image distortion in rotational panoramic radiography: V. Object morphology; inner structures. *Acta Radiol Diagn* 1982; 23: 59–62. doi: https://doi.org/10.1177/028418518202300109